ATLAS-PA Technical Manual

August 2023

Please email skibbelo@msu.edu with any questions.
Principal Investigator: Lori E. Skibbe, Ph.D.
Co-Principal Investigators: Gary A. Troia, Ph.D.; Ryan P. Bowles, Ph.D.
Prepared by Sarah Goodwin, Ph.D., Research Associate

Table of contents

Acknowledgments

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through grant R324A150063 (PI: Skibbe). The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

The research team also wishes to acknowledge everyone who made the project possible:

- Children and their families who participated
- Project coordinators who trained RAs and served as liaisons with school sites
- Research assistants
- Item developers and reviewers who served as content experts

- Web developers and illustrators
- School and childcare center personnel:
  - Classroom teachers
  - Paraprofessionals
  - Speech-language pathologists
  - Occupational therapists
  - Principals, superintendents, and all key school administrators

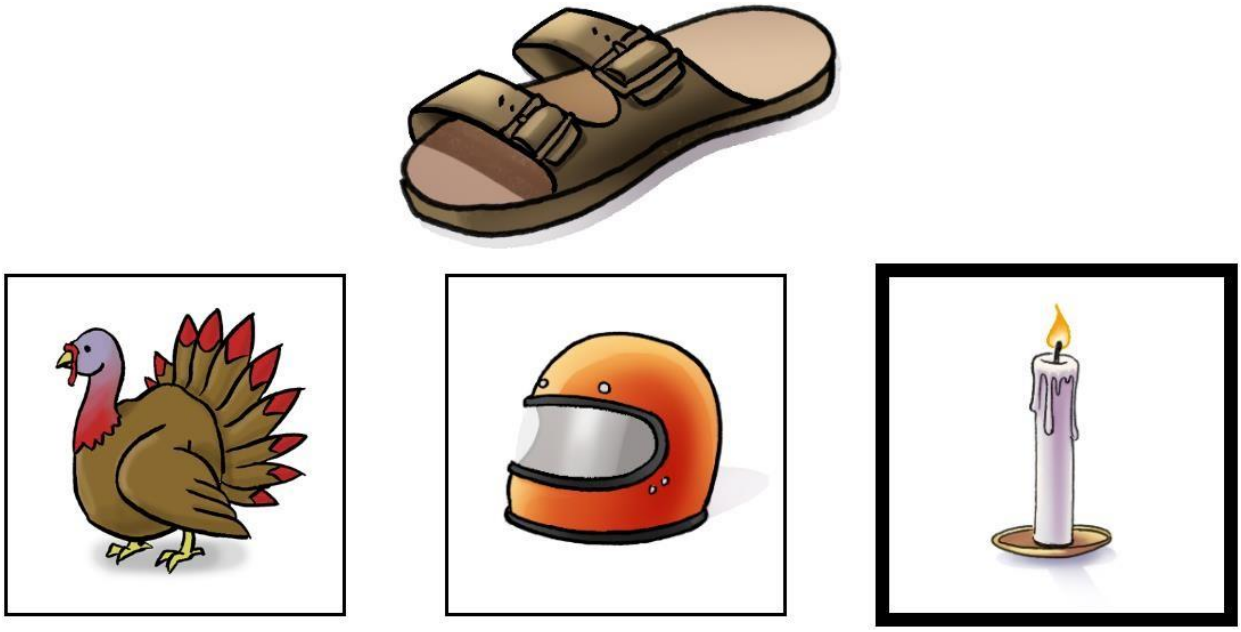
Overview of ATLAS

The ATLAS Phonological Awareness (ATLAS-PA) section is a computer-adaptive assessment of phonological awareness (PA), or the underlying understanding of the sound structure of language. ATLAS-PA was developed for children ages 3;0 to 7;11, including those who have disabilities related to speech-language production. It requires only receptive item responses, so children do not have to speak to provide their answers.

ATLAS-PA requires an internet connection and runs in the Google Chrome browser on a desktop or laptop computer or Android tablet. Developers at Matrix, the Center for Digital Humanities and Social Sciences at MSU, worked with the research team to program ATLAS-PA. It was written in PHP 7, HTML 5, Javascript, and MySQL. Because the assessment is web based, there is no need for software installation, aside from ensuring that Google Chrome is installed and operating system software is up to date.

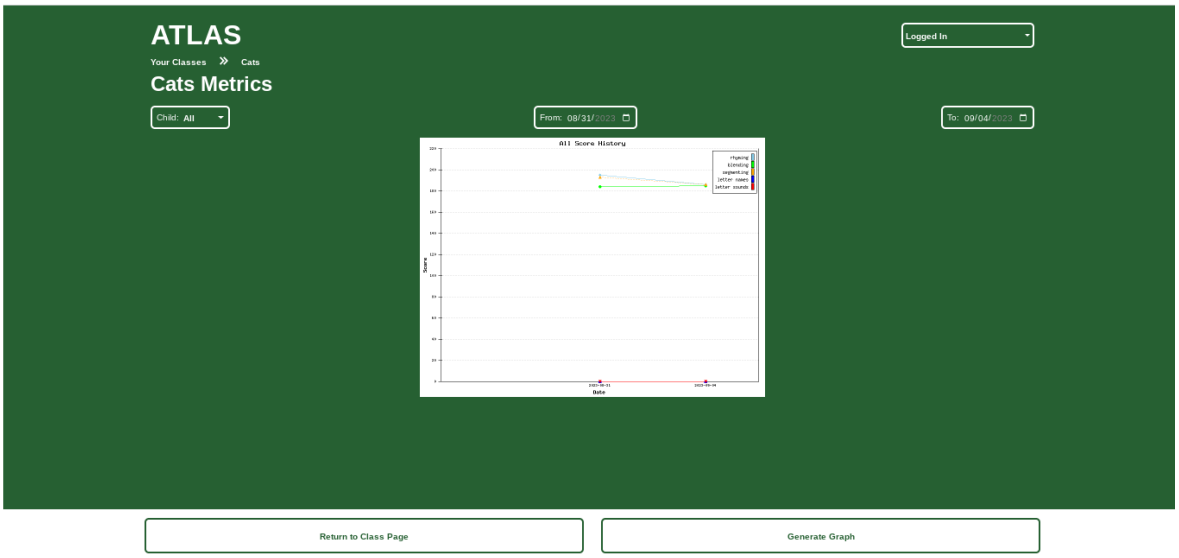
There are three subtests included as part of ATLAS-PA: rhyming (e.g., What rhymes with *cat*?), blending (e.g., Put *pan… cake* together), and segmentation (e.g., What's *keyboard* without -*board*?). There are no nonsense words included for any tasks. All items present three concrete picture representations of English words, side by side, on a plain white background. The tablet or computer plays a recording where a speaker of Midwestern U.S. English reads the prompt (e.g., "What rhymes with *bee*?"), followed by the name of each option (e.g., "blue, key, mop"); a solid black border appears around each image as its name is spoken. Children can use a touchscreen or mouse to tap or click one of the three options, or an assessor may help them select their choice. Children receive corrective feedback during the practice items, but no feedback is provided once the testing module is launched. The audio stimulus plays twice total: it first plays once, followed by a pause, then encouraging language ("Now *you* try it!"). After that, the stimulus question repeats, and if a child gives no response, the next item is given.

