

# Exploring Whether a Difference is a Meaningful Difference

Tim Jacobbe, Southern Methodist University

Chris Franklin, University of Georgia

Gary Kader, Appalachian State University

Kaycie Maddox, Northeast Georgia Regional Educational Service Agency

Published: March 2023

## Overview of Lesson

This investigation focuses on students conducting a comparative experiment to explore whether there are meaningful differences between the number of times people can write their name with their dominant hand and the number of times people can write their name with their non-dominant hand. This lesson explores what might be considered a meaningful difference when comparing two groups.

Students will be randomly assigned to one of two groups. The first group will be asked to write their name as many times as they can with their dominant hand. The second group will be asked to write their name as many times as they can with their non-dominant hand. Students will collect data about the number of times names are written by each member of the two groups. They will display the data in comparative boxplots. Analysis of the data will include data displays and calculations of measures of center and spread so comparisons between the two groups can be made.

## Type of Data

- One categorical variable and one quantitative variable
- Data generated or collected as a class

## Learning Objectives

- CCSSM.6.SP.B.4. Display numerical data in plots on a number line, including dotplots, histograms, and boxplots.
- CCSSM.6.SP.B.5. Summarize numerical data sets in relation to their context, such as by the following:
  - Reporting the number of observations
  - Describing the nature of the attribute under investigation, including how it was measured and its units of measurement
  - Giving quantitative measures of center (median and/or mean) and variability (interquartile range and/or mean absolute deviation), as well as describing any overall pattern and any striking deviations from the overall pattern with reference to the context in which the data were gathered
  - Relating the choice of measures of center and variability to the shape of the data distribution and the context in which the data were gathered
- CCSSM.7.SP.B.3. Informally assess the degree of visual overlap of two numerical data distributions with similar variabilities measuring the difference between the centers by expressing it as a multiple of a measure of variability.

## **Audience**

- This lesson is intended for use with 7<sup>th</sup> grade students as it explores a difference between two groups.
- This lesson would be appropriate for 6<sup>th</sup> grade students to engage in distributional reasoning as well as an introductory lesson for high school students who are beginning to explore the concept of comparing groups as a precursor to formal inference.
- Prerequisites: Prior to this lesson, students should have experience with formulating questions and collecting data. They should be familiar with common summary statistics for distributions such as mean, median, IQR, and the MAD. Students should be familiar with dotplots and boxplots. This lesson can be used as a means to introduce all of these topics with students; however this lesson does not specifically address how to approach teaching these specific concepts.

## **Time Required**

It is expected this lesson will take approximately 60 – 90 minutes.

You may want to consider collecting this data from all of your students a day (or more) before you plan to teach the lesson, combining sections so that the sample size in each group is larger.

## **Technology and Other Materials**

- *Technology:* Statistical software can be used to create graphical displays of quantitative data and calculate summary statistics. Instructions are provided for CODAP but the lesson can be done by hand or with post-it notes to represent the data and create data displays. If you do not use technology for this lesson, then careful attention must be given to the scale of any displays utilized especially when making comparisons.
- Paper and writing utensil for each student
- Stopwatch
- Computer connected to projector to display data and statistical software
- Copy of student handout for each student. This is optional if teachers are comfortable with having students engage with the online tools; however the student handout will provide documentation of student engagement.

## Lesson Plan

This investigation focuses on students conducting a comparative experiment to explore whether there are meaningful differences between the number of times people can write their name with their dominant hand and the number of times people can write their name with their non-dominant hand. This lesson explores what might be considered a meaningful difference when comparing two groups.

### Formulate a Statistical Question

- Ask your students if they are able to do some things better with one hand than the other. Have you ever seen someone that can do something well with both hands? Have you heard of switch hitters in baseball?
- Discuss some reasons why students might be able to do something better with one hand than the other. The following are some possible reasons students may come up with: biological/natural, practice, injury.
- After students have discussed some of their ideas, ask them if there is something we can do to see whether one hand is better or more efficient than the other. They may suggest doing lots of things with one hand and something else with another hand. The context could vary as long as it can be measured in some way for comparison purposes. Some other examples could be throwing and catching a ball as many times as possible or bouncing a ball as many times as possible. This investigation is written with the assumption that the context will be a comparison of the number of times students can write their names in 60 seconds.
  - Is there a meaningful difference in the number of times students can write their name in 60 seconds with their dominant vs. their non-dominant hand?
- Prompt students to respond to question 1 on the student handout, which asks why the statistical question is phrased in terms of dominant and non-dominant hands rather than left and right hands.

### Collect Appropriate Data

- Before collecting data, there are procedures that need to be discussed with your students. It is important that your students are placed randomly into a group, that each student writes their name the same way (e.g., print vs. cursive; full name vs. nickname), and that the same timing standards are used. All students should start and stop at the same time.

Please note: It is preferred that students do not write their names both ways, because the focus of this investigation is the comparison of two groups. The samples would not be independent of one another if students write their names both ways. If they write their names both ways, then the best way to do this investigation would be to have them subtract the number of times they can write their name with their non-dominant hand from the number of times they can write their name with their dominant hand. This would be perfectly fine to address 6<sup>th</sup> grade learning objectives, but not the 7<sup>th</sup> grade objectives we aim to address here. In order to get a larger sample of students, you may want to consider collecting this data from all of your students a day (or more) before you plan to teach the lesson and combine two classes so each

group will have approximately 25 students. This is not necessary to teach the lesson, but it is also helpful to have time to enter data without students present.

- Ask students how the two groups should be formed. Students might suggest that there should be an equal number of boys and girls in each group, and some students will want to make sure the best athletes in class are spread between both groups. However, these designs do not ensure randomness. It is important that the groups are formed in a random manner. Random assignment helps ensure that the two groups are similar in any attributes that might make a difference in writing their names.
- Discuss with your students how you might assign them randomly. One way to select students randomly is to write each of their names on an index card and then, after thoroughly mixing, draw one card at a time from the bag. The student named on the first card is assigned to the Dominant Hand group; the student named on the second card drawn is assigned to the Non-dominant Hand group. Assignment of students continues to alternate until all the names have been drawn. Alternatively, you can number students from 1 to  $x$  and then use a random number generator to select half for each group. Random.org is a great resource to select random numbers. Note: Teachers may want to have the two groups approximately the same size to make the comparison across groups a bit easier; however it should be noted that the groups do not have to be equal sizes in experiments as long as random assignment is used.
- Once all the procedures and groups are established, have all students complete question 2 on the handout to ensure everyone is aware of the data collection procedures.
- Have all students begin writing their name with their assigned hand as many times as possible. Suggested language, “Get Ready; 3, 2, 1, Start; 45 seconds left; 30 seconds left; 15 seconds left; 5, 4, 3, 2, 1; Stop.”
- Have students count the number of times they are able to write their name and put that number at the top of their page. Make sure students also label what group they are in (dominant or non-dominant).
- Collect the class data. You can have the numbers written on the board or entered into a spreadsheet (saved as a .csv file) under the headings “Hand” and “Frequency.” An example is shown in the table below. Question 3 on the student handout also provides a table for recording the class data.

*Note.* These are example data. The lesson expects you to collect your own data with your students. This is provided to help understand what to do with the data you collect.

Number of Times for Dominant Hand Group												
20	26	22	15	20	21	34	21	20	35	27	20	19
Number of Times for Non-Dominant Hand Group												
8	14	11	9	9	12	15	10	6	15	10	8	10

Example Table for .csv file to use in CODAP

Hand	Frequency
Dominant	20
Dominant	26
Dominant	22
Dominant	15
Dominant	20
Dominant	21
Dominant	34
Dominant	21
Dominant	20
Dominant	35
Dominant	27
Dominant	20
Dominant	19
Non-dominant	8
Non-dominant	14
Non-dominant	11
Non-dominant	9
Non-dominant	9
Non-dominant	12
Non-dominant	15
Non-dominant	10
Non-dominant	6
Non-dominant	15
Non-dominant	10
Non-dominant	8
Non-dominant	10

