How Many Letters Should Preschoolers in Public Programs Know? The Diagnostic Efficiency of Various Preschool Letter-Naming Benchmarks for Predicting First-Grade Literacy Achievement

Shayne B. Piasta, **Yaacov Petscher**, and **Laura M. Justice**
The Ohio State University, Florida State University

Abstract

Review of current federal and state standards indicates little consensus or empirical justification regarding appropriate goals, often referred to as benchmarks, for preschool letter-name learning. The present study investigated the diagnostic efficiency of various letter-naming benchmarks using a longitudinal database of 371 children who attended publicly funded preschools. Children's uppercase and lowercase letter-naming abilities were assessed at the end of preschool, and their literacy achievement on 3 standardized measures was assessed at the end of 1st grade. Diagnostic indices (sensitivity, specificity, and negative and positive predictive power) were generated to examine the extent to which attainment of various preschool letter-naming benchmarks was associated with later risk for literacy difficulties. Results indicated generally high negative predictive power for benchmarks requiring children to know 10 or more letter names by the end of preschool. Balancing across all diagnostic indices, optimal benchmarks of 18 uppercase and 15 lowercase letter names were identified. These findings are discussed in terms of educational implications, limitations, and future directions.

Keywords

emergent literacy; alphabet knowledge; letter naming; federal and state standards; benchmark; early learning

Research in the past two decades has greatly increased knowledge concerning the beginnings of successful academic trajectories for children. Such work has established that young children's emerging understandings of numbers, letters, and sounds are important predictors of later academic achievement (e.g., Badian, 1995; Hammill, 2004; Jordan, Kaplan, Locuniak, & Ramineni, 2007; Reynolds & Bezruczko, 1993). As a result, benchmarks for knowledge of these concepts in preschool and kindergarten are being set forth by federal, state, and professional organizations (e.g., National Association for the Education of Young Children, 1998; Neuman & Roskos, 2005; U.S. Department of Health and Human Services, Administration for Children and Families, 2003), and these are often

Correspondence concerning this article should be addressed to: Shayne B. Piasta, Children's Learning Research Collaborative, The Ohio State University, 356 Arps Hall, 1945 North High Street, Columbus, OH 43207. piasta.1@osu.edu.

used to make important educational and fiscal decisions. One example of such a benchmark comes from the Office of Head Start, in which programs are legislatively mandated to collect data on children's progress at least three times annually toward identifying "at least ten letters of the alphabet, especially those in their own name" (U.S. Department of Health and Human Services, Administration for Children and Families, 2003). Benchmarks such as these, however, have not always been examined in terms of their utility for ensuring later academic success. The latter is especially important given that teachers may be inclined to focus instructional attention and assessment activities toward benchmarked skills and criteria (Powell, Diamond, Bojczyk, & Gerde, 2008).

Benchmarks focusing on a specific measurable skill, as opposed to broad competencies, may particularly influence teachers' instructional foci (Powell et al., 2008). The present study investigated one such specific indicator of emergent literacy ability, namely, the number of letters a child ought to know (i.e., letter-naming benchmarks). We examined the extent to which various preschool letter-naming benchmarks were associated with successful literacy outcomes in first grade. Although only one of a plethora of benchmarks in early childhood education, letter-naming benchmarks are particularly common in early learning standards and largely familiar to preschool teachers (Powell et al., 2008). In fact, Powell et al.'s (2008) work suggests that letter knowledge and related benchmarks are emphasized above other literacy goals in many preschool teachers' classrooms. Moreover, empirical attention to establishing the diagnostic efficacy of letter-naming benchmarks can provide a model for how other benchmarks can be rigorously examined.

The Importance of Early Letter-Naming Ability

Research on letter naming and its role in preparing children for literacy success has a long history in the United States (e.g., see Chall, 1967/1983; Durrell, 1980; Ehri, 1983; Foulin, 2005; Groff, 1984; Mason, 1984; Piasta & Wagner, 2010a; Treiman & Kessler, 2003, for descriptions and discussions of such research). As evidenced by this literature, the ability to name the letters of the alphabet during preschool and kindergarten is a well-established predictor of children's later literacy skills (Hammill, 2004; National Early Literacy Panel, 2008; Scarborough, 1998; Schatschneider, Fletcher, Francis, Carlson, & Foorman, 2004). The recent synthesis conducted by the National Early Literacy Panel (2008), for example, indicated significant predictive correlations of .48 to .54 between early letter knowledge and later decoding, spelling, and reading comprehension skills. Such relations between early letter knowledge and later literacy skills appear to be independent of children's age, socioeconomic status, IQ, and other emergent literacy skills, such as oral language and phonological awareness (National Early Literacy Panel, 2008; Snowling, Gallagher, & Frith, 2003; Storch & Whitehurst, 2002). Letter-naming ability has thus been asserted as the best early predictor of children's later literacy success.

Importantly, the relation between children's early alphabet abilities and later literacy skills is likely causal in nature (Ehri, 1987; Levin, Shatil-Carmon, & Asif-Rave, 2006; Piasta & Wagner, 2010a; Treiman & Kessler, 2003; cf. Foulin, 2005). In learning about individual letters, children develop their initial understanding of the symbolic nature of written language. Ultimately, they learn that written letters represent the sounds of spoken language

and can be used to map print to speech (i.e., the alphabetic principle). This beginning understanding is evident in children's early reading and writing attempts. Young children's emergent writing, for example, often includes invented spellings in which letters are used to represent syllables or sounds, such as writing *JF* for *giraffe* or *BBL* for *baseball*, and often includes reliance on letter names when representing sounds, such as writing *MT* for *empty* (Read, 1971; Richgels, 1986; Treiman & Tincoff, 1997; Treiman, Tincoff, & Richmond-Welty, 1996; Treiman, Weatherston, & Berch, 1994). Moreover, research suggests that letter names aid children in learning associated letter sounds; children are able to extract the sound cues contained at the beginning (e.g., /b/ in *B*) or end (e.g., /f/ in *F*) of letter names (Evans, Bell, Shaw, Moretti, & Page, 2006; Piasta & Wagner, 2010b; Treiman, Pennington, Shriberg, & Boada, 2008; Treiman, Tincoff, Rodriguez, Mouzaki, & Francis, 1998). These findings, along with similar evidence concerning the early development of decoding abilities (e.g., Byrne & Fielding-Barnsley, 1989; Ehri, 1987; Treiman & Kessler, 2003), have led to the conclusion that letter knowledge is essential for children's development of literacy skills (Ehri, 1998).

The implications of these findings are that children with high letter-naming abilities in preschool and kindergarten are likely to experience success in literacy learning, whereas children with low letter-naming abilities are likely to experience later literacy difficulties. Indeed, children considered at risk for literacy difficulties, whether due to socioeconomic, cognitive, or genetic factors, tend to have lower letter-naming abilities than peers who are not at risk (Bowey, 1995; Duncan & Seymour, 2000; Elbro & Petersen, 2004;

Lyytinen et al., 2004; Snowling et al., 2003; Torppa, Poikkeus, Laakso, Eklund, & Lyytinen, 2006). Longitudinal studies indicate that low letter-naming abilities are characteristic of those children who exhibit reading disabilities (Catts, Fey, Zhang, & Tomblin, 2001; Puolakanaho et al., 2007; Torppa, Lyytinen, Erskine, Eklund, & Lyytinen, 2010; Torppa et al., 2006). Catts et al. (2001), for instance, followed an epidemiological sample of 604 kindergarteners into second grade. Of a large battery of language- and literacy-related measures, letter-naming ability was the best early predictor of whether children exhibited reading comprehension deficits in second grade. Similar results were found by Puolakanaho et al. (2007) when letter-naming abilities were assessed prior to kindergarten entry in a longitudinal study of 198 children. In fact, letter-naming ability as assessed at ages 3.5 years, 4.5 years, and 5.5 years was the most consistent predictor of second-grade word reading and spelling disabilities.

These correlational results are complemented by additional studies showing group differences in children's early letter knowledge. In a prospective study that followed children from 36 months to 8 years of age, for example, Scarborough (1990, 1991) found significant delays in kindergarten letter knowledge for those children with reading disabilities in Grade 2 as compared with those children with typical literacy skills. Gallagher, Frith, and Snowling (2000) and Snowling et al. (2003) reported similar results when retrospectively examining the preschool letter knowledge of children with and without reading disabilities in first and second grade; children with identified reading disabilities exhibited significantly lower levels of letter-naming abilities than children with normal literacy skill development. On the basis of this research, many commonly used kindergarten

readiness assessments (e.g., Brigance Inventory of Early Development–II: Brigance, 2004; Early Screening Inventory: Meisels & Wiske, 1983) and early literacy screeners (e.g., Florida Assessments for Instruction in Reading: Florida Department of Education, 2009; Dynamic Indicators of Basic Early Literacy Skills: Good, Kaminski, Smith, Laimon, & Dill, 2001; Phonological Awareness Literacy Screening for Preschool: Invernizzi, Sullivan, Meier, & Swank, 2004; Texas Primary Reading Inventory: Texas Education Agency, 1998) include letter-naming components.

Letter Naming as a Learning Goal

Given the centrality of alphabet understanding for children's literacy acquisition, it is unsurprising that letter-naming ability is an important learning goal for young children. It is recognized as an important component of school readiness (Hair, Halle, Terry-Humen, Lavelle, & Calkins, 2006) and acknowledged by professional organizations and government agencies as an essential skill for children to acquire (e.g., National Association for the Education of Young Children, 1998; U.S. Department of Health and Human Services, Administration for Children and Families, 2003). Accordingly, many preschool and kindergarten curricula include emphases on letter-name learning specifically (e.g., Let's Begin With the Letter People: Abrams Learning Trends, 2001; Creative Curriculum: Dodge, Colker, & Heroman, 2002; Open Court Reading Pre-K: SRA/McGraw-Hill, 2003; see also Justice, Pence, Bowles, & Wiggins, 2006) as do many early intervention programs (see Piasta & Wagner, 2010a, for a review). Both teachers and parents are urged to include letter naming as part of typical learning routines (Armbruster, Lehr, & Osborn, 2001, 2003).

The emphasis on letter-name learning during early childhood has been formalized in the standards adopted by states and federal programs. Initiatives to create early learning standards and the associated benchmarks embedded within them have taken hold over the past 10 years, largely because of the requirements of federal programs and federal funding provided to states (Neuman & Roskos, 2005). For example, as we noted earlier, the federal Head Start legislation set a benchmark of naming at least 10 letters for its preschool graduates (U.S. Department of Health and Human Services, Administration for Children and Families, 2003). The Early Reading First and Reading First programs, created as a result of the No Child Left Behind Act of 2002, also set ambitious goals for promoting children's letter-naming abilities (U.S. Department of Education, 2002, 2003). During 2005 through 2008, for instance, Early Reading First performance targets indicated that the average child completing the program ought to know 16 to 19 letter names; grantees were required to report how many children participating in Early Reading First projects met this goal annually (U.S. Department of Education, 2009). The newly developed Common Core State Standards (National Governors Association Center for Best Practices and Council of Chief State School Officers, 2010) also sets forth letter-naming benchmarks, indicating that kindergarten children ought to be able to "name all upper and lowercase letters of the alphabet" (p. 15).

General reviews of state early learning standards (e.g., Bracken & Crawford, 2010; Neuman & Roskos, 2005) indicate that the majority of states include standards related to children's alphabet learning. States appear to vary, however, in the specificity of these standards and

whether explicit benchmarks for letter naming are set forth. In our own informal review of current state standards, conducted in winter 2010, we found 10 states that set specific standards or benchmarks for children's letter naming at the end of preschool. Table 1 provides an overview of these standards and lines by setting benchmarks of knowing 10 letter names, and 21 states (Arkansas, Delaware, Georgia, Illinois, Maine, Michigan, Minnesota, Missouri, North Carolina, Nebraska, New Jersey, New Mexico, Nevada, New York, Ohio, Oregon, Pennsylvania, South Carolina, Tennessee, Washington, West Virginia) set more general guidelines that children learn "some" or "several" letter names. Thirteen states (Colorado, Connecticut, Iowa, Kansas, Kentucky, Louisiana, Massachusetts, Mississippi, New Hampshire, Oklahoma, Rhode Island, Utah, Vermont) indicate that children should be able to recognize some letters but do not specify the need to know letter names. Finally, two states (Maryland, Wisconsin) do not mention letter names but indicate standards for learning letter sounds, and five states (Florida, Hawaii, Idaho, Montana, North Dakota) make general references to beginning to understand or be aware of letters. Notably, the states also differ in whether benchmarks are set only for uppercase letter naming (e.g., Indiana, Pennsylvania, Virginia), are specified for both uppercase and lowercase letter naming (e.g., California, Ohio, Texas, Washington), or are ambiguous with respect to case (e.g., Arizona, Wyoming, Oregon).

The plethora of benchmarks (see Table 1) that concern young children's letter knowledge indicate that many state and federal policymakers as well as constituents within professional organizations (e.g., Head Start) recognize the importance of early alphabet knowledge to later literacy achievement. However, it is also noteworthy that there is little consensus as to what are appropriate benchmarks for preschool letter-naming abilities.

Purpose of the Present Study

As is apparent from the literature reviewed thus far, it is clear that early letter-naming abilities are an important skill for young children to acquire. Nonetheless, although the federal government and all 50 states appear to recognize the importance of early alphabet knowledge, there is little agreement as to appropriate benchmarks for preschool letter-naming abilities. To our understanding, no justification is provided in any standards document concerning the letter-naming benchmarks set forth, nor have these benchmarks been empirically investigated to determine their merit. The purpose of the present study was to empirically investigate the utility of these various uppercase and lowercase letter-naming benchmarks in terms of the extent to which these predict children's successful acquisition of literacy skills in elementary school. Two specific research questions were addressed:

1. For children served in public programs, to what extent are existing benchmarks for end-of-preschool letter-naming abilities associated with subsequent risk status on first-grade measures of word reading, spelling, and reading comprehension?

2. To what extent can optimal end-of-preschool letter-naming benchmarks be identified for these children?

We anticipated that the results of this work would be theoretically informative for understanding how children's early letter knowledge may be associated with future risks in

reading achievement, contributing to a growing literature on this topic (e.g., Badian, 1995; Catts et al., 2001; Elbro & Petersen, 2004). More salient, however, study findings should have direct bearing on the myriad educational policies stipulating how many letters young children ought to know at specific educational junctures.

METHOD

Participants

Data were collected as part of a larger study of preschool shared reading practices conducted in Virginia and Ohio. The larger study involved a random selection of children (N = 551) enrolled in 85 preschool classrooms during the 2004–2005 or 2005–2006 academic years. All classrooms were supported by public funding (33 Head Start classrooms, 39 Title I or state-subsidized classrooms, 13 private preschool centers accepting vouchers). The majority of teachers (76%) reported using commercially available curricula commonly used in publicly funded preschool settings (e.g., see Clifford et al., 2005; Justice, Mashburn, Hamre, & Pianta, 2008): 38 teachers reporting use of HighScope (Hohmann & Weikart, 1995), 18 teachers reporting use of Creative Curriculum (Dodge, Colker, & Heroman, 2002), and nine teachers reporting use of the Language-Focused Curriculum (Bunce, 1995). Although it is unknown whether teachers who used no formal curriculum (24%) adhered to a systematic scope and sequence of letter instruction in their classrooms, none of the three curricula specifically named by teachers include explicit benchmarks for teaching letter names. Because teachers' use of a comprehensive curriculum is not generally associated with children's letter knowledge (Mashburn et al., 2008) and state and federal early learning standards are agnostic with respect to classroom curricula, curricula use is not considered a confound in this study.

Children eligible for the study (a) were enrolled in classrooms with participating teacher volunteers, (b) were between 3 years 6 months and 4 years 11 months of age at study entry, (c) were expected to enroll in kindergarten the following academic year, (d) did not have an individualized education plan for a cognitive or social/emotional disability that would prevent completion of assessments, and (e) were able to be assessed in English. Parent consent forms were distributed to all eligible children in each classroom. From those children for whom caregivers completed and returned consent forms, an average of six were randomly selected per classroom.

Children contributing data to the present study (n = 371) were those who completed a letter-naming assessment in the spring of preschool as well as follow-up assessments 2-years later, when most were completing first grade.[1] Of the original 551 children, 94 withdrew or did not complete preschool assessments, and 86 withdrew or could not be located for first-grade assessments. The 371 remaining children were 52 months of age, on average, at preschool entry (SD = 4.48 months). The majority of students were White, non-Hispanic (44%) or African American (36%); other children were Hispanic/Latino (5%), multiracial (10%), or of other races (3%); 2% were unreported. The majority of children's families reported average yearly incomes between $20,000 and $30,000, with 63% reporting incomes at or

---

[1]In the present sample, 64 children remained in kindergarten, and three were promoted to second grade at the follow-up assessment.

below $30,000 (range of $5,000 or less to over $85,000). The majority of children's mothers had high school diplomas as their highest degree (76%), with 8% holding associates degrees, 6% holding bachelor degrees, and 2% holding graduate-level degrees (8% unreported). These indicators of relatively low socioeconomic background are consistent with children's enrollment in publicly-funded preschools.

Information on children's prior educational history (prior to their enrollment in preschool during the project year) was not collected by project staff. Daily attendance records were, however, collected for the year of the study in which children attended preschool. For each child, the number of days present at school was divided by the total number of school days to arrive at a percentage daily attendance rate. On average, children attended 89% of days ($SD = .09$; range = 32%–100%).

Procedure

For the purposes of the present study, children were assessed in the spring of preschool and again 2 years later, during the spring of first grade for the majority of the sample. All children were assessed individually by trained research staff in quiet locations at their respective schools. All assessment activities were conducted within a 4-week assessment window.

Letter-naming ability—In the spring of preschool, children completed the Uppercase and Lowercase Alphabet Recognition subtests of the Phonological Awareness Literacy Screening for Preschool (Cronbach's α = .84; Invernizzi, Sullivan, et al., 2004). In each of these subtests, children are asked to name each of the 26 letters as presented in a random order on a single printed sheet. Children were first presented with and asked to respond to the sheet showing the 26 uppercase letters; on completion, children were presented with the sheet showing 26 lowercase letters. The number of correct responses was tallied separately for each of these subtests, such that scores reflected the number of uppercase letters and the number of lowercase letters that a child correctly named. Children were then classified as to whether or not they met various letter-naming benchmarks, to be discussed subsequently (see Analyses section).

Literacy achievement—Children's literacy achievement was assessed 2 years later using three subtests from the Woodcock– Johnson Tests of Achievement (3rd ed.; WJ-III; split-half reliabilities ranging from $r = .77$ to $r = .99$; Woodcock, McGrew, & Mather, 2001). Word reading was assessed using the WJ-III Letter–Word Identification subtest, in which children are asked to identify letters and read words of increasing difficulty. Although the Letter–Word Identification subtest does include an initial seven items that involve letter identification (children view an array of letters and are asked to identify or name individual letters), the tasks administered to first-grade children largely involve reading words of increasing complexity.[2] Spelling was assessed using the WJ-III Spelling subtest, in which children are asked to write letters and spell words of increasing difficulty. Similar to the Letter–Word Identification subtest, the spelling tasks include several initial tasks for young children involving printing individual letters but then shift to spelling of simple and increasingly complex individual words for first-grade age children. Reading comprehension

was assessed using the WJ-III Passage Comprehension subtest. Initial items on this subtest ask children to indicate which of several pictures are related in meaning; subsequent items follow a cloze procedure in which children are asked to select a picture or produce a word that completed a given phrase or written passage.

Raw scores for each of these subtests were converted to standard scores using the WJ-III CompuScore software (Schrank & Woodcock, 2001). These standard scores ($M = 100$, $SD = 15$) reflect the performance of children on these subtests relative to a normal distribution and were used to determine children's risk status. Children with standard scores at or below 96 were classified as being at risk for literacy difficulties; children with standard scores above 96 were classified as not at risk. A score of 96 or less is commensurate with scoring at or below the 40th percentile, which is the proficiency criterion used by the majority of the states (American Institutes for Research, 2007).

Analyses

The two research questions were addressed by examining diagnostic efficiency indices as generated by a series of 2×2 contingency matrices, similar to those used to evaluate the efficiency of educational screening assessments (e.g., Petscher, Kim, & Foorman, 2011; Streiner, 2003). Several alternative methods have been used to set benchmarks with the goal of identifying students who are at risk for later difficulties, including (a) bivariate correlations, (b) interrater reliability, (c) percentile rank cutpoints, (d) IQ–achievement discrepancy, and (e) expert panel analysis (see Schatschneider, Petscher, & Williams, 2008). Of these methods, Silberglitt and Hintze (2005) found that diagnostic efficiency provides the best balance between Type I and II errors and is more flexible with predictive power. Thus, this approach was used in the present study.

For the present purposes, separate matrices were generated for each letter-naming benchmark examined, with uppercase and lowercase letters considered independently, and for each literacy achievement measure (i.e., Letter–Word Identification, Spelling, Passage Comprehension). The matrices described children's dichotomized performance according to a selected letter-naming benchmark (e.g., fewer than 10 letters correctly known or 10 or more letters correctly known) as well as whether children were at risk on a given first-grade literacy outcome (e.g., below or at the 40th percentile on the WJ-III Letter–Word Identification or above the 40th percentile on the WJ-III Letter–Word Identification). Essentially, these matrices classified children into one of four categories, as illustrated in the Appendix: (a) not meeting the benchmark and being at risk on the first-grade literacy outcome (true positive); (b) not meeting the benchmark but not being at risk on the first-grade literacy outcome (false positive); (c) meeting the benchmark and being at risk on the

---

[2]The first 14 items of the WJ Letter–Word Identification subtest assess a mix of letter knowledge and word reading; subsequent items assess word reading only. In the current sample, only eight children (less than 2% of the sample) had raw scores of 14 or below on this subtest, indicating that the vast majority of our first-grade sample completed items that assessed word reading. To assuage concerns that associations between letter naming and WJ Letter–Word Identification scores were driven by the inclusion of a few letter-related items, we adjusted the WJ Letter–Word Identification raw scores to remove the 14 points awarded for such items and found that correlations between letter knowledge and the adjusted WJ Letter–Word Identification scores (rs .54 and .57, respectively, with uppercase and lowercase letter naming) mirrored those between letter knowledge and the unadjusted WJ Letter–Word Identification scores (r .54 and .54). We thus chose to use the unadjusted WJ Letter–Word Identification scores and corresponding standardized scores throughout the article.

first-grade literacy outcome (false negative); or (d) meeting the benchmark and not being at risk on the first-grade literacy outcome (true negative).

Using the formulas presented in the Appendix, these matrices were used to generate several diagnostic efficiency indices: sensitivity, specificity, positive predictive power, and negative predictive power. These indices represent those typically used to describe classification accuracy (Streiner, 2003). In the present study, sensitivity represented the proportion of children who did not meet the preschool benchmark out of all children who were at risk on first-grade literacy outcomes. Specificity represented the proportion of children who met the benchmark out of all children who were not at risk on first-grade literacy outcomes. Positive predictive power indicated that the proportion of children who were at risk on first-grade literacy outcomes out of all children who did not meet the preschool benchmark. Negative predictive power indicated the proportion of children who were not at risk on first-grade literacy outcomes out of all children who met the preschool benchmark. In addition to these indices, a phi coefficient was computed to provide a classification agreement index that corrected for chance.

To address the first research question, a set of letter-naming benchmarks were identified from current state and federal standards. The indices just described were generated for each of the following benchmarks: 10, 13, 16, 18, 19, and 20 letter names known (see Table 1). In addition, indices were generated for knowing all (26) of the letter names. We focused on the negative predictive power when examining these results. We reasoned that the implicit assumption behind standards is that children who meet particular benchmarks are thought to be on the path to success; in other words, if a child meets a given benchmark, he or she should not exhibit risk for literacy difficulties in first grade. This is captured in the negative predictive power. We additionally reasoned that although the converse might also hold (i.e., children who do not meet a benchmark should be more likely to exhibit risk in first grade, captured through positive predictive power), this may not be true, because children who do not meet benchmarks may be afforded supplemental instruction or other opportunities to increase their chances of success.

For the second research question, we examined sensitivity, specificity, positive predictive power, and the phi coefficient, in addition to negative predictive power for every possible benchmark (i.e., 1 through 26), broadening beyond those appearing across the variety of standards. Although we again emphasized the negative predictive power as explained earlier, we reasoned that optimal benchmarks should also exhibit minimal tradeoffs among sensitivity, specificity, and predictive power. We anticipated that sensitivity, specificity, and positive predictive power might not meet the traditional rules of thumb for evaluating these indices (Petscher et al., 2011), because benchmarks were not expressly designed to screen children. Moreover, we optimistically expected that children who did not meet benchmarks would be given other opportunities by their teachers and parents to continue developing their early literacy skills. Thus, optimal benchmarks were identified as those that maximized negative predictive power while also balancing the other indices.

An additional component of checking the validity of optimal benchmarks involved testing whether a differential prediction of risk identification occurred as a function of the selected

benchmark and a small set of demographic characteristics, to include child race (Black/non-Black), ethnicity (Hispanic/non-Hispanic), age, and attendance rates. This analysis involved a series of multiple logistic regressions predicting risk on each of the outcome measures (i.e., Letter–Word Identification, Spelling, and Passage Comprehension). The independent variables included a variable that represented whether children were identified as meeting or not meeting the benchmark (coded as 1 and 0, respectively), a variable that represented a selected demographic characteristic, and the interaction between these two variables. A statistically significant interaction term would suggest that a differential prediction of risk existed for the different demographic characteristics based on the selected benchmarks.

RESULTS

Descriptive statistics for letter-name knowledge and literacy achievement measures are provided in Table 2. On average, the children knew nearly 18 (17.6) uppercase letters in the spring of preschool, which converges well with findings collected using the same instrument for all beginning kindergartners (N 83,099) in Virginia as part of universal screening (M 17.5; Invernizzi, Justice, Landrum, & Booker, 2004). Interestingly but not surprisingly, the full range of alphabet knowledge was demonstrated for both uppercase and lowercase letter-naming abilities for these children in the spring of the preschool year, making it clear that even after a year of preschool, children show substantial individual differences in this aspect of development. For uppercase letter names, 3% of children knew no letter names, 25% knew fewer than 10 letter names, and 29% knew all 26 letter names. For lowercase letters, 8% of children knew no letter names, 35% knew fewer than 10 letter names, and 8% knew all 26 names. With respect to literacy achievement in first grade, application of our risk-status criteria resulted in 14%, 17%, and 35% of children being identified as at or below the 40th percentile on word reading, spelling, and reading comprehension, respectively. Uppercase and lowercase letter knowledge as assessed in the spring of preschool moderately correlated with the three outcomes, ranging from .40 to .42, and strong correlations were observed among the three outcomes (Table 3).

The various letter-naming benchmarks identified from state and federal standards are listed in Table 4 for uppercase letters and in Table 5 for lowercase letters along with corresponding estimates for all indices; negative predictive power is listed in the third column from the right. As can be seen, patterns were similar across both uppercase and lowercase letters. Negative predictive power was generally high, particularly when predicting risk on Letter–Word Identification and Spelling. For these two outcomes, negative predictive power was .89 or greater for benchmarks of 10 or above. Thus, 89% of children who knew at least 10 letter names by the end of preschool were not identified as at risk for reading or spelling difficulties in first grade. For higher benchmarks, negative predictive power increased only slightly, ranging up to .97 for knowing all 26 letter names (i.e., 97% of children who knew all of the letter names by the end of preschool did not exhibit risk in first grade). Negative predictive power was lower when predicting risk on Passage Comprehension and ranged from .74 to .85 for uppercase letters and .77 to .93 for lowercase letters. The greater range demonstrates the larger increases in negative predictive power for this outcome when higher benchmarks were set. Phi coefficients, which can be

interpreted as a measure of strength between two variables, ranged from weak (.21) to strong (.63) for uppercase letters and from weak (.11) to moderate (.46) for lowercase letters.

Subsequent analyses examined all possible benchmarks (i.e., 1 through 26 letters) to determine whether there might be optimal benchmarks for end-of-preschool letter-naming abilities. Given the strong negative predictive power across most potential benchmarks, sensitivity, specificity, and positive predictive power were examined to investigate benchmarks at which all four diagnostic indices were best balanced for each individual literacy outcome. These optimal benchmarks were then compared across literacy outcomes to examine consistency in predicting first-grade risk status. Results for uppercase letters showed that a benchmark of 17 letters was optimal when predicting risk status on Letter–Word Identification, compared with optimal benchmarks of 18 letters for Spelling and 20 letters for Passage Comprehension (see the top half of Table 6). Examining these benchmarks across all three literacy outcomes suggested minimal tradeoffs in the overall balance of indices. Negative predictive power was high, ranging from .79 to .95. Sensitivity and specificity were moderate, ranging from .61 to .79 and from .71 to .73, respectively. As anticipated, positive predictive power was low, ranging from .25 to .53. A benchmark of 18 uppercase letters best balanced these indices across all literacy outcomes. Although a benchmark of 20 showed the highest sensitivity and negative predictive power, the incremental loss in the other two indices suggests that 18 letters is a more optimal benchmark.

For lowercase letters, benchmarks of 13, 15, and 16 letters were found to best balance all four efficiency indices when predicting risk on Letter–Word Identification, Spelling, and Passage Comprehension, respectively (see lower half of Table 6). The patterns of indices were similar to those for uppercase benchmarks. Negative predictive power ranged from .76 to .95. Sensitivity ranged from .60 to .79. Specificity ranged from .58 to .73. Finally, positive predictive power ranged from .23 to .53. Across all three literacy outcomes, a benchmark of 15 lowercase letters appeared optimal, minimizing decreases in specificity and positive predictive power while also maximizing sensitivity and negative predictive power.

Using the identified optimal benchmarks of 15 for lowercase letters and 18 for uppercase letters, multiple logistic regressions were conducted to test for any differential identification of risk on each of the outcome measures based on the selected benchmarks. These analyses explored four demographic characteristics for which there was reasonable variability among the sample: child race (Black/non-Black), ethnicity (Hispanic/non-Hispanic), age, and attendance rates. Results from this analysis, depicted in Table 7, showed there to be no significant interaction term for any of the demographic characteristics in relation to any of the outcomes, particularly once p values were adjusted to account for inflated Type I error rates.

DISCUSSION

The goal of the present study was twofold. First, we sought to document the extent to which the various current state and federal preschool letter-naming benchmarks were associated

with successful literacy outcomes or, alternatively, children's risk status, in first grade. Second, we sought to examine all possible letter-naming benchmarks to determine whether optimal benchmarks could be identified. This study builds on a rich research tradition regarding children's letter knowledge development and, to our knowledge, represents the first empirical investigation of the relations between early childhood education benchmarks and later academic outcomes using diagnostic efficiency indices. The importance of such a study is demonstrated by recent findings showing that (a) preschool educators appear to privilege benchmark-type standards or learning indicators (e.g., children should know 10 or more letters) over more general indicators (e.g., children should show an increased awareness of print) and (b) benchmarks might be particularly influential to what is taught and assessed within preschool classrooms (Powell et al., 2008). These findings underscore concerns that so few benchmarks, including those addressing letter-name knowledge, have been empirically investigated for their predictive power or their relation to future learning achievement.

Our results showed that current state and federal letter-naming benchmarks appeared adequate when evaluated solely on negative predictive power. Negative predictive power was consistently high for all benchmarks at or above 10 letters. This indicates that the vast majority of children who met the various letter-naming benchmarks were indeed on the path to literacy success, with very few of such children exhibiting risk on first-grade literacy outcomes. In addition, higher benchmarks were always associated with greater negative predictive power, consistent with past findings indicating moderate-to-strong correlations between letter-naming skill and later literacy abilities (e.g., Hammill, 2004; National Early Literacy Panel, 2008; Scarborough, 1998; Schatschneider et al., 2004). The increase in negative predictive power, however, was small as incrementally higher letter-naming benchmarks were set. For Letter–Word Identification and Spelling, in particular, there appeared to be minimal utility in setting higher benchmarks, with all benchmarks attaining the traditional criteria of values greater than .80 (Petscher et al., 2011). In general, little negative predictive power was gained by setting benchmarks greater than 10.

On the other hand, when all diagnostic indices are considered, many current letter-naming benchmarks may be low. Our data support optimal benchmarks of 18 uppercase letters and 15 lowercase letters when considering all three literacy outcomes. These optimal benchmarks continued to have high negative predictive power, as described earlier, but also maximized classification accuracy in balancing negative predictive power with sensitivity, specificity, and positive predictive power. Use of these optimal benchmarks thus affords the most confidence that children who meet the benchmarks will continue to succeed in literacy tasks and that children who do not meet the benchmarks are those most likely to continue to struggle with literacy learning.

Ensuring such accuracy is not a trivial matter. For instance, a good deal of research substantiates the efficacy of providing early prevention and intervention to aid young learners at risk for academic difficulties (e.g., Barnett, 1995; Campbell, Pungello, Miller-Johnson, Burchinal, & Ramey, 2001; Conyers, Reynolds, & Ou, 2003; National Early Literacy Panel, 2008; Schweinhart, Berrueta-Clement, Barnett, Epstein, & Weikart, 1985), and many teachers and schools use benchmarks as a means of determining which children

are eligible for these types of services. These eligibility criteria work under the assumption that benchmarks reliably differentiate children who are and are not likely to succeed in literacy learning and that children meeting the benchmarks are not in need of extra support to ensure literacy success. However, if the benchmark is set too low, even children who meet the benchmark may be at risk for later literacy difficulties. These children would be overlooked and fail to receive the instructional services which might have led to superior outcomes. Increasing the accuracy of benchmarks is also important for limiting overidentification of children in need of services. Given the limited amount of financial and other resources available to schools, such services should be targeted toward those children who are truly at risk for experiencing literacy difficulties.

Setting higher benchmarks, such as the 18 or 15 that we recommend, may be particularly important in light of results for the Passage Comprehension outcome. This outcome consistently showed the lowest negative predictive power relative to the other literacy outcomes. A benchmark of 10 letters, for instance, resulted in negative predictive power of only .74 as well as particularly low sensitivity. These indices increased substantially as higher benchmarks were set. Moreover, the risk classification for Passage Comprehension was most aligned with normative results, which is important for generalizability. The base rate in the normative population for each literacy outcome was defined as 40% (i.e., 40% of children assessed on these measures are expected to be at risk for literacy difficulties). Yet, rates of 14% and 17% for Letter–Word Identification and Spelling were found in the current data. This may indicate a tendency for children in the present sample to have better-than-typical basic print and sound knowledge. On the other hand, the base rate for Passage Comprehension (35%) closely mirrors the normative rate. Because base rates influence estimates of predictive power (Schatschneider, Petscher, & Williams, 2008) and the normative base rate was best represented by Passage Comprehension, we are most confident that the results for this measure provide accurate estimates of the diagnostic utility of letter-naming benchmarks for the population at large.

In conducting the present work, our primary purpose was generally pragmatic in nature, given that there are a plethora of benchmarks available related to early letter knowledge (see Table 1), yet for none of these is any justification provided. For instance, we are not aware of any research suggesting that future literacy outcomes would be positively affected if a child knew 10 rather than 9 letters, as one might expect given the emphasis on children acquiring knowledge of 10 letter names in the Head Start Outcomes Framework (U.S. Department of Health and Human Services, Administration for Children and Families, 2003). However, the present work is also important for theoretical reasons, because it provides additional evidence of the continuity among young children's early literacy skills and their later reading abilities. The results converge well with other treatments of this topic (e.g., Hammill, 2004; National Early Literacy Panel, 2008; Scarborough, 1998; Schatschneider et al., 2004). In considering the negative predictive power of the letter-knowledge measures, both uppercase and lowercase, in relation to children's future reading achievement in decoding, spelling, and reading comprehension, our results showed that there were lower predictive indices for reading comprehension as compared with the other reading outcomes. Such findings suggest that possibly other early literacy skills, such as vocabulary, might have stronger predictive value when considering children's future

comprehension outcomes, particularly for identifying children at risk. It is also possible that an overreliance on more readily measurable skills in early literacy benchmarks, such as alphabet knowledge, could deter teachers from addressing more difficult-to-measure achievements, such as vocabulary (Powell et al., 2008). Our investigation of letter-naming benchmarks should not be misconstrued to argue for attending to early alphabet knowledge during preschool at the cost of attending to other emergent literacy skills, and it is entirely possible to provide instruction that facilitates children's learning of code-focused and meaning-focused skills alike (e.g., through shared book reading; see Pentimonti, Justice, & Piasta, in press). The extent to which educational benchmarking is associated not only with accurate identification of children's current and future risks but also with teachers' instructional priorities is an important direction for research in educational psychology.

Limitations and Future Directions

Several limitations warrant discussion. First, we note that the values for sensitivity, specificity, and positive predictive power were lower than typically reported in the assessment literature (e.g., National Center on Response to Intervention, 2011). This finding was not unexpected for a number of reasons. First, the criterion of .80 is a rule of thumb meant to apply to the diagnostic efficiency index of most interest and utility, and there is always a tradeoff in balancing the other indices (Petscher et al., 2011). In the assessment literature, many argue that sensitivity and specificity ought to meet this minimum criterion (e.g., Compton, Fuchs, Fuchs, & Bryant, 2006; Jenkins, 2003; cf. Petscher et al., 2011). As we have argued, however, the purpose of benchmarks differs in subtle yet important ways from screening assessments, making negative predictive power a more useful index. This decision is also supported by alternative perspectives on assessment and screening (Petscher et al., 2011; Schatschneider et al., 2008). Second, although the .80 rule of thumb exists, it is important to note that there are not universally agreed-on criteria that are consistently used in the literature. This is particularly true when examining diagnostic efficiency in the education realm because of difficulties in creating dichotomous risk categories. Unlike the medical field in which many outcomes are naturally discrete (i.e., one either has or does not have a given disease), educational risk involves dichotomizing continuous outcomes by setting theoretically based cutpoints. Thus, there remains a good deal of debate concerning diagnostic efficiency criteria when applied to educational issues (Petscher et al., 2011). Third, the lower values for sensitivity, specificity, and positive predictive power may be due to greater numbers of false positives or to children who did not meet the benchmark but were not at risk on later literacy assessments. We intentionally allowed for this possibility given (a) the purpose of benchmarks and the resultant, conservative focus on negative predictive power and (b) our hope that children who did not meet benchmarks would be afforded extra support and instructional opportunities to increase their literacy learning. Fourth, the choice of proficiency for the distal outcome measures will have an impact on the magnitude of the diagnostic indices. In the present investigation, the 40th percentile was selected as the proficiency criterion, given its common use by states and other agencies; however, different results will be observed on the basis of how proficiency is quantified. As an example using the present data, if the 20th percentile is targeted for proficiency, the sensitivity increases, specificity remains approximately the same, negative predictive power increases, and positive predictive power decreases.[3] Intuitively, this makes sense: When risk

status on a distal measure is defined more stringently, there is greater accuracy in delineating those children truly at risk (i.e., true positives) but not necessarily those children truly not at risk (i.e., true negatives). If a more liberal cutpoint is used, chances for false positives increase, resulting in lower sensitivity and higher specificity. These differences reinforce the notion that one must use a defensible estimate for proficiency.

Additional limitations include the characteristics of our sample. The children in the present sample were all drawn from publicly funded preschool centers serving socioeconomically disadvantaged families, and there was a disproportionate overrepresentation of children who are African American. Moreover, all children were proficient English speakers. It is unclear the extent to which these findings would hold for other populations of children, to include children who come from homes in which literacy practices and priorities differ from the present sample. It might be that results would differ for children who come from more advantaged homes or who are not proficient in English, for instance. However, it is not entirely clear that the limitations regarding the sample preclude the external validity and generalizability of the findings discussed here. Results of a statewide literacy screening involving more than 80,000 kindergarteners assessed in the fall of the academic year using the same uppercase task used in this study strongly converged with our results (Invernizzi, Justice, et al., 2004), in terms of the number of uppercase letter names known (i.e., 17.6 in the present sample vs. 17.5 in the statewide sample). This finding suggests that the alphabet knowledge of children in this study was similar to that of a larger population of children about the same age. Moreover, the normative or lower-than-normative base rates for first-grade risk of literacy difficulties also suggest that children in the present sample's literacy performance were representative of a broader population of children. In other words, when considering the first-grade performance of our participants, they did not generally appear to be a high-risk sample, thus suggesting that the present findings may generalize to more advantaged children. This should be addressed, however, in future research. Relatedly, we also want to note a limitation related to the classrooms in which the children were enrolled. The teachers generally reported using no curriculum to guide instruction or using one of several comprehensive curricula that do not follow a scope and sequence of letter instruction. Results might differ for children who attend preschool programs that focus more systematically or explicitly on letter-knowledge development.

Future research might also consider whether the present results generalize beyond the literacy measures used in this study. Although we focused on literacy achievement as measured by the WJ-III subtests, this assessment has recently been criticized for its heavy dependence on children's decoding skills (e.g., Keenan, Betjemann, & Olson, 2008), and investigations of the utility of letter-naming benchmarks in predicting other types of literacy risk are warranted. Borrowing from the response-to-instruction literature (e.g., Fuchs, Fuchs, & Compton, 2004; Vaughn & Fuchs, 2003), it would also be interesting to examine the comparative utility of static, end-of-year letter-naming benchmarks versus benchmarks based on children's growth in letter-naming abilities across the preschool year.

---

[3]Results are available on request from the second author.

Finally, we acknowledge that meeting one specific benchmark by the end of preschool, including knowing a particular number of letter names, does not guarantee literacy success. Similarly, failure to reach a particular benchmark should not be used as the sole means for determining eligibility for early literacy prevention or intervention. Educators ought to continue to use a number of indicators when making such decisions. Nonetheless, given the significance of letter knowledge in children's acquisition of literacy skills and the tendency for early childhood educators to attend to highly specified benchmarks in their instruction, we believe the present results represent an important contribution in establishing letter-naming benchmarks that are valid and optimal, and are thus worthy of attention from practitioners and policymakers.

Acknowledgments

References

Abrams Learning Trends. Let's Begin With the Letter People. Waterbury, CT: Author; 2001.

American Institutes for Research. Reading First state APR data. Washington, DC: Author; 2007.

Armbruster, BB.; Lehr, F.; Osborn, J. Put reading first: The research building blocks for teaching children to read. Washington, DC: Eunice Kennedy Shriver National Institute of Child Health and Human Development; 2001.

Armbruster, BB.; Lehr, F.; Osborn, J. A child becomes a reader: Birth through preschool. Washington, DC: Eunice Kennedy Shriver National Institute of Child Health and Human Development; 2003.

Badian NA. Predicting reading ability over the long term: The changing roles of letter naming, phonological awareness and orthographic processing. Annals of Dyslexia. 1995; 45:79–96.10.1007/BF02648213 [PubMed: 24234189]

Barnett WS. Long-term effects of early childhood programs on cognitive and school outcomes. Future of Children. 1995; 5:25–50.10.2307/1602366

Bowey JA. Socioeconomic status differences in preschool phonological sensitivity and first-grade reading achievement. Journal of Educational Psychology. 1995; 87:476–487.10.1037/0022-0663.87.3.476

Bracken BA, Crawford E. Basic concepts in early childhood educational standards: A 50-state review. Early Childhood Education Journal. 2010; 37:421–430.10.1007/s10643-009-0363-7

Brigance, AH. Brigance Inventory of Early Development–II. North Billerica, MA: Curriculum Associates; 2004.

Bunce, BH. Building a language-focused curriculum for the preschool classroom. Vol. II. Baltimore, MD: Brookes; 1995.

Byrne B, Fielding-Barnsley R. Phonemic awareness and letter knowledge in the child's acquisition of the alphabetic principle. Journal of Educational Psychology. 1989; 81:313–321.10.1037/0022-0663.81.3.313

Campbell FA, Pungello EP, Miller-Johnson S, Burchinal M, Ramey CT. The development of cognitive and academic abilities: Growth curves from an early childhood educational experiment. Developmental Psychology. 2001; 37:231–242.10.1037/00121649.37.2.231 [PubMed: 11269391]

Catts HW, Fey ME, Zhang X, Tomblin JB. Estimating the risk of future reading difficulties in kindergarten children: A research-based model and its clinical implementation. Language, Speech, and Hearing Services in Schools. 2001; 32:38–50.10.1044/01611461(2001/004)

Chall, JS. Learning to read: The great debate. New York, NY: McGraw-Hill; 1967. Original work published 1983

Clifford RM, Barbarin O, Chang F, Early D, Bryant D, Howes C, Pianta R. What is pre-kindergarten? Characteristics of public pre-kindergarten programs. Applied Developmental Science. 2005; 9:126–143.10.1207/s1532480xads0903_1

Compton DL, Fuchs D, Fuchs LS, Bryant JD. Selecting at-risk readers in first grade for early intervention: A two-year longitudinal study of decision rules and procedures. Journal of Educational Psychology. 2006; 98:394–409.10.1037/0022-0663.98.2.394

Conyers LM, Reynolds AJ, Ou SR. The effect of early childhood intervention and subsequent special education services: Findings from the Chicago child–parent centers. Educational Evaluation and Policy Analysis. 2003; 25:75–95.10.3102/01623737025001075

Dodge, DT.; Colker, LJ.; Heroman, C. The Creative Curriculum for preschool. 4. Washington, DC: Teaching Strategies; 2002 p.

Duncan LG, Seymour PHK. Socio-economic differences in foundation-level literacy. British Journal of Psychology. 2000; 91:145–166.10.1348/000712600161736 [PubMed: 10832511]

Durrell DD. Commentary: Letter-name values in reading and spelling. Reading Research Quarterly. 1980; 16:159–163.10.2307/747353

Ehri, LC. A critique of five studies related to letter-name knowledge and learning to read. In: Gentile, LM.; Kamil, ML.; Blanchard, JS., editors. Reading research revisited. Columbus, OH: Merrill; 1983. p. 131-153.

Ehri LC. Learning to read and spell words. Journal of Reading Behavior. 1987; 19:5–31.

Ehri, LC. Grapheme–phoneme knowledge is essential to learning to read words in English. In: Metsala, JL.; Ehri, LC., editors. Word recognition in beginning literacy. Mahwah, NJ: Erlbaum; 1998. p. 3-40.

Elbro C, Petersen DK. Long-term effects of phoneme awareness and letter sound training: An intervention study with children at risk for dyslexia. Journal of Educational Psychology. 2004; 96:660–670.10.1037/0022-0663.96.4.660

Evans MA, Bell M, Shaw D, Moretti S, Page J. Letter names, letter sounds and phonological awareness: An examination of kindergarten children across letters and of letters across children. Reading and Writing. 2006; 19:959–989.10.1007/s11145-006-9026-x

Florida Department of Education. Florida Assessments for Instruction in Reading. Tallahassee, FL: Author; 2009.

Foulin JN. Why is letter-name knowledge such a good predictor of learning to read? Reading and Writing. 2005; 18:129–155.10.1007/s11145-004-5892-2

Fuchs D, Fuchs LS, Compton DL. Identifying reading disabilities by responsiveness-to-instruction: Specifying measures and criteria. Learning Disability Quarterly. 2004; 27:216–227.10.2307/1593674

Gallagher A, Frith U, Snowling MJ. Precursors of literacy delay among children at genetic risk of dyslexia. Journal of Child Psychology and Psychiatry. 2000; 41:203–213.10.1017/S0021963099005284 [PubMed: 10750546]

Good, RH.; Kaminski, RA.; Smith, S.; Laimon, D.; Dill, S. Dynamic Indicators of Basic Early Literacy Skills. 5. Eugene, OR: Institute for Development of Educational Achievement, University of Oregon; 2001.

Groff PJ. Resolving the letter name controversy. Reading Teacher. 1984; 37:384–388.

Hair E, Halle T, Terry-Humen E, Lavelle B, Calkins J. Children's school readiness in the ECLS-K: Predictions to academic, health, and social outcomes in first grade. Early Childhood Research Quarterly. 2006; 21:431–454.10.1016/j.ecresq.2006.09.005

Hammill DD. What we know about correlates of reading. Exceptional Children. 2004; 70:453–468.