A sample Rhyme item can be seen here. The speaker prompts, "What rhymes with *sandal*? Turkey. Helmet. Candle," while the on-screen images are displayed:



Sample visual display of Rhyming item

Administration consists of 15 or fewer items per subtest, with most children needing fewer than 10 items to complete a subtest. The exact number of items depends on the adaptive test algorithms, as items are targeted at a child's PA ability level (see scoring section for more information). Results are converted to scaled scores, along with scoring interpretations showing ATLAS-PA results linked to other published PA assessments. Results are provided for individual children as well as for classes or groups of students. Testing dates are indicated with results presented graphically so that users can monitor progress over time.



Sample progress monitoring page of score results

Construct/Domain assessed: Phonological awareness

The phonological awareness (PA) section of ATLAS is designed to assess English PA. PA is the underlying sensitivity to and ability to manipulate the sound structure of language. It can involve analysis (breaking apart spoken stimuli into smaller components), synthesis (putting together smaller units into larger components), or epilinguistic skills (word rhyming, alliteration). In English, tasks related to rhyming, blending, and segmentation represent a unidimensional construct (Anthony et al., 2002; Anthony & Lonigan, 2004; Schatschneider et al., 1999; Stahl & Murray, 1994). Developmentally, children tend to acquire skills of rhyming earlier, followed by

syllable-level or compound-word synthesis. Later, most children form analysis skills, with phoneme-level blending or segmentation being the most difficult to acquire and the latest learned (McConnell et al., 2019; Roth et al., 2002; Roth et al., 2006; Torgesen et al., 1992).

PA is an important predictor of later reading skills (McDowell et al., 2007; Lonigan et al., 2008). This predictive relationship of PA to later word recognition (Lonigan, 2004; Perfetti et al., 1987; Torgesen et al., 1994) holds even when controlling for overall IQ, socioeconomic status, verbal memory, and vocabulary knowledge (Bryant et al., 1990; MacLean et al., 1987; Wagner & Torgesen, 1987; Wagner et al., 1994). There is evidence that low levels of PA are associated with short- and long-term difficulties in learning to read (Byrne & Fielding-Barnsley, 1995; Card & Dodd, 2006; Dahlgren Sandberg, 2006; Dahlgren Sandberg & Hjelmquist, 1997; Larrivee & Catts, 1999; Stothard et al., 1998). Preschoolers with speech sound disorders have lower PA knowledge than their peers with typical language development (Anthony et al., 2011). Moreover, PA not only facilitates, but also has indirect effects on, growth in letter name and sound knowledge and reading growth (Burgess & Lonigan, 1998). When children play with sounds, their PA development is boosted (Peeters et al., 2009).

Many children have difficulties with speech production in early childhood. In 2018-19, 14% of U.S. public school students between the ages of 3 to 21 received services for special education under the Individuals with Disabilities Education Act; of that population, approximately one in five (19%) services were for speech or language impairment (U.S. Department of Education, 2020). Children who begin kindergarten with underdeveloped language skills may struggle with school literacy demands. Moreover, kindergarteners with language impairment whose spoken language does not improve show evidence of lagging behind their peers in second- and fourth-grade oral language skills that build their reading ability (Hulme et al., 2015; Lervåg et al., 2018; Snowling et al., 2020; Storch & Whitehurst, 2002). A focus on PA early in instruction, including explicit attention to it in preschool or Head Start curricula (Skibbe et al., 2016), may help reduce the risk of later speech production and/or reading difficulties. However, several PA assessments prior to 2015 were not validated with children with disabilities and have testing barriers such as requiring children to respond out loud to test items. ATLAS-PA was validated with both children with typical development and children with disabilities that affect their speech-language production, with key assessment development considerations.

Item development and piloting

ATLAS-PA underwent a rigorous development process. First, three consultants with expertise related to children with disabilities provided input on item design. Items are presented on a plain white background, given that many children with disabilities have challenges attending to

relevant stimuli when presented electronically (Thompson et al., 2019). In order to help manage children's attention during the task, item prompts were highlighted with a thick black frame when labeled. However, examiners noted that some children lost attention when required to wait for all picture options to be labeled.

Ten early childhood experts designed and reviewed ATLAS-PA items and provided input on how to structure instructions associated with entry into the test. The panel consisted of specialists in language and literacy development, speech-language pathology, and psychometrics. All panel members had a connection with one of two university preschools that serve around 200 children annually and train preservice early childhood teachers. The preschools include children whose families pay tuition, those who receive early childhood special education services, and those eligible for Head Start or Great Start Readiness programs. Following their recommendations, ATLAS-PA was developed with three types of practice opportunities to allow for entry into the test: Basic, Basic+ ("Basic Plus"), and Enhanced.

The Basic level of instructions can be used with children who have few behavioral issues and likely know basic PA concepts. The second level, Basic+, is for if a child has one or two behavioral or cognitive challenges but is a relatively independent test taker. The most supportive level, Enhanced, helps children who need individualized presentation of tasks with behavioral support from a caregiver or teacher.

The Basic level of instructional support presents practice items to children with brief feedback. Basic+ and Enhanced levels of instruction both offer more support on practice items. For example, instructions for the Basic level ask children to identify the correct answer without touching other on-screen options. However, for Basic+ and Enhanced levels of instruction on practice items, children are asked to tap or click on the different options on the screen (e.g., "Touch snowman… Touch lunchbox"), then focus on identifying the correct answer ("What do you get when you put lunch… box… together to make one word?"). Additional examples are provided to help children understand the concept of rhyming, blending, or segmentation, depending on the subtest. For the Enhanced level of instruction, including during all of the scored items, the next question does not appear until the assessor selects the green arrow in the upper right to advance test items.

After items were crafted, two small usability studies were conducted, one with 59 children and another with 64 participants. Based on the results of the user testing, a design edit was made so that children did not have to wait for all of the response options to be spoken by the interface before they could choose their answer. In the final version of ATLAS-PA, children can select their response immediately after the question prompt, even before all three options for the picture names are spoken. This allows children to move more quickly to the next item and maintain engagement, focus, and attention.

Calibration and item selection, participants, procedures

Children ages 3;0 to 7;11 took ATLAS-PA from late summer 2016 to fall 2018. The research project was classified exempt under category 1, normal educational practices, by MSU's Institutional Review Board. Recruitment took place across schools and childcare centers located in the U.S. Midwest. Parents or guardians completed a demographic questionnaire and provided written consent approximately one week before research assistants visited schools. Testing was conducted in a quiet space at the school or childcare center. At testing time, children gave verbal and/or nonverbal assent to participate. Families received a $10.00 USD gift card to a local large retail establishment and a children's storybook to thank them for their participation.

The two groups of children included those exhibiting typical development ($n_1 = 998$) and those with disabilities affecting speech/language production ($n_2 = 280$). The group with typical development also contained multilingual children whose English proficiency was sufficient for them to take ATLAS-PA, as measured by PPVT receptive vocabulary ability or parent report. Appendix A provides the demographic information for participants, including the numbers of participants in each category of disabilities affecting speech-language production. Some families of children with typical development (n = 228) reported languages in addition to English that they used in the home. Languages included Arabic, Chinese, Spanish, and 47 other languages with fewer than 10 participants who used each; reporting guidelines for State of Michigan school districts require us to suppress sample sizes of minority language groups because that may be identifying information.

Membership in the group of children with disabilities was determined by whether a child had an individualized education or family service plan (IEP/IFSP) for speech and/or language support. The speech-language impairment information was collected from parents and occasionally corroborated by a speech teacher or other school system administrative contact. It is unknown whether children with an IEP/IFSP also had a clinical diagnosis of speech-language impairment (Ireland & Conrad, 2016; Selin et al., 2019). During testing, we observed how children interacted with the early literacy tasks, and those observations were noted by the project manager and research assistants, all who have experience assessing, and in several cases teaching, children with special needs. We brought to the research site any notes from the parent questionnaire and used those when we arrived, to help build rapport with the child, such as supports the parents

recommended help their children learn (e.g., a toy the child enjoyed, a space with few auditory/visual distractions, etc.).

For test administration, assessments were given during one to three separate sessions, depending on class time and to ensure that children were not fatigued. The starting item on the measures of PA other than ATLAS were selected based on the age or grade range appropriate for each child with typical development.

Children received the following measures of phonological awareness. Children exhibiting typical development received two of the three ATLAS-PA subtests and at least two other published PA measures. Children with speech difficulties took all three ATLAS-PA subtests and were not administered other published batteries that required language production.

The other PA measures administered to children with typical development included:

- the *Test of Preschool Early Literacy*: *TOPEL* Phonological Awareness subtest (Lonigan et al., 2007),
  or
- the *Comprehensive Test of Phonological Processing*, 2nd ed.: *CTOPP-2* Phonological Awareness subtest (Wagner et al., 2013);

and

- the *Preschool Early Literacy Indicators*: *PELI* Word Parts and First Sound for preschoolers (Kaminski et al., 2018), with expressive tasks for 3- and 4-year-olds,
  or
- the *Dynamic Indicators of Basic Early Literacy Skills*: *DIBELS Next* for grades K, 1, and 2 (Good et al., 2011), with expressive tasks that are timed: First Sound Fluency for kindergarteners and Phoneme Segmentation Fluency: all kindergarten, grade 1, and grade 2 children.

ATLAS-PA data analysis procedures

ATLAS-PA items were analyzed using Rasch measurement, probabilistic modeling that allows for the mapping of examinee ability and item difficulty on a shared interval scale (Bond & Fox, 2015; Rost, 2001). The model we employed was the Rasch (1960) approach to dichotomous outcomes:

$$\ln (P_{ni} / (1 - P_{ni})) = \theta_n - \beta_i$$

where P_{ni} is the probability of examinee n with trait level θ_n (i.e., PA ability) succeeding on item i which has difficulty level β_i. We used the Rasch measurement software Winsteps (Linacre, 2020b) to analyze item response data. To mitigate guessing, because children have a 1-in-3 chance of a correct response to these three-option multiple-choice items, we employed an item response cutoff technique to treat item responses of below .27 expected probability as missing values (in Winsteps, using the command CUTLO = -1).

Initially, 120 ATLAS-PA items were designed: 40 each in rhyming, blending, and segmentation subtests. During fall 2017 data collection with children exhibiting typical development, after items had approximately 300 responses each, we ran preliminary Rasch IRT analyses. This was done to determine how well the items were targeted to these children's ability levels and whether items fit the model well. Two items had highly unpredictable responses and were removed from the item pool: rhyming *lace* (with "face") and segmenting *plate* (to "play"). *Lace* may not have been an illustration easily recognized by young children, and we could not determine a clear reason why the item *plate* was so unpredictable. All 118 items otherwise had an acceptable range of difficulty levels and model fit.

With the remaining ATLAS-PA items, data collection was completed with typically-developing children and children with speech-language IEP/IFSPs. After these two data collection phases ended in fall 2018, item performance was reanalyzed to determine the active item pool for the final measures.

The computer-adaptive testing (CAT) algorithms were designed to accommodate both active items and experimental items. Active items contribute to ability estimation, while experimental items are administered to gauge item difficulty and do not factor into ability estimation. The experimental items are embedded at approximately every fifth position. Appendix B describes the CAT procedures.

Validity, reliability, item analyses

We used a series of processes to analyze the validity and reliability of ATLAS-PA. First, we conducted a principal components analysis (PCA) of the residuals to determine whether items were as related to each other as expected, measuring a unidimensional construct of phonological awareness (PA) ability. Second, we checked summary statistics, including Rasch reliability (https://www.winsteps.com/winman/reliability.htm) to determine the internal consistency of ATLAS-PA. Third, we investigated infit and outfit statistics to check that items fit the Rasch model well. Fourth, we ran two separate differential item functioning (DIF) analyses for 1) gender category and 2) ability group to ensure items did not measure PA differently for girls and

boys, or for children with typical development compared to children with speech-language IEPs. Lastly, in R programming language (R Core Team, 2020), we correlated ATLAS-PA with the other published assessments of PA, where appropriate for children's ages: correlations ranged from .36 to .65 between ATLAS-PA and TOPEL, CTOPP-2, DIBELS Next, and PELI. See Appendix C for correlation matrix.

