A PSYCHOMETRIC ANALYSIS OF THE GOALS AND OUTCOMES ASSOCIATED WITH LEARNING STATISTICS (GOALS) INSTRUMENT

Anelise Sabbag and Andrew Zieffler
University of Minnesota, United States
sabb0013@umn.edu

Several randomization-based approaches to teaching statistical inference have been developed and taught since the publication of Cobb's (2007) landmark article. The CATALST curriculum is one example of this approach. The proposed poster provides results from a study conducted to gather validity evidence to support the use of the *Goals and Outcomes Associated with Learning Statistics* (GOALS) instrument in evaluating the effectiveness of the CATALST curriculum. The study examined two primary research questions: (1) Which measurement model is appropriate to model the data from the GOALS instrument; and (2) Based on the measurement model, how can the GOALS instrument be improved?

GOALS was administered to 289 undergraduate students from six universities in the United States enrolled in statistics courses that used the CATALST curriculum. The data collected were analyzed under both the classical test theory and item response theory frameworks. Three measurement models were fitted to the data: the *Two-Parameter Logistic Model*, the *Mixed Model* (comprised of both the Two-Parameter Logistic Model and the Graded-Response Model), and the *Bi-Factor Model*. These models were compared at the item- and model-levels. The Mixed Model was found to most appropriately model students' responses to GOALS. Results from fitting the Mixed Model suggested that there are some items that need to be revised. In addition, we found a need for more discriminating items in order to improve the score precision and total test information.

REFERENCES

American Educational Research Association and the National Council on Measurement in Education (AERA/APA/NCME), (1999). *Standards for Educational and Psychological Testing*. Washington, D.C., American Educational Research Association.

American Statistical Association (2005). College Report. *Guidelines for assessment and instruction in statistics education*. Washington DC: Author

De Ayala, R. J. (2009). *The theory and practice of item response theory*. New York: Guilford Press.

DeMars, C. E. (2012). Confirming Testlet Effects. *Applied Psychological Measurement*, *36*(2), 104-121.

Garfield, J., & Franklin, C. (2011). Assessment of learning, for learning, and as learning in statistics education. In *Teaching Statistics in School Mathematics-Challenges for Teaching and Teacher Education* (pp. 133-145). Springer Netherlands.

Garfield, J. B., & Gal, I. (1999). Assessment and statistics education: Current challenges and directions. *International Statistical Review / Revue Internationale De Statistique, 67*(1), pp. 1-12. http://www.jstor.org/stable/1403562

Thorndike, R.M., & Thorndike-Christ, T. (2010). *Measurement and evaluation in psychology and education* (8th ed.). Upper Saddle River, NJ: Pearson Education.