Introducing Statistical Literacy: A Lesson Plan

Milo Schield, Augsburg University
Schield@Augsburg.edu

ECOTS Challenge: Create an activity to address this topic:
> Statistical and Data Literacy for the Modern Student – what every student needs to know in [their] personal and professional lives.[1]

RESPONSE:
Every student should know how to read and interpret the statistics in the everyday media: the statistics presented in words, tables and graphs. Every student needs to be introduced to statistical literacy.

Statistical literacy studies statistics in the everyday media (how they are constructed and manipulated).

This lesson plan is designed to introduce students to confounder-based statistical literacy.
This is designed for two to four 50 minutes sessions.

The students are either college students or high-school juniors or seniors.
Parts of this can be used with students in the lower grades.

LEARNING OBJECTIVES:

Students will be able to:

1. distinguish statistics from numbers by giving examples.
2. state why statistics are different from numbers.
3. describe (by giving examples) how statistics can be influenced by
    - Confounding: how a related factor can influence a connection or comparison.
    - Assembly: how they are defined, counted, measured, summarized, presented.
    - Randomness
    - Error or bias.
4. explain why "Take care" is good advice in dealing with statistics.
5. relate the stories to the ideas they were used to illustrate
6. connect the four letters in CARE with the four kinds of influence on a statistic.

REQUIRED RESOURCES:
- Coins (quarters preferred): One per student. See page 6.

CONNECTIONS TO POLICIES OR GUIDELINES:
- Review the recommendations in the 2016 update to the GAISE guidelines.[2]
  "Teach statistical thinking". "Give students experience with multivariable thinking"
  Understand the … "principles of statistical design and tools to assess and account for
  the possible impact of other measured and unmeasured confounding variables."

---

[1] https://causeweb.org/cause/ecots/ecots22/activity-contest/submit
[2] https://www.amstat.org/docs/default-source/amstat-documents/gaisecollege_full.pdf

LAUNCH ACTIVITY:

Teacher: Adults who shave their faces tend to be taller than adults who shave their legs.
Does this mean you should shave your face if you want to get taller?
Students: No.

Teacher: What else can explain this comparison?
Students: Gender. Men tend to be taller than women. Men more likely to shave their faces.

Teacher: Welcome to statistical literacy.
Statistical comparisons can be true but still conceal something important.
Statistical comparisons can be true but omit the story behind the statistics.

INTRODUCTION:

Teacher: *Today we are going to learn about statistical literacy.*
Raise your hand if you know what literacy means? *[Listen to answers]*
Raise your hand if you know what a statistic is? [Listen to answers]

Are some statistics controversial? *Does anyone know of any examples?*

Teacher: *Statistical literacy studies:*
Statistics in the everyday media. How social statistics are created and manipulated.

Teacher: *After these classes, you can tell your folks that you learned about statistical literacy.*
You can share some of the stories.
They may be very surprised, because they never studied what you are going to study.

Here is a roadmap of where we are going:

Introducing Statistical Literacy

1. Statistics are different from numbers
2. Statistics are numbers in context (reality)
3. Statistics can be influenced
by Confounders
by Assembly
by Randomness
by Error (bias)
4. "Take CARE" when dealing with statistics.
5. Letters of CARE match the 4 kinds of influence

Look at how many words are new: context, confounder, assembly, random and bias.
We will introduce these new words by using stories. You need to remember the stories.

We will come back to this roadmap after our journey.

>1 Mathematics vs. statistics.

Teacher: *Statistics are different from numbers. Let's see how that works.*
You all know that 1+1 is 2. That is mathematics. Today we are going to study statistics.

Teacher: What is one bunny plus one bunny? [Long pause]
Hold up your hand if you answered 2 bunnies
Hold up your hand if you answered differently than 2 bunnies?
What was your answer? [Asks individual students who held up hands.]
Student answers: *One bunny; the other died or got away.*
Three bunnies (or 2 and a half bunnies); they had a baby.