The principal components analysis (PCA) of residuals investigates whether some items share a secondary measurement dimension, or similar patterns of unexpectedness, after accounting for the first dimension. To determine whether ATLAS-PA showed multidimensionality, we checked the size of eigenvalues, elbow of the scree plot, and disattenuated correlations for person ability. We also considered the item content and subskill (rhyme, blend, segment) assessed. There was weak evidence of multidimensionality, with the first eigenvalue of 3.1 but no clear elbow in the scree plot with the next four eigenvalues being 2.8, 2.4, 2.0, and 1.8. Eigenvalues less than 2.0 are ideal; however, with the first eigenvalue being 3.1, or about the strength of 3 items, we checked the contrast plots to determine whether items on each contrast shared content that differed dramatically from other items. The first principal component appeared to separate rhyming items from blending items, but the disattenuated correlation was 1.00 between theta as estimated separately from blending items alone versus from rhyming items alone. Moreover, no other content differences were identified in other principal components, and no disattenuated correlation was below .89, so we are considering ATLAS-PA to be unidimensional enough for our purpose. Summary statistics, including dimensionality, fit, and other Rasch-based output for the 110 items in the ATLAS-PA complete pool as of June 2020, can be found in Appendix D.

Next, we checked Rasch item fit, considering values from 0.6 to 1.4 (1.0 is ideal) to indicate good fit to the Rasch model. We examined both infit and outfit mean-square values: outfit is more sensitive to those persons who display response patterns very different from the rest of examinees. Values below 0.6 indicate that the item does not provide as much measurement information as expected (overfit), and items with values higher than 1.4 do not fit well and/or have responses that show more variation than anticipated (misfit). All 118 remaining items had acceptable **infit** values between 0.6 and 1.4, while four of these items had **outfit** values above 1.4, displaying misfit. Responses were unpredictable on one blending (/k/ + /aʊ/ = *cow*, outfit mean-square statistic = 1.58) and three segmentation items ("cable" to *bull*, outfit = 1.45; "walrus" to *wall*, outfit = 1.43; and "cape" to *ape*, outfit = 1.41). We removed these four items from the pool, leaving 114 PA items.

For gauging whether items were measuring PA in the same way for different groups of interest controlling for overall PA level, we then ran differential item functioning (DIF) analyses. We conducted two sets of analyses, one for gender and another for IEP status (membership in group of children with typical development, or membership in group with disabilities impacting speech-language production). As per Educational Testing Service recommendations in a Rasch

measurement framework (Linacre, 2020a; Zwick et al., 1999), we identified whether items showed DIF if they had at least .64 logits' difference in difficulty. This difference in item functioning indicates that, when person ability and item facility are well matched, the probability of success on that item differs by at least .12 between the two groups. After conducting the following DIF analyses, we removed four items, leaving a final ATLAS-PA item pool of 110 items (39 rhyming, 35 blending, and 36 segmentation).

Scoring and interpretation

The ATLAS-PA results in this section refer to the version children took before the test became adaptive. Estimated reliability of the June 2020 ATLAS-PA item pool is .97; however, since an adaptive version is now active, a reliability of .90 or above is automatically targeted by the CAT algorithms.

Children's PA ability ranged from -4.09 to 5.91 logits, while items ranged from -1.25 to 1.80, indicating a wider span of children than of items. ATLAS-PA items are best matched to PA ability levels of 3-, 4-, and 5-year-olds with typical development and 3- to 7-year-olds with individualized education plans targeting speech-language needs, so a wider range of difficulty of items is in development in order to tap PA ability for typically-developing 6- and 7-year-olds.

Interval-scale scores in logits from Winsteps were converted to an M-score, with a mean of 200 and standard deviation of 20. ATLAS reports these M-scores to users for easier interpretability. Scaled scores are used rather than raw scores because examinees did not see every item, and items vary in difficulty. As anticipated, children with speech-language IEPs do score lower overall (19.59 *M*-score) than children with typical development (20.11 *M*-score). Girls and boys did not score significantly differently on ATLAS-PA. We checked this before and after we ran analyses for DIF. The 110 PA items, as of June 2020, do not display DIF for gender category or disability status; in other words, phonological awareness is being measured in the same way for different groups. PA scores by child groups can be seen in Appendix E: by children's age group and mother's education level. Older children do outperform younger children, but results do not differ among children whose mothers have different educational experiences.

Please see our published manuscript in *Language, Speech, and Hearing Services in Schools* at https://pubs.asha.org/doi/10.1044/2020_LSHSS-19-00006 (Skibbe et al., 2020), which was conducted after data collection with 938 children with typical development and 227 with speech-language disabilities.
ATLAS also contains prototype Letter Knowledge subtests, with receptive-skills items currently being validated. This extends the work with expressive-response letter knowledge assessment

validation by Tortorelli et al. (2017). Further information on these subtests will be provided as they become available.

Further updates to the ATLAS will be available at accesstoliteracy.com.

References

Anthony, J. L., & Lonigan, C. J. (2004). The nature of phonological sensitivity: Converging evidence from four studies of preschool and early grade-school children. *Journal of Educational Psychology, 96,* 43-55.

Anthony, J. L., Aghara, R. G., Dunkelberger, M. J., Anthony, T. A., Williams, J. M., & Zhang, Z. (2011). What factors place children with speech sound disorders at risk for reading problems? *American Journal of Speech-Language Pathology, 20*, 146-160.

Anthony, J. L., Lonigan, C. J., Burgess, S. R., Driscoll, K., Phillips, B. M., & Cantor, B.G. (2002). Structure of preschool phonological sensitivity: Overlapping sensitivity to rhyme, words, syllables, and phonemes. *Journal of Experimental Child Psychology, 82,* 65-92.

Bishop, D. V., Brown, B. B., & Robson, J. (1990). The relationship between phoneme discrimination, speech production, and language comprehension in cerebral-palsied individuals. *Journal of Speech, Language, and Hearing Research, 33*(2), 210-219.

Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd ed.). Routledge.

Bryant, P. E., MacLean, M., Bradley, L. L., & Crossland, J. (1990). Rhyme and alliteration, phoneme detection, and learning to read. *Developmental Psychology, 26,* 429-438.

Burgess, S. R., & Lonigan, C. J. (1998). Bidirectional relations of phonological sensitivity and prereading abilities: Evidence from a preschool sample. *Journal of Experimental Child Psychology*, *70*, 117-141.

Card, R., & Dodd, B. (2006). The phonological awareness abilities of children with cerebral palsy who do not speak. *Augmentative and Alternative Communication, 22*, 149-159.