### Analyze the Data

- With the class data displayed on the board, ask your students if they think one group was able to write their names more than the other. Explain to your students that even though it is easy to see the difference with this extreme example it is often difficult to compare groups by just looking at the numbers; it is helpful to organize the data in a graph.
- Students can work individually, in pairs, or groups up to 3 to engage in this portion of the lesson.
- Have your students find the five-number summary for each group. They should know that the data must be first put in order for each group (as in the example below). They should record their answers on question 4 on the student handout.

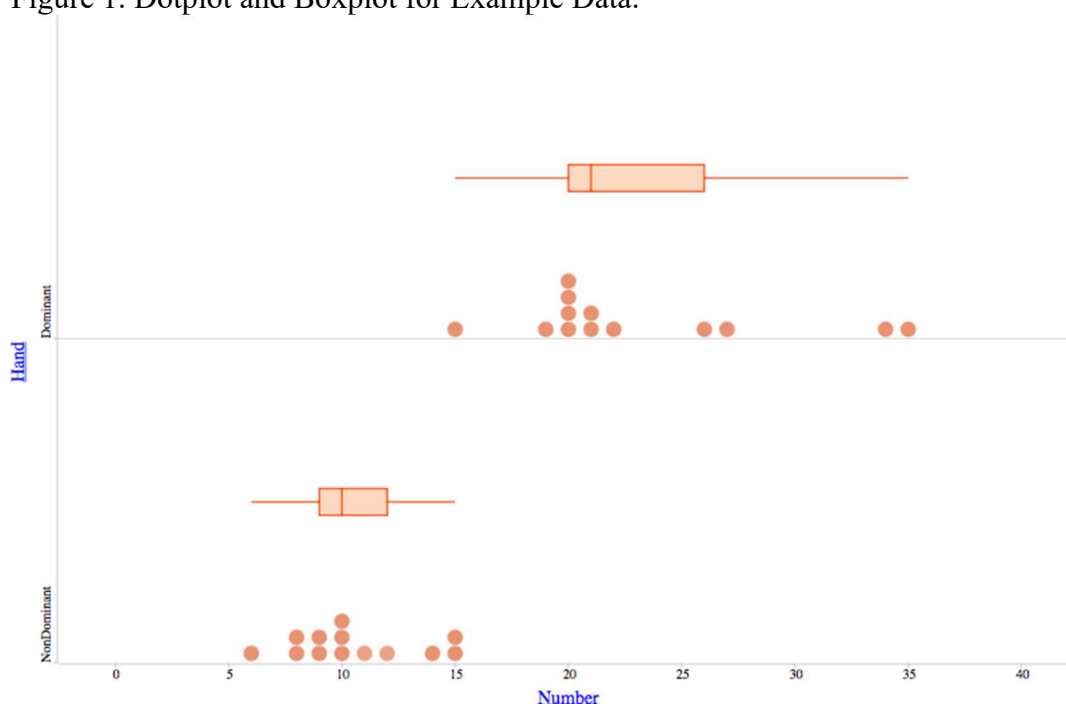
*Note.* These are example data. The lesson expects you to collect your own data with your students. This is provided to help understand what to do with the data you collect.

Number of Times for Dominant Hand Group												
15	19	20	20	20	20	21	21	22	26	27	34	35
Number of Times for Non-Dominant Hand Group												
6	8	8	9	9	10	10	10	11	12	14	15	15

	Minimum	1st Quartile	Median	3rd Quartile	Maximum
Dominant Hand	15	20	21	26	35
Non-Dominant Hand	6	9	10	12	15

- Have students make a boxplot and dotplot of the data on their student handout individually, in pairs, or small groups. There is space to draw the graphs on question 5. The figure below shows a dotplot and boxplot for the example data. Separate instructions on how to make displays in CODAP are given in the attached materials at the end of this lesson.

