Making Connections and Understanding Statistics: Students' Ratings of the Utility of Key Concepts in the Introductory Statistics Course

Rossi A. Hassad, PhD, MPH
Professor, School of Social & Behavioral Sciences
Mercy College, Dobbs Ferry, New York, USA

Abstract
There is almost universal consensus that the goal of the introductory statistics course at the college level should be to facilitate statistical literacy. This study sought to identify which concepts students believe facilitated their understanding of statistics and its practical applications, and the extent to which each concept was helpful. College students (N = 52) taking an introductory statistics course were asked to list 5 course concepts which they found helpful. Their responses were then sorted, and 13 concepts were identified and presented to them for rating on a scale of 1 to 5 (least to most helpful). Macro or unifying concepts, in particular, correlation and causation, statistical significance, reliability, and confidence interval, were the highest rated, followed by variability, relationship, sampling, standard error, and formulas. The lowest rated concepts were skewness, generalizing, interaction, and confounding. Deep and meaningful learning may be fostered by organizing the curriculum to progress from the macro (or integrating) concepts to the more abstract concepts. These macro concepts can be introduced early and developed later. Further research is required.

Key Words: Statistics, Introductory, Literacy, Concepts, Curriculum, Teaching

1. Background

Improving the introductory statistics course has been the central focus of the statistics education reform movement from its inception, given its foundational role in fostering statistical and quantitative literacy, a core requirement for most academic majors (Forbes, 2014; Tishkovskaya & Lancaster, 2012). Underpinning this reform is the constructivist learning philosophy, and in this regard, the emphasis has been on shifting from a mathematized statistics curriculum to one that is more concept-based and applied, with attention to the context and interpretation of data (Hassad, 2011, 2013). To date, significant strides have been recorded regarding effective pedagogical strategies, assessment methods, and integration of technology. In particular, active learning strategies, including cooperative (or small group) learning, and authentic assessment have emerged as best practices for facilitating statistical literacy (Hassad, 2014).

Consistent with the Guidelines for Assessment and Instruction in Statistics Education (GAISE) report (Franklin et al., 2007), the goal of the introductory course is to engender learning that is deep and meaningful, leading to transferrable knowledge and skills. However, there is a lack of consensus among educators regarding optimal course content and sequencing of topics; a long-standing debate (Chance & Rossman, 2001; Malone et al., 2010). The status quo can be characterized as widespread conformity to the material (and order of topics) as is presented in textbooks (a textbook-centric approach). Indeed, there is a dearth of research in this area, and the published literature is comprised of mostly expert commentaries, case reports, and a few classroom-based studies using concept maps, which have been helpful in providing insight into students' understanding and misconceptions. Also lacking, is information about students' relational thinking and learning trajectories regarding introductory statistics, and what aspects of the course they find most helpful in facilitating conceptual understanding.

2. Objective and Rationale

This study sought to identify which concepts students believe facilitated their understanding of statistics and its practical applications, and the extent to which each concept was helpful. Such information can help instructors to determine how to best scaffold the curriculum so as to facilitate students to make connections between concepts and create meaningful and robust cognitive models and networks.

3. Methodology

This was a multi-method classroom study. College and university majors in psychology, the humanities, and social sciences (N = 52) taking an introductory statistics course were asked (as the end of the semester) to list five course concepts which they found helpful in facilitating their understanding of statistics and its practical applications. Their responses were then sorted, and thirteen (13) concepts were identified and presented to them for rating on a scale of 1 to 5 (least to most helpful). Descriptive and correlation analyses were performed.

The introductory statistics course was designed and administered in accordance with the GASIE report. The course material encompasses common statistical methods and their applications within the disciplines, as well as considerations for collecting, organizing, analyzing, interpreting and presenting data. The material covers (in this sequence) descriptive and inferential statistics; including frequency distributions, graphs, measures of central tendency, measures of variability, sampling, z-score and the normal distribution, as well as tests of hypothesis such as: t-tests, ANOVA, linear correlation and regression, and chi-squared analysis. Effect size, study designs (observational and experimental) and research concepts (association and causation; confounding and interaction) are also addressed just prior to hypothesis testing, and the IBM-SPSS software is used for data analysis. In order to demonstrate the integration and application of knowledge and skills in a meaningful way, students are required to complete a small-group project in which they explore and analyze secondary data using the IBM-SPSS software, interpret and contextualize the variables and results, and submit a coherent written report.

4. Results

Macro or integrating concepts (Table 1), in particular, correlation and causation, statistical significance, reliability, confidence interval, variability, and relationship, were the highest rated (in terms of helping to understand statistics and its practical applications), followed by sampling, standard error, and formulas. The lowest rated concepts were skewness, generalizing, interaction, and confounding. Meaningful relationships between concepts were observed (Table 2). Of note is that higher ratings of the concept of "correlation and causation" as being helpful were associated with lower ratings of "formulas" as being helpful.