Dahlgren Sandberg, A., & Hjelmquist, E. (1997). Language and literacy in nonvocal children with cerebral palsy. *Reading and Writing, 9*, 107–133.

Dahlgren Sandberg, A. (2006). Reading and spelling abilities in children with severe speech impairments and cerebral palsy at 6, 9, and 12 years of age in relation to cognitive development: a longitudinal study. *Developmental Medicine & Child Neurology*, *48*, 629-634.

Dunn, L. M., & Dunn, D. M. (2007). *PPVT-4: Peabody picture vocabulary test*. Pearson Assessments.

Foley, B. E., & Pollatsek, A. (1999). Phonological processing and reading abilities in adolescents and adults with severe congenital speech impairments. *AAC: Augmentative and Alternative Communication*, *15*, 156–173.

Good, R. H., III, Kaminski, R. K., & Cummings, K. (2011). *DIBELS Next*. Dynamic Measurement Group.

Hulme, C., Nash, H. M., Gooch, D., Lervåg, A., & Snowling, M. J. (2015). The foundations of literacy development in children at familial risk of dyslexia. *Psychological Science*, *26*(12), 1877-1886.

Ireland, M., & Conrad, B. J. (2016). Evaluation and eligibility for speech-language services in schools. *Perspectives of the ASHA Special Interest Groups*, *1*(16), 78-90.

Kaminski, R. A., Abbott, M., Bravo-Aguayo, K., & Good, R. H. (2018). *Preschool Early Literacy Indicators (PELI)*. Dynamic Measurement Group.

Larrivee, L. S., & Catts, H. W. (1999). Early reading achievement in children with expressive phonological disorders. *American Journal of Speech-Language Pathology, 8,* 118-128.

Lervåg, A., Hulme, C., & Melby-Lervåg, M. (2018). Unpicking the developmental relationship between oral language skills and reading comprehension: It's simple, but complex. *Child Development, 89*(5), 1821-1838.

Linacre, J. M. (2020a). *Table 30.1 Differential item functioning DIF pairwise*. Retrieved from <www.winsteps.com/winman/table30_1.htm>.

Linacre, J. M. (2020b). *Winsteps® Rasch measurement computer program*. <www.winsteps.com>.

Lonigan, C. J. (2004, October). *Development & promotion of emergent literacy: An evidence-based perspective*. Paper presented at the Emergent Literacy Development Conference, Niagara Falls, NY.

Lonigan, C. J., Schatschneider, C., & Westberg, L. (2008). Identification of children's skills and abilities linked to later outcomes in reading, writing and spelling. In *Developing early literacy: A report of the National Early Literacy Panel* (pp. 55-106). National Institute for Literacy.

Lonigan, C. J., Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (2007). *Test of Preschool Early Literacy (TOPEL)*. Pro-Ed.

MacLean, M., Bryant, P. E., & Bradley, L. L. (1987). Rhymes, nursery rhymes, and reading in early childhood. *Merrill-Palmer Quarterly, 33,* 255-282.

McConnell, S. R., Wackerle-Hollman, A. K., Scheppmann, K., Bradfield, T., & Rodriguez, M. (2019). *myIGDIs: Individual growth and development indicators of early literacy* [assessment instrument]. Renaissance Learning, Inc.

McDowell, K. D., Lonigan, C. J., & Goldstein, H. (2007). Relations among socioeconomic status, age, and predictors of phonological awareness. *Journal of Speech, Language, and Hearing Research, 50*, 1079-1092.

Peeters, M., Verhoeven, L., de Moor, J., & Van Balkom, H. (2009). Importance of speech production for phonological awareness and word decoding: The case of children with cerebral palsy. *Research in Developmental Disabilities*, *30*(4), 712-726.

Perfetti, C., Beck, I., Bell, L. & Hughes, C. (1987). Phonemic knowledge and learning to read are reciprocal. *Merrill-Palmer Quarterly*, *33*, 283-319.

Rasch, G. (1960, 1980). Probabilistic models for some intelligence and attainment tests. University of Chicago Press (original work published 1960 by the Danish Institute for Educational Research).

R Core Team (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, https://www.R-project.org/.

Rost, J. (2001). The growing family of Rasch models. In A. Boomsma, M. A. J. van Duijn, & T. A. B. Snijders (Eds.), *Essays on item response theory* (pp. 25-42). Springer.

Roth, F. P., Troia, G. A., Worthington, C. K., & Dow, K. A. (2002). Promoting awareness of sounds in speech: An initial report of an early intervention program for children with speech and language impairments. *Applied Psycholinguistics, 23,* 535-565.

Roth, F., Troia, G. A., Worthington, C. K., & Handy, D. (2006). Promoting awareness of sounds in speech (PASS): The effects of intervention and stimulus characteristics on the blending

performance of preschool children with communication impairments. *Learning Disability Quarterly, 29*, 67-88.

Schleppegrell, M. J. (2004). *The language of schooling: A functional linguistics perspective*. Routledge.

Selin, C. M., Rice, M. L., Girolamo, T., & Wang, C. J. (2019). Speech-language pathologists' clinical decision making for children with specific language impairment. *Language, Speech, and Hearing Services in Schools*, *50*(2), 283-307.

Skibbe, L. E., Bowles, R. P., Goodwin, S., Troia, G. A., & Konishi, H. (2020). The Access to Literacy Assessment System for Phonological Awareness (PA): An adaptive measure of PA appropriate for children with speech and/or language impairment. *Language, Speech, and Hearing Services in Schools*, *51*(4), 1124-1138.

Skibbe, L. E., Gerde, H. K., Wright, T. S., & Samples-Steele, C. R. (2016). A content analysis of phonological awareness and phonics in commonly used Head Start curricula. *Early Childhood Education Journal*, *44*(3), 225-233.

Snowling, M. J., Hayiou-Thomas, M. E., Nash, H. M., & Hulme, C. (2020). Dyslexia and Developmental Language Disorder: Comorbid disorders with distinct effects on reading comprehension. *Journal of Child Psychology and Psychiatry*, *61*(6), 672-680.

Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: evidence from a longitudinal structural model. *Developmental Psychology*, *38*(6), 934-947.

Stothard, S. E., Snowling, M. J., Bishop, D. V. M., Chipchase, B. B., & Kaplan, C. A. (1998). Language impaired preschoolers: A follow-up into adolescence. *Journal of Speech, Language, and Hearing Research, 41*, 407-418.

Torgesen, J. K., Morgan, S., & Davis, C. (1992). The effects of two types of phonological awareness training on word learning in kindergarten children. Journal of Educational Psychology, 84, 364-370.

