Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

*Acad Pediatr*. Author manuscript; available in PMC 2023 September 01.

Published in final edited form as:

*Acad Pediatr*. 2022 ; 22(7): 1133–1141. doi:10.1016/j.acap.2021.12.024.

# Psychometric Assessment of Pilot Language and Communication Items on the 2018 and 2019 National Survey of Children's Health

**Helena Hutchins**[1,2], **Lara Robinson**[1], **Sana Charania**[3], **Reem Ghandour**[4], **Kathy Hirsh-Pasek**[5,6], **Jennifer Zubler**[7,8]

[1]Child Development Studies Team, Division of Human Development and Disability, National Center on Birth Defects and Developmental Disabilities, Centers for Disease Control and Prevention, 4770 Buford Hwy S106-4, Atlanta, GA 30341-3717, USA.

[2]Oak Ridge Institute for Science and Education, Centers for Disease Control and Prevention Research Participation Programs, P.O. Box 117, Oak Ridge, TN 37831-0117.

[3]Early Hearing Detection and Intervention Team, Division of Human Development and Disability, National Center on Birth Defects and Developmental Disabilities, Centers for Disease Control and Prevention, 4770 Buford Hwy S106-4, Atlanta, GA30341-3717, USA.

[4]Office of Epidemiology and Research, Maternal and Child Health Bureau, Health Resources and Services Administration, 5600 Fishers Lane, Room 18N122, Rockville, MD 20857, USA.

[5]Department of Psychology, Temple University, 6th Floor, Weiss Hall, 1701 N 13th St, Philadelphia, PA 19122, USA.

[6]Brookings Institution, 1775 Massachusetts Ave., NW, Washington, DC 20036, USA.

[7]Eagle Global Scientific,1826 N Loop 1604 W, Suite 336E, San Antonio, TX 78248, USA.

[8]Learn the Signs. Act Early Program, Division of Human Development and Disability, National Center on Birth Defects and Developmental Disabilities, Centers for Disease Control and Prevention, 4770 Buford Hwy S106-4, Atlanta, GA 30341-3717, USA.

## Abstract

**Objective:** Until recently, normative data on language and communication development among children in the United States have not been available to inform critical efforts to promote language development and prevent impairments. This study represents the first psychometric assessment of nationally representative data derived from a National Survey of Children's Health (NSCH) pilot measure of language and communication development among children ages 1–5 years.

**Corresponding Author:** Helena J. Hutchins, MPH, BSEd; Centers for Disease Control and Prevention; 4770 Buford Hwy S106-4, Atlanta, GA 30341-3717, USA; telephone: (317) 270-7584; pne8@cdc.gov.

**Conflicts of interest/Competing interests** (include appropriate disclosures): Authors have no conflicts of interests to report.

**Disclaimer:** The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

**Methods:** We analyzed 14,573 parent responses to language and communication items on the 2018 and 2019 NSCH to evaluate whether the newly added 11 items represent a single latent trait for language and communication development and to determine normative age of success on each item. We applied weighted, one-parameter Item Response Theory to rate and cluster items by difficulty relative to developmental language ability. We examined Differential Item Functioning (DIF) using weighted logistic regression by demographic factors.

**Results:** Together, exploratory factor analysis resulting in a single factor > 1 and explaining 93% of the variance and positive correlations indicated unidimensionality of the measure. Item characteristic curves indicated groupings were overall concordant with proposed milestone ages and representative of an approximate 90% success cut-point by child age. Indicated normative age cut-points for three of the items differed slightly from proposed milestone ages. Uniform DIF was not observed and potential nonuniform DIF was observed across 5 items.

**Conclusions:** Results have the potential to enhance understanding of risk and protective factors, inform efforts to promote language and communication development, and guide programmatic efforts on early detection of language delays.

### Keywords

language development; psychometrics; Item Response Theory; language milestones; survey data

## INTRODUCTION

The prevalence of expressive and receptive language delay among children 3–5 years old in the United States ranges from 5–8%.[1] Mastery of language and communication during early childhood is among the strongest predictors of later academic achievement and intellectual ability;[2–4] for example, data indicates medium to large effect sizes.[5] Failure to meet language development milestones during early childhood can be an early indication of developmental disorder or learning disability, and may be a secondary characteristic of other disability or hearing loss.[4, 6] Furthermore, deficits in language development have been associated with medium to large effect sizes related to poor cognitive, educational, and occupational outcomes in adulthood.[7] Population-level data on early language and communication development can be used to improve the accuracy of screening and surveillance tools, identify the prevalence of delays, and characterize risk and protective factors to inform language and communication promotion efforts. However, until recently, nationally representative data on children's early language and communication (referred to herein as language) development have not been available to inform these efforts.[8]

Potential delays in language development can be evident during early childhood.[9] The American Academy of Pediatrics (AAP) recommends pediatric developmental surveillance at each health supervision visit and periodic screening using validated screeners at recommended ages in order to detect developmental delays and intervene early to improve long-term developmental outcomes.[10] Language is included as a developmental domain in the AAP's surveillance and screening recommendations.[10] The Centers for Disease Control and Prevention's (CDC) *Learn the Signs. Act Early*. program promotes developmental surveillance among caregivers and providers and encourages screening if concerns are

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

identified for early identification of delays.[11] Following the detection of possible language delays through surveillance and screening, pediatric providers can refer children for further evaluation and treatment.[10]

A number of developmental surveillance and screening tools incorporating measures of language development have been developed.[11–13] However, accurate and efficient surveillance of population level, early language development poses challenges. There is variability in the accuracy of screening instruments for language development[1] and in timing of language skill acquisition among typically developing children.[14] Screening instruments for language development can be lengthy and time-consuming,[1, 15] inhibiting translation to more concise and widely implemented surveillance measures. Parent-reported developmental screening tools are common, though variabilities in parents' cultural norms, knowledge, and awareness of their children's abilities can impact validity of parent-report instruments.[15] Nevertheless, parent responses to language instruments demonstrate high concurrent and predictive validity[15, 16] and comparable accuracy when compared to direct assessments.[1] Psychometric analyses of representative, normative data on individual milestones has been recommended to inform developmental surveillance.[12, 17]

