

*The Impact of Vertical Scaling Decisions on Value-Added Models of Student Achievement and School Effectiveness*

Derek C. Briggs, PhD – University of Colorado, Boulder

Jonathan P. Weeks, PhD – University of Colorado, Boulder

Edward Wiley, PhD – University of Colorado, Boulder

Wednesday, April 23

10:00am – Noon

---

The purpose of this study was to evaluate the sensitivity of value-added models to the way an underlying vertical scale has been established. Longitudinal item-level data with both student and school-level identifiers over time for the entire state of Colorado was analyzed. We use this data to address two principal research questions: (1) What is the sensitivity of a longitudinal score scale to the way the test scores have been vertically scaled? (2) What impact do different IRT-based vertical scaling approaches have on value-added estimates of student growth trajectories and school effects? Eight different vertical scales were established on the basis of choices made for three key variables: Item Response Theory (IRT) modeling approach, calibration approach and student proficiency estimation approach. Each scale represented a methodological approach that was in some sense defensible. Longitudinal values of each scale were used as the outcome variable in two commonly used value-added models. In one analysis different fixed effect and empirical Bayes estimates of student-level and school-level growth are compared; in another empirical Bayes estimates of school-level effects are compared. Our findings suggest that these sorts of value-added estimates can in fact be quite sensitive to choices made in the development of a vertical scale.