Figure 1. Dotplot and Boxplot for Example Data.



- Bring the class together for a discussion of the data as well as to ensure they have all the information they need to fill in the median, IQR, mean, and MAD for each group on question 6 of the student handout.

Answers based on the example data are below. It is recommended that you provide these values to your students in this lesson instead of having them calculate the values by hand.

	<b>Median</b>	<b>IQR</b>	<b>Mean</b>	<b>MAD</b>
Dominant Hand	21	6	23.08	4.6
Non-Dominant Hand	10	3	10.54	2.2

- Have students answer questions 7 and 8 on the student handout individually, in pairs, or in small groups. They should write statements that compare center (median of dominant vs median of non-dominant; mean of dominant vs. mean of non-dominant); variability (IQR of dominant vs. IQR of non-dominant; MAD of dominant vs. MAD of non-dominant), and shape (symmetric vs. skewed). (*Note.* Mode is not an appropriate measure of center to use for numerical data. Shape should be determined from the dotplots, because the boxplots only show 5 values.) They will also comment on the overlap and separation between the groups and finally make a judgment about whether the difference is meaningful.
  - Example: The mean for dominant hand (23.08) is greater than the mean for non-dominant hand (10.54). This is a difference of 12.54. The median for dominant hand (21) is greater than the median for non-dominant hand (10). This is a difference of 11. The spread is greater for dominant hand (IQR of 6 and MAD of 4.6) than non-dominant hand (IQR of 3 and MAD of 2.2). This means there is less spread in the non-dominant hand. The dominant-hand distribution seems to be skewed, whereas the non-dominant hand distribution seems bell shaped.
  - For overlap and separation, they should say something like, “100% of the data for the dominant hand is greater than or equal to all of the data for the non-dominant hand.”
- Bring the class together for a discussion to explore question 8, which asks about whether the difference between the two groups is meaningful. Students will likely need to be guided through question 8 on the student handout. The goal of this question is to express the difference between two measures of center as a multiple of a measure of variability. If you want to use the medians, you must use the IQR. If you want to use the means, you must use the MAD. A common rule is that if the difference between two centers is more than 2 multiples of a measure of variability, then it is a meaningful difference. For example, the difference between two medians should be 2 times the IQR in order to be considered meaningful. The difference between two means should be 2 times the MAD in order to be considered meaningful.
  - The difference in medians is 11. The question is “How many IQRs away is the median for the non-dominant hand from the median for the dominant hand? Or “How many IQRs does it take to get from one median to the other?”



- The question is which IQR should you use. For exploration let's use both:
  - Use 6: The medians being 11 apart would mean they are 1 and  $\frac{5}{6}$  IQRs away from each other.
  - Use 3: The medians being 11 apart would mean they are 3 and  $\frac{3}{4}$  IQRs away from each other.
  - In statistics, we try to be more conservative (or careful) when saying things are different. Is it more conservative to say the medians are 1 and  $\frac{5}{6}$  IQRs or 3 and  $\frac{3}{4}$  IQRs apart? It is more conservative to say they are 1 and  $\frac{5}{6}$  IQRs apart so that's why we always use the larger of the two IQRs. This is also true for the MAD.
- The difference in means is 12.54. Being conservative, we will use the larger MAD of 4.6 to see how many MADs apart the means are.
  - Start at the mean for non-dominant, which is 10.54. Add 4.6 repeatedly to see how many times it must be added to get to 23.08.
    - $10.54 + 4.06 = 14.6$
    - $14.6 + 4.06 = 18.66$
    - $18.66 + 4.06 = 22.72$
    - $22.72 + ? = 23.08$ , so it's  $23.08 - 22.72 = 0.36$  (not very much)
    - So, it's about 3 MADs apart. To be precise, you can also figure out how much  $0.36/4.06$  is. This happens to be 0.09 MADs, so technically the means are 3.09 MADs apart.
  - To provide a visualization of this difference, consider displaying the comparative boxplots on a projector:
    - Fold a piece of paper to represent the larger IQR.
    - Start at the median for the non-dominant hand.
    - See how many times you need to "add on" the length of the piece of paper to get to the median of the dominant hand.
    - This should be very close to the calculation.