## The Inclusion of Language Development Items on a National Survey

In 2016, a multiple-domain school readiness measure of early childhood development was incorporated into the National Survey of Children's Health (NSCH) as part of a proposed Healthy and Ready to Learn (HRTL) National Outcome Measure.[18, 19] The NSCH is an annual, parent-report survey of children's health and wellbeing, representative of non-institutionalized US children, ages 0–17 years, funded and directed by Health Resources and Services Administration's Maternal Child Health Bureau (HRSA's MCHB) and conducted by the United States Census Bureau. Expert review and psychometric assessment of the 2016 HRTL data revealed a gap in measurement of early language development.[20]

To address this gap, HRSA's MCHB, CDC, Child Trends, and language experts in academia collaborated to develop early language items for the NSCH. Drawing upon existing language items that could be reported on reliably by a caregiver, 11 language items were developed for parents/caregivers of children ages 1–5. The 11 language items were first included in the 2018 NSCH, with data collected from June 2018 through January 2019.[21] A second year of data was collected from June 2019 through January 2020.[22]

## Item Response Theory

Item Response Theory (IRT) has been previously applied to assess psychometric properties and relative difficulty of child development instrument items. [12, 17, 23] IRT assumes that a child's reported success on an item is dependent upon the child's overall developmental language ability, the latent trait predicted by the full IRT model. IRT utilizes item characteristic curves (ICCs) to illustrate the relative difficulty of each item using probability of success as a function of the predicted developmental ability level. IRT can also be used to model a second parameter, item discrimination, i.e., the degree to which the item discriminates between children of different developmental ability levels.[24] For example, an item on which most children achieve success at a similar level of developmental ability will

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

have higher discrimination compared to an item with lower discrimination on which children achieve success over a wide range of developmental ability levels.[24]

Analyses to detect differential item functioning (DIF), or item bias, is a common follow-up to IRT. DIF is present when responses to an item differ between two groups of respondents and the predicted latent trait (developmental ability) is held constant.[24, 25] Non-uniform DIF refers to change in the relationship between groups across developmental ability levels, indicated by non-parallel group ICCs for an item. For example, girls may be more likely than boys to speak two words in a sentence at earlier stages of development, but less likely than boys to speak two words in a sentence at later stages of development. Whereas if uniform DIF is present, differential likelihood of speaking two words in a sentence by sex would be similar across the full range of developmental ability, and group ICCs would be parallel.[12, 25]

### Research Questions

To examine the first nationally representative data on children's language development, the aim of the present study was to conduct psychometric analyses of language items piloted in the 2018 and 2019 NSCH. We examined two primary research questions. First, we sought to evaluate whether items are best described by one underlying latent trait — developmental language ability. Second, we utilized IRT to examine whether the data support the proposed, *a priori* age milestones based on relative item difficulty.

## METHODS

### Survey Design, Procedures, and Data

We analyzed data from the 2018 and 2019 NSCH. The NSCH was redesigned in 2016 as an annual survey, merging the original NSCH and the National Survey of Children with Special Health Care Needs (NS-CSHCN). The 2016 redesign was a transition away from the State and Local Area Integrated Telephone Survey (SLAITS) random-digit dialing phone survey format, reflecting national changes in landline usage. The redesigned NSCH invites caregivers to participate in the survey via a secure, confidential website, or mail-in paper survey.[21, 22]

The 2018 and 2019 NSCH used an address-based random sampling of households[21, 22] which resulted in a total of 59,963 questionnaires of which 14,701 were for children ages 1–5 years and 14,573 had complete data on the language items. No exclusion criteria beyond complete data on language items was applied. Surveys are completed by a parent or guardian (herein referred to as parents) and include information on child health and development, healthcare and service utilization, and community and family characteristics. In households with multiple children, one child was randomly selected to be the subject of the questionnaire. Survey items are weighted to account for nonresponse and unequal probability of selection. Weights were adjusted to reflect the demographic distribution of non-institutionalized U.S. children. The overall weighted response rate (for ages 0–17 years) was 43.1% and 42.4% for 2018 and 2019, respectively. Surveys were completed in Spanish or English. The data from two survey years were combined using guidance provided by the

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Census Bureau on producing multi-year estimates.[26] Additional information about NSCH survey methods is available elsewhere.[21, 22]

## Measure development – Table 1

Child development and language experts collaborated to develop a list of language items that could be administered in a survey format based on relevant literature and existing developmental milestones for early childhood language. The NSCH items were adapted from three sources.[11, 13, 27] The items were refined through an iterative, collaborative process, and final items were reviewed by the Census Bureau's cognitive testing team prior to survey inclusion. The final list of items included five items for receptive language (e.g., "Understand the meaning of the word 'no'") and six for expressive language (e.g., "Say at least one word, such as 'hi,' or 'dog'").

Parent respondents to the 2018 and 2019 surveys were asked for child's age in years. No additional age data, (e.g., date of birth or age in months) were available in the public use data files. Age 1-year can be interpreted as inclusive of ages 12–23 months. Parents of children ages 1–5 years were prompted to respond to all dichotomous language items, with response options "yes" and "no." The phrasing of each item, item source, receptive or expressive designation, and the hypothesized milestone age (expected age of "yes" response on the item) are listed in Table 1. Complete survey content is accessible via the NSCH website: https://www.nschdata.org/learn-about-the-nsch/survey-instruments.

## Data Analysis

The current study focused on: 1) evaluating the pilot items for unidimensionality; and 2) determining the normative age of success (yes responses) for each item based on relative item difficulty and proportion of reported success by age. Descriptive summary statistics (weighted and unweighted) were calculated to assess item success. Clopper-Pearson 95% confidence intervals (CI) were calculated for all percentages. Weighted results were stratified by child's age in years, 1–5.

We first generated a correlation matrix (Supp. Table 1) to determine the strength of item correlations across language items and conducted Exploratory Factor Analysis to evaluate whether the items represent a single, underlying latent trait. Finally, we employed weighted Item Response Theory (IRT)[24] to assess relative difficulty of the items and to identify potential cut-points defining normative age of success. Clustering of items based on IRT parameter estimates were also compared to the percentage of success on each item, stratified by child age to ensure the proportion of success across indicated age cut-points was comparable. [17] Before accounting for complex survey sampling, we conducted 1-parameter (holding the discrimination parameter constant) and 2-parameter (modeling discrimination for each item) IRT models to determine the best fit for our data. We assessed model fit by comparing ICC curves and corresponding parameter estimates for unweighted 1- and 2-parameter IRT models. We also assessed model fit with likelihood ratio tests, statistical significance defined by $p$-value $< 0.05$.

