Research Article

The Access to Literacy Assessment System for Phonological Awareness: An Adaptive Measure of Phonological Awareness Appropriate for Children With Speech and/or Language Impairment

Lori E. Skibbe,[a] Ryan P. Bowles,[a] Sarah Goodwin,[a]
Gary A. Troia,[b] and Haruka Konishi[c]

Purpose: The Access to Literacy Assessment System–Phonological Awareness (ATLAS-PA) was developed for use with children with speech and/or language impairment. The subtests (Rhyming, Blending, and Segmenting) are appropriate for children who are 3–7 years of age. ATLAS-PA is composed entirely of receptive items, incorporates individualized levels of instruction, and is adaptive in nature.
Method: To establish the construct validity of ATLAS-PA, we collected data from children with typical development (*n* = 938) and those who have speech and/or language impairment (*n* = 227).
Results: Rasch analyses indicated that items fit well together and formed a unidimensional construct of phonological awareness. Differential item functioning was minimal

between the two groups of children, and scores on ATLAS-PA were moderately to strongly related to other measures of phonological awareness. Information about item functioning was used to create an adaptive version of ATLAS-PA.
Conclusions: Findings suggest that ATLAS-PA is a valid measure of phonological awareness that can be used with children with typical development and with speech and/or language impairment. Its adaptive format minimizes testing time and provides opportunities for monitoring progress in preschool and early elementary classrooms.
Supplemental Material: https://doi.org/10.23641/asha. 12931691

C hildren with a primary speech and/or language impairment account for 43% of those receiving special education services within schools (U.S. Department of Education, 2017). In addition, many more children require services due to a different primary disability such as cerebral palsy, Down syndrome, or autism, in

which there are often associated speech and/or language impairments. Children with speech and/or language impairments regularly struggle to meet educational goals related to literacy achievement beginning in preschool and kindergarten (Anthony et al., 2011; Justice et al., 2009; Pentimonti et al., 2016), often in the area of phonological awareness (Catts et al., 2002; Pentimonti et al., 2016). Phonological awareness (PA), the understanding of the sound structure of language, is an important skill that predicts children's later literacy knowledge (e.g., Lerner & Lonigan, 2016; Lonigan et al., 2008; Wagner et al., 1994). The National Early Literacy Panel (2008) identified PA as one of the most consistent predictors of later literacy achievement for preschoolers, even when considering the contributions of IQ and socioeconomic status. A challenge for understanding individual differences in PA knowledge is the lack of assessments designed for children with speech and/or language impairments. Children with speech and/or language impairments may need assessments that include

[a]Department of Human Development and Family Studies, Michigan State University, East Lansing
[b]Department of Counseling, Educational Psychology, and Special Education, Michigan State University, East Lansing
[c]Department of Education, Missouri Western State University, St. Joseph

Correspondence to Lori E. Skibbe: skibbelo@msu.edu

adaptations for accessibility, including the use of explicit instructions and minimal testing time, using a framework that minimizes unnecessary distractors during the testing process. In this article, we describe the development and validation of the Access to Literacy Assessment System–Phonological Awareness (ATLAS-PA), a new, adaptive measure of PA tailored for children with speech and/or language impairment that is administered using a web-based browser.

Development of PA

PA skills emerge early, develop rapidly throughout early childhood, and have strong implications for later literacy achievement, making them an ideal candidate for frequent assessment (Moyle et al., 2013). Individual differences in PA ability can be observed during early childhood and remain fairly consistent over time (e.g., Lonigan et al., 1998; Wagner et al., 1994). For children with typical development, PA seems to benefit from children's tendency to play with language, experiences with print and print-related concepts, and high-quality formal reading instruction (Snow et al., 1998; Torgesen et al., 1994; Troia et al., 1998).

Children exhibit their knowledge of PA initially by manipulating larger units of sound (e.g., words) and then progressing until they are also able to perform tasks that require parsing words at the level of the phoneme (Anthony & Lonigan, 2004). There is evidence to suggest that further advances in PA are supported by other early reading skills, such that PA provides a pathway through which these abilities build upon themselves (e.g., Cassar & Treiman, 2004; Ehri & Snowling, 2004; Torgesen et al., 1994). Children are commonly asked to perform PA tasks during preschool and the early elementary grades that focus on rhyming (e.g., Anthony & Lonigan, 2004), as well as syllable and sound blending and segmentation (Lonigan et al., 2009). Importantly, these different types of tasks appear to reflect the same latent trait, given prior work in this area suggesting that PA is unidimensional for young children who exhibit typical development (Anthony & Lonigan, 2004; Anthony et al., 2002; Schatschneider et al., 1999).

Similar to children with typical development, PA is a necessary precursor for functional reading for all students with disabilities, even when the disability is associated with moderate to severe developmental delays that can influence speech output (Browder et al., 2009). There is an abundance of research showing that children with speech and/or language impairment, autism, Down syndrome, and cerebral palsy commonly exhibit lower levels of PA when compared to their peers with typical development (Dessemontet et al., 2017; Dynia et al., 2019; Næss et al., 2012; Peeters et al., 2009; Thatcher, 2010). A range of reasons may contribute to lower PA skills in children with speech and/or language impairment relative to children with typical development. Challenges may reflect difficulties in figuring out the meaningful sound patterns represented in speech (Preston & Edwards, 2010). Lower speech abilities may also contribute to the differences in PA, as verbal speech allows children to play out loud with the sounds of language, facilitating

the development of PA; such experiences may be more limited for children with some types of disabilities, thus minimizing their opportunities to develop skills in this area (Peeters et al., 2009). This highlights the need for more accessible PA assessments that can evaluate knowledge for children with a range of linguistic needs and capabilities.

As with children who develop typically, PA is a significant correlate of early literacy skills for preschoolers with speech sound disorders (Rvachew & Grawburg, 2006) and a predictor of later decoding for many children with disabilities (Dynia et al., 2017; Tambyraja et al., 2015). Even students with remediated speech sound disorders continue to have lower literacy scores in late elementary and early middle school when compared to students who were developing typically (Farquharson, 2015). Findings point to the enduring importance of being able to store and manipulate phonological information using working memory, a skill that appears to be challenging for those who have speech sound disorders (Anthony et al., 2011; Farquharson et al., 2018).

Importantly, students with disabilities appear to profit from reading instruction that includes attention to PA (Lemons & Fuchs, 2010), including young children with speech and/or language impairment (Skibbe et al., 2011). Experts generally agree that preventing reading difficulties is easier and more cost-effective than working to remediate reading challenges later in a student's career (Francis et al., 1996); ATLAS-PA is thus designed for children attending preschool and early elementary grades to reflect the need to assess children's PA earlier in their school careers.

Issues With Current Assessments of PA

Despite the significance of this core early literacy skill, there is a dearth of standardized and validated tools of PA for children with speech and/or language impairment (e.g., Barker et al., 2014; Iacono & Cupples, 2004). The lack of PA assessments that have considered the needs of children with speech and/or language impairment can create barriers to testing within schools (Thurlow, 2010). Some researchers have simply utilized existing assessments of PA even though they may not be valid for children with speech and/or language impairment (e.g., Hesketh, 2004). Others have created their own assessment items with unknown psychometric properties and unclear score interpretations (Dahlgren Sandberg, 2006; Iacono & Cupples, 2004; Vandervelden & Siegel, 2001), or adapted existing assessments without addressing the psychometric characteristics of the adapted version (e.g., Card & Dodd, 2006). These ad hoc adaptations do not necessarily function identically to the original versions, and the adapted formats can have important ramifications for research findings (Dahlgren Sandberg, 2001; Peeters et al., 2009, 2008). For example, Card and Dodd (2006) compared the phonological abilities of children with cerebral palsy who are unable to speak with those who do speak and found between-group differences using some, but not all, types of testing formats. Inconsistent findings on the role of speech in PA ability

have led some researchers to consider other aspects of development, such as IQ, to explain findings (Peeters et al., 2008). As a result, without a measure of PA that is validated for children with speech and/or language impairments, we do not have a clear understanding about PA for this group of children and may in fact be misestimating their skills and underestimating their overall cognitive abilities. An accurate assessment of PA skills for students with speech and/or language impairment is critical to hold schools accountable for providing effective literacy instruction to all children (Lemons et al., 2012), especially since the early years of special education often put little emphasis on literacy instruction (Browder et al., 2006). Accurate early language measures help answer the call for educators to monitor academic progress for those receiving specialized services within schools (Lemons et al., 2018), but this can be challenging without the appropriate tools.

The purpose of ATLAS-PA is to provide a general measure of children's PA skill while also allowing professionals to monitor children's progress in this area over time. In addition to measuring PA skills for children with typical development, ATLAS-PA is designed to be used with children who have an educational identification requiring speech-language services within their schools; however, it is not intended to be used to make a clinical determination of a speech and/or language impairment (see Ireland & Conrad, 2016, for more information about the distinction between these two purposes). This is accomplished by measuring skills commonly incorporated into curricula and targeted as part of interventions in this area (e.g., blending activities included in the Promoting Awareness of Speech Sounds curricula; Roth et al., 2006). Curriculum-based measures, such as ATLAS-PA, are useful progress-monitoring tools that are easy to administer, cost-effective, and able to show development over time. There are several other curriculum-based measures targeting PA currently on the market (e.g., Individual Growth and Development Indicators [IGDIs], Dynamic Indicators of Basic Early Literacy Skills [DIBELS], Phonological Awareness Literacy Screening-PreK), but these were designed for children exhibiting typical patterns of development. This can be problematic, particularly as many curriculum-based measures do not accurately capture PA ability levels for students with speech and/or language impairment; in addition, many are not appropriate for children younger than 4 years of age (Invernizzi et al., 2010; Missall et al., 2006). ATLAS-PA is an adaptive measure of PA, which is made possible by administration using a web-based browser (see Chapelle & Douglas, 2006). Children with disabilities are often taught early reading skills, including areas related to PA, using technology-supported methods of instruction (Grindle et al., 2013; Koppenhaver et al., 2007), so this approach aligns well with current instructional practices in the field.

Research Questions

The goal of this study is to describe the development process of ATLAS-PA and to provide validity evidence for ATLAS-PA with a large-scale validation study involving children with speech and/or language impairment as well as children exhibiting typical development. In particular, using a Rasch measurement approach, we considered four research questions to examine construct validity through internal structure and relations to other variables as sources of validity evidence (AERA/APA/NCME, 2014):

1. Is the ATLAS-PA assessing a unidimensional construct? We hypothesize that, similar to other work in this area (Anthony & Lonigan, 2004; Anthony et al., 2002; Schatschneider et al., 1999), ATLAS-PA would represent a unidimensional construct of PA for children within this age range.

2. Do the ATLAS-PA items function as expected within the Rasch measurement modeling framework? We anticipate that, based on typical Rasch fit statistics, most items will fit the model well.

3. Are there any differences in ATLAS-PA item performance attributable to ability group? It is hypothesized that children with speech and/or language impairment would display lower levels of PA when compared to children with typical development, but that the items would function similarly across the groups.

4. How do scores on the ATLAS-PA relate to scores on other measures of PA? We expect the correlations between the ATLAS-PA and other PA assessments to be consistent with reported correlations among current PA assessments as reported in other validation studies (e.g., .5–.6 as reported by Lonigan et al., 2007).

Method

Instrument Development

ATLAS-PA was developed by a research team consisting of experts in early childhood language and literacy development, speech-language pathology, and psychometrics, using a rigorous iterative process involving a panel of early educators, extensive pilot testing, and a large-scale validation study. The first step was to identify what aspects of PA the assessment would target and identify critical features to include in the measure. We convened a panel of 10 early childhood educators to include a number of features that make it useful and valid for children with speech and/or language impairment. Everyone on the panel was associated with one of two university preschools, which serve approximately 200 children and train preservice teachers; participants included the director of the preschools, the associate director of one of the preschools, and eight early childhood educators. All educators had at least a bachelor's degree. Classrooms served children eligible for Head Start or the Great Start Readiness Program, those whose families paid tuition, and those who were receiving early childhood special education services.

Consistent with the panel's reported classroom practice, we chose to focus ATLAS-PA on three PA skills:

rhyming, blending, and segmenting. In consultation with the panel and best practices in special education, we identified four critical features to include in ATLAS-PA: items with concrete content that do not require speech output, design features that increase accessibility and reduce self-regulatory demands, explicit individualized instructions, and the use of adaptive algorithms to minimize testing time. Below, we provide more information about each of the unique design features of ATLAS-PA. We then present an item calibration and construct validation study to ensure that the items work together to validly measure PA (Embretson, 1983); as part of this process, we provide preliminary evidence that scores from ATLAS-PA relate to other measures of PA. Finally, we describe the adaptive algorithms implemented to minimize testing time.

Item Development

We developed 120 items to represent the three target areas of PA: rhyming, blending, and segmenting. In order to be appropriate for children with speech and/or language impairment, ATLAS-PA relies entirely on nonverbal response options and only includes items requiring selection among alternative responses (i.e., multiple choice). Although some existing measures employ multiple choice (e.g., Test of Preschool Early Literacy [TOPEL]; Lonigan et al., 2007) or forced choice (e.g., beginning sounds subtest of the Phonological Awareness Literacy Screening; Invernizzi et al., 2003) on some items or subtests relevant for young children, no well-validated and normed measure is based on selection in its entirety for our intended age range. Besides allowing for nonverbal responses, the multiple-choice format offers many advantages such as efficiency and standardization of administration and scoring; however, there are well-known issues with guessing or chance responding on such a test format. To address this problem, ATLAS-PA accounts for guessing using a response cutoff technique available in a Rasch measurement approach (Andrich et al., 2012; Gershon, 1992; Linacre, 2018c).

For each test item, there are three illustrations along the bottom representing three response options. For rhyming items, a pictorial representation of the target word is also displayed in the center top of the screen. Items are presented using a plain, blank background; prior work has shown that some children with speech and/or language impairment may have challenges attending to literacy materials without getting distracted by irrelevant stimuli (Thompson et al., 2019), so we made the item response options as salient on the display as possible. All test items were implemented electronically using a tablet, similar to what has been done with other adaptive measures of language (Chapelle & Douglas, 2006).

Test items were created by a panel of four experts, including two speech-language pathologists, one expert on early childhood language and literacy development, and one psychometrician with an applied focus on early childhood language and literacy. Possible response options were balanced for consonant and vowel diversity. The target

words were selected to be concrete and familiar to preschool children. A review of existing assessments and curricula of PA were evaluated to create a pool of possible words to use for target words and response options. All possible words were evaluated using the MRC Psycholinguistic Database (Coltheart, 1981), which rates words on a scale from 100 to 700, with larger values indicating a higher level of concreteness. We intentionally chose more concrete words that also were highly imageable based on the judgment of seven scholars, including those with expertise in disabilities related to speech and/or language development, and removed any words that were considered to be less salient from the overall pool of words to be used when creating items. ATLAS-PA includes 279 words (target and response options) ranging from 365 to 670 in levels of concreteness ($M = 589.65$, $SD = 39.56$). Words were also analyzed in accordance with the age at which they are typically acquired using ratings from Kuperman et al. (2012). Words utilized were acquired, on average, during preschool ($M = 4.83$ years old, $SD = 1.18$ years).

The audio files for the items were recorded in a sound studio by a voice actor with a Midwestern accent. The stem for each item type (e.g., What rhymes with cat?) is presented, and then each of the three response options are named as the associated pictures are highlighted with a thick black outline. Response time is not considered when calculating a child's score, as some persons with disabilities need additional time to process and respond to test items (J. N. Kaufman et al., 2014). If the child does not respond to a particular item after 5 s, the entire item, including stem and response options, is repeated. A slower item presentation process is available at the behest of the test administrator, with slower speech and greater intervals between response options.

ATLAS-PA allows for children to respond to items immediately after they hear the test question to minimize testing time and maximize engagement; pilot data with a small sample of children ($n = 20$) indicated that performance did not change when children were allowed to respond before response options were labeled. However, this approach required us to ensure that the pictorial representations of response options were as clear as possible. To test the degree to which our illustrations elicited the intended vocabulary, we piloted the images in four locations across the United States (California, Michigan, Pennsylvania, Texas). Twenty children with typical development at each site ($n = 80$ overall) were asked to label all of the illustrations verbally. If more than four children across sites provided the same unintended response (e.g., "leg" for *knee*), the illustration was edited for clarity and retested.

Instructions and Practice

In accordance with recommendations for best practice and feedback from our panel of educators, we created instructions that are explicit, individualized, and include many opportunities for practice (Coyne et al., 2006). Instructions were created using an iterative process that involved piloting directions and practice items with two children

exhibiting typical development and four children with speech and/or language impairment related to the verbal production of speech, including two children with autism, one with Down syndrome, and one identified with a speech-language impairment. In addition, field notes were taken during the data collection process to identify whether additional revisions needed to be made to the instructions and practice items.

