

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Estimating demographic bias on tests of children's early vocabulary

Permalink

<https://escholarship.org/uc/item/1ws2z2b0>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 44(44)

Authors

Kachergis, George

Francis, Nathan

Frank, Michael C.

Publication Date

2022

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Estimating demographic bias on tests of children's early vocabulary

George Kachergis

Stanford University, Stanford, California, United States

Nathan Francis

Stanford University, Stanford, California, United States

Michael Frank

Stanford University, Stanford, California, United States

Abstract

Children's early language skill has been linked to later educational outcomes, making it important to measure early language accurately. Parent-reported instruments such as the Communicative Development Inventories (CDIs) have been shown to provide reliable and valid measures of children's aggregate early language skill. However, CDIs contain hundreds of vocabulary items, some of which may not be heard (and thus learned) equally often by children of varying backgrounds. This study used a database of American English CDIs to identify words demonstrating strong bias for particular demographic groups of children, on dimensions of sex (male vs. female), race (white vs. non-white), and maternal education (high vs. low). For each dimension, we identified dozens of strongly biased items, and showed that eliminating even some of these items can reduce the magnitude of differences between groups. Additionally, we investigated how well the relative frequency of words spoken to young girls vs. boys predicted sex-based word learning bias, and discuss possible sources of demographic differences in early word learning.