

BACKGROUND

What is messy data? Data that has variation, outliers, and unclear trends

What is clean data? Data that has been streamlined to emphasize trends, typically the data seen in educational materials.

Why use case studies? Case studies provide context to data students interact with and allows students to see the many disciplines that use data.

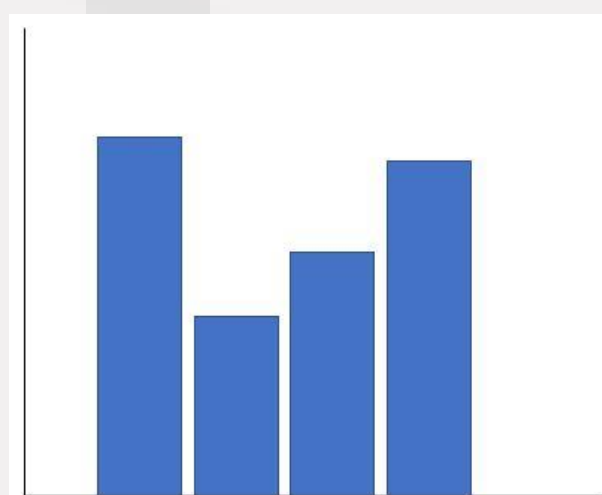
Primary research question - How does the use of messy data affect students understanding of data, as seen in their ability to read and interpret graphs?

METHOD

Case studies were used to interact with data displays. Student data literacy skills were assessed with different display types.

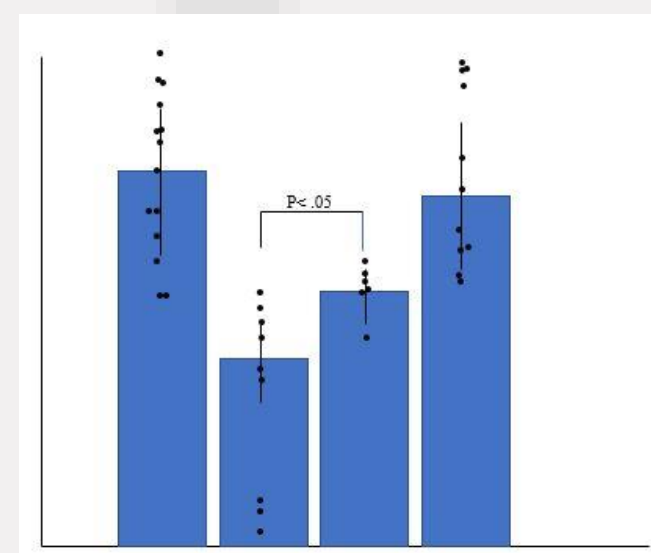
Non-Treatment Group

Displays were altered to be free of messy elements.



Treatment Group

Displays were from original source.



Data Collection Instruments

Likert surveys on student confidence, assessments, and focus groups were used.

RESULTS

Date Literacy Skills

Post-assessment comparison between the two groups showed no significant difference in skill level ($p > .05$).

Growth was seen within non-treatment group for 2 of the 3 displays ($p < .05$) and in treatment group for error bars (2XSE) (Figure 1)

