

## Sentence Repetition as a Diagnostic Tool for Developmental Language Disorder: A Systematic Review and Meta-Analysis

Leah Ward<sup>1</sup>, Kamila Polisenska<sup>2,1</sup>, Colin Bannard<sup>1</sup>

<sup>1</sup>University of Manchester, Manchester, United Kingdom. <sup>2</sup>City, University of London, London, United Kingdom

### Abstract

*Background:* Sentence repetition (SR) tasks are popular for use in language assessment as well as research. Performance on a SR task is viewed as a promising clinical marker of developmental language disorder (DLD). However, the ways these tasks have been designed and evaluated has varied considerably.

*Purpose:* We conducted a systematic review and multilevel meta-analysis examining the accuracy of SR tasks in distinguishing between typically developing (TD) children and children with DLD. It explores variation in the way that SR tasks are administered and/or evaluated and examines whether variability in the reported ability of SR to detect DLD is related to these differences.

*Method:* Four databases were searched to identify studies which had used a SR task on groups of monolingual children with DLD and TD children. Searches produced 3,459 articles of which, after screening, 66 were included in the systematic review. A multilevel meta-analysis was then conducted using 46 of these studies. Multiple preregistered subgroup analyses were conducted in order to explore the sources of heterogeneity.

*Results:* The systematic review found a great deal of methodological variation, with studies spanning 19 languages, 39 SR tasks, and four main methods of production scoring (sentence binary, sub-sentence binary, target binary, and error scoring). There was also variation in study design, with different sampling (clinical and population sampling) and matching methods (age- and language-matching). The overall meta-analysis found that on average TD children outperformed children with DLD on the SR tasks by 2.08 SDs. Subgroup analyses found that effect size only varied as a function of matching method and language of task.

*Conclusions:* Our results indicate that SR tasks can distinguish children with DLD from both age- and language- matched samples of TD children. The usefulness of SR as a clinical marker appears robust to most kinds of task and study variation.