To address the differential needs of test takers, the instructions and practice trials on ATLAS-PA were tailored to individual children using a systematic, three-tiered system. Test administrators are able to select one of three levels of instructions and practice: Basic, Basic+, or Enhanced. Most children with typical development will receive instructions at a Basic level, in which children respond to two practice items and are given corrective feedback if needed. This method for introducing the test is typical of commercialized assessments (e.g., TOPEL), but may not be sufficient for some students with speech and/or language impairment. Our goal was to remove the construct-irrelevant variance resulting from students not understanding the demands of the test while maintaining the integrity of our scores. Thus, we provided varied prompts and opportunities for practice to those students to ensure that item responses were based on a child's level of PA rather than a lack of understanding of the task, but kept the testing module the same across students.

The next most supportive instructions for ATLAS-PA are the Basic+ level. Basic+ is designed for children who need more thorough instructions and more opportunities for practice, but who can take a test relatively independently (i.e., sit at a computer for 5–10 min with minimal behavioral prompts). Using support strategies found to be successful in prior research (J. Kaufman et al., 2009; Shank et al., 2010; Warschausky, 2009), the Basic+ level provides corrective feedback for up to three trial items if children get the initial practice item incorrect. Teachers in the focus group also cautioned that children may be unfamiliar with certain word choices or technical terms (e.g., *rhyme*), so corrective feedback varies the language used to support children's skills when possible (e.g., "ends with the same sounds" to supplement *rhyme*). If necessary, children can also practice with their chosen method for response (e.g., eye gaze). Finally, the Enhanced level of support is intended for children who need assistance from a test administrator, as prior work indicates that children with moderate to severe disabilities may not be able to use electronic literacy materials independently (Thompson et al., 2019). It is recommended that children who are not able to focus on a classroom task for at least 5 min take advantage of the Enhanced level of support. When opting for this type of support, test administrators will encounter a welcome page that will provide examples of behavioral supports that the test administrator may provide during the assessment (e.g., physical guidance, verbal prompts to refocus attention, regular positive reinforcements; Watling & Schwartz, 2004), although the specific supports utilized are at the discretion of the test administrator. Children who do not answer any practice items correctly are automatically moved to a higher level of instructional support (i.e., from Basic to Basic+ to Enhanced). Children must answer at least one practice item correctly by the Enhanced level in order to move onto the testing phase.

Item Calibration and Construct Validation Study

To examine the effectiveness of our approach to designing items and individualized instructions, we undertook an evaluation process where we administered all items from ATLAS-PA to a group of children with typical development. We removed all misfitting items and subsequently administered ATLAS-PA to a group of children with speech and/or language impairment.

Participants

Two groups of children (*N* = 1,165 overall) ages 3;0 (years;months) to 7;11 took ATLAS-PA. The study design and materials were reviewed by the institutional review board at Michigan State University. The study (IRB X15-599e) was determined to be exempt under Category 1, as it only involved normal educational practices. Parents of participants provided written consent before participating, and children provided verbal and/or nonverbal assent before working with research assistants. Participants received a $10 gift card and a children's book for participating in the study.

The first group of children exhibited typical development as reported by parents and had no current Individualized Education Program (IEP) for speech and/or language impairments (*n* = 938 [445 girls], M_{age} = 62.55 months, *SD* = 14.62 months). Recruitment occurred within 57 schools located in the Midwest. Children with typical development in this study were predominantly White/Caucasian (490 children [52.24%]), followed by Black/African American (223 [23.77%]), multiracial (119 [12.69%]), Asian/Pacific Islander (46 [4.90%]), Hispanic or Latino (40 [4.26%]), Native American (1 [0.11%]), or other (14 [1.49%]). Most parents reported that English was the primary language spoken within their homes (783 [83.48%]); inclusion criteria required parents to affirm that the participating child spoke English fluently. Of the parents who reported that they spoke another language at home, 46 languages were represented. Maternal education varied: some high school (68 mothers [7.25%]), high school diploma or equivalent (146 [15.57%]), some college (234 [24.95%]), undergraduate degree (220 [23.45%]), and graduate/professional school (230 [24.52%]). Annual household income also varied: less than $25,000 (330 households [35.18%]); $25,000–$49,999 (174 [18.55%]); $50,000–$74,000 (108 [11.51%]); $75,000–$99,999 (90 [9.59%]); and more than $100,000 (177 [18.87%]).

The second group of children had a reported speech and/or language impairment (*n* = 227 [77 girls], M_{age} = 66.46 months, *SD* = 16.31 months). To be eligible for inclusion in the group of children with speech and/or language impairment, children needed to meet the following eligibility

criteria: be between 3 and 7 years of age, have goals related to speech and/or language in an IEP, and have parents report that they could understand English in a way that is similar to a native speaker. Recruitment was done through flyers distributed and collected at 90 schools by special education coordinators, teachers, and speech-language pathologists who were familiar with our eligibility criteria. Parents were asked to confirm that children had IEP goals related to speech and/or language and were receiving services related to these areas. Similar to participants with typical development, children with speech and/or language impairment were predominantly White/Caucasian (130 [57.27%]), followed by Black/African American (48 [21.15%]), multiracial (27 [11.89%]), Hispanic or Latino (11 [4.85%]), or other (4 [1.76%]); however, there were fewer Asian/Pacific Islander children (1 [0.44%]; $\chi^2_1 = 8.29$, $p < .01$) than in the first group, with no Native American children represented. Additionally, most parents reported that English was the primary language spoken within their homes (205 [90.31%]); four other languages were reported to be spoken within participants' homes (i.e., Spanish, Burmese, Albanian, and American Sign Language). Maternal education and annual household income brackets were generally similar among both groups, though, for the second group, there were more mothers who reported having attended some college ($\chi^2_1 = 7.79$, $p < .01$) or came from households with incomes of under $25,000 ($\chi^2_1 = 4.68$, $p = .03$), and fewer mothers in the second group held graduate degrees ($\chi^2_1 = 5.81$, $p = .02$) or had household incomes of $100,000 or greater ($\chi^2_1 = 17.63$, $p < .01$). Relatively, more boys were in the group with IEPs compared to the group of children with typical development ($\chi^2_1 = 9.17$, $p < .01$). For age distribution, we attempted to recruit similar numbers of children within each age bracket of 3- to 7-year-olds. The group with typical development had proportionally more 4-year-olds than the group with IEPs ($\chi^2_1 = 19.16$, $p < .01$), and there were proportionally more 6-year-olds in the second group than in the first ($\chi^2_1 = 6.43$, $df = 1$, $p < .05$), but the proportions of 3-, 5-, and 7-year-olds were not significantly different between the two groups.

The children with speech and/or language impairments represented a broad range of children, who had varying levels of speech production capabilities. In addition to having an IEP for speech/language, children were reported to have the following disabilities and/or impairments: autism spectrum disorder (20 [8.81%]), attention deficit (11 [4.85%]), intellectual disability (11 [4.85%]; includes cognitive impairment, global developmental delay, Down syndrome, and fetal alcohol syndrome), vision impairment (eight children [3.52%]), high social/emotional needs (seven [3.08%]), learning disability (six [2.64%]), hearing impairment (five [2.20%]), movement/coordination problem (three [1.32%]), physical disability (two [0.88%]), and/or cerebral palsy (two [0.88%]). Fifty-seven of the 227 children (25.11%) were reported to have at least one other developmental disability in addition to a speech and/or language impairment. See Table 1 for demographic characteristics broken down by ability group.

ATLAS-PA Item Pool

The initial ATLAS-PA item pool consisted of 120 items: 40 rhyming, 40 blending, and 40 segmenting. In order to keep total testing time under 1 hr while ensuring that all items were administered to a large number of children, we employed a planned missingness design. Children with typical development were randomly assigned two out of the three subtests, and within each subtest, children were randomly assigned 30 of the 40 items, for a total of 60 items administered to each child. After an initial period of data collection, we identified two items as showing substantial statistical misfit, indicating examinee responses were overly unpredictable: rhyming *lace* (with "face") and segmenting *plate* (to "play"). The issue with *lace* may have arisen from challenges illustrating the word in a way familiar to young children. We did not identify a clear source of the issue with *plate*. These two items were removed from the item pool early in the data collection process, leaving 118 items in the item pool. Children with speech and/or language impairment were administered all three subtests, and within each subtest were randomly assigned 30 of the 39 or 40 remaining items. An example of each type of item (rhyming, blending, segmentation) is provided in Supplemental Materials S1, S2, and S3, respectively.

Other Measures of PA

In addition to taking ATLAS-PA, children exhibiting typical development were administered two other measures of PA chosen based on the child's age relative to the valid age range of the measure. The other PA measures are not appropriate for all of the children with speech and/or language impairment and thus were not administered to this group.

TOPEL. For the children with typical development, all 3-year-olds and half of the 4- to 6-year-olds ($n = 442$) were administered the PA subtest of the TOPEL (Lonigan et al., 2007). This subtest requires children to put sounds together to form a new word (blending) and to remove sounds from a word to form a new word (elision). Children were given prompts such as, "Point to the word you get when you say 'tooth'–'brush' together." As reported in the test manual, the internal consistency of TOPEL PA is .87 and test–retest stability over a 2-week period was .83 (Lonigan et al., 2007).

Comprehensive Test of Phonological Processing–Second Edition. For the children with typical development, half of the 4- to 6-year-olds were administered the ages 4–6 years version of the Blending and Elision subtests of the Comprehensive Test of Phonological Processing–Second Edition (Wagner et al., 2013). All 7-year-olds with typical development were administered the ages 7–24 years version of the Blending and Elision subtests. The Elision subtest requires children to identify or name the word that remains when a part of the word has been removed. Items ranged between elision of compound words (e.g., toothpaste without tooth), of syllables (e.g., catcher without -er), and of phonemes (e.g., grain without -n). Blending involves combining words into

Table 1. Demographic information.

Demographic variables	Children with typical development (*n* = 938)	Children with speech and/or language impairment (*n* = 227)
Gender	445 girls, 493 boys	77 girls, 150 boys
Race/ethnicity		
White/Caucasian	490 (52.24%)	130 (57.27%)
Black/African American	223 (23.77%)	48 (21.15%)
Hispanic or Latino	40 (4.26%)	11 (4.85%)
Asian/Pacific Islander	46 (4.90%)	1 (0.44%)
Native American	1 (0.11%)	0 (0.00%)
Multiracial	119 (12.69%)	27 (11.89%)
Other	14 (1.49%)	4 (1.76%)
Maternal education		
Some high school	68 (7.25%)	20 (8.81%)
High school diploma or equivalent	146 (15.57%)	33 (14.54%)
Some college	234 (24.95%)	78 (34.36%)
Undergraduate degree	220 (23.45%)	46 (20.26%)
Graduate/professional school	230 (24.52%)	38 (16.74%)
Annual household income		
Less than $25,000	330 (35.18%)	98 (43.17%)
$25,000–$49,999	174 (18.55%)	44 (19.38%)
$50,000–$74,999	108 (11.51%)	37 (16.30%)
$75,000–$99,999	90 (9.59%)	21 (9.25%)
More than $100,000	177 (18.87%)	16 (7.05%)

compound words, syllables into words, and phonemes into words. The Blending and Elision subtest raw scores are converted to percentile ranks and scaled scores via score lookup tables to yield a composite PA score. Internal consistency for the Comprehensive Test of Phonological Processing–Second Edition PA composite score is .92 for the ages 4–6 years version and .93 for the ages 7–24 years version (Wagner et al., 2013).

Preschool Early Literacy Indicators. For the children with typical development, all 3- and 4-year-olds (*n* = 584) were administered the PA subtest of the Preschool Early Literacy Indicators (Kaminski et al., 2018). This subtest assesses preschool-age children's ability to identify or say the first part or the first sound of a word (e.g., the first part of rainbow is rain). Interrater reliability ranges from .90 to .98.

DIBELS Next. Kindergarteners with typical development (*n* = 136; the DIBELS Next is grade based rather than age based) received the First Sound Fluency (FSF) subtest of the DIBELS Next. FSF provides 1 min for children to say the first sounds of orally introduced words. In addition, for the children with typical development, all kindergarten, Grade 1, and Grade 2 children (*n* = 340) received the Phoneme Segmentation Fluency subtest, which asks children to listen to orally introduced words and say all of the sounds in the word. The alternate-form reliability for the FSF and Phoneme Segmentation Fluency subtests is .72 and .88, respectively (Dewey et al., 2012).

Analyses

We used a Rasch measurement approach to examine the construct validity of the item pool of the ATLAS-PA. Rasch measurement is a form of item response analysis that offers a strong approach to validation at the item level (Bond & Fox, 2015; Rost, 2001), yielding scores that have

good evidence of interval scaling (Perline et al., 1979). The specific model we used was the Rasch (1960) model for dichotomous outcomes:

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

where P_{ni} is the probability of examinee n with trait level θ_n (i.e., PA level) succeeding on item i, which has difficulty level β_i. Data were analyzed using the Rasch measurement software Winsteps (Linacre, 2018c). Because of the potential for guessing with multiple-choice items, we used the CUTLO = −1 option, which treats individual item responses with expected probability of correct response below .27 as missing. Alternative values for CUTLO yielded identical results.

To address the first research question regarding dimensionality, we used a Rasch principal components analysis of residuals (Linacre, 1998). To address the second research question regarding item functioning, we examined the fit of the items to the Rasch model using standard Rasch fit statistics, Infit and Outfit. To address the third research question about differences in item functioning across gender and ability group (typically developing and speech and/or language impairment), we performed a differential item functioning (DIF) analysis. Finally, to address the fourth research question about nomothetic span, we examined correlations between scores from the ATLAS-PA and other measures of PA.

Results
Dimensionality

We first tested the dimensionality of ATLAS-PA using a principal components analysis of the residuals, which

considers the pattern of discrepancies between observed and predicted scores (Linacre, 1998). Conceptually, this analysis examines whether, after accounting for the primary measurement dimension, some items remain more related than expected, suggesting those items share a second dimension. Using the recommendations of Linacre (2018a), we identified potential multidimensionality if (a) at least one component beyond the primary measurement dimension had an eigenvalue above 2, (b) a scree plot of the eigenvalues had a clear elbow, (c) the disattenuated correlation between θ as estimated separately on potential components was substantially less than 1, and/or (d) the potential components included interpretable differences in item content.

For ATLAS-PA, the first eigenvalue was 3.1, suggesting potential multidimensionality. However, there was no clear elbow, as the next four eigenvalues were 2.8, 2.4, 2.0, and 1.8. The first principal component separated rhyming items from blending items, although the disattenuated correlation between θ as estimated separately from the contrasted items (i.e., from blending items alone vs. from rhyming items alone) was 1.00. No other principal component had interpretable content, and no disattenuated correlation was below .89. Thus, there was, at most, weak evidence of multidimensionality, and we concluded that ATLAS-PA was essentially unidimensional.

Item Functioning

We next examined item fit to establish that the items measure PA validly within a Rasch measurement framework. Following standard practice, we consulted both Infit and Outfit mean-square values; each has an expected value of 1.0, with higher values indicating the item fits poorly or has excess noise and lower values indicating the item offers less information than expected. We considered values between 0.6 and 1.4 to indicate good fit (Wright & Linacre, 1994). Of the initial item pool of 118 items (120 original items less two that were removed before the validation process), all 118 items' infit values fell within this range, showing excellent fit to the model. However, four items displayed outfit mean-square values above 1.4, suggesting some unpredictable responses: one blending item (/k/ + /a / = cow, outfit = 1.58) and three segmenting items ("cable" to bull, outfit = 1.45; "walrus" to wall, outfit = 1.43; and "cape" to ape, outfit = 1.41). We removed these items from the item pool, which then contained 114 items. There was no clear pattern in content for the misfitting items, suggesting that the item pool as a whole is validly measuring PA.

DIF

As additional evidence of validity, we conducted DIF analyses to examine whether an item behaves differently for different groups, controlling for overall level of PA. We considered DIF across gender and, separately, DIF across disability status (typical development and disability related to speech and/or language production). We followed conventions used by the Educational Testing Service (Zwick et al., 1999) for identifying moderate to strong DIF converted into a Rasch difficulty metric (Linacre, 2018b): statistical

significance ($p < .05$) using a Rasch–Welch t test with a difficulty difference (DIF contrast) of at least .64 logits, which implies that when person ability and item difficulty are well matched, the probability of a correct response differs by about .12 across the two groups. As preliminary analyses, we compared overall levels of PA for these groups. Girls ($n = 522$; $M = 1.28$ logits, $SD = 1.95$) and boys ($n = 637$; $M = 1.10$ logits, $SD = 1.96$) did not perform significantly differently on the ATLAS-PA ($d = 0.09$, $t(1157) = 1.4845$, $p = .1379$). As expected, children with speech and/or language impairment ($M = 0.89$ logits, $SD = 1.77$) had lower overall PA ability than children with typical development ($n = 938$; $M = 1.25$ logits, $SD = 1.96$), $d = 0.19$, $t(371.48) = 2.72$, $p = .007$. Although differences between the two groups were relatively small in magnitude, results should be interpreted in light of the fact that the children with speech and/or language impairment were, on average, about 4 months older than the children with typical development in the present work.

Four of the 114 remaining items met the DIF criteria for either gender or disability status. One item had DIF associated with gender: Blending snake (/sne/ + /k/) was significantly easier for girls (DIF contrast = .72). Two items displayed DIF associated with disability status: Blending dog [/dɔ/ + /g/] was easier for children with typical development (DIF contrast = .71), while blending toothbrush [tooth + brush] was easier for children with speech and/or language impairment (DIF contrast = .67). One item, segmenting "cartoon" (to yield car), had DIF on both gender and disability: easier for girls (DIF contrast = .68) and for children with speech and/or language impairment (DIF contrast = .91). These four items were removed from the item pool, leaving a final item pool consisting of 110 items, with 39 rhyming, 35 blending, and 36 segmenting items.

Final Item Pool

Figure 1 shows the Wright map reflecting examinee ability and item difficulty on a shared scale. Child PA ability ranged widely from −4.60 to +5.96 ($M = 1.16$) logits, but with most of the sample falling near the range of the item difficulty, which ranged from −0.99 (easiest) to +1.77 logits (most difficult). However, there was a substantial proportion of individuals whose level of PA fell above the range of the items. PA ability, as measured by the ATLAS-PA, was associated with age ($r = .53$, $p < .001$; $M_{3-year-olds} = -0.09$, $M_4 = 0.39$, $M_5 = 1.15$, $M_6 = 2.52$, $M_7 = 2.96$). The PA levels of three-, four-, and five-year-old children were best matched to the difficulty of the ATLAS-PA items. By subtest, blending items were the easiest on average (−0.27 logits), with rhyming (+0.20) and segmenting (+0.17) being slightly more difficult. However, rhyming, blending, and segmenting item difficulty ranges overlapped substantially, and no one item type clustered into one area of the difficulty range; hence, all subtests provided information across the range of examinee ability levels. Estimated Rasch-based reliability was .91 for the children exhibiting typical development and .94 for children with speech and/or language impairment; however, we used a measurement precision

stopping rule for the adaptive version, which allows us to select the intended reliability level of .90 for the final version of the ATLAS-PA (see below).

Nomothetic Span

Finally, we looked at the nomothetic span (Embretson, 1983) of the ATLAS-PA to examine the relations between the ATLAS-PA and other measures of PA. For this analysis, we restricted our sample to children evidencing typical development, as the other measures of PA were not designed for children with speech and/or language impairment and therefore scores on these assessments could not be considered valid for these children. Furthermore, children only completed assessments intended for their current age or grade (e.g., only 3- and 4-year-olds took the Preschool Early Literacy Indicators), so there was some restriction of range in PA skills. Table 2 presents the correlation matrix of ATLAS-PA with other PA batteries. Consistent with our expectations, the ATLAS-PA was moderately to strongly correlated with all other PA measures ($r = .49–.65$), at roughly the same magnitude as the other measures were correlated with each other. This suggests that ATLAS-PA is measuring the construct of PA in a similar way as validated, published batteries.

Discussion

Results demonstrate that ATLAS-PA is a reliable and valid measure of PA for children with and without speech and/or language impairment. The items utilized for ATLAS-PA represented a unidimensional construct of PA, consistent with several other studies in this area (Anthony & Lonigan, 2004; Anthony et al., 2002; Schatschneider et al., 1999). It is likely that PA skills are separate from other areas of language functioning (Anthony et al., 2014), thus warranting measures that focus explicitly on this skill set. Furthermore, evidence from our data suggests that ATLAS-PA is moderately to strongly related to other commonly used measures of PA, indicating the ways in which captured PA knowledge is reflective of other measures in the field.

Item Functioning and Validity for Children With Speech and/or Language Impairment

ATLAS-PA includes items that work well for a broad range of children. For those with typical development, ATLAS-PA appears to be best suited for those who are between 3 and 6 years of age, as the items were not well targeted to the PA levels of children with typical development who were 7 years of age. However, children with speech and/or language impairment often display lower levels of PA (Dessemontet et al., 2017; Dynia et al., 2019; Peeters et al., 2009), so educators should consider children's instructional level, in addition to their age, when determining whether ATLAS-PA could prove informative for their students with speech and/or language impairment. In the present work, although there were ceiling effects for the 7-year-olds with typical development, ATLAS-PA measured PA skills

Figure 1. Wright map of Access to Literacy Assessment System for Phonological Awareness examinee ability and item difficulty, in logits. PA = phonological awareness; M = mean; S = 1 standard deviation from mean; T = 2 standard deviations from mean. Each pound sign (#) indicates four children; each dot indicates one to three children. Each X indicates one item.

well for the 6- and 7-year-olds with speech and/or language impairment included in our sample, justifying a larger age range for this group of children.

For PA, testing format can affect performance and score interpretation (e.g., Card & Dodd, 2006), making it

Table 2. Correlations among measures of phonological awareness for children with typical development.

	ATLAS-PA	TOPEL	CTOPP 4–6	CTOPP 7+	DIBELS-FSF	DIBELS-PSF	PELI
ATLAS-PA	1.00 (n = 938)						
TOPEL	.56 (n = 441)	1.00 (n = 442)					
CTOPP 4–6	.65 (n = 371)	NA	1.00 (n = 371)				
CTOPP 7+	.58 (n = 97)	NA	NA	1.00 (n = 97)			
DIBELS-FSF	.49 (n = 136)	.63 (n = 40)	.54 (n = 96)	NA	1.00 (n = 136)		
DIBELS-PSF	.56 (n = 340)	.60 (n = 42)	.57 (n = 199)	.25 (n = 97)	.69 (n = 133)	1.00 (n = 340)	
PELI	.52 (n = 583)	.60 (n = 397)	.57 (n = 168)	NA	NA	NA	1.00 (n = 584)

Note. For all correlations, *p* < .001. ATLAS-PA = Access to Literacy Assessment System–Phonological Awareness; TOPEL = Test of Preschool Early Literacy; CTOPP 4–6 = Comprehensive Test of Phonological Processing–Second Edition, ages 4–6 years version; CTOPP 7+ = Comprehensive Test of Phonological Processing–Second Edition, ages 7+ years version; DIBELS-FSF = Dynamic Indicators of Basic Early Literacy Skills–First Sound Fluency; DIBELS-PSF = Dynamic Indicators of Basic Early Literacy Skills– Phoneme Segmentation Fluency; PELI = Preschool Early Literacy Indicators; NA = not applicable.

important to examine item functioning directly for children with and without speech and/or language impairment. Analyses here demonstrated that the vast majority of items worked in a similar fashion for all children, suggesting that we are able to assess PA skills accurately for those children not able to answer the expressive items often used in other tests of PA. The accurate assessment of PA for children with speech and/or language impairment should increase educators' capacity to provide early reading instruction individualized to each student's needs.

Unique Features of ATLAS-PA

There are several innovative features that make ATLAS-PA unique among tests of PA. First, measures of PA often require children to respond to items verbally, which can be challenging for many children with speech and/or language impairment (e.g., TOPEL; Lonigan et al., 2007). ATLAS-PA uses only receptive items to capture children's knowledge of PA, allowing it to be used by a greater number of children than many other measures. Second, ATLAS-PA includes three tiers of individualized instructions and practice items, increasing the probability that children with speech and/or language impairment can access the measure in ways that are accessible to them. Third, the measure was designed using a plain background and simple format that drew children's attention to relevant parts of the item, without including commonly used interactive features (e.g., hot spots or games) that require children to multitask or that could distract from children's performance on the assessment (Bus et al., 2015). Fourth, it is well recognized that many disabilities are associated with processing speed (Calhoun & Mayes, 2005), yet a number of early literacy assessments include speed of response as part of test administration (e.g., DIBELS 8th Edition; University of Oregon Center for Teaching and Learning, 2018). ATLAS-PA allows for variation in the speed of administration to increase confidence that scores reflect knowledge of PA specifically, rather than construct irrelevant variance associated with the way in which the task is administered.

