Making the positive foundation for Statistical Learning

Sanjeewani Weerasingha, PhD[1], Steephanson Anthonymuthu, PhD[1]

[1] Department of Statistics, The Ohio State University

Story of the Study

Statistics has proven its status as the most powerful and essential interdisciplinary science because of its reach in all areas of science, engineering, agriculture, business, and education. This reflects there is a growing importance of teaching statistics in a careful and thoughtful manner. It is important to make good and positive impression about the learning of Statistics especially among the students who take Statistic as general education. One thing that all instructors have in common, regardless of who and what they teach in statistics, is having students who are skeptical about the effective usage of Statistics learning (eg., Loannidis, 2005; Shrout1 & Rodgers 2018).

What is it all about?

- As instructors, we routinely get a chance to assess students' understanding of course material through graded assignments and exams.
- However, it is equally important to evaluate the conceptual growth students achieve by the end of the semester.
- Measuring this conceptual shift provides valuable information about how effectively students have generalized the core principles beyond procedure knowledge.
- The incorrect use of statistics often happens because researchers feel persuaded to meet certain statistical criteria for their work to be accepted/published, rather than simply making honest mistakes in applying the methods
- How can our classroom activities motivate and inspire students to overcome such misunderstanding?

PURPOSE – Right usage of Statistics

The objective of this study was to engage students in examining both correct and incorrect applications of statistical methods in scientific research and to evaluate whether such engagement fosters a positive shift toward the proper use of statistics.

PROCEDURE

- A key contributor to this issue is the application of statistical tools by individuals who may not have formal training in statistics.
- At the beginning of the semester, approximately 300 students in an Introductory Statistics class participated in a classroom survey activity designed to gather information on their general feelings toward Statistics.
- Throughout the semester, as students learned various topics and tools in statistical analysis, they engaged in discussions regarding the correct and incorrect application of these concepts to real-world scenarios.
- At the end of the semester, another classroom survey activity was given, to evaluate students conceptual shift towards the correct understanding of the Statistics area.

DATA/TOOLS

To investigate this phenomenon, pre- and post-intervention comments from a randomly selected sample of 45 students were evaluated. Each response was scored on a scale from 1 to 10 and categorized for analysis.

1-3 No change

4-7 Surface to conceptual shift

8-10 Deeper application understanding.

1. Andrew Lang: Most people use statistics like a drunk man uses a lamppost, more for support than illumination.

2. Mark Twain: Facts are stubborn things, but statistics are pliable.

3. Thomas Sowell: One of the first things taught in introductory statistics textbooks is that correlation is not causation. It also one of the first things forgotten.

4. Ernest Rutherford: If your experiment needs a statistician, you need a better experiment.

5. Mark Twain : There are lies, damned lies and statistics.

6. Nenia Campbell: All statistics have outliers.

7. Ron DeLegge: 99 percent of all statistics only tell 49 percent of the story.

8. Edwards Deming: Without data you're just another person with an opinion.

9. Edward Tufte: If the statistics are boring, you've got the wrong numbers.

10. Charles Babbage: Errors using inadequate data are much less than those using no data at all.

11. Karl Pearson: Statistics is the grammar of science.

12. Andrejs Dunkels: It's easy to lie with statistics. It's hard to tell the truth without statistics.

13. Stephen Few: Numbers have an important story to tell. They rely on you to give them a voice.

14. George Box: Statisticians, like artists, have the bad habit of falling in love with their models.

15. Henry Clay: Statistics are no substitute for judgment.

Part I: Please go through the above quotes and find whether you are interested in one or few of them. Please indicate those numbers in your answer.

Part II: Please write a sentence about what you are thinking about Statistics. This can be your own idea or one you choose from the above quotes that you think right, or any statement based on two or three quotes.

Note : We will be returning to your opinion back after sometime in this class and I will ask you to reinvestigate your statement after you learn some topics in Statistics in a future quiz.

This question is related to the Question 08 in Quiz 2. At the beginning of the class, you shared your initial thought about what statistics is idea about what Statistics is. Can you now write about current understanding of Statistics? Are your ideas the same as before, or have they changed?

Sample comment for category 1

Pre:

Out of the quotes listed I really liked Edwards Deming, "Without data you're just another person with an opinion." I am a huge believer in numbers and data in just about everything. Being into sports, numbers and probability were such a huge factor in predicting player stats and teams' success.

Post:

Out of the quotes listed, I really liked Edwards Deming, "Without data you're just another person with an opinion." I referenced how I am a huge believer in numbers and data in just about everything. Being into sports, numbers and probability were such a huge factor in predicting player stats and teams' success. After taking this statistics course, I can confidently say my grasp of stats and overall data/numbers is much better. I strongly agree with the statement above as without data in statistics, it really is just an opinion.

Sample comment for category 3

Pre:

I like this quote as many people use stats ab a sort of buttress? to their believes but not the bigger picture. They do not see the sample sizes or that the roulette table doesn't owe you a win because you lost 1000 times, the odds are the same every time

13.Numbers do not mean much unless we critically think about them. It is humans who give statistics their meaning.

Post:

My ideas have deepened as I have seen truly what statistics is. Before this class, I really just thought of statistics as this dark magic that could explain away anything with the magic of math. Now I see that it is merely the math that we all are familiar with, and the observer is the one that gives those numbers power and meaning. We have created these genius equations and methods to closely follow reality and predict the future, and they are beautiful not because they are magic but because we give them meaning and put them in the correct place.

RESULTS

Wilcoxon signed rank exact test

data: score1

V = 1009, p-value = 3.038e-11

alternative hypothesis: true location is greater than 1

95 percent confidence interval:

2.000004 Inf

$H_0: \Delta = 1 \; VS \; H_1: \Delta > 1$

A one -sided hypothesis testing shows a significant improvement for deeper understanding of the application of statistics

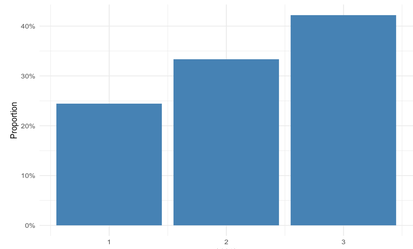

Fig. 1 Improvement of right usage of Statistics at the end of the semester for all students

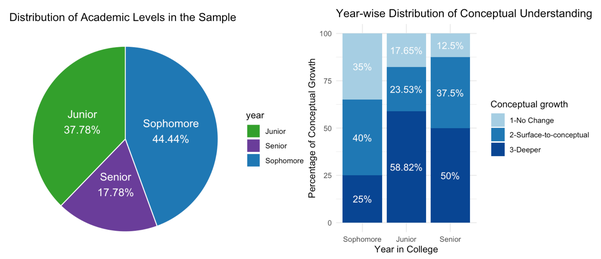

Fig. 2 Class level wise conceptual understanding level.

CONCLUSIONS

- This study indicates that students were able to overcome misconceptions and develop a deeper understanding of the application of statistics.
- Furthermore, the results underscore the importance of fostering a positive perception of statistics learning, particularly among students enrolled in statistics courses as part of their general education requirements.

FUTURE WORK

- The data were obtained through classroom assignments, where students were required to provide responses. As a result, the findings may not fully capture an authentic shift in their understanding.
- For future research, it is recommended to collect data via voluntary surveys rather than assignment-based submissions to enhance the validity and reliability of the results.