Research Week 2020

Equivalency of Automated vs Conventional Language Assessment in Children with and without Neurodevelopmental disorders

Kristin Hinton, M.Ed., Jill Dolata

OHSU

Keywords

Autism, ADHD, Language Disorder in children

Abstract

Background

Language assessment for children has predominately utilized face-to-face testing methods (i.e., conventional administration); however, automated administration could increase clinician efficiency, access to services, standardization of administration, and perhaps child interest (Laborda, 2007; Noland, 2017). Though automated assessment may have advantages, this new method of assessment has limited evidence for score equivalence (Carson, Gillon & Boustead, 2011). In this project, we evaluate the equivalence of scores across conditions for four expressive subtests of a modified common language assessment: Expressive Vocabulary (EV), Recalling Sentences (RS), Word Structure (WS), and Formulated Sentences (FS).

Methods

77 children with Autism Spectrum Disorder (ASD), Attention-Deficit Hyperactivity Disorder (ADHD) and Developmental Language Disorder (DLD), and Typical Development represent the subset of participants who received both automatic and conventional administrations in a larger study designed to develop automated test scoring using speech recognition. Participants were randomly assigned to balanced conditions (automatic or conventional), and tests were administered four-six weeks apart.

Results

Data entry is ongoing; preliminary analyses are from 41 participants (average age 7.1 years (sd=1.1); 63% male, 75% Caucasian) with ASD (10), ADHD (10), DLD (5) and TD (16). Pearson's product-moment correlations were run to assess relationships between raw scores on four conventionally or automatically administered expressive language tasks. Preliminary analyses showed the relationship to be linear with both variables on each task normally distributed, as assessed by Shapiro-Wilk's test ($p > .05$), and there were no outliers. There were statistically significant, strong correlations between visit one and

visit two raw scores, $r(39) = .97$ (RS), .88 (EV) .92 (WS), .90 (FS), $p < .0005$ for all correlations.

Conclusions

Initial results are promising support for reliability of testing in automatic conditions. Additional analyses will be completed prior to Research Week 2020 and will include the rest of the subject pool and further analysis of condition effects.