Torgesen, J. K., Wagner, R. K. & Rashotte, C. A. (1994). Longitudinal studies of phonological processing and reading. *Journal of Learning Disabilities*, *27*, 276-286.

Tortorelli, L. S., Bowles, R. P., & Skibbe, L. E. (2017). Easy as AcHGzrjq: The Quick Letter Name Knowledge Assessment. *The Reading Teacher*, *71*(2), 145-156.

U.S. Department of Education (2020). *The Condition of Education - Preprimary, Elementary, and Secondary Education - Elementary and Secondary Enrollment - Students With Disabilities - Indicator May (2020)*. Institute of Education Sciences, National Center for Education Statistics. http://nces.ed.gov/programs/coe/indicator_cgg.asp

Vandervelden, M. C., & Siegel, L. S. (1999). Phonological processing and literacy in AAC users and students with motor speech impairments. *Augmentative and Alternative Communication, 15*, 191-211.

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin*, *101*, 192-212.

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1994). The development of reading-related phonological processing abilities: New evidence of bi-directional causality from a latent variable longitudinal study. *Developmental Psychology*, *30*, 73-78.

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., & Pearson, N. A. (2013). *Comprehensive Test of Phonological Processing* (2nd ed.; *CTOPP-2*). Pro-Ed.

Zhang, X., & Tomblin, J. B. (2000). The association of intervention receipt with speech-language profiles and social-demographic variables. *American Journal of Speech-Language Pathology, 9*(4), 345-357.

Zwick, R., Thayer, D. T., & Lewis, C. (1999). An empirical Bayes approach to Mantel-Haenszel DIF analysis. *Journal of Educational Measurement, 36*(1), 1-28.

Appendices

Appendix A

ATLAS-PA Participant Demographics and Administration Procedures

Table 1

Demographic Information for ATLAS-PA Participants

	Children with typical development (n = 998)	Children with speech and/or language impairment (n = 280)
Age		
3-year-olds	163 (16.33%)	41 (14.64%)
4-year-olds	330 (33.07%)	69 (24.64%)
5-year-olds	226 (22.65%)	65 (23.21%)
6-year-olds	172 (17.23%)	63 (22.50%)
7-year-olds	107 (10.72%)	42 (15.00%)
Gender		
Female	473 (47.39%)	100 (35.71%)
Male	525 (52.61%)	179 (63.93%)
Race/Ethnicity		
White/Caucasian	497 (49.80%)	165 (58.93%)
Black/African American	233 (23.35%)	58 (20.71%)
Hispanic or Latino	50 (5.01%)	17 (6.07%)
Asian/Pacific Islander	64 (6.41%)	2 (0.71%)
Native American/ Alaskan	6 (0.60%)	3 (1.07%)
Multiracial	120 (12.02%)	28 (10.00%)
Other	28 (2.81%)	7 (2.50%)

Maternal Education		
Some high school	74 (7.41%)	20 (7.14%)
High school diploma or equivalent	161 (16.13%)	42 (15.00%)
Some college	240 (24.05%)	92 (32.86%)
Undergraduate degree	224 (22.44%)	51 (18.21%)
Graduate/professional school	246 (24.65%)	46 (16.43%)
NA	53 (5.31%)	29 (10.36%)
Annual household income		
Less than $25,000	350 (35.07%)	115 (41.07%)
$25,000 to $49,999	189 (18.94%)	58 (20.71%)
$50,000 to $74,999	109 (10.92%)	42 (15.00%)
$75,000 to $99,999	94 (9.42%)	22 (7.86%)
More than $100,000	186 (18.64%)	20 (7.14%)
NA	70 (7.01%)	23 (8.21%)

Table 2

Disability Categories for Participants

Disability	N
Autism spectrum disorder	29
Early childhood developmental delay	2
Emotional impairment	13
Hearing impairment	9
Intellectual disability	11
Learning disability	16
Motor/ coordination problems	13
Other health impairment	21
Physical impairment	4
Speech & language impairment	241
Severe multiple impairment	1
Traumatic brain injury	1

Corrected vision	11
Visual impairment	6
Uncorrected vision	1

Notes:
The ELL group included one participant with ASD and 2 participants with speech & language impairment.
"Intellectual disability" includes children with cognitive impairments, Down syndrome, and fetal alcohol syndrome.
79 children in the IEP group had multiple disabilities and were thus represented in multiple groups. As a result, the sum of all categories is greater than the number of participants.

Table 3

Instruments Administered, by Data Collection Phase

For data collection phase listed at right, children received the following assessments:	**Typically-developing children**	**Children with disabilities relating to speech production**
ATLAS-PA: all children ages 3;0 to 7;11	2 of the 3 subtests: rhyming and blending, blending and segmentation, or rhyming and segmentation	All 3 PA subtests: rhyming, blending, and segmentation
TOPEL: all 3-year-olds and approx. half of 4- to 6-year-olds	Phonological Awareness subtest, receptive and expressive tasks	NA
CTOPP 4-6: all 3-year-olds and approx. half of 4- to 6-year-olds	Phonological Awareness subtests of Elision, Blending Words, and Sound Matching	NA
CTOPP 7+: 7-year-olds	Phonological Awareness subtests of Elision, Blending Words, and Phoneme Isolation	NA
PELI: all preschoolers	Word Parts and First Sound subtests	NA
DIBELS Next: all kindergarteners, 1st graders, and 2nd graders	First Sound Fluency (kindergarteners only) and Phoneme Segmentation Fluency subtests	NA

Appendix B

CAT Procedure for Initial Item Selection

1) Select item if current item is first item (otherwise item has already been selected)
2) Administer item
3) While the examinee is selecting a response to the item
    a) Write out line to outfile for previous item, if current item is not first item
    b) Assuming the current item response will be correct,
        i) Calculate the new provisional ability estimate and standard error
        ii) Select the next item
        iii) If item does not meet exposure controls, repeat ii (select new item)
        iv) If item does not meet content controls, repeat iii
    c) Repeat assuming the current item response will be incorrect
4) Obtain the current item response. Based on the item response
    a) Check stopping rule
    b) If stop=yes, then stop test and go to step 5
    c) If stop=no, administer the next item (go to step 2)
5) If stop=yes
    a) Administer additional experimental items if necessary
    b) Write out line to outfile for last active item
    c) Rescale ability
    d) Report score

Appendix C

Table 4

Correlations Among Measures of Phonological Awareness for Children with Typical Development

