ABSTRACT

Multivariate thinking is an increasingly recommended and important skill for developing statistical thinking. Currently, few studies have explored how students develop multivariate thinking. This study was conducted to learn more about developing this skill particularly when using visualization. It explored the following research questions: *(1) How does students' multivariate thinking develop as they take part in a series of activities designed to introduce and promote reasoning with multiple variables? How do student responses to questions requiring multivariate thinking change throughout the semester? (2) What challenges surrounding multivariate thinking persist after taking part in the intervention? Do any new challenges emerge after the completion of these activities?*

For this study, a unit on multivariate thinking was created for a data visualization course that consisted of ten activities and three assignments, implemented in Fall 2021. The students' responses on assignments were qualitatively analyzed for evidence of multivariate thinking pertaining to seven learning outcomes. Two students were observed from different sections of the course to gain insight into students' multivariate reasoning throughout the unit. Additionally, three students were interviewed at the end of the unit to provide rationale for their answers on the last assignment.

Results indicated that over the course of the multivariate thinking unit, students improved in their ability to create multivariate graphs using R. Overall students' reasoning with multiple variables improved throughout the unit, until the assignments and activities asked them to reason with more than three variables. At the end of the unit, most students still did not know if it was appropriate to make causal claims with their data. However, they remained consistently apt in their ability to create and update directed acyclic graphs, propose relationships among their variables of interest, and provide logical potential causal

variables.

Analysis of responses across the three assignments helped identify trends in the students' performance on each learning outcome and identified similar challenges as seen in the literature, such as confusion about observational data, making causal claims, and potential bias in responses due to the context of the data. Finally, the cognitive interviews provided insight into some challenges and misconception students held and gave a sense of their final multivariate reasoning skills at the end of this unit. Future work is needed to define the skills needed for multivariate thinking, the sequence of those skills for a learning trajectory, and to determine additional ways to support students' development of multivariate thinking.