Introduction

Background

The use of counterbalancing in scientific research is well established regarding treatment order as a method of reducing threats to internal validity. In educational and psychological research, the need for counterbalancing self-report measures has been examined (Cunningham, Preacher & Banaji, 2001), but assessment presentation order is often overlooked and rarely incorporated into the design of studies utilizing raters. As a result, the effect of order of multiple test administrations to a single rater in behavior protocols has not been thoroughly examined (Lucas, 1992).

Objective

The current study presents the results of an analysis of order effects utilizing data from a larger validation study of the Direct Behavior Rating Single Item Scales (DBR-SIS).

Method

Participants and Setting

For this study, a total of 1976 students were rated by teachers on all three measures in the fall, winter and spring. Students were mostly Caucasian (82%), with 7% Hispanic, and 13% students with disabilities. Students were nested within teachers, with 202 teachers participating in 23 urban, suburban, and rural school districts in 3 states: Connecticut, New York, and Missouri.

Measures

The Direct Behavior Rating – Single Item Scale (DBR-SIS) reflects the teacher’s perception of the proportion of time a student is observed engaged in a target behavior (academic engagement, respectful, disruptive) from 0 (never) to 10 (always). Students were rated twice daily for five days. The composite rating is the sum of the AE, RS, and 10% of the BESS (partial SSIS), or the BESS combined T score, while controlling for student behavior characteristics within classroom. The level 1 and level 2 equations are as follows:

Level-1 Model

\[ Y_{ij} = \beta_0 + \beta_1 \text{(GRADE)} + \beta_2 \text{(BESE)} + \beta_3 \text{(SSISSP)} + \epsilon_{ij} \]

Level-2 Model

\[ \beta_0 = \tau_00 + \tau_01 \text{(ORDER)} + u_0 \]
\[ \beta_1 = \tau_10 + \tau_11 \text{(GRADE)} + u_1 \]
\[ \beta_2 = \tau_20 + \tau_21 \text{(BESE)} + u_2 \]
\[ \beta_3 = \tau_30 + \tau_31 \text{(SSISSP)} + u_3 \]

Variables were created to reflect the order of administration of other tests relative to the dependent variable. Eighty-one models were run to test effects of change in test order on lower elementary (grade 1 & 2), upper elementary (grade 4&5), and middle school (grade 7&8) separately for fall, winter and spring.

Results

While test order was found to be associated with teacher’s global perceptions of students in the MANOVA analyses, these results are somewhat mitigated by the lack of random assignment of students to classrooms. Therefore, it is possible that the characteristics of students within clusters impact these results. The multilevel models, which do control for student behavior, only examine one behavioral measure at a time and seem to indicate that test order does not impact teacher ratings.

Summary and Conclusions

Preparation of this poster was supported by a grant from the Institute for Education Sciences (IES), U.S. Department of Education (R324B080014). For additional information, please direct all correspondence to Sandra Chafouleas at sandra.chafouleas@uconn.edu