The data gathered to validate the item pool for ATLAS-PA were used to make it an adaptive test. Adaptive tests tend to yield the same measurement precision as nonadaptive tests, but with approximately half as many items (Madsen, 1991; Weiss, 1982), making adaptive testing an effective way to reduce both physical and attentional fatigue associated with many disabilities. Each subtest in the adaptive version is administered separately with items administered until a minimum level of measurement precision is achieved. Most commonly, a child will be administered eight items on each subtest to yield measurement precision equivalent to a subtest reliability between .75 and .80. If a child completes all three subtests, the reliability of the overall PA score is above .90. Testing time for all three subtests on the adaptive version usually takes under 10 min.

Clinical Implications

ATLAS-PA provides general information about children's PA using a web-based platform that is easy to access and inexpensive. As with other curriculum-based measures (e.g., IGDIs), ATLAS-PA can be used to monitor children's progress over time. However, unlike other similarly structured measures, ATLAS-PA relies entirely on receptive responses and includes individualized levels of instruction. PA is one of the key targets for early literacy assessment (Lonigan et al., 2009), instruction (Lemons & Fuchs, 2010), and intervention (Hund-Reid & Schneider, 2013; Skibbe et al., 2011). Given the strong predictive value of PA for later reading development (National Early Literacy Panel, 2008), ATLAS-PA represents an important tool that educators and

practitioners can use to evaluate and monitor children's PA skill development during early childhood.

By utilizing an adaptive format, ATLAS-PA will minimize the amount of time students need to spend being assessed. Attention is linked to language throughout early childhood (Gooch et al., 2016) and is a concern for many children with speech and/or language impairment (Maher et al., 2015), making a shorter testing time an especially important consideration. In addition, the adaptive nature of this measure also facilitates its use as a progress monitoring tool, as children will be exposed to different items on each testing occasion. General outcomes progress monitoring tools, such as the IGDIs (McConnell et al., 2002), have long been touted as the best way to capture young children's development during preschool and kindergarten, but are often not inclusive of children with speech and/or language impairment.

Limitations and Future Directions

Our goal was to create a measure of PA that could be used by children who had an educational determination of speech and/or language impairment within their schools. By design, this included children whose challenges resulted from a variety of disabilities, including speech and/or language impairment, autism, physical disabilities, and intellectual disabilities. Also, since we did not measure speech or language abilities directly, it is possible that a small percentage of the sample identified as having typical development actually had an undiagnosed speech and/or language impairment. Across etiologies, items functioned well; however, we did not have a large enough sample to detect whether particular items are problematic for specific disability groups. Some research, for example, has suggested that color memory and perception are more challenging for children with autism (Franklin et al., 2008), yet some of our items included pictures representing color words (e.g., red). In addition, the variability in motor, cognitive, and vocal capabilities within our sample precluded the use of a gold standard measure of PA for our children with speech and/or language impairment. Understanding whether construct validity varies for a particular subsample will be an ongoing area of future research.

Item responses from the ATLAS-PA are captured in a central data repository, which will allow us to consider, as more data become available, whether individual differences in performance can be attributed to particular types of disabilities. ATLAS-PA was designed to be appropriate for a broad range of children with speech and/or language impairment, although we did not confirm any reported diagnoses with direct assessment. We also recognize that, although ATLAS-PA is accessible to a broader range of children than other assessments, additional work may be needed to increase its accessibility. In the present work, 20 children with speech and/or language impairments were excluded from the present work because they could not complete ATLAS-PA; of these, nine were minimally verbal and all had behaviors that interfered with the testing process, a common barrier for assessment of children with higher needs (Tager-Flusberg et al., 2017). Note that three children who were minimally verbal were able to complete ATLAS-PA without difficulty, so the absence of spoken language did not appear to be a barrier for completion in isolation of challenging behaviors. Additional work is needed to consider how to expand the population of students for whom ATLAS-PA is relevant and valid.

Conclusions

ATLAS-PA is a new adaptive measure of PA with strong psychometric properties, available to interested users at www.accesstoliteracy.com (website launch anticipated spring 2020). By designing items and instructions to be appropriate for children who have speech and/or language impairment, this measure allows researchers and educational professionals to capture a more accurate assessment of PA skills for children with speech and/or language impairment than other PA measures currently available in the field. Specifically, ATLAS-PA provides educational professionals and researchers with the following: (a) a more effective tool to assess this critical literacy skill among children across the ability spectrum, (b) opportunities to include a broader population of children in regular early elementary screening systems rather than relying on modified assessments or tests with often nonvalidated and thus questionable accommodations, and (c) a PA assessment that relies on more rigorous approaches to validation for children with speech and/or language impairment than typically employed with such assessments.

Author Contributions

Lori E. Skibbe: Conceptualization (Equal), Data curation (Equal), Funding acquisition (Lead), Writing - Original Draft (Lead), Writing - Review & Editing (Lead). **Ryan P. Bowles:** Conceptualization (Equal), Data curation (Lead), Formal analysis (Lead), Funding acquisition (Supporting), Methodology (Lead), Writing - Original Draft (Supporting), Writing - Review & Editing (Supporting). **Sarah Goodwin:** Data curation (Supporting), Formal analysis (Supporting), Validation (Supporting), Visualization (Lead), Writing - Original Draft (Supporting), Writing - Review & Editing (Supporting). **Gary A. Troia:** Conceptualization (Supporting), Funding acquisition (Supporting), Writing - Original Draft (Supporting), Writing - Review & Editing (Supporting). **Haruka Konishi:** Methodology (Supporting), Project administration (Supporting), Writing - Original Draft (Supporting), Writing - Review & Editing (Supporting).

Acknowledgments

References

American Education Research Association/American Psychological Association/National Council on Measurement in Education. (2014). Standards for educational and psychological testing. American Education Research Association.

Andrich, D., Humphry, S. M., & Marais, I. (2012). Quantifying local, response dependence between two polytomous items using the Rasch model. *Applied Psychological Measurement, 36*(4), 309–324. https://doi.org/10.1177/0146621612441858

Anthony, J. L., Aghara, R. G., Dunkelberger, M. J., Anthony, T. I., 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*(2), 146–160. https://doi.org/10.1044/1058-0360(2011/10-0053)

Anthony, J. L., Davis, C., Williams, J. M., & Anthony, T. I. (2014). Preschoolers' oral language abilities: A multilevel examination of dimensionality. *Learning and Individual Differences, 35,* 56–61. https://doi.org/10.1016/j.lindif.2014.07.004

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*(1), 43–55. https://doi.org/10.1037/0022-0663.96.1.43

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*(1), 65–92. https://doi.org/10.1006/jecp.2002.2677

Barker, R. M., Bridges, M. S., & Saunders, K. J. (2014). Validity of a non-speech dynamic assessment of phonemic awareness via the alphabetic principle. *Augmentative and Alternative Communication, 30*(1), 71–82. https://doi.org/10.3109/07434618.2014.880190

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

Browder, D., Gibbs, S., Ahlgrim-Delzell, L., Courtade, G. R., Mraz, M., & Flowers, C. (2009). Literacy for students with severe developmental disabilities: What should we teach and what should we hope to achieve. *Remedial and Special Education, 30*(5), 269–282. https://doi.org/10.1177/0741932508315054

Browder, D. M., Wakeman, S. Y., Spooner, F., Ahlgrim-Delzell, L., & Algozzine, R. (2006). Research on reading instruction for individuals with significant cognitive disabilities. *Exceptional Children, 72*(4), 392–408. https://doi.org/10.1177/001440290607200401

Bus, A. G., Takacs, Z. K., & Kegel, C. A. T. (2015). Affordances and limitations of electronic storybooks for young children's emergent literacy. *Developmental Review, 35,* 79–97. https://doi.org/10.1016/j.dr.2014.12.004

Calhoun, S. L., & Mayes, S. D. (2005). Processing speed in children with clinical disorders. *Psychology in the Schools, 42*(4), 333–343. https://doi.org/10.1002/pits.20067

Card, R., & Dodd, B. (2006). The phonological awareness abilities of children with cerebral palsy who do not speak. *Augmentative and Alternative Communication, 22*(3), 149–159. https://doi.org/10.1080/07434610500431694

Cassar, M., & Treiman, R. (2004). Developmental variations in spelling: Comparing typical and poor spellers. In C. A. Stone, E. R. Silliman, B. Ehren, & K. Apel (Eds.), *Handbook of language and literacy: Development and disorders.* Guilford.