In order to detect potential non-uniform and uniform DIF, we used weighted logistic regression methods to compare models with the predicted latent trait as an independent

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

variable.[12, 25, 28, 29] We tested for DIF by groups using demographic variables indicated by the literature: highest parental education (graduated from high school or less vs. education beyond high school), child race/ethnicity (non-Hispanic/white vs. other race/ethnicity), child sex (male vs. female), and household language (English vs. other).[12, 17] All analyses were conducted using Stata17.[30]

## RESULTS:

Results are based on 14,573 children ages 1–5 years with complete language data. There were missing data on all 11 language items, for a total of 262 responses (134 in 2018 and 128 in 2019) with no additional missing responses. No systematic demographic differences were detected among children with missing caregiver responses to the 11 language items (personal communication with U.S. Census Bureau, [December 18, 2019]). Descriptive statistics and proportion of item success can be found in Table 2. Overall, the item "Tell a story with a beginning, middle, and end" (59.7, 95% CI: 58.0, 61.4) had the smallest proportion of success ("yes" responses), while the item "Say at least one word" (96.9, 95% CI: 96.2, 97.5) had the greatest proportion of success. Correlation coefficients, in order of frequency/proportion of item success, can be found in Supplemental Table 1. All observed correlations were positive and statistically significant based on $p$-values < 0.05, with coefficients ranging from 0.14 for "Tell a story" and "Say at least one word" to 0.77 for "Ask questions like 'why,' 'how'" and "Ask questions like 'who,' 'what.'" Items with the lowest frequencies/proportions of success, were highly correlated, while those with the highest frequencies/proportions of success were less correlated. Correlation coefficients were smallest between items proposed as 1-year items and the item proposed as a 5-year item. Exploratory factor analysis of item data (unweighted) resulted in a single factor > 1 and explaining 93% of the variance.

Comparison of unweighted 1 and 2-parameter IRT models indicated similar ordering and clustering of items by difficulty parameters (Supplemental Figures 1 and 2). ICC curves remained generally parallel across models. Discrimination parameters increased with item difficulty, with lowest discrimination for items "Say at least one word" and "Understand 'no,'" and highest discrimination for the item "Tell a story." The likelihood ratio test indicated better model fit with separate discrimination parameters for each item, as afforded by the 2-parameter model (p-value < 0.001). However, higher discrimination of the more difficult items in the 2-parameter model may be due to the age range of children for whom these survey questions were answered. For example, "Tell story" is an indicated age 5-year item, and these data do not include children aged over 5 years. Given these observations, and statistical limitations to accounting for complex survey sampling in the 2-parameter model, we focus our interpretation on the weighted 1-parameter IRT model.

Figure 1 displays results of the 1-parameter IRT analysis, weighted to account for complex NSCH sampling and clustering procedures; ICC curves are ordered and clustered by difficulty relative to developmental language ability from left to right. Corresponding item parameter values are presented in Table 3. Delineations between clusters of items, indicated by spatial proximity of ICC curves and by numerical item parameter values, point to potential age cut-points for normative attainment of items. Based on the weighted,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

1-Parameter model, difficulty parameter values for indicated age 1 year items range from −1.95 to −1.89 and included “Say at least one word” and “Understand the meaning of “no”, from −1.58 to −1.29 for the five indicated age 2 year items (“Follow a verbal direction”, “Point to things in a book”, “Follow 2-step directions”, “Use 2 words together”, “Understand words such as ‘in,’ ‘on,’ or ‘under’”), and −0.98 and −0.94 respectively for indicated age 3 year items (“Uses 3 words together in a sentence” and “Ask questions like ‘who,’ ‘what’”). The difficulty parameter value was −0.75 for the indicated age 4 year item “Ask questions like ‘why,’ ‘how’”, and was −0.37 for the indicated age 5 year item “Tell a story.” Item parameter estimates indicated age cut-points represented approximately 90% item success by age (Table 2). For one item, “Ask questions like ‘why,’ ‘how’”, 90% success was achieved one year earlier than proposed and indicated by ordering and clustering of IRT parameters, corresponding 95% Confidence Intervals, and ICCs (i.e., age 3 years versus age 4 years). The ordering and clustering of items by item difficulty parameters was similar between weighted and unweighted 1-Parameter models. For comparison, we also present results of the unweighted 1-Parameter model in the Supplemental Table 2 and Supplemental Figure 1.

Results of logistic regression tests of DIF (Table 4) suggest possible non-uniform DIF by sex among three items (“Follow a verbal direction”, “Put 2 words together”, and “Ask questions like ‘why,’ ‘how”), caregiver education for two items (“Follow a verbal direction” and “Understand words such as ‘in,’ ‘on,’ and ‘under’”), and household language for one item (“Say at least one word”) based on $p$-values < 0.05 and Wald 95% Confidence Intervals. Results did not provide evidence of potential uniform DIF based on > 10% change to coefficients from Model 1 to Model 2.[12, 25]

## DISCUSSION

We conducted a psychometric analysis of the 11 pilot items for children’s language development included on the 2018 and 2019 NSCH. Analyses indicated that the items were representative of a single underlying trait, language ability. Based on a representative sample of US children ages 1–5 years, we used IRT to order and cluster items by difficulty and to inform age cut-points suggesting when parents report that most children ( ~ 90% or more) attain the language skills described by each of the pilot items. Weighted percentage of parent-reported success on the items by child age in years were similar across indicated age cut-points (range: 90.3%−96.4%), consistent with prior recommendations for normative data on developmental milestones.[17] Results indicated only minor differences in proposed and reported attainment of 2- and 3-year items and few indications of item bias by demographic factors.

These results offer information on normative, population level estimates of skill success, with important implications for programmatic efforts to address and promote language development. While overall concordant with proposed age milestones, our results indicate that further examination of normative ages of parent-reported success might be warranted for some 2- and 3-year-old items. “Use 3 words together in a sentence” was originally proposed as a 2- to 3-year-old item and was more closely associated with the 3-year-old cluster. However, “Follow 2-step directions” was proposed as a 2- to 3-year-old item and

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

was more closely associated with the 2-year-old cluster. One 3-year item, "Understand words such as 'in,' 'on,' and 'under'" was more closely associated with the 2-year-old items when clustered by difficulty. Earlier indicated age cut-points for items "Understand words such as 'in,' 'on,' and 'under'" and "Follow 2-step directions" suggest that parents may perceive that receptive language items are attained at or before age 2 years, with success on most expressive language items occurring at later ages. This finding is consistent with current theory that development of expressive language builds upon and follows receptive language development[31] and delays in expressive language may be more observable than receptive language.[32] Reported success on some receptive language items, such as the item "Follow 2-step directions," could reflect parent perceptions of child compliance, which likely increases with age.[33] However, the wording of some NSCH pilot items might have influenced parent responses. For example, responses to the item "Use 3 words together *in a sentence*" might have been based on parental expectations for grammatically correct sentence structure, leading to less success among 2-year-old children despite examples of simple "sentences" included with the item. These results suggest additional research to establish normative language milestones for national surveys may be warranted, such as efforts to identify parental interpretation of questions, and refinement of child age data (e.g., age in months ) to allow more detailed assessments of item responses by age.

We detected indications of non-uniform DIF among five of the pilot items. Wide confidence intervals with lower-limit values close to zero prevent us from drawing conclusions about the presence of DIF among these items. Furthermore, we are unable to definitively distinguish clinically meaningful differences across groups from potential item bias. However, these findings can inform cognitive testing on NSCH items to further assess potential response bias among the 11 pilot items. Social, environmental,[34] and biological[35] influences may have contributed to marginal indications of DIF among some of the items. Detection of possible non-uniform DIF by child sex among three of the items may have been influenced by differences in parenting communication behaviors by child gender[36] or by gender differences in vocabulary growth that may have a biological basis.[35] Non-uniform DIF by sex on some receptive language items, including "Follow a verbal direction without hand gestures," may also reflect sex differences in parents' perceptions of child compliance behaviors.[33] Further research can inform understanding of why two responses differed by parent educational attainment and one item by household language; these differences may be attributed to child language disparities due to differences in exposure to a language-rich environment[37] and/or the wording of the items. These data did not indicate the presence of DIF by child race/ethnicity among any of the items.

Limitations to this study are noted. First, language data from NSCH are reliant upon parent report, and, as items have not been validated against professional language assessments, they may be subject to response bias. However, the accuracy of parent-report child-development instruments has been demonstrated elsewhere[1, 15, 16] and our test of DIF points to minimal potential for item bias among the groups assessed. Second, response rates may have resulted in a non-representative sample; however, sample weights applied to the analyses were designed to attenuate bias resulting from nonresponse.[21] Third, the 2018 NSCH did not collect birthdate or age in months data, and this data was not released for 2019, preventing investigation of proportion of item success by child age in months. Given the rapid growth

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

in typical language development in early childhood, child age in months may allow for more precision on the age cut-points for items. Specifically, ICC curve spacing and lack of overlap among 95% confidence intervals for item difficulty parameter estimates among indicated 2-year items suggest the potential for a significant delineation between the language abilities of younger and older two-year-olds. Fourth, convergence failure of the weighted 2-parameter model limited our test of model fit to unweighted data. However, difficulty parameter estimates did not differ with and without accounting for complex survey procedures in 1-Parameter models which may indicate comparability across weighted and unweighted models (Table 3 and Figure 1 [weighted]; Supplemental Table 2 and Supplemental Figure 1[unweighted]). Finally, the 2018 and 2019 NSCH language items are not intended to identify risk of or diagnose developmental delays as a developmental screening or diagnostic tool, respectively; rather, they are intended to provide information on population-level estimates of skill success.

The results of this study can inform population level monitoring of language development among young children through refinement of the NSCH data collection related to children's language development. These data may also inform collection and analysis of normative data for other developmental milestones. Age cut-points indicated here, representing approximately 90% success among U.S. children, could be applied to future research with NSCH data to describe risk and protective factors for normative language development. In future NSCH research, the results of this study can assist in defining population level age-specific standards for estimating the prevalence of children "on track" (i.e., meeting expected milestones) for early language development [19, 20, 38, 39] Over time, these data may be used to inform and monitor national and state policy and promotion efforts, such as by tracking the prevalence of children "on track" and comparing state data to national estimates.[20] Prior studies have employed normative NSCH data on other measures of early development for national and state prevalence estimates of children "on track" for school readiness and for self-regulation. [20, 38, 39]

The NSCH language items were designed to address a gap in nationally representative data on language development. This study is responsive to the call to enhance the utility of developmental surveillance tools by using large, representative samples to assess individual, normed milestone data.[17] These data offer the first evaluation of population-level early language skills measured using a nationally-representative parent-reported survey format. Future research could examine the validity of the NSCH language items against validated individualized normed milestones.

## CONCLUSION

This study represents the first psychometric analysis of language items on the 2018 and 2019 NSCH using data representative of US children ages 1–5 years. Using weighted IRT methods, these data suggest normative age cut-points were overall concordant with previously established age milestones, with minor adjustments indicated for three of the items. Our psychometric assessment offers new, normative data on parent-reported language development in a representative sample of US children which can be used to inform national and state efforts to support language development in young children and the monitoring of

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

the success of such efforts. These findings may inform future research on risk and protective factors for language and communication development among US children and inform health promotion efforts at a population level. With future validation of the items and the indicated age cut-points using data from subsequent survey years, these findings may also inform efforts to refine parent-report developmental surveillance and screening tools, enhancing public health strategies to improve early identification and intervention.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## ACKNOWLEDGEMENTS

The authors would like to extend special thanks to Melissa Danielson, MSPH, and Benjamin Zablotsky, PhD, for analytical guidance and support.

This research was supported in part by an appointment to the Research Participation Program at the CDC administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the US Department of Energy and CDC (for Helena Hutchins).

### Funding

(information that explains whether and by whom the research was supported): This research was supported in part by an appointment to the Research Participation Program at the Centers for Disease Control and Prevention administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the U.S. Department of Energy and CDC.

### Role of funder/sponsor:

The funder did not participate in the work.

