IASE 2020 ROUNDTABLE DISCUSSANT SUMMARY: TOPIC 2

Discussant: Gail Burrill & Chair: Daniel Frischemeier
Michigan State University, United States, & Paderborn University, Germany
burrill@msu.edu

Topic 2: **The emerging role of multivariate thinking in inferential reasoning**

	Title	Presenter/Co-Author(s)
Long paper	*Now is the Time for Causal Inference in Introductory Statistics*	Karsten Luebke

PRELIMINARY RESEARCH QUESTIONS

- What are the important ideas and skills in relation to multivariate thinking in statistical inference?
- How can the use of technology promote multivariate thinking with large datasets?
- How does the use of multivariate data nurture students to become data detectives?
- What are innovative tasks or sequence of instructional activities that can be used to support students' multivariate thinking?

KEY DISCUSSION THEMES

In the one paper in Topic 2, Luebke argued that it is more necessary than ever to integrate elements of causal inference into the curriculum and to help students think beyond data and beyond the issue that "correlation is not causation". He described how directed acyclic graphs can support students and learners in developing cognitive inferential reasoning processes and provided some ideas on how to integrate these elements into a syllabus.

The discussion might be summarized in a reasoning chain: Statistics is about understanding the world; the world is multivariable, which suggests a rethinking of curriculum to bring in multivariate analyses; if statistics is not changed to gain insights through data, statistics runs the risk of becoming obsolete. Some key ideas that emerged from the discussion:

- Causality is the heart of what statisticians want to know about the world, but a tension exists between this view and formal inference. On one hand, we can't isolate ourselves to randomized experiments but on the other, we have to be careful about claiming causation because of observation and also be aware that correlations are particularly problematic with time dependent data. There is a real need to think about approaching the concept of causality from a different perspective and to consider what meanings might underlie a correlation.
- Recognizing the reality that most data contexts are multivariate, students should learn about compounding and bias and be aware of oversimplified conclusions

There seemed to be a consensus that technology can serve as a bridge from middle school to tertiary by providing opportunities for students to think about multivariate relationships, for example by using color in software such as Tinkerplots. Participants offered ideas for innovative tasks or sequence of instructional activities that might be used to support students' multivariate thinking:

- Multivariate analysis can begin with exploratory data analysis and modeling by deemphasizing some content (i.e., formal inference) and omitting other content to bring in causal inference.
- Given that more secondary schools provide a background in some elementary statistical concepts, a promising approach might be to reshape the current university introductory course in the direction of statistical literacy and offer a new course focusing on gaining insights into data.
- A directed acyclic graph can be used to encode assumptions about the knowledge generating process. A question, however, is whether the graphs provide new insights or are they our a-priori assumptions - not new insights but what we start out with? The response: although qualitative assumptions are used in the graphs, there are testable implications, and, conditional on model and assumptions, an intervention effect can be calculated.