Table 1: Students' Ratings of the Utility of Key Concepts in the Introductory Statistics Course

Concepts	**Mean**
1. Correlation and Causation	4.29
2. Statistical Significance	4.27
3. Reliability	4.00
4. Confidence Interval	3.90
5. Variability	3.88
6. Relationship	3.87
7. Sampling	3.83
8. Standard Error	3.81
9. Formulas	3.81
10. Skewness	3.56
11. Generalizing	3.21
12. Interaction	3.08
13. Confounding	2.67
Rated on a scale of 1 to 5 (least to most helpful)	

Table 2: Relationship between Statistical Concepts (as rated by students)

Concepts	**Pearson's r**
Sampling, and Skewness	.501
"Correlation and Causation", and Reliability	.495
Confidence Interval, and Statistical Significance	.477
Confidence Interval, and Standard Error	.465
Confounding, and Reliability	.414
"Correlation and Causation", and Confounding	.361
Sampling, and Standard Error	.340
Confounding, and Interaction	.341
"Correlation and Causation", and Formulas	-.306
All reported coefficients are statistically significant (p<.05).	

5. Discussion and Implications

Students identified key unifying concepts (particularly 1-6, Table 1) as being most helpful to their understanding of statistics. Together these six concepts suggest the theme of "covariational reasoning" (Zieffler & Garfield, 2009), that is, critical thinking about the relationship between two variables, which is considered an important cognitive activity for statistical literacy. Specifically, evidence from cognitive research supports that the brain is predisposed to connecting and categorizing (Barrett, 2009), therefore, attention to the concept of "correlation and causation" can be beneficial and rewarding to learning statistics in general. This concept is important for explaining research outcomes, particularly in the disciplines of psychology and the other behavioral sciences, which tend to have a preponderance of non-experimental research, and hence the need to make this distinction (correlation versus causation).

Issues of "correlation and causation" are omnipresent, therefore, curricular material that is relatable to students should be used, as this can lead to "discovery" and "aha" moments, and hence deep and meaningful learning. Another integrating concept is "statistical significance", which is generally viewed as counterintuitive and abstract, however, it may be better understood when explained in a constructivist manner, with reference to decision-making, including the implications of fluke or spurious, versus reliable or systematic outcomes.

In general, this study reinforces that the GAISE approach can be effective in promoting conceptual understanding and statistical literacy. In particular, more meaningful and engaged learning may be fostered by organizing the curriculum to progress from the macro (or unifying) concepts to the more abstract concepts. The integration and clustering of related concepts may also prove helpful. Use of authentic curricula and assessments (including real and interesting data) is a best practice, in this regard. Also, a subtle but important message from these findings seems to be that emphasizing formulas and calculations in an introductory statistics course may prove counterproductive to deep and meaningful learning; a long-held perspective among statistics educators (Moore, 2007).

Further research should address student characteristics, potential confounding factors, and formal assessment of student learning. Research strategies such as focus groups and concept maps could help to elucidate these results. Course content and sequencing of topics need to be considered as priority areas for research in statistics education. Above all, we must take the sensitive issue of faculty preparation more seriously. Finally, this study utilized a small convenience sample, which could limit the external validity of the findings.

Acknowledgements

Special thanks to the administration of Mercy College for supporting my participation in the American Statistical Association, Joint Statistical Meetings (JSM 2016, Chicago).

References

Barrett, L. F. (2009). The future of psychology: Connecting mind to brain. *Perspectives on Psychological Science*, *4*(4), 326-339.

Chance, B., & Rossman, A. (2001). Sequencing topics in introducing statistics: A debate on what to teach when. *The American Statistician*. *Volume 55*, 140-144.

Forbes, S. (2014). The coming of age of statistics education in New Zealand, and its influence internationally. *Journal of Statistics Education*, *22*(2).

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2007). Guidelines for assessment and instruction in statistics education (GAISE) report. *Alexandria: American Statistical Association*.

Hassad, R.A. (2011). Constructivist and behaviorist approaches: Development and initial evaluation of a teaching practice scale for introductory statistics at the college level. *Numeracy*, *4*(2), 7.

Hassad, R.A. (2014). The Status of Reform in Statistics Education: A Focus on The Introductory Course. In K. Makar, B. de Sousa, & R. Gould (Eds.), Sustainability in statistics education. *Proceedings of the Ninth International Conference on Teaching Statistics* (ICOTS9, July 2014), Flagstaff, Arizona, USA. Voorburg, The Netherlands: International Statistical Institute.

Hassad, R.A. (2013). Teaching Introductory Statistics for Evidence-Based Practice: Integration of Context. In S. Forbes and B. Phillips (Eds.) *Proceedings of the Joint IASE/IAOS Satellite Conference*, Macao, China. August 2013.

Malone, C., Gabrosek, J., Curtiss, P., & Race, M. (2010). Resequencing topics in an introductory applied statistics course. *The American Statistician*, *64*(1), 52–58.

Moore, D. S. (2007). *The basic practice of statistics* (Vol. 2). New York: WH Freeman.

Tishkovskaya, S., & Lancaster, G. A. (2012). Statistical education in the 21st century: a review of challenges, teaching innovations and strategies for reform. *Journal of Statistics Education*, *20*(2), 1-55.

Zieffler, A., & Garfield, J. (2009). Modeling the growth of students' covariational reasoning during an introductory statistics course. *Statistics Education Research Journal*, *8*(1), 7-31.