	ATLAS-PA	*TOPEL*	*CTOPP 4-6*	*CTOPP 7+*	*DIBELS-FSF*	*DIBELS-PSF K*	*DIBELS-PSF Grade 1*	*PELI*
ATLAS-PA	1.00 (*n* =998)							
TOPEL	0.36 (*n* = 438)	1.00 (*n* = 453)						
CTOPP 4-6	0.65 (*n* = 371)	NA	1.00 (*n* = 371)					
CTOPP 7+	0.57 (*n* = 97)	NA	NA	1.00 (*n*= 97)				
DIBELS FSF	0.44 (*n* = 155)	0.57 (*n* = 40)	0.52 (*n* = 96)	NA	1.00 (*n*= 155)			
DIBELS PSF K	0.49 (*n* = 154)	0.57 (*n* = 41)	0.51 (*n* = 94)	NA	0.68 (*n*= 153)	1.00 (*n*= 154)		
DIBELS PSF Grade 1	0.38 (*n* = 217)	NA	0.51 (*n* = 105)	0.25 (*n* = 97)	NA	NA	1.00 (*n* = 217)	
PELI	0.51 (*n* = 599)	0.48 (*n* = 394)	0.57 (*n* = 168)	NA	NA	NA	NA	1.00 (*n*= 600)

Note: Table includes only children with typical development. For all correlations, $p < .001$

Appendix D

Test, Item-Level, and Person-Level Summaries

Summaries of ATLAS results (logits)

SUMMARY OF 1174 MEASURED (NON-EXTREME) PERSON

	TOTAL			MODEL	INFIT		OUTFIT	
	SCORE	COUNT	MEASURE	S.E.	MNSQ	ZSTD	MNSQ	ZSTD
MEAN	34.9	53.0	.78	.40	1.00	.06	.99	.05
SEM	.5	.4	.04	.01	.00	.03	.01	.03
P.SD	16.8	15.0	1.51	.19	.10	1.02	.26	1.08
S.SD	16.8	15.0	1.51	.19	.10	1.02	.26	1.08
MAX.	83.0	87.0	4.65	1.08	1.39	4.58	3.68	4.63
MIN.	1.0	7.0	-2.82	.23	.71	-3.88	.19	-3.83

REAL RMSE	.45	TRUE SD	1.44	SEPARATION	3.23	PERSON RELIABILITY	.91
MODEL RMSE	.44	TRUE SD	1.44	SEPARATION	3.28	PERSON RELIABILITY	.91
S.E. OF PERSON MEAN = .04							

MAXIMUM EXTREME SCORE: 101 PERSON 7.9%
MINIMUM EXTREME SCORE: 2 PERSON .2%
LACKING RESPONSES: 1 PERSON
CUTLO= -1.0 CUTHI= .0 logits

SUMMARY OF 1277 MEASURED (EXTREME AND NON-EXTREME) PERSON

	TOTAL			MODEL	INFIT		OUTFIT	
	SCORE	COUNT	MEASURE	S.E.	MNSQ	ZSTD	MNSQ	ZSTD
MEAN	36.6	53.3	1.14	.51				
SEM	.5	.4	.05	.01				
P.SD	17.6	14.9	1.93	.43				
S.SD	17.6	14.9	1.93	.43				
MAX.	85.0	87.0	5.91	1.88				
MIN.	.0	5.0	-4.09	.23				

REAL RMSE	.67	TRUE SD	1.81	SEPARATION	2.68	PERSON RELIABILITY	.88
MODEL RMSE	.67	TRUE SD	1.81	SEPARATION	2.70	PERSON RELIABILITY	.88
S.E. OF PERSON MEAN = .05							

PERSON RAW SCORE-TO-MEASURE CORRELATION = .80
CRONBACH ALPHA (KR-20) PERSON RAW SCORE "TEST" RELIABILITY = .93 SEM = 4.57
STANDARDIZED (50 ITEM) RELIABILITY = .77

SUMMARY OF 110 MEASURED (NON-EXTREME) ITEM

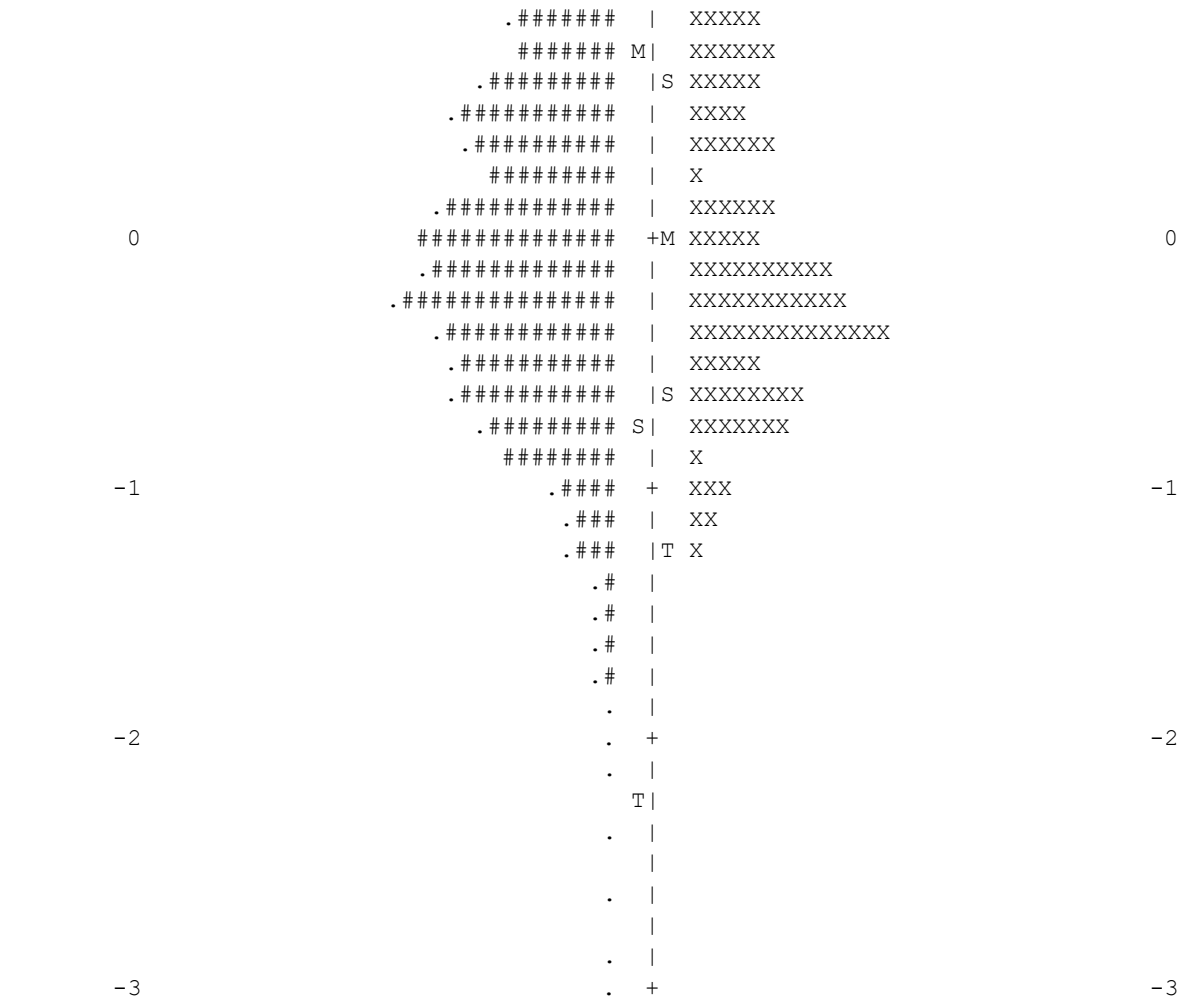
	TOTAL			MODEL	INFIT		OUTFIT	
	SCORE	COUNT	MEASURE	S.E.	MNSQ	ZSTD	MNSQ	ZSTD
MEAN	425.2	619.0	.00	.10	1.00	.00	.99	-.04
SEM	7.9	8.3	.06	.00	.01	.18	.01	.13
P.SD	82.2	86.2	.65	.01	.08	1.89	.15	1.38
S.SD	82.6	86.6	.65	.01	.08	1.89	.15	1.39

```
| MAX.     557.0     727.0        1.80     .13      1.22   4.69   1.42   3.61 |
| MIN.     199.0     350.0       -1.25     .09       .86  -3.87    .74  -3.06 |
|-------------------------------------------------------------------------------------------------------|
| REAL RMSE    .10 TRUE SD     .64  SEPARATION  6.15  ITEM   RELIABILITY  .97 |
|MODEL RMSE    .10 TRUE SD     .64  SEPARATION  6.26  ITEM   RELIABILITY  .98 |
| S.E. OF ITEM MEAN = .06                                                     |
-------------------------------------------------------------------------------------------------------
                 DELETED:      8 ITEM
ITEM RAW SCORE-TO-MEASURE CORRELATION = -.95
Global statistics: please see Table 44.
UMEAN=.0000 USCALE=1.0000
```