## Abbreviations:

| | |
|---|---|
| **NSCH** | National Survey of Children's Health |
| **CDC** | Centers for Disease Control and Prevention |
| **HRTL** | Healthy and Ready to Learn |
| **HRSA's** | MCHB Human Resources and Security Administration's Maternal and Child Health Bureau |
| **IRT** | Item Response Theory |
| **ICCs** | Item Characteristic Curves |
| **DIF** | Differential Item Functioning |
| **CI** | Confidence Interval |

## REFERENCES

1. Berkman ND, Wallace I, Watson L, et al. Screening for Speech and Language Delays and Disorders in Children Age 5 Years or Younger: A Systematic Review for the U.S. Preventive Services Task Force Rockville (MD): Agency for Healthcare Research and Quality (US); July 2015.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

2. Pace A, Luo R, Hirsh-Pasek K, Golinkoff R. Identifying pathways between socioeconomic status and language development. Annual Review of Linguistics 2017;3:285–308. Doi: 10.1146/annurev-linguistics-011516-034226

3. Pace A, Alper R, Burchinal MR, Golinkoff RM, Hirsh-Pasek K. Measuring success: Within and cross-domain predictors of academic and social trajectories in elementary school. Early Childhood Research Quarterly 2019;46:112–125. doi:10.1016/j.ecresq.2018.04.001

4. Wilks T, Gerber RJ, Erdie-Lalena C. Developmental milestones: cognitive development. Pediatrics in Review 2010;31:364–367. doi:10.1542/pir.31.9.364 [PubMed: 20810700]

5. Young AR, Beitchman JH, Johnson C, et al. Young adult academic outcomes in a longitudinal sample of early identified language impaired and control children. Journal of Child Psychology and Psychiatry 2002;43:635–645. doi:10.1111/1469-7610.00052 [PubMed: 12120859]

6. McLaughlin MR. Speech and language delay in children. American Family Physician 2011;83:1183–1188. [PubMed: 21568252]

7. Johnson CJ, Beitchman JH, Brownlie EB. Twenty-year follow-up of children with and without speech-language impairments: family, educational, occupational, and quality of life outcomes. American Journal of Speech-Language Pathology 2010;19:51–65. doi:10.1044/1058-0360(2009/08-0083) [PubMed: 19644128]

8. Siu AL, U. S. Preventive Services Task Force. Screening for speech and language delay and disorders in children aged 5 years or younger: US Preventive Services Task Force Recommendation Statement. Pediatrics 2015;136:e474–481. doi:10.1542/peds.2015-1711. [PubMed: 26152670]

9. Council on Children With Disabilities, Section on Developmental Behavioral Pediatrics, Bright Futures Steering Committee, Medical Home Initiatives for Children With Special Needs Project Advisory Committee. Identifying Infants and Young Children With Developmental Disorders in the Medical Home: an algorithm for developmental surveillance and screening. Pediatrics 2006;118:405–420. doi:10.1542/peds.2006-1231 [PubMed: 16818591]

10. Lipkin PH, Macias MM; Council On Children With Disabilities, Section on Developmental and Behavioral Pediatrics. Promoting optimal development: identifying infants and young children with developmental disorders through developmental surveillance and screening. Pediatrics 2020;145. doi: 10.1542/peds.2019-3449

11. Daniel KL, Prue C, Taylor MK, Thomas J, Scales M. 'Learn the signs. Act early': a campaign to help every child reach his or her full potential. Public Health 2009;123 Suppl 1:e11–e16. doi:10.1016/j.puhe.2009.06.002

12. Sheldrick RC, Perrin E. Evidence-based milestones for surveillance of cognitive, language, and motor development. Academic Pediatrics 2013;13:577–586. doi: 10.1016/j.acap.2013.07.001 [PubMed: 24238685]

13. Hagan JF, Shaw JS, Duncan PM, eds. Bright Futures: Guidelines for Health Supervision of Infants, Children, and Adolescents [pocket guide] 4th ed. Elk Grove Village, IL: American Academy of Pediatrics; 2017.

14. Conti-Ramsden G, Durkin K. Language development and assessment in the preschool period. Neuropsychology Review 2012;22:384–401. doi:10.1007/s11065-012-9208-z [PubMed: 22707315]

15. Sachse S, Von Suchodoletz W. Early identification of language delay by direct language assessment or parent report? Journal of Developmental Behavioral Pediatrics 2008;29:34–41. doi:10.1097/DBP.0b013e318146902a [PubMed: 18300723]

16. Sim F, Thompson L, Marryat L, Ramparsad N, Wilson P. Predictive validity of preschool screening tools for language and behavioural difficulties: A PRISMA systematic review. PLoS One 2019;14:e0211409. doi:10.1371/journal.pone.0211409 [PubMed: 30716083]

17. Sheldrick RC, Schlichting LE, Berger B, et al. Establishing new norms for developmental milestones. Pediatrics 2019;144. doi:10.1542/peds.2019-0374

18. Moore KAP K; Pina G; Anderson S Being Healthy and Ready to Learn is linked with Family and Neighborhood Characteristics for Preschoolers www.childtrends.org: ChildTrends; 2020.

19. Paschall K, Anderson Moore K, Piña G, Anderson S. Comparing the National Outcome Measure of Healthy and Ready to Learn with Other Well-Being and School Readiness

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Measures: Child Trends; April 06,2020. https://www.childtrends.org/wpcontent/uploads/2020/04/FamilyCharacteristics_ChildTrends_April2020.pdf. Accessed December 17, 2021.

