Title: Using Generalizability Theory to Validate an Existing Measure of Social Communication in Neurogenic Syndrome Populations

Authors: Lisa R. Hamrick,1 Alison M. Haney,1 Sean P. Lane,1 & Bridgette L. Tonnsen,1

1Purdue University

Introduction: Children with neurogenic syndromes (NGS) experience severe and cascading developmental delays that can benefit from early, targeted treatment. However, few outcome measures have been identified as sensitive enough to detect small, meaningful developmental changes in these populations to appropriately evaluate treatment effectiveness. Importantly, many measures that assess early developmental features are designed for use in typically developing (TD) populations and do not account for the immense heterogeneity, slower developmental trajectories, and unique profiles of strengths and weaknesses that characterize NGS phenotypes. This study leverages generalizability theory (GT), an extension of classical test theory that provides nuanced information about sources of measurement variance and reliability, to determine whether an existing measure of social communication could be used reliably in populations with NGS. We first examine the degree to which variability in scale scores is due to individual differences, chronological age, or item-level content in three NGS groups (Angelman syndrome [AS], Prader Willi syndrome [PWS], Williams syndrome [WS]) and a TD comparison group. We then use this information to calculate the reliability of between-person differences and change in social communication skills over time for all groups.

Method: Participants included 108 NGS (38 AS, 28 PWS, 42 WS) and 59 typically developing (TD) children, aged 1 to 71 months, from an ongoing longitudinal study of early development in NGS populations. Participants’ mothers completed the Communication and Symbolic Behavior Scale – Infant-Toddler Checklist (CSBS-ITC; Wetherby & Prizant, 2003), a 24-item screening measure intended to detect social communication delays in children with mental ages under 24 months. The CSBS-ITC was collected longitudinally at 6-to-12-month intervals. Number of observations per participant ranged from 1 to 5 (M=2.91), with a total of 421 total observations across participants. Multilevel modeling was used to identify variance estimates of individual differences, chronological age, and item-level content, and the two-way interactions of these predictors. Percent variance was calculated for each predictor and interaction. Reliability estimates for the ability of the CSBS-ITC to differentiate between individuals at a fixed timepoint (RI1) and the ability to reliably detect change in skills over time (RC) were calculated based on formulas published in Cranford et al. (2006).

Results: Age accounted for the most variance in TD scores (46.40%), while individual differences (21.69%) and item content (21.56%) primarily accounted for variance in NGS scores. Thus, in the TD group, a child’s age was the best predictor of their score on any given item of the CSBS-ITC. In contrast, person-level factors and the particular skill measured by the item were the best predictors of the child’s score on CSBS-ITC items in the NGS group. Variance estimates calculated separately for each NGS group revealed item content (47.57%), individual differences (14.35%), and the interaction of these predictors (15.13%) accounted for the most variance in AS CSBS-ITC scores, while item content, individual differences, and age accounted for the most variance in PWS (16.84%, 17.94%, and 20.46%, respectively) and WS (19.03%, 14.23%, and 27.04%) groups. Reliability estimates suggested that the CSBS-ITC can reliably assess skill levels in TD (RI1=.84) and NGS populations (RI1=.94 for all 3 syndromes). Reliability to detect change was best for the PWS (RC=.78) and TD (RC=.64) groups, but still fair in the AS (RC=.42) and WS (RC=.59) groups.

Discussion: Results suggest that although the variance in skills measured by the CSBS-ITC are explained by different factors in NGS and TD, this screening measure can reliably detect individual differences in skill levels in both groups. The ability of the CSBS-ITC to detect reliable change over time was best in the PWS group, suggesting this measure is sensitive to natural developmental trajectories observed in this NGS group and could be used to measure within-individual change over time. This level of reliability was not observed for the AS or WS groups, suggesting that the CSBS-ITC may be less sensitive at detecting change in social communication skills over time that is characteristic of these phenotypes. Future directions include using GT analysis to identify subsets of items that optimize reliability estimates for particular syndrome groups, thus improving this scales utility among NGS groups with nuanced and unique phenotypes. This study provides a practical example of how GT approaches may be used to evaluate scale reliability across heterogeneous groups with developmental delays.

References/Citations: