Validity of Automated vs. Hand-Scored Written Expression Curriculum-Based Measurement Samples



Keller-Margulis M., Matta, M., & Mercer, S.

BACKGROUND:

- Automated Written Expression-Curriculum Based Measurement (i.e., aWE-CBM) represents a promising alternative to the traditional WE-CBM for the evaluation of writing in the context of universal screening.
- Scope of the poster is to investigate the validity of three automated scoring methods for the prediction of general writing skills in relation to traditional WE-CBM metrics.

METHODS AND ANALYSES

- 163 fourth-grade students completed a 3-minute writing sample in the Winter and their writing state test score (i.e., STAAR) was collected from the school district.
- 2. Composite scores for WE-CBM and aWE-CBM were calculated (see scoring procedures in the right bar), and simple and multiple regression models were then developed to predict general writing skills.

RESULTS

- Correct minus Incorrect Word Sequences (CIWS) was the best predictor of writing skills among the traditional WE-CBM metrics.
- Both ReaderBench and Coh-Metrix predicted quality scores, paired with variables evaluating structural and mechanics errors (i.e., percentage of misspelled words, typos, and grammar errors) showed validity coefficients of similar magnitude (see results in the right bar).

DISCUSSION

• aWE-CBM enables to rapidly compute a variety of metrics relative to multiple levels of language (i.e., word-, sentence-, and discourse-level) and eliminates the threat of low agreement among different raters.

Automated text evaluation tools offer feasible methods for the assessment of writing quality in the context of universal screening.











The University of Houston is an EEO/AA institution.

UNIVERSITY of HOUSTON COLLEGE OF EDUCATION

SCORING PROCEDURES OF WRITING SAMPLES



RESULTS OF REGRESSION MODELS

Model	Predictors	STAAR scores R ²	STAAR Not proficient		
			AUC ¹	CI_{low}^{2}	CI _{upp} ²
Traditio	onal WE-CBM scores				
1	TWW	0.00	0.48	0.34	0.62
2	WSC	0.02	0.55	0.42	0.68
3	CWS	0.19	0.75	0.65	0.85
4	CIWS	0.25	0.86	0.79	0.93
aWE-C	BM predicted quality scores				
5	ReaderBench	0.14	0.67	0.55	0.79
6	Coh-Metrix	0.16	0.64	0.53	0.76
7	PEG Total	0.18	0.84	0.76	0.92
8	ReaderBench + GAMET	0.26	0.81	0.71	0.92
9	Coh-Metrix + GAMET	0.28	0.82	0.72	0.91

¹ AUC: Area Under the Curve calculated on whether or not the students met the grade level on the STAAR test.

 2 Confidence Intervals (CI) calculated at the 95% level.