20. Ghandour RMM AK; Murphy K; Bethell C; Jones JR; Harwood R; Buerlein J; Kogan M; Lu M School readiness among U.S. children: development of a pilot measure. Child Indicators Research 2018;12:1389–1411. doi:10.1007/s12187-018-9586-8
21. United States Census Bureau. 2018 National Survey of Children's Health Methodology Report Washington, DC. 2018. https://www2.census.gov/programs-surveys/nsch/technical-documentation/methodology/2018-NSCH-Methodology-Report.pdf. Accessed December 17, 2021.
22. United States Census Bureau. 2019 National Survey of Children's Health Methodology Report Washington, DC. 2019. https://www2.census.gov/programs-surveys/nsch/technical-documentation/methodology/2019-NSCH-Methodology-Report.pdf. Accessed Decebmer 17, 2021.
23. Visser-Bochane MI, Reijneveld SA, Krijnen WP, van der Schans CP, Luinge MR. Identifying Milestones in Language Development for Young Children Ages 1 to 6 Years. Academic Pediatrics 2020;20:421–429. doi:10.1016/j.acap.2019.07.003 [PubMed: 31306799]
24. Lewis-Beck MS, Bryman A, Liao TF, eds. Encyclopedia of Survey Research Methods: "Item Response Theory": Sage Publications, Inc.; 2011. doi:10.4135/9781412950589
25. Swaminathan HR, Rogers H.J. Detecting Differential Item Functioning Using Logistic Regression Procedures. Journal of Educational Measurement 1990;27:361–370.
26. United States Census Bureau. National Survey of Children's Health, Guide to Multi-Year Analysis Washington, DC. 2019. https://www2.census.gov/programs-surveys/nsch/technical-documentation/methodology/NSCH-Guide-to-Multi-Year-Estimates.pdf. Accessed December 17, 2021.
27. Tennessee Department of Education. Speech or Language Impairment Evaluation Guidance Nashville, TN. 2018. https://www.tn.gov/content/dam/tn/education/special-education/eligibility/se_speech_or_language_impairment_evaluation_guidance.pdf. Accessed December 17, 2021.
28. Camilli G, Congdon P. Application of a Method of Estimating DIF for Polytomous Test Items. Journal of Educational and Behavioral Statistics 1999;24:323–341.
29. Kim S-H, Cohen AS, Alagoz C, Kim S. DIF Detection and Effect Size Measures for Polytomously Scored Items. Journal of Educational Measurement 2007;44:93–116.
30. StataCorp. 2021. Stata Statistical Software: Release 17 College Station, TX: StataCorp LLC.
31. Beard A Speech, language and communication: a public health issue across the lifecourse. Symposium: Social Paediatrics Vol 28: Paediatrics and Child Health; 2017:126–131. doi: 10.1016/j.paed.2017.12.004
32. Feldman HM. How Young Children Learn Language and Speech. Pediatrics in Review 2019;40:398–411. doi: 10.1542/pir.2017-0325 [PubMed: 31371633]
33. Kalb LM, Loeber R. Child disobedience and noncompliance: a review. Pediatrics 2003;111:641–652. doi:10.1542/peds.111.3.641 [PubMed: 12612249]
34. Hoff E How social context support and shape language development. Developmental Review 2006;26:55–88. doi:10.1016/j.dr.2005.11.002
35. Burman DD, Bitan T, Booth JR. Sex differences in neural processing of language among children. Neuropsychologia 2008;46:1349–1362. doi:10.1016/j.neuropsychologia.2007.12.021 [PubMed: 18262207]
36. Leaper C, Anderson KJ, Sanders P. Moderators of gender effects on parents' talk to their children: a meta-analysis. Developmenta Psychology 1998;34:3–27.
37. Rowe ML. Understanding socioeconomic differences in parents' speech to children. Child Development Perspectives 2017;12:122–127. doi:10.1111/cdep.12271
38. Ghandour RM, Hirai AH, Moore KA, et al. Healthy and Ready to Learn: Prevalence and Correlates of School Readiness among United States Preschoolers. Academic Pediatrics 2021;21:818–829. doi:10.1016/j.acap.2021.02.019 [PubMed: 33667721]
39. Claussen AH, Robinson LR, Kaminski JW, et al. Factors Associated with Self-regulation in a Nationally Representative Sample of Children Ages 3–5 Years: United States, 2016. Maternal and Child Health Journal 2021;25:27–37. doi:10.1007/s10995-020-03039-6 [PubMed: 33219911]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

**"What's New":**

A concise, parent-report pilot measure recently added to the National Survey of Children's Health offers the first nationally representative, normative data on children's language and communication development, with minimal evidence of item bias and indicating overall alignment with existing age milestones for language development.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

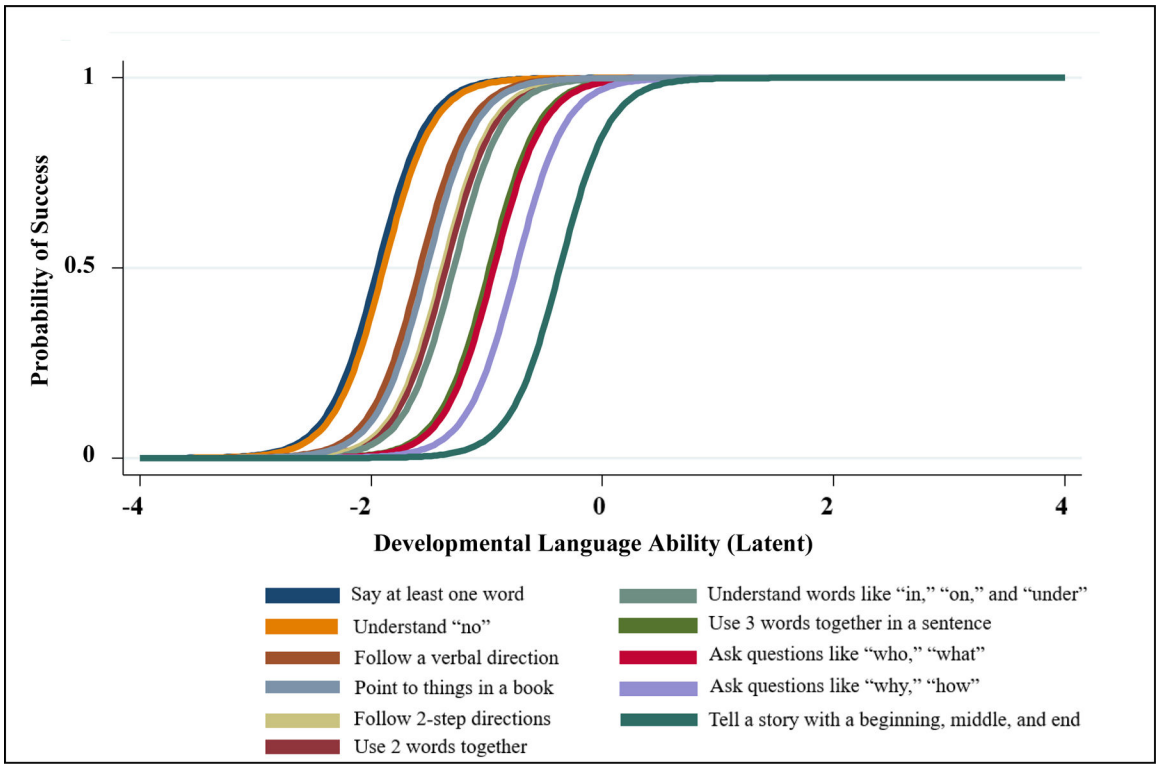


**Figure. 1-**

Weighted 1-parameter Item Response Theory Item Characteristic Curves of the Language and Communication Items for Children Age 1–5 Years—National Survey of Children's Health 2018 and 2019[a]