Teacher: *Here's the next problem. What is one ice-cube plus one ice-cube?*
Hold up your hand if you answered 2 ice cubes.
Hold up your hand if you answered differently than 2 ice-cubes?
Teacher: *What was your answer?* [Asks individual students who held up hands for differently.]
Student answers: *One ice-cube; the other fell on the floor. Zero ice-cubes: they both melted.*
Some fraction between zero and two; they each melted a bit.

Teacher: *Based on these two examples, what have you learned?*
Students: ***Statistics are different from numbers**. Repeat this two more times. .*
Statistics are different from numbers. Statistics are different from numbers.

2. Teacher: *What is 6 plus 7? What is 60 plus 70? What is 60% plus 70%?*
Suppose I sell a product nationwide. I divide the country into two districts: East and West.
Suppose I have a 60% market share on the eastern district; a 70% market share in the western?
What is my market share for the entire country? [Long pause]
Hold up your hand if you think the answer is 130%?
Hold up you hand if you think that answer is wrong?
Teacher: Ask students holding up their hands, *"Why is 130% wrong?"*
Students: *You can't have more than a 100% share of the market.*

Conclusion:

Teacher: *So what have we learned?*
In math, numbers are just numbers. The words don't matter that much.
In statistics, the numbers involve the real world. Statistics are numbers involving reality.

Teacher: *Statistics are different from numbers: **statistics are numbers in context**.*
Repeat after me:
Statistics are numbers in context.
Statistics are numbers in context
Statistics are numbers in context.

#3-Confounding: Statistics can be influenced by Confounding

Teacher: In this activity, let's see how statistics can be influenced by related factors.

Teacher: I want you to study three different cases. All three involve statistical connections.

Teacher #1: *People living in cold climates have noticed an amazing connection.*
As ice cream sales increase, burglaries increase.
Do you see any way that one could cause the other? [Pause]
Is there something else that might cause both to vary? [Pause]
ANSWER: *Weather: More of each in summer; less of each in winter*?

Teacher #2: *Two teams were competing in basketball. They both took the same number of shots.*
One team scored on 20% of its shots; the other team scored on 25% of its shots.
The first team – the 20% team – won the game.
How is this possible? How can the best team lose the game? [Pause]
Assume the teams matched in free-throws taken and made.
ANSWER: *Best team scored mainly on two-point shots (close in).*
The worst team scored mainly on three-point shots (further out).

Teacher #3: *Suppose that the state research/emergency hospital had the highest patient death rate.*
The hospital with the lowest patient death rate in the state was located in a small rural town.
This rural hospital didn't have an on-site staff; it didn't even have an emergency room. [Pause]
Teacher: *How could the best hospital in the state have the highest patient death rate?*
How could one of the poorest hospitals have the lowest patient death rate?
Student: *Best hospital had more patients, so they could have more deaths.*
Teacher: *That's an excellent observation and itis true.*
But there is something in the problem that handles the difference in size?
Can anyone see what word adjusts the deaths for the number of patients?
Student: *The word "rate".*
Teacher: *Yes, the patient death rate is always per patient or per 100 patients.*
The "per patient" takes into account the number of patients involved.
Teacher: *So what else can explain how the best hospital has the highest patient death rate?*
Student: *The sickest patients go to the best hospital.*
ANSWER: *Yes. The sicker the patient, the more likely they are to die.*

Teacher: *How can we tie these three cases together?*
Notice that in each case there was a third factor that caused the confusion.
There is a word in English that indicates confusing: it is 'confound'
You find it often in sports stories. "The pitcher's knuckle ball confounded the batter"
Here's how we can use 'confound' to summarize the three cases:

- *Weather confounded the connection between ice cream sales and burglaries.*
- *Type of shot confounded the connection between team success rate and winning the game.*
- *Patient condition confounded the connection between patient death rate and the hospital.*

Teacher: *We need a name for these third factors that confuse a connection or a comparison.*
Statisticians call these confusing factors "confounders".

Teacher: *So what have we learned about statistics?* [Pause]
Student: **Statistics can be influenced by CONFOUNDING (CONFOUNDERS)**. Repeat three times.