Hohmann, M.; Weikart, DP. Active learning practices for preschool and child care programs: Educating young children. Ypsilanti, MI: HighScope Press; 1995.

Invernizzi MA, Justice LM, Landrum T, Booker K. Early literacy screening in kindergarten: Widespread implementation in Virginia. Journal of Literacy Research. 2004; 36:479–500.10.1207/s15548430jlr3604_3

Invernizzi, MA.; Sullivan, A.; Meier, JD.; Swank, L. Phonological Awareness Literacy Screening for Preschool: Teacher's manual. Charlottesville, VA: University of Virginia; 2004.

Jenkins, JR. Candidate measures for screening at-risk students; Paper presented at the National Research Center on Learning Disabilities Responsiveness-to-Intervention Symposium; 2003 Dec. Retrieved from http://www.nrcld.org/symposium2003/jenkins/index.html

Jordan NC, Kaplan D, Locuniak MN, Ramineni C. Predicting first-grade math achievement from developmental number sense trajectories. Learning Disabilities Research and Practice. 2007; 22:36–46.10.1111/j.1540-5826.2007.00229.x

Justice LM, Mashburn AJ, Hamre BK, Pianta RC. Quality of language and literacy instruction in preschool classrooms serving at-risk pupils. Early Childhood Research Quarterly. 2008; 23:51–68.10.1016/j.ecresq.2007.09.004 [PubMed: 22773887]

Justice LM, Pence K, Bowles RB, Wiggins A. An investigation of four hypotheses concerning the order by which 4-yearold children learn the alphabet letters. Early Childhood Research Quarterly. 2006; 21:374–389.10.1016/j.ecresq.2006.07.010

Keenan JM, Betjemann RS, Olson RK. Reading comprehension tests vary in the skills they assess: Differential dependence on decoding and oral comprehension. Scientific Studies of Reading. 2008; 12:281–300.10.1080/10888430802132279

Levin I, Shatil-Carmon S, Asif-Rave O. Learning of letter names and sounds and their contribution to word recognition. Journal of Experimental Child Psychology. 2006; 93:139–165.10.1016/j.jecp.2005.08.002 [PubMed: 16249004]

Lyytinen H, Aro M, Eklund K, Erskine J, Guttorm T, Laakso ML, Torppa M. The development of children at familial risk for dyslexia: Birth to early school age. Annals of Dyslexia. 2004; 54:184–220.10.1007/s11881-004-0010-3 [PubMed: 15741935]

Mashburn AJ, Pianta RC, Hamre B, Downer JT, Barbarin OA, Bryant D, Howes C. Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. Child Development. 2008; 79:732–749.10.1111/j.14678624.2008.01154.x [PubMed: 18489424]

Mason, JM., editor. Early reading: A developmental perspective. Vol. 1. New York, NY: Longman; 1984.

Meisels, S.; Wiske, M. The Early Screening Inventory. New York, NY: Teachers College Press; 1983.

National Association for the Education of Young Children. Learning to read and write: Developmentally appropriate practices for young children. A joint position statement of the International Reading Association and the National Association for the Education of Young Children. 1998. Retrieved from http://www.naeyc.org/about/positions/pdf/psread98.pdf

National Center on Response to Intervention. Screening tool chart. 2011. Retrieved from http://www.rti4success.org/tools_charts/screening.php

National Early Literacy Panel. Developing early literacy. Washington, DC: National Institute for Literacy; 2008.

National Governors Association Center for Best Practices and Council of Chief State School Officers. Common core state standards for English language arts and literacy in history/social studies, science, and technical subjects. Washington, DC: Authors; 2010.

Neuman SB, Roskos K. The state of state pre-kindergarten standards. Early Childhood Research Quarterly. 2005; 20:125–145.10.1016/j.ecresq.2005.04.010

Pentimonti, J.; Justice, LM.; Piasta, SB. Sharing books with children. In: Shanahan, T.; Lonigan, C., editors. Literacy in preschool and kindergarten children: The National Early Literacy Panel and beyond. Baltimore, MD: Brookes; in press

Petscher Y, Kim YS, Foorman BR. The importance of predictive power in early screening assessments: Implications for placement in the RTI framework. Assessment for Effective Intervention. 2011; 36:158–166.10.1177/1534508410396698

Piasta SB, Wagner RK. Developing emergent literacy skills: A meta-analysis of alphabet learning and instruction. Reading Research Quarterly. 2010a; 45:8–38.10.1598/RRQ.45.1.2 [PubMed: 20671801]

Piasta SB, Wagner RK. Learning letter names and sounds: Effects of instruction, letter type, and phonological processing skill. Journal of Experimental Child Psychology. 2010b; 105:324–344.10.1016/j.jecp.2009.12.008 [PubMed: 20097352]

Author Manuscript

Powell DR, Diamond KE, Bojczyk KE, Gerde HK. Head Start teachers' perspectives on early literacy. Journal of Literacy Research. 2008; 40:422–460.10.1080/10862960802637612

Puolakanaho A, Ahonen T, Aro M, Eklund K, Leppnen PHT, Poikkeus AM, Lyytinen H. Very early phonological and language skills: Estimating individual risk of reading disability. Journal of Child Psychology and Psychiatry. 2007; 48:923–931.10.1111/j.14697610.2007.01763.x [PubMed: 17714377]

Read C. Pre-school children's knowledge of English phonology. Harvard Educational Review. 1971; 41:1–34.

Reynolds AJ, Bezruczko N. School adjustment of children at risk through fourth grade. Merrill–Palmer Quarterly. 1993; 39:457–480.

Richgels DJ. An investigation of preschool and kindergarten children's spelling and reading abilities. Journal of Research and Development in Education. 1986; 19:41–47.

Scarborough HS. Very early language deficits in dyslexic children. Child Development. 1990; 61:1728–1743.10.2307/1130834 [PubMed: 2083495]

Scarborough HS. Antecedents to reading disability: Preschool language development and literacy experiences of children from dyslexic families. Reading and Writing. 1991; 3:219–233.

Scarborough, HS. Early identification of children at risk for reading disabilities. In: Shapiro, BK.; Accardo, PJ.; Capute, AJ., editors. Specific reading disability: A view of the spectrum. Timonium, MD: York Press; 1998. p. 75-120.

Schatschneider C, Fletcher JM, Francis DJ, Carlson CD, Foorman BR. Kindergarten prediction of reading skills: A longitudinal comparative analysis. Journal of Educational Psychology. 2004; 96:265–282.10.1037/0022-0663.96.2.265

Schatschneider, C.; Petscher, Y.; Williams, KM. How to evaluate a screening process: The vocabulary of screening and what educators need to know. In: Justice, LM.; Vukelic, C., editors. Every moment counts: Achieving excellence in preschool literacy instruction. New York, NY: Guilford Press; 2008. p. 304-316.

Schrank, FA.; Woodcock, R. WJ III CompuScore and profiles program [Computer software]. Itasca, IL: Riverside; 2001.

Schweinhart LJ, Berrueta-Clement JR, Barnett WS, Epstein AS, Weikart DP. Effects of the Perry Preschool Program on youths through age 19: A summary. Topics in Early Childhood Special Education. 1985; 5:26–35.10.1177/027112148500500204

Silberglitt B, Hintze J. Formative assessment using CBM-R cut scores to track progress toward success on state-mandated achievement tests: A comparison of methods. Journal of Psychoeducational Assessment. 2005; 23:304–325.10.1177/073428290502300402

Snowling MJ, Gallagher A, Frith U. Family risk of dyslexia is continuous: Individual differences in the precursors of reading skill. Child Development. 2003; 74:358–373.10.1111/1467-8624.7402003 [PubMed: 12705560]

SRA/McGraw-Hill. Open Court Reading Pre-K. Columbus, OH: Author; 2003.

Storch SA, Whitehurst GJ. Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. Developmental Psychology. 2002; 38:934–947.10.1037/0012-1649.38.6.934 [PubMed: 12428705]

Streiner DL. Diagnosing tests: Using and misusing diagnostic and screening tests. Journal of Personality Assessment. 2003; 81:209–219.10.1207/S15327752JPA8103_03 [PubMed: 14638445]

Texas Education Agency. Texas Primary Reading Inventory. Austin, TX: Author; 1998.