Catts, H. W., Fey, M. E., Tomblin, J. B., & Zhang, X. (2002). A longitudinal investigation of reading outcomes in children with language impairments. *Journal of Speech, Language, and Hearing Research, 45*(6), 1142–1157. https://doi.org/10.1044/1092-4388(2002/093)

Chapelle, C. A., & Douglas, D. (2006). *Assessing language through computer technology.* Cambridge University Press.

Coltheart, M. (1981). The MRC psycholinguistic database. *The Quarterly Journal of Experimental Psychology, 33*(4), 497–505. https://doi.org/10.1080/14640748108400805

Coyne, M. D., Zipoli, R. P., Jr., & Ruby, M. F. (2006). Beginning reading instruction for students at risk for reading disabilities: What, how, and when. *Intervention in School and Clinic, 41*(3), 161–168. https://doi.org/10.1177/10534512060410030601

Dahlgren Sandberg, A. (2001). Reading and spelling, phonological awareness, and working memory in children with severe speech impairments: A longitudinal study. *Augmentative and Alternative Communication, 17*(1), 11–26.

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*(8), 629–634. https://doi.org/10.1111/j.1469-8749.2006.tb01331.x

Dessemontet, R. S., de Chambrier, A.-F., Martinet, C., Moser, U., & Bayer, N. (2017). Exploring phonological awareness skills in children with intellectual disability. *American Journal on Intellectual and Developmental Disabilities, 122*(6), 476–491. https://doi.org/10.1352/1944-7558-122.6.476

Dewey, E. N., Latimer, R. J., Kaminski, R. A., & Good, R. H. (2012). *DIBELS Next Development: Findings from Beta 2 Validation Study* (Technical Report No. 10). Dynamic Measurement Group.

Dynia, J. M., Bean, A., Justice, L. M., & Kaderavek, J. N. (2019). Phonological awareness emergence in preschool children with autism spectrum disorder. *Autism & Developmental Language Impairments, 4,* 1–15. https://doi.org/10.1177/2396941518822453

Dynia, J. M., Brock, M. E., Justice, L. M., & Kaderavek, J. N. (2017). Predictors of decoding for children with autism spectrum disorder in comparison to their peers. *Research in Autism Spectrum Disorders, 37,* 41–48. https://doi.org/10.1016/j.rasd.2017.02.003

Ehri, L. C., & Snowling, M. J. (2004). Developmental variation in word recognition. In C. Addison Stone, E. R. Silliman, B. J. Ehren, & K. Apel (Eds.), *Handbook of language and literacy: Development and disorders* (pp. 433–460). Guilford.

Embretson, S. E. (1983). Construct validity: Construct representation versus nomothetic span. *Psychological Bulletin, 93*(1), 179–197. https://doi.org/10.1037/0033-2909.93.1.179

Farquharson, K. (2015). After dismissal: Examining the language, literacy, and cognitive skills of children with remediated speech sound disorders. *Perspectives on School-Based Issues, 16*(2), 50–59. https://doi.org/10.1044/sbi16.2.50

Farquharson, K., Hogan, T., & Bernthal, J. E. (2018). Working memory in school-age children with and without a persistent speech sound disorder. *International Journal of Speech-Language Pathology, 20*(4), 422–433. https://doi.org/10.1080/17549507.2017.1293159

Francis, D. J., Shaywitz, S. E., Stuebing, K. K., Shaywitz, B. A., & Fletcher, J. M. (1996). Developmental lag versus deficit models of reading disability: A longitudinal, individual growth curves analysis. *Journal of Educational Psychology, 88*(1), 3–17. https://doi.org/10.1037/0022-0663.88.1.3

Franklin, A., Sowden, P., Burley, R., Notman, L., & Alder, E. (2008). Color perception in children with autism. *Journal of Autism and Developmental Disorders, 38,* 1837–1847. https://doi.org/10.1007/s10803-008-0574-6

Gershon, R. (1992). Guessing and measurement. *Rasch Measurement Transactions, 6*(2), 209–210.

Gooch, D., Thompson, P., Nash, H. M., Snowling, M. J., & Hulme, C. (2016). The development of executive function and language skills in the early school years. *The Journal of Child Psychology and Psychiatry, 57*(2), 180–187. https://doi.org/10.1111/jcpp.12458

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

Grindle, C. F., Hughes, J. C., Saville, M., Huxley, K., & Hastings, R. P. (2013). Teaching early reading skills to children with autism using Mimiosprout Early Reading. *Behavioral Interventions, 28*(3), 203–224. https://doi.org/10.1002/bin.1364

Hesketh, A. (2004). Early literacy achievement of children with a history of speech problems. *International Journal of Language & Communication Disorders, 39*(4), 453–468. https://doi.org/10.1080/13682820410001686013

Hund-Reid, C., & Schneider, P. (2013). Effectiveness of phonological awareness intervention for kindergarten children with language impairment. *Canadian Journal of Speech-Language Pathology & Audiology, 37*(1), 6–25.

Iacono, T., & Cupples, L. (2004). Assessment of phonemic awareness and word reading skills of people with complex communication needs. *Journal of Speech, Language, and Hearing Research, 47*(2), 437–449. https://doi.org/10.1044/1092-4388 (2004/035)

Invernizzi, M., Juel, C., Swank, L., & Meier, J. (2003). *Phonological Awareness Literacy Screening–Kindergarten*. University of Virginia.

Invernizzi, M., Landrum, T. J., Teichman, A., & Townsend, M. (2010). Increased implementation of emergent literacy screening in pre-kindergarten. *Early Childhood Education Journal, 37,* 437–446. https://doi.org/10.1007/s10643-009-0371-7

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. https://doi.org/10.1044/persp1.SIG16.78

Justice, L. M., Bowles, R. P., Pence Turnbull, K. L., & Skibbe, L. E. (2009). School readiness among children with varying histories of language difficulties. *Developmental Psychology, 45*(2), 460–476. https://doi.org/10.1037/a0014324

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

Kaufman, J. N., Donders, J., & Warschausky, S. (2014). A comparison of visual inspection time measures in children with cerebral palsy. *Rehabilitation Psychology, 59*(2), 147–154. https://doi.org/10.1037/a0036019

Kaufman, J., Warschausky, S., Van Tubbergen, M., Asbell, S., & Donders, J. (2009, February). *Modified assessment of visual inspection time in children with cerebral palsy*. Paper presented at the annual meeting of the International Neuropsychological Society, Atlanta, GA, United States.

Koppenhaver, D. A., Hendrix, M. P., & Williams, A. R. (2007). Toward evidence-based literacy interventions for children with severe and multiple disabilities. *Seminars in Speech and Language, 28*(1), 79–89. https://doi.org/10.1055/s-2007-967932

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods, 44*(4), 978–990. https://doi.org/10.3758/s13428-012-0210-4

Lemons, C. J., & Fuchs, D. (2010). Phonological awareness of children with Down syndrome: Its role in learning to read and the effectiveness of related interventions. *Research in Developmental Disabilities, 31*(2), 316–330. https://doi.org/10.1016/j.ridd.2009.11.002

Lemons, C. J., Kloo, A., Zigmond, N., Fulmer, D., & Lupp, L. (2012). Implementing an Alternate Assessment based on Modified Academic Achievement Standards: When policy meets practice. *International Journal of Disability, Development, and Education, 59*(1), 67–79. https://doi.org/10.1080/1034912X.2012.654963

Lemons, C. J., Vaughn, S., Wexler, J., Kearns, D. M., & Sinclair, A. C. (2018). Envisioning an improved continuum of special education services for students with learning disabilities: Considering intervention intensity. *Learning Disabilities Research & Practice, 33*(3), 131–143. https://doi.org/10.1111/ldrp.12173

Lerner, M. D., & Lonigan, C. J. (2016). Bidirectional relations between phonological awareness and letter knowledge in preschool revisited: A growth curve analysis of the relation between two code-related skills. *Journal of Experimental Child Psychology, 144,* 166–183. https://doi.org/10.1016/j.jecp.2015.09.023

Linacre, J. M. (1998). Structure in Rasch residuals: Why principal components analysis (PCA)? *Rasch Measurement Transactions, 12*(2), 636. https://www.rasch.org/rmt/rmt122m.htm

Linacre, J. M. (2018a, December). *Dimensionality: Contrasts & variances.* https://www.winsteps.com/winman/principalcomponents.htm

Linacre, J. M. (2018b, December). *Table 30.1 Differential item functioning DIF pairwise.* https://www.winsteps.com/winman/table30_1.htm

Linacre, J. M. (2018c). *Winsteps® Rasch measurement computer program*. Winsteps.com

Lonigan, C. J., Anthony, J. L., Phillips, B. M., Purpura, D. J., Wilson, S. B., & McQueen, J. (2009). The nature of preschool phonological processing abilities and their relations to vocabulary, general cognitive abilities, and print knowledge. *Journal of Educational Psychology, 101*(2), 345–358. https://doi.org/10.1037/a0013837

Lonigan, C. J., Burgess, S. R., Anthony, J. L., & Barker, T. A. (1998). Development of phonological sensitivity in 2- to 5-year-old children. *Journal of Educational Psychology, 90*(2), 294–311. https://doi.org/10.1037/0022-0663.90.2.294

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.