#3 Assembly: Statistics can be influenced by Assembly

Teacher: *In this activity we will see how statistics can influenced by assembly: how they are constructed.*

Teacher: *What is bullying?*
Students: *Someone is being mean. Someone hit someone. Someone hurt someone deliberately.*
Teacher: *Suppose I define bullying as deliberately hitting someone else.*
Suppose I count up how many times students were bullied last week.
Suppose that averages once a day. How many times is that for last week?

Teacher: *How could we define bullying to change the count for last week? [Pause]*
Suppose that I change the definition of bullying. I use a new definition.
Now bulling is deliberately hitting someone else or threatening to hit someone else.
If I look back at that week, could the count of bullying be different?
Hold up you hand if you think it could be more?
How could the new count be more than five?
Students: *You have more ways to bully.*
Teacher: *Hold up your hand if you think it could it be less than five using our new definition?* [Pause]
Student: *Probably not (unless someone made a mistake)*

Teacher: *How could we define bullying so the count for last week was less than five?* [Long pause]
Suppose I define bullying as hitting someone else so hard it marks the skin.
Hold up your hand if you think this new definition could decrease the count?

Teacher: *How could we define bullying so that the count is zero? [Pause]*
Students: *The wound is so bad you have to go to Emergency in an ambulance?*
The wound is so bad, that you die.

Teacher: *So what have we learned about the statistics of bullying?*
Student: *The count depends on how we define bullying. The size of a statistic depends on its definition.*

Teacher: So statistics can be influenced by how they are defined, counted or measured.
Statistics can be influenced by how they are ASSEMBLED.

Teacher: Repeat after me:
Statistics can be influenced by how they are Assembled.
Statistics can be influenced by how they are Assembled.
Statistics can be influenced by how they are Assembled.

#3 Randomness: Statistics can be influenced by Randomness [Coins]

Material: Teacher brings a jar of quarters. More quarters than students.

Teacher: *Hold up your hand if you think some people are luckier than others.*
Students: Some hold up their hands.
Teacher: *Is it possible that someone in this class has special powers?* [Pause]
Teacher: *We are going to find out who that person is right now.*
Students: Get excited about learning something new about one of their members.

Teacher: *I am giving each one of you a coin: a quarter. It has two sides.*
Head is the side that has a face.
Tail is the side that does not have a face.

Teacher: *Take out a sheet of paper and a pencil.*
Flip your coin 10 times.
After each flip record the result, head or tail, in order. You must note the order.
If you use H for Heads and T for tails, you might get H T T H T H H
[Wait while students flip their coin 10 times.]

Teacher: *Now look at your sheet. What is the longest run?*
A run is an unbroken series of heads or tails.
So H T H H would have a longest run of two. T T H T T T H would have a longest run of 3.
Write down the count of the longest run of heads or tails in your data.

Teacher: Writes the numbers 1 through 9 on the board vertically with 1 at the top.
Hold up your hand if you had a longest run of one. Record that number next to one.
Repeat for two, three, etc. Record the # of students for each run length.
Typically there are just one or two students with the longest run in the class.

Teacher: *Do you think getting a run of ___ heads is very likely?*
Student: *No. All the rest of us had shorter runs.*
Teacher: *What could explain these extreme results?* [Long pause]
Student: *Maybe they have special powers?*
Teacher: *What else?* [Much longer pause]
Student: *Maybe it is just a coincidence…*

Teacher: *How would we tell the difference between special powers and coincidence? [Long pause]*
Student: *Repeat the game.*
Teacher: *Yes, let's repeat the game.*
Almost always someone else will have the longest run, indicating it was just a coincidence.

Teacher: *So why did we flip coins?*
Student: To see if someone had special powers?
Teacher: *Yes, but what caused the longest run?*
Student: Chance
Teacher: *Yes, chance (randomness) can influence the connection between the student and their performance.*

Teacher: *Speaking generally,* ***statistics can be influenced by RANDOMNESS****. Repeat three times.*

#3 Randomness: Statistics can be influenced by Randomness [Sports Illustrated]

Teacher: This part of the lesson is on chance: how chance can influence a statistic.

Teacher: *Every year, a sports magazine (Sports Illustrated) features their athlete of the year.*
Readers of this magazine have noticed the following result:
The athlete of the year typically does worse the next year.
What might explain this connection? [Pause]

Student: They get the big head so they don't focus like they used to.
They don't practice as hard.