Wright map of all ATLAS participants

```
INPUT: 1278 PERSON  118 ITEM  REPORTED: 1277 PERSON  110 ITEM  2 CATS WINSTEPS 4.8.2.0
-------------------------------------------------------------------------------------------------------

MEASURE                                  |                                  MEASURE
  <more> --------------------------- PERSON  -+- ITEM    ---------------------- <rare>
    5       .##########################  +                                        5
                                        |
                                        |
                                     .  |
                                     .  |
                                   ###  |
                                        |
                                 .#####  |
    4                            .#####  +                                        4
                                     .#  |
                                   .## T|
                                     .#  |
                                   .##  |
                                  .####  |
                                   .###  |
                                     .#  |
    3                            .####  +                                        3
                                 .####  |
                                   .##  |
                                    ##  |
                                     .#  |
                                  .####  |
                                    .## S|
                                 .####  |
    2                            .####  +                                        2
                                  .###  |
                                 .####  |  X
                                 .#### |
                                 #####  |  X
                                .#####  |  XX
                              .#######  |T X
                                .#####  |  XXX
    1                          .########  +  XX                                  1
```



Letters appearing beside the pipes (vertical lines in center of Wright map):
M = average child ability (left side) or item difficulty (right side)
S = 1 standard deviation from mean
T = 2 SD from mean

Appendix E

ATLAS-PA Scores by Age and Maternal Education Level

Table 5

Descriptive Statistics by Child Age

	Typically Developing					IEPs					**Total Sample**				
(yr;mo)	M	SD	Min	Med-ian	Max	M	SD	Min	Med-ian	Max	**M**	**SD**	**Min**	**Med-ian**	**Max**
3;0-3;11	19.59	1.70	14.59	19.20	24.89	19.65	1.38	17.56	19.41	24.26	**19.61**	**1.76**	**15.90**	**19.16**	**24.93**
4;0-4;11	19.58	1.69	15.36	19.19	24.93	19.37	1.76	16.31	18.90	24.89	**19.86**	**2.06**	**14.59**	**19.18**	**24.89**
5;0-5;11	20.05	2.16	14.88	19.23	24.74	20.70	2.18	17.49	20.57	24.68	**20.41**	**2.18**	**16.31**	**19.75**	**24.93**
6;0-6;11	20.22	2.34	16.79	19.31	24.87	21.43	2.05	17.38	21.83	24.59	**20.03**	**1.95**	**15.36**	**19.44**	**24.66**
7;0-7;11	20.59	1.93	17.06	20.27	24.64	20.26	1.78	17.30	19.98	24.64	**20.04**	**1.73**	**17.30**	**19.63**	**24.64**

Table 6

Descriptive Statistics by Maternal Education Level

	Typically Developing					IEPs					**Total Sample**				
	M	SD	Min	Median	Max	M	SD	Min	Median	Max	**M**	**SD**	**Min**	**Median**	**Max**
Some High School	19.64	2.00	14.88	19.29	24.66	19.83	1.84	17.47	19.30	24.29	**20.09**	**2.06**	**16.65**	**19.46**	**24.87**
High School Diploma	20.01	2.01	17.01	19.33	24.93	20.01	2.23	17.07	19.30	24.74	**19.97**	**2.04**	**15.36**	**19.31**	**24.89**
Some College	19.91	2.06	15.36	19.20	24.87	19.67	1.93	16.66	18.98	24.68	**19.84**	**1.95**	**14.88**	**19.24**	**24.89**
Undergraduate Degree	19.91	1.93	14.59	19.38	24.93	19.88	1.62	17.58	19.41	24.23	**20.18**	**2.01**	**14.59**	**19.56**	**24.93**
Graduate/ Professional School	19.92	1.98	15.90	19.31	24.89	19.37	1.37	16.96	18.94	23.08	**20.04**	**1.98**	**15.90**	**19.52**	**24.93**

Access to Literacy Assessment System
accesstoliteracy.com

Paper in *LSHSS* on the
Access to Literacy Assessment System-Phonological Awareness
https://pubs.asha.org/doi/10.1044/2020_LSHSS-19-00006

Contact: Dr. Lori Skibbe
skibbelo@msu.edu