### Interpret the Results in the Context of the Original Question

- Have your students recall the original question, "Is there a meaningful difference in the number of times students can write their name in 60 seconds with their dominant vs. their non-dominant hand?" Ask your students to write a summary of the experiment that starts with stating an answer to the question and then supporting their answer with their analysis. They should focus their summary on using center and spread measures, but also include a discussion about the shapes of the graphs they drew.
  - The examples in the discussion above provide the answers that should be part of this discussion (answers to questions 7 and 8 on the student handout).
- This question is not on the student handout but should be part of a closing discussion. Have your students describe what they think the distribution of the number of times teachers could write their names with their dominant hand vs. the number of times



teachers could write their names with their non-dominant hand would look like. Do they think it would be different if this experiment were done with teachers?

- They should not expect a different result with teachers because it is reasonable to expect teachers are equally not as comfortable or skilled at writing their names with their non-dominant hands but students should realize that generalizations to another group cannot be made without doing the experiment with that group. They should ask their teachers to do it next 😊.

### Attached Materials

- Student handouts
- Students assessment materials with solutions
- Directions for using CODAP to carry out the data analysis for this investigation

#### Student Assessment Materials with Solutions

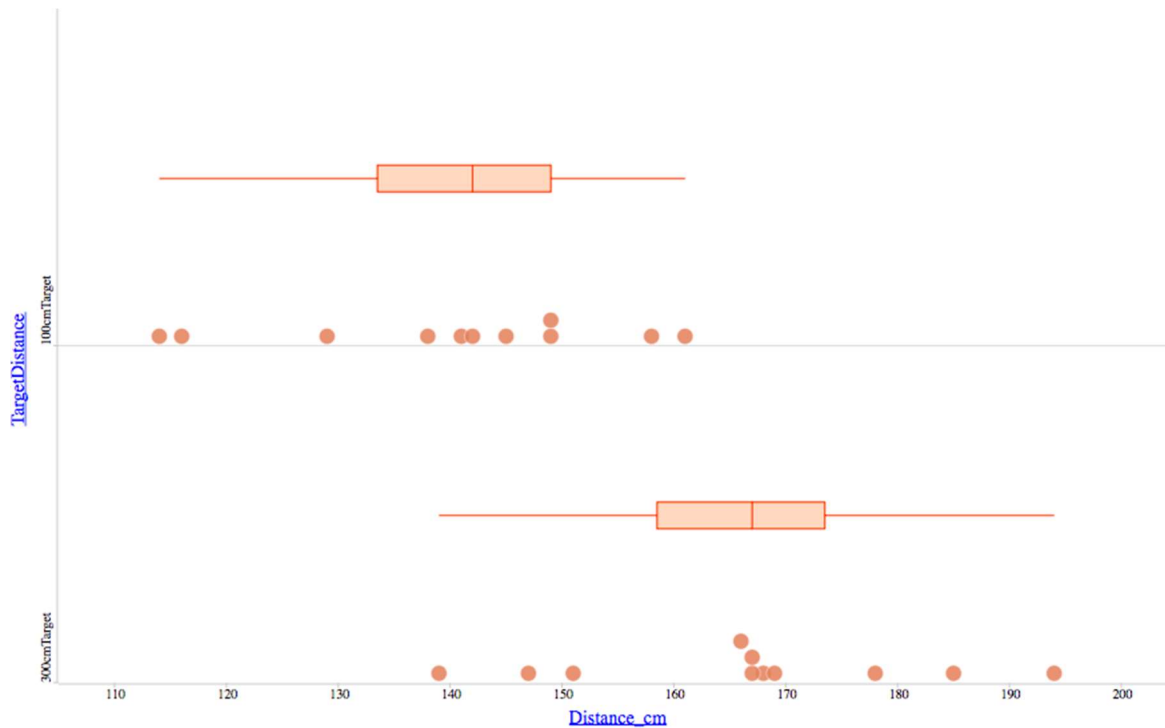
A group of students conducted an experiment to compare the effect of where a target line is placed for the standing long jump. Target lines were placed at 100 cm and 300 cm.