[a]weighted to account for complex survey sampling procedures

Author Manuscript Author Manuscript Author Manuscript Author Manuscript



**Table 1.**

Pilot Language and Communication Items for Children Age 1–5 Years on the 2018 and 2019 National Survey of Children's Health

| **Question:** ***Is this child able to do the following…*** | **Item Origin:** | **Language Type:** | **Proposed Child Age:** |
|---|---|---|---|
| Say at least one word, such as "hi" or "dog"? | LTSAE | Expressive | 1 year |
| Understand the meaning of the word "no"? | LTSAE[a] | Receptive | 1 year |
| Follow a verbal direction without hand gestures, such as "Wash your hands."? | LTSAE/BF[b] | Receptive | 2 years |
| Point to things in a book when asked? | LTSAE | Receptive | 2 years |
| Use 2 words together, such as "car go"? | LTSAE | Expressive | 2 years |
| Use 3 words in a sentence together, such as "Mommy come now."? | LTSAE/BF | Expressive | 2–3 years |
| Follow 2-step directions, such as "Get your shoes and put them in the Basket."? | LTSAE/BF | Receptive | 2–3 years |
| Understand words such as "in," "on," and "under"? | LTSAE/BF | Receptive | 3 years |
| Ask questions like "who," "what," "when," "where"? | TN/VA Dept of Ed[c] | Expressive | 3 years |
| Ask questions like "why" and "how"? | TN/VA Dept of Ed | Expressive | 4 years |
| Tell a story with a beginning, middle, and end? | LTSAE | Expressive | 5 years |

[a] Learn the Signs. Act Early.[11]

[b] Bright Futures Guidelines[13]

[c] Tennessee & Virginia Dept. of Education Guidance[27]

*Acad Pediatr*. Author manuscript; available in PMC 2023 September 01.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript



*Acad Pediatr*. Author manuscript; available in PMC 2023 September 01.

**Table 2.**

Unweighted Frequencies and Weighted Percentages of Reported Success Stratified by Age for Each of the Language and Communication Items for Children Age 1–5 Years on the 2018 and 2019 National Survey of Children's Health

| | Unweighted | | Weighted[b] | | | | |
|---|---|---|---|---|---|---|---|
| **Item** | **Success (Frequency) N = 14,573** | **Success % (95% CI)[a]** | **Success Among 1 Year Olds %(95% CI)** | **Success Among 2 Year Olds (%) (95% CI)** | **Success Among 3 Year Olds (%) (95% CI)** | **Success Among 4 Year Olds (%) (95% CI)** | **Success Among 5 Year Olds (%) (95% CI)** |
| Say at least one word | 14,274 | 96.9 (96.2 – 97.5) | 90.3 (87.1 – 92.9) | 98.2 (97.0 – 99.0) | 98.8 (98.1 – 99.3) | 98.4 (97.4 – 99.1) | 98.8 (97.9 – 99.3) |
| Understand "no" | 14,210 | 96.5 (95.7 – 97.2) | 90.3 (87.3 – 92.8) | 96.9 (95.0 – 98.3) | 98.4 (97.3 – 99.2) | 98.6 (97.6 – 99.2) | 98.2 (96.7 – 99.1) |
| Follow a verbal direction | 13,821 | 92.7 (91.7 – 93.7) | 72.4 (68.3 – 76.2) | 96.4 (95.1 – 97.5) | 97.7 (96.4 – 98.7) | 98.8 (98.0– 99.3) | 97.9 (96.3 – 98.9) |
| Point to things in a book | 13,662 | 91.8 (90.9 – 92.7) | 69.0 (65.0 – 72.7) | 95.9 (94.7 – 96.9) | 98.1 (97.2 – 98.8) | 98.3 (97.4 – 99.0) | 97.3 (95.8 – 98.5) |
| Follow 2-step directions | 13,387 | 89.6 (88.5 – 90.6) | 60.3 (56.2 – 64.2) | 94.3 (92.4 – 95.8) | 97.8 (96.5 – 98.7) | 97.8 (96.4 – 98.8) | 97.0 (95.3 – 98.2) |
| Use 2 words together | 13,254 | 88.8 (87.7 – 89.9) | 55.2 (51.1 – 59.2) | 94.1 (92.3 – 95.6) | 97.8 (96.7 – 98.6) | 98.5 (97.6 – 99.2) | 97.9 (96.4 – 98.8) |
| Understand words such as "in," "on," and "under" | 13,285 | 87.4 (86.0 – 88.6) | 55.7 (51.5 – 59.9) | 90.3 (87.4 – 92.8) | 95.4 (93.4 – 96.9) | 97.3 (95.8 – 98.4) | 97.4 (95.9 – 98.5) |
| Use 3 words together in a sentence | 12,244 | 79.8 (78.3 – 81.2) | 24.8 (21.5 – 28.4) | 82.0 (78.5 – 85.1) | 95.4 (94.6 – 96.1) | 98.2 (97.4 – 98.9) | 97.5 (96.1 – 98.5) |
| Ask questions like "who," "what," | 12,029 | 78.8 (77.3 – 80.2) | 25.9 (22.5 – 29.5) | 80.7 (77.9 – 83.2) | 94.6 (93.2 – 95.8) | 95.4 (93.1 – 97.1) | 96.4 (94.7 – 97.7) |
| Ask questions like 'why," "how" | 11,233 | 73.1 (71.5 – 74.6) | 15.5 (12.8 – 18.6) | 66.5 (62.9 – 69.9) | 91.0 (88.7 – 93.1) | 95.5 (93.7 – 97.0) | 95.7 (93.8 – 97.1) |
| Tell a story with a beginning, middle, and end | 9,476 | 59.7 (58.0 – 61.4) | 2.4 (1.3 – 4.0) | 35.4 (32.1 – 38.8) | 79.5 (76.2 – 82.5) | 88.2 (85.1 – 90.8) | 92.5 (90.2 – 94.4) |

Unweighted sample sizes by age: 1-year-olds: 2,197; 2-year-olds: 3,220; 3-year-olds: 3,113; 4-year-olds: 3,027; 5-year-olds: 3,016.

[a] CI =Clopper-Pearson Confidence Interval

[b] Weighted to account for complex survey sampling procedures.