Teacher: *What else could explain how they became the athlete of the year?*

Student: They worked really hard. They really focused on their sport.

Teacher: *There are a lot of athletes that do really well every year.*
What might explain why a particular athlete did really well?

Students: Silence.
Teacher: *Have you ever had some days where everything goes right?*
Have you ever had some days where everything seems to go wrong?

Teacher: *What might explain being athlete of the year and you have a perfect day?*
Student: Luck? Coincidence?
Teacher: *Yes.*

Teacher: *So if having an exceptional day or an exceptional season is just luck or chance,*
do we expect the next day or season to be equally good or exceptional?
Student: No. If its coincidence it isn't likely to repeat.

Teacher: *So what could explain the Sports Illustrated Jinx?*
Students: Chance.
Teacher: *Statisticians might call it Randomness.*
So, ***statistics can be influenced by RANDOMNESS***

Teacher: Repeat after me:
Statistics can be influenced by Randomness.
Statistics can be influenced by Randomness.
Statistics can be influenced by Randomness.

#3 Error/Bias: Statistics can be influenced by Error or Bias

Teacher: *Let's talk about errors, mistakes or lies can change statistics.*
I am going to use a new word: bias.
Cutting on the bias means cutting cloth at an angle when threads run horizontally or vertically.
Being biased means to be for or against a person, group or idea that is unfair or closed-minded.
Saying that society should ban something (olives, onions) just because you don't like it is biased.
*Let's see how t**hree kinds of bias** can influence the size of a statistic.*

#1: Subject bias:
Teacher: *You all learned about lying at an early age.*
Hold up your hands if you think that adults sometimes lie.
Hold up your hand if you always know when someone is lying.
Hold up your hand if you think there is a fool-proof test to detect lying.

Teacher: *Do you think some people will lie when asked these questions?*
Have you ever stolen something belonging to someone else?
Have you ever used an illegal drug?
Have you ever been arrested?
Have you ever spent a night in jail?

Teacher: *Why would some people say "No" when the correct answer was "Yes"?*
Why would some people say "Yes" when the correct answer was "No"?
Is there any math formula to measure these things? [No]
If we are counting people based on how they answer these very personal questions,
we have good reason to expect the counts to be in error due to **subject bias**.

#2: Researcher bias:
Teacher: *Is there a way to ask a question to encourage someone to give a particular answer?*
How would you do it?
Is there any way to tell how much influence the question has on the answer?

Is there a way to weigh things to make them heavier (lighter)?

*We say these errors are due to **measurement bias.***

#3: Sampling bias:
Can the choice of people surveyed influence the resulting statistics?
Suppose I wanted to see what fraction of students supported having an indoor pool.
What students should I survey to get the highest (lowest) percentage agreeing?
If I survey shoppers leaving Walmart, will a high percentage say they prefer Walmart?
*We say these errors are due to **sampling bias.***

Teacher: Notice that all three of these biased results involve error.

Teacher: repeat after me:
Statistics can be influenced by BIAS OR ERROR.
Statistics can be influenced by BIAS OR ERROR.
Statistics can be influenced by BIAS OR ERROR.

4. Review what we learned

Teacher: *Now let's pull all of this together.*

Teacher: *What is the difference between numbers and statistics?*
Recall the one bunny plus one bunny, or the one ice-cube plus one ice-cube.
Student: Statistics count or measure things in reality where the reality can make a difference.