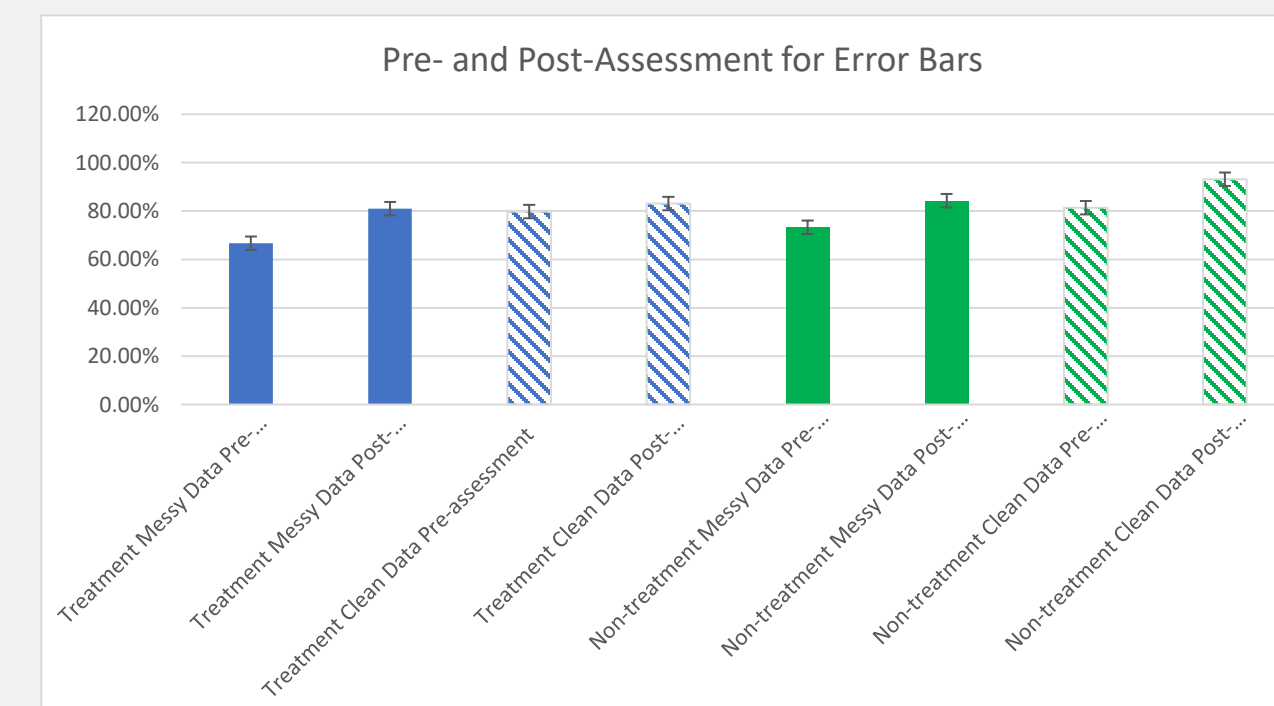


Figure 1: The pre- and post-assessment the focus period on error bars.

Student Confidence

Student confidence in their skills significantly increased in both groups. The greatest growth occurred in areas that pertained to data interpretation.

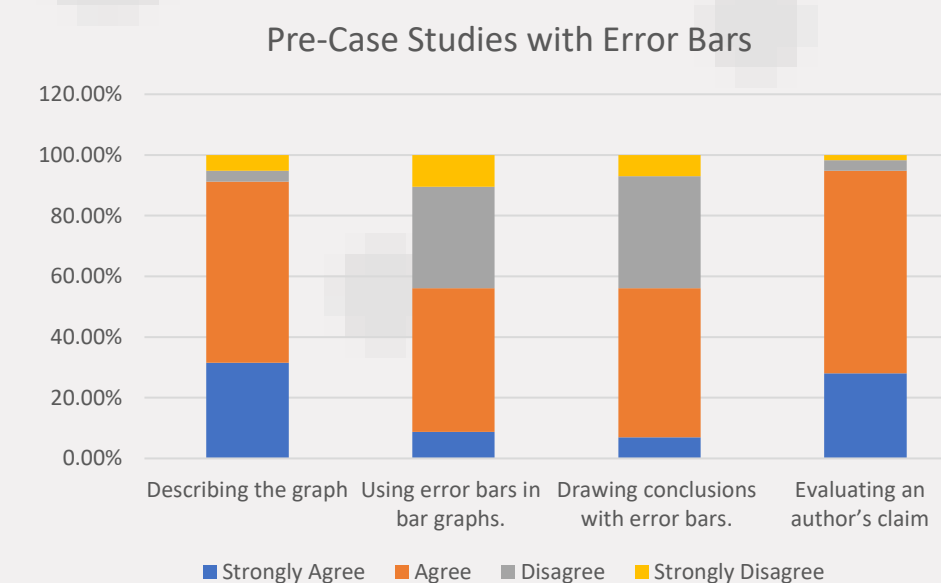
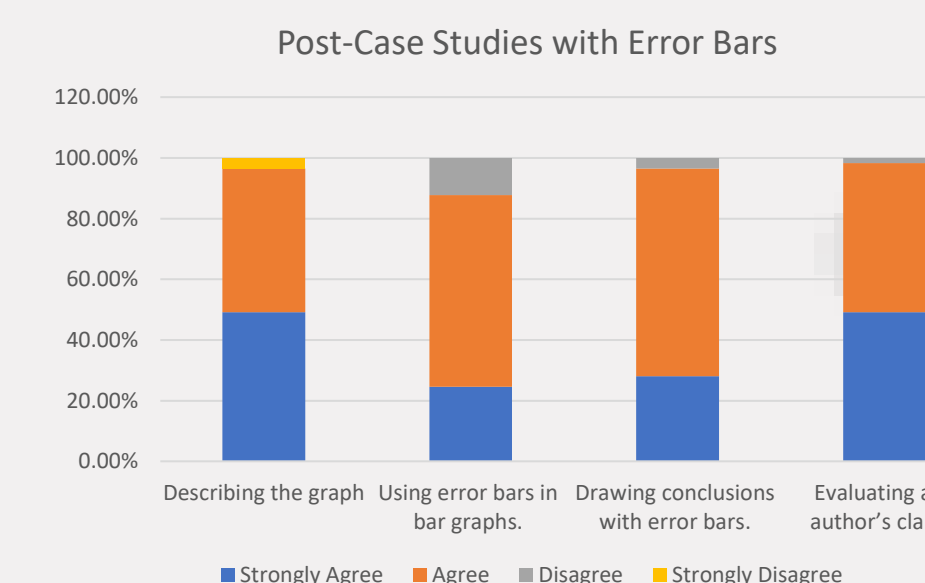


Figure 2: Student confidence with error bars before treatment. Figure 3: Student confidence with errors post treatment, with greatest growth in forming conclusions.



Critical Thinking

Both groups stated that case studies helped them form conclusions about the data. Those in the treatment group mentioned the need to question data. Asking, can data become too clean, removing important information.

CONCLUSIONS

Using Messy Data

Use of messy data didn't impact student data literacy skills, but its exposure may have resulted in students thinking more critically about data.

The treatment group discussed the need to ask about how data is "treated" when displayed. The use of messy data raised awareness that data being presented may not always be in its original form.

Case Studies

The case studies showed the importance of context when interpreting data. Knowing how data was collected, when and for how long, along with information on the subject being studied plays a role in forming conclusions about the data

Acknowledgments

I would like to thank my Research 1 students, Marcie Reuer, and Terrill Patterson