[c] Horizontal borders represent indicated cut-points by child age (1–5 years) based on IRT difficulty parameter estimates and corresponding ICC curves, consistent with approximately 90% success on each item.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

**Table 3.**

1-Parameter IRT Item Parameter Values of the Language and Communication Items for Children Age 1–5 Years—National Survey of Children's Health 2018 and 2019[a]

| | **Discrimination** | **(95% CI)[b]** |
|---|---|---|
| Constant | 4.67 | (4.36, 4.99) |
| | **Difficulty** | **(95% CI)** |
| Say at least one word | −1.95 | (−2.04, −1.83) |
| Understand "no" | −1.89 | (−1.98, −1.81) |
| Follow a verbal direction | −1.58 | (−1.64, −1.51) |
| Point to things in a book | −1.52 | (−1.58, −1.46) |
| Follow 2-step directions | −1.39 | (−1.45, −1.34) |
| Use 2 words together | −1.36 | (−1.41, −1.31) |
| Understand words such as "in," "on," and "under" | −1.29 | (−1.34, −1.23) |
| Use 3 words together in a sentence | −0.98 | (−1.02, −0.94) |
| Ask questions like "who," "what" | −0.94 | (−0.98, −0.90) |
| Ask questions like 'why," "how" | −0.75 | (−0.78, −0.71) |
| Tell a story | −0.37 | (−0.40, −0.34) |

[a]Weighted to account for complex survey sampling procedures.

[b]CI = Wald Confidence Interval

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

**Table 4.**

Differential Item Functioning Analysis of Language and Communication Items for Children Age 1–5 Years—National Survey of Children's Health 2018 and 2019[a]

| | | Uniform DIF Coefficients[b] | | Non-uniform DIF Model 3 Interaction Term | |
|---|---|---|---|---|---|
| **Item** | **Group** | **Model 1** | **Model 2** | ***P*-value[c]** | **(95% CI[d])** |
| Say at least one word | Sex | 3.19 | 3.19 | 0.28 | (−0.50, 1.73) |
| | Race/ethnicity | 3.19 | 3.19 | 0.56 | (−0.79, 1.47) |
| | Household Language | 3.19 | 3.19 | 0.03* | (0.37, 6.04) |
| | Caregiver Education | 3.19 | 3.19 | 0.88 | (−1.34, 1.55) |
| Understand "no" | Sex | 2.81 | 2.83 | 0.77 | (−0.79, 1.07) |
| | Race/ethnicity | 2.81 | 2.80 | 0.94 | (−0.96, 0.90) |
| | Household Language | 2.81 | 2.80 | 0.64 | (−1.67, 1.03) |
| | Caregiver Education | 2.81 | 2.84 | 0.57 | (−1.40, 0.77) |
| Follow a verbal direction | Sex | 4.87 | 4.91 | 0.04* | (0.09, 2.56) |
| | Race/ethnicity | 4.87 | 4.86 | 0.55 | (−1.60, 0.85) |
| | Household Language | 4.87 | 4.86 | 0.34 | (−2.26, 0.79) |
| | Caregiver Education | 4.87 | 4.86 | 0.01* | (−3.03, −0.52) |
| Point to things in a book | Sex | 4.51 | 4.50 | 0.06 | (−0.06, 2.02) |
| | Race/ethnicity | 4.51 | 4.51 | 0.46 | (−1.42, 0.64) |
| | Household Language | 4.51 | 4.60 | 0.17 | (−2.27, 0.40) |
| | Caregiver Education | 4.51 | 4.52 | 0.58 | (−1.83, 1.03) |
| Follow 2-step directions | Sex | 5.15 | 5.23 | 0.46 | (−0.85, 1.86) |
| | Race/ethnicity | 5.15 | 5.16 | 0.51 | (−1.79, 0.89) |
| | Household Language | 5.15 | 5.14 | 0.24 | (−2.54, 0.64) |
| | Caregiver Education | 5.15 | 5.15 | 0.77 | (−2.02, 2.72) |
| Put 2 words together | Sex | 4.98 | 4.98 | 0.04* | (0.73, 2.89) |
| | Race/ethnicity | 4.98 | 5.01 | 0.15 | (−2.35, 0.36) |
| | Household Language | 4.98 | 4.97 | 0.62 | (−2.89, 1.72) |
| | Caregiver Education | 4.98 | 4.98 | 0.69 | (−1.96, 1.29) |
| Understand words such as "in," "on," and "under" | Sex | 4.62 | 4.67 | 0.42 | (−0.54, 1.28) |
| | Race/ethnicity | 4.62 | 4.70 | 0.08 | (−1.73, 0.09) |
| | Household Language | 4.62 | 4.69 | 0.90 | (−1.16, 1.32) |
| | Caregiver Education | 4.62 | 4.66 | 0.03* | (−1.97, −0.84) |
| Use 3 words together in a sentence | Sex | 7.08 | 7.08 | 0.70 | (−1.99, 1.33) |
| | Race/ethnicity | 7.08 | 7.08 | 0.07 | (−2.83, 0.13) |

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

| | | Uniform DIF Coefficients[b] | | Non-uniform DIF Model 3 Interaction Term | |
|---|---|---|---|---|---|
| **Item** | **Group** | **Model 1** | **Model 2** | ***P*-value[c]** | **(95% CI[d])** |
| | Household Language | 7.08 | 7.10 | 0.55 | (−3.16, 1.68) |
| | Caregiver Education | 7.08 | 7.09 | 0.23 | (−3.09, 0.73) |
| Ask questions like “who,” “what” | Sex | 7.26 | 7.25 | 0.53 | (−1.07, 2.08) |
| | Race/ethnicity | 7.26 | 7.26 | 0.16 | (−2.52, 0.41) |
| | Household Language | 7.26 | 7.26 | 0.26 | (−3.07, 0.83) |
| | Caregiver Education | 7.26 | 7.30 | 0.58 | (−2.39, 1.34) |
| Ask questions like “why,” “how” | Sex | 7.41 | 7.41 | 0.045 * | (0.03, 3.00) |
| | Race/ethnicity | 7.41 | 7.46 | 0.07 | (−2.75, 0.11) |
| | Household Language | 7.41 | 7.43 | 0.48 | (−2.45, 1.14) |
| | Caregiver Education | 7.41 | 7.53 | 0.26 | (−0.81, 3.02) |
| Tell a story | Sex | 7.18 | 7.24 | 0.28 | (−0.45, 1.56) |
| | Race/ethnicity | 7.18 | 7.21 | 0.20 | (−1.64, 0.34) |
| | Household Language | 7.18 | 7.20 | 0.39 | (−1.89, 0.74) |
| | Caregiver Education | 7.18 | 7.19 | 0.52 | (−0.92, 1.80) |

[a] All analyses were weighted.

[b] A cutoff of >10 % change to coefficients from Model 1 to Model 2 was used to indicate potential uniform DIF.

[c] A cutoff of $P < 0.05$ for the interaction term in Model 3 was used to indicate potential non-uniform DIF.

[d] CI = Wald Confidence Interval

* Possible non-uniform DIF detected