Teacher: *Remember the cases on ice-cream sales and burglaries, basketball completion percentages and winning, patient death rates and hospital quality. What was the point of these cases?*
Student: A comparison or connection can be influenced by a third factor.
Teacher: *What did we call that third factor?*
Student: A confounder.
Teacher: So? **Statistics can be influenced by CONFOUNDING.**

Teacher: *Remember the exercise on bullying? What was the point of that exercise?*
Student: The count depended on how bullying was defined.
Teacher: ***So? Statistics can be influenced by ASSEMBLY.***

Teacher: *Remember using coins to detect special powers. Remember the Sports Illustrated Jinx.*
What was the point of these two stories?
Student: A connection can be just a coincidence, influenced by chance.
Teacher: *So?* ***Statistics can be influenced by RANDOMNESS.***

Teacher: Remember noticing that people can lie?
Remember noticing that we can weigh things to make them heavier or lighter?
Remember how we could choose people to survey to get the answer we wanted?
Teacher: So? **Statistics can be influenced by BIAS or ERROR.**

Teacher: *Pick just one word to describe each of these four situations:*
Student: **Confounding, Assembly, Randomness and Error.**

Teacher: *What words can we create from these four letters: C, A, R and E?*
Student: CARE, RACE and ACRE.

Teacher: *Using CARE, what advice can we give someone who is dealing with statistics?*
Student: Take CARE.
Teacher: *Repeat this several times!!* ***When dealing with statistics, Take CARE."***
Now we have a way of connecting this good advice with the four kinds of influence on statistics.

Teacher: *And why do we take CARE of statistics – but not of numbers.*
Student: Statistics are numbers in context where the context matters.
Statistics deal with reality where the reality matters.

Teacher: *We have covered a lot of material. Let's see how much you remember.*

Tell me the details of each of these 11 stories;

1. *Bunny math*
2. *Sports Illustrated jinx*
3. *Ice-cream sales and burglaries*
4. *Creating survey questions*
5. *Counting bullying*
6. *Worst basketball team wins.*
7. *Answering sensitive questions*
8. *Ice-cube math*
9. *Best hospital has highest patient death rate*
10. *Longest run of heads or tails*
11. *Asking Walmart shoppers if they like Walmart*

Now can you connect these 11 stories to the statistics vs. numbers and to the four kinds of influence?

1. *What stories connect with the difference between statistics and numbers?*
   Bunny math and ice-cube math

2. *What stories connect with how statistics can be influenced by confounding?*
   Ice-cream sales and burglaries, worst team wins basketball game, and
   best hospital has highest patient death rate.

3. *What stories connect with how statistics can be influenced by assembly?*
   Bullying: how to increase and how to decrease the count.

4. *What stories connect with how statistics can be influenced by randomness?*
   Longest run when flipping coins; Sport Illustrated jinx.

5. *What stories connect with how statistics can be influenced by error/bias?*
   Counting people that answer sensitive questions.
   Influencing people's answers by wording the question.
   Influencing the results by selecting those surveyed.

If you can't recall the connections, this lesson is a failure.

Now that you have recalled the stories and their connections, let's summarize this lesson.

6. Review the stories

Now let's review our journey.
Here is a roadmap of where we have been:

Introducing Statistical Literacy

1. Statistics are different from numbers
2. Statistics are numbers in context (reality)
3. Statistics can be influenced
by Confounders
by Assembly
by Randomness
by Error (bias)
4. "Take CARE" when dealing with statistics.
5. Letters of CARE match the 4 kinds of influence

1. We noticed that statistics are different from numbers.
2. We recognized that statistics are number in context.
   that statistics are the counts and measures or things in the real world (in reality),
3. We recognized that statistics can be influenced.
   We saw this in four different situations:
   Confounding: How a related factor or characteristic can influence a statistic
   Assembly: How definitions can increase or decrease a statistic.
   Randomness: How flipping coins and being the Sports Illustrated pick can involve chance
   Error (bias): How subject bias, researcher bias and sampling bias can influence statistics.

Teacher: With Take CARE, you have a way of thinking about statistics.
Now you can ask four questions about any statistic.
1. Could a statistic be influenced by Confounding?
2. Could a statistic be influenced by Assembly?
3. Could a statistic be influenced by Randomness?
4. Could a statistic be influenced by Error or bias?

These four questions may open the doorway to other questions.
And these questions all remind you that statistics are different from numbers.
Statistics are numbers in context – in reality. And the reality matters.

ACKNOWLEDGEMENT:
Thanks to Travis Weiland (9/22/2022, personal communication) for suggesting the addition of the "Required Resources" and "Connections to Policies or Guidelines" on page 1 and the "Launch Activity" on page 2.