Madsen, H. S. (1991). Computer-adaptive test of listening and reading comprehension: The Brigham Young University approach. In P. Dunkel (Ed.), *Computer-assisted language learning and testing: Research issues and practice* (pp. 237–257). House.

Maher, C., Crettenden, A., Evans, K., Thiessen, M., Toohey, M., Watson, A., & Dollman, J. (2015). Fatigue is a major issue for children and adolescents with physical disabilities. *Developmental Medicine & Child Neurology, 57*(8), 742–747. https://doi.org/10.1111/dmcn.12736

McConnell, S. R., McEvoy, M. A., & Priest, J. S. (2002). "Growing" measures for monitoring progress in early childhood education: A research and development process for individual growth and development indicators. *Assessment for Effective Intervention, 27*(4), 3–14. https://doi.org/10.1177/073724770202700402

Missall, K. N., McConnell, S. R., & Cadigan, K. (2006). Early literacy development: Skill growth and relations between classroom

variables for preschool children. *Journal of Early Intervention, 29*(1), 1–21. https://doi.org/10.1177/105381510602900101

Moyle, M. J., Heilman, J., & Berman, S. S. (2013). Assessment of early developing phonological awareness skills: A comparison of the Preschool Individual Growth and Development Indicators and the Phonological Awareness and Literacy Screening–PreK. *Early Education & Development, 24*(5), 668–686. https://doi.org/10.1080/10409289.2012.725620

National Early Literacy Panel. (2008). *Developing early literacy: Report of the National Early Literacy Panel*. National Institute for Literacy.

Næss, K. B., Melby-Lervåg, M., Hulme, C., & Lyster, S. H. (2012). Reading skills in children with Down syndrome: A meta-analytic review. *Research in Developmental Disabilities, 33*(2), 737–747. https://doi.org/10.1016/j.ridd.2011.09.019

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. https://doi.org/10.1016/j.ridd.2008.10.002

Peeters, M., Verhoeven, L., van Balkom, H., & de Moor, J. (2008). Foundations of phonological awareness in pre-school children with cerebral palsy: The impact of intellectual disability. *Journal of Intellectual Disability Research, 52*(1), 68–78. https://doi.org/10.1111/j.1365-2788.2007.00986.x

Pentimonti, J. M., Murphy, K. A., Justice, L. M., Logan, J. A. R., & Kaderavek, J. N. (2016). School readiness of children with language impairment: Predicting literacy skills from pre-literacy and social–behavioural dimensions. *Journal of Language & Communication Disorders, 51*(2), 148–161. https://doi.org/10.1111/1460-6984.12193

Perline, R., Wright, B. D., & Wainer, H. (1979). The Rasch model as additive conjoint measurement. *Applied Psychological Measurement, 3*(2), 237–255. https://doi.org/10.1177/014662167900300213

Preston, J., & Edwards, M. L. (2010). Phonological awareness and types of sound errors in preschoolers with speech sound disorders. *Journal of Speech, Language, and Hearing Research, 53*(1), 44–60. https://doi.org/10.1044/1092-4388(2009/09-0021)

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

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. https://doi.org/10.1007/978-1-4613-0169-1_2

Roth, F., Troia, G., Worthington, C., & 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*(2), 67–88. https://doi.org/10.2307/30035536

Rvachew, S., & Grawburg, M. (2006). Correlates of phonological awareness in preschoolers with speech sound disorders. *Journal of Speech, Language, and Hearing Research, 49*(1), 74–87. https://doi.org/10.1044/1092-4388/06/4901-0074

Schatschneider, C., Francis, D. J., Foorman, B. R., Fletcher, J. M., & Mehta, P. (1999). The dimensionality of phonological awareness: An application of item response theory. *Journal of Educational Psychology, 91*(3), 439–449. https://doi.org/10.1037/0022-0663.91.3.439

Shank, L. K., Kaufman, J., Leffard, S., & Warschausky, S. (2010). Inspection time and attention-deficit/hyperactivity disorder symptoms in children with cerebral palsy. *Rehabilitation Psychology, 55*(2), 188–193. https://doi.org/10.1037/a0019601

Skibbe, L. E., Justice, L. M., & Bowles, R. P. (2011). Implementation processes associated with a home-based phonological awareness intervention for children with specific language impairment. *International Journal of Speech-Language Pathology, 13*(2), 110–124. https://doi.org/10.3109/17549507.2011.524246

Snow, C. E., Burns, S., & Griffith, P. (Eds.) (1998). *Preventing reading difficulties in young children*. National Academy Press.

Tager-Flusberg, H., Skwere, D. P., Joseph, R. M., Brukilaccio, B., Decker, J., Eggleston, B., Meyer, S., & Yoder, A. (2017). Conducting research with minimally verbal participants with autism spectrum disorder. *Autism, 21*(7), 852–861. https://doi.org/10.1177/1362361316654605

Tambyraja, S. R., Farquharson, K., Logan, J. A. R., & Justice, L. M. (2015). Decoding skills in children with language impairment: Contributions of phonological processing and class experiences. *American Journal of Speech-Language Pathology, 24*(2), 177–188. https://doi.org/10.1044/2015_AJSLP-14-0054

Thatcher, K. L. (2010). The development of phonological awareness with specific language-impaired and typical children. *Psychology in the Schools, 47*(5), 467–480. https://doi.org/10.1002/pits.20483

Thompson, J. L., Plavnick, J. B., & Skibbe, L. E. (2019). Eye-tracking analysis of attention to an electronic storybook for minimally verbal children with autism spectrum disorder. *The Journal of Special Education, 53*(1), 41–50. https://doi.org/10.1177/0022466918796504

Thurlow, M. L. (2010). Steps toward creating fully accessible reading assessments. *Applied Measurement in Education, 23*(2), 121–131. https://doi.org/10.1080/08957341003673765

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1994). Longitudinal studies of phonological processing and reading. *Journal of Learning Disabilities, 27*(5), 276–286. https://doi.org/10.1177/002221949402700503

Troia, G. A., Roth, F. P., & Graham, S. (1998). An educators' guide to phonological awareness: Measures and intervention activities for children. *Focus on Exceptional Children, 31*(3), 1–12. https://doi.org/10.17161/foec.v31i3.6917

University of Oregon Center for Teaching and Learning. (2018). *DIBELS 8th Edition.*

U.S. Department of Education. (2017). *Office of Special Education Programs, Individuals with Disabilities Education Act (IDEA) database*. Retrieved July 10, 2017, from https://www2.ed.gov/programs/osepidea/618-data/state-level-data-files/index.html#bcc

Vandervelden, M. C., & Siegel, L. S. (2001). Phonological processing in written word learning: Assessment for children who use augmentative and alternative communication. *Augmentative and Alternative Communication, 17*(1), 37–51. https://doi.org/10.1080/aac.17.1.37.51

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1994). Development of reading-related phonological processing abilities: New evidence of bidirectional causality from a latent variable longitudinal study. *Developmental Psychology, 30*(1), 73–78. https://doi.org/10.1037/0012-1649.30.1.73

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

Warschausky, S. (2009, February). *Assessment in pediatric rehabilitation psychology: Access and psychometrics*. Invited address, American Psychological Association Division 22 Midyear Conference, Jacksonville, FL, United States.

Watling, R., & Schwartz, I. S. (2004). Understanding and implementing positive reinforcement as an intervention strategy for children with disabilities. *The American Journal of Occupational Therapy, 58,* 113–116. https://doi.org/10.5014/ajot.58.1.113

Weiss, D. J. (1982). Improving measurement quality and efficiency with adaptive testing. *Applied Psychological Measurement, 6*(4), 473–492. https://doi.org/10.1177/014662168200600408

Wright, B. D., & Linacre, J. M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions, 8,* 370. https://www.rasch.org/rmt/rmt83b.htm

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. https://doi.org/10.1111/j.1745-3984.1999.tb00543.x