Table 4.1.4. The length of the jumps (in cm) for each group.

100 cm Target	149	141	161	114	116	142	129	149	138	158	145
300 cm Target	168	185	194	167	147	151	169	178	167	166	139

1. Does the distance a target line is from the start line affect the distance students jump in the standing long jump? **Yes, students tended to jump farther when the target line was set at 300 cm.**
2. Use words, numbers, and graphs to justify your answer by using at least one graph, a measure of center, and a measure of spread.

	100 cm target	300 cm target
Mean	140.2	166.5
Minimum	114	139
Q1	129	151
Median	142	167
Q3	149	178
Maximum	161	194
IQR	20	27
MAD	11.6	11.4



Students tended to jump farther when the target line was set at 300 cm than at 100 cm. The mean jumping distance for the 300 cm target was 166.5 cm, while the mean for the 100 cm target was 140.2. The boxplot of the 300 cm target group is shifted much further right than the 100 cm target group. About 75% of the data in the 300 cm target group are greater than about 75% of the 100 cm group.

- Using your answer to question 2, comment on whether you consider there to be a meaningful difference between the distance students tended to jump when the target line was set at 300 cm versus 100 cm.

There seems to be a meaningful difference between the distance students tended to jump when the target line was set at 300 cm versus 100 cm. The difference in the mean jump length was 26.3 cm (166.5 cm - 140.2 cm) which is 2.26 MADs (26.3/11.6).

## Startup Guide to Using CODAP

### How to Enter Data into Spreadsheet so it can be used for comparing groups in CODAP

*Note:* Data can also be entered directly in CODAP. These instructions assume data is entered in a spreadsheet (e.g., in Excel) and that is uploaded to CODAP.

- Enter a variable in the first column. For example, for the dominant-hand investigation, the name of the variable could be “Hand.” CODAP will not recognize spaces in variable names. If your variable has more than 1 word, make sure you do not put spaces or special

characters. If you want to record units, you can use an underscore ( \_ ). For example, if the data were height in cm, you could have the title of the column be Height\_cm.

2. Enter a variable name in the second column.
3. Record each data value with the type of hand (Dominant or Non-Dominant) and the Number of Times that person was able to write their name.
4. Once you finish entering your data, click File, Save As, Put your File Name, \*\*\*Make sure you change the File Format to Comma Separated Values (.csv)\*\*\*. There will be a warning that appears saying some features will not work. Click “Yes” that you want to keep using that format. CODAP will not read the data otherwise.

### **How to Import that Data into CODAP**

1. Visit: <https://codap.concord.org>
2. Click “Launch CODAP” in upper right-hand corner.
3. Click “Create New Document”
4. Click the 3 horizontal lines in the upper left-hand corner.
5. Select “Import”
6. Click in the middle of the box (where it says “Drop file here or click here to select file.”)
7. Select your .csv file you just created.
8. Your data will appear in a table.

### **How to Make Comparative Dotplots/Boxplots in CODAP**

Now that your data is in a table, you can make a graph.

1. Click “Graph” in upper left-hand corner (next to Tables). This will display dots to represent the data in an unorganized manner (this is not a scatterplot).
2. Go to the table and drag the variable “Number” to the horizontal axis. This organizes the data into one, combined dotplot.
3. Go to the table and drag the variable “Hand” to the vertical axis. You now have comparative dotplots.
4. To add comparative boxplots, click the ruler icon on the right-hand side of your graph. Check the boxplot box. You now have comparative boxplots.
5. You can explore the numerical summaries available on CODAP by checking various boxes under the ruler icon as well.

### **References**

Common Online Data Analysis Platform [Computer software]. (2014). Concord, MA: The Concord Consortium.