Torppa M, Lyytinen P, Erskine J, Eklund K, Lyytinen H. Language development, literacy skills, and predictive connections to reading in Finnish children with and without familial risk for dyslexia. Journal of Learning Disabilities. 2010; 43:308–321.10.1177/0022219410369096 [PubMed: 20479461]

Torppa M, Poikkeus AM, Laakso ML, Eklund K, Lyytinen H. Predicting delayed letter knowledge development and its relation to Grade 1 reading achievement among children with and without familial risk for dyslexia. Developmental Psychology. 2006; 42:1128–1142.10.1037/0012-1649.42.6.1128 [PubMed: 17087547]

Piasta et al. Page 20

Treiman, R.; Kessler, B. The role of letter names in the acquisition of literacy. In: Reese, HW.; Kail, R., editors. Advances in child development and behavior. Vol. 31. San Diego, CA: Academic Press; 2003. p. 105-135.

Treiman R, Pennington BF, Shriberg LD, Boada R. Which children benefit from letter names in learning letter sounds? Cognition. 2008; 106:1322–1338.10.1016/j.cognition.2007.06.006 [PubMed: 17692304]

Treiman R, Tincoff R. The fragility of the alphabetic principle: Children's knowledge of letter names can cause them to spell syllabically rather than alphabetically. Journal of Experimental Child Psychology. 1997; 64:425–451.10.1006/jecp.1996.2353 [PubMed: 9073380]

Treiman R, Tincoff R, Richmond-Welty ED. Letter names help children to connect print and speech. Developmental Psychology. 1996; 32:505–514.10.1037/0012-1649.32.3.505

Treiman R, Tincoff R, Rodriguez K, Mouzaki A, Francis DJ. The foundations of literacy: Learning the sounds of letters. Child Development. 1998; 69:1524–1540. [PubMed: 9914638]

Treiman R, Weatherston S, Berch D. The role of letter names in children's learning of phoneme–grapheme relations. Applied Psycholinguistics. 1994; 15:97–122.10.1017/S0142716400006998

U.S. Department of Education. Guidance for the Reading First program. Washington, DC: Author; 2002.

U.S. Department of Education. Guidance for the Early Reading First program. Washington, DC: Author; 2003.

U.S. Department of Education. Early Reading First performance. 2009. Retrieved from http://www2.ed.gov/programs/earlyreading/performance.html

U.S. Department of Health and Human Services, Administration for Children and Families. The Head Start path to positive child outcomes. 2003. Retrieved from http://www.hsnrc.org/CDI/outcontent.cfm

Vaughn S, Fuchs LS. Redefining learning disabilities as inadequate response to instruction: The promise and potential problems. Learning Disabilities Research and Practice. 2003; 18:137–146.10.1111/1540-5826.00070

Woodcock, R.; McGrew, KS.; Mather, N. Woodcock–Johnson Tests of Achievement. 3. Itasca, IL: Riverside; 2001.

Appendix. Calculation of Diagnostic Efficiency Indices

When children are assessed relative to a benchmark and risk status on another outcome is subsequently assessed at a distal point in time, four possible results may occur. These are illustrated in the 2×2 contingency presented as Table A1: (a) children may not meet the benchmark and may be identified as at risk on the outcome (true positives), (b) children may not meet the benchmark and may not be identified as at risk on the outcome (false positives), (c) children may meet the benchmark and may be identified as at risk on the outcome (false negatives), and (d) children may meet the benchmark and may not be identified as at risk on the outcome (true negatives). Table A2 shows a 2×2 contingency table for optimal uppercase benchmark, and Table A3 shows a 2×2 contingency table for optimal lowercase benchmark.

Table A1

Sample Result From a Diagnostic Validity Test

Benchmark	Outcome	
	At risk	Not at risk
Does not meet (fail)	True positive (TP)	False positive (FP)

	Outcome	
Benchmark	**At risk**	**Not at risk**
Meets (pass)	False negative (FN)	True negative (TN)

Note. From this matrix, diagnostic efficiency indices may be calculated using the following formulas: sensitivity = TP/(TP + FN); specificity = TN/(FP + TN); positive predictive power = TP/(TP + FP); negative predictive power = TN/(FN + TN). We illustrate this process using the 2×2 contingency tables for the optimal uppercase (Table A2) and lowercase (Table A3) benchmarks identified in the present work and the Letter–Word Identification subtest of the Woodcock–Johnson Tests of Achievement (3rd ed.). Risk status on the latter was defined as scoring at or below the 40th percentile.

Table A2

A 2×2 Contingency Table for Optimal Uppercase Benchmark

	Letter-Word Identification	
Uppercase letter-naming benchmark (18)	**At risk**	**Not at risk**
Does not meet (fail)	38	113
Meets (pass)	14	204

Note. Using Table A2, sensitivity for the uppercase letter benchmark is calculated with 38/(38 + 14) = .73; specificity is 204/(113 + 204) = .64; positive predictive power is 38/(38 + 113) = .25; and negative predictive power is 204/(14 + 204) = .94.

Table 3

A 2×2 Contingency Table for Optimal Lowercase Benchmark

	Letter-Word Identification	
Lowercase letter-naming benchmark (15)	**At risk**	**Not at risk**
Does not meet (fail)	37	123
Meets (pass)	16	195

Note. Using Table A3, sensitivity for the lowercase letter benchmark is calculated with 37/(37 + 16) = .70; specificity is 195/(123 + 195) = .61; positive predictive power is 37/(37 + 123) = .23; and negative predictive power is 195/(16 + 195) = .92.

Table 1

Benchmarks for Letter-Naming Abilities From State, Federal, and Professional Standards

Document	Benchmark
	State preschool standards[a]
State of Alaska Early Learning Guidelines	Correctly identifies 10 or more letters of the alphabet Identifies a letter for a given letter name, for most letters
Alabama Performance Standards for 4-year-olds	Identifies at least 10 letters of the alphabet, especially those in child's own name
Arizona Early Childhood Standards	Recognizes and names at least 10 letters of the alphabet
California Desired Results Outcomes; Preschool Learning Foundations	Progresses from recognizing some letters to knowing at least 10 letters by sight and name to knowing most letters Matches more than half of uppercase letter names and more than half of lowercase letter names to their printed form
Early Learning Standards for Children Entering Kindergarten in the District of Columbia	Identifies 10 or more letters
Indiana Academic Standards for Young Children from Birth to Age 5	Names 13 uppercase letters Points to and names six letters
South Dakota Early Learning Guidelines	Identifies at least 10 letters of the alphabet, especially those in child's own name
Revised Texas Prekindergarten Guidelines	Names at least 20 uppercase and at least 20 lowercase letters
Virginia's Foundation Blocks for Early Learning: Comprehensive Standards for Four-Year-Olds	Correctly identifies 10 to 18 alphabet (uppercase) letters by name in random order
Early Learning and Development Benchmarks: Washington State	Identifies a letter for a given letter name, for most letters Recognizes several uppercase and lowercase letters
Wyoming Department of Education Early Childhood Readiness Standards	Associates at least 10 letters with their shapes or sounds
	Other standards
Head Start Outcomes Framework	Identifies at least 10 letters of the alphabet, especially those in their own name
Early Reading First Program Performance Reports	Identifies the average number of letters for preschool-age children in Early Reading First programs as measured by the Upper Case Alphabet Knowledge subtask on the PALS-PreK assessment (performance targets of 16 to 19 letters from 2005 to 2008)
National Association for the Education of Young Children: Learning to Read and Write	Identifies some letters
Common Core State Standards for English Language Arts	Recognizes and names all uppercase and lowercase letters of the alphabet (kindergarten)

Note. All standards documents were current as of March 2010. References for standards documents are available from the first author on request. PALS Phonological Awareness Literacy Screening for Preschool (Invernizzi, Sullivan, et al., 2004).

[a] Standards for an additional 34 states include reference to being able to recognize or identify *some* or *several* letters.

Table 2

Descriptive Statistics for Children's Letter-Naming Abilities and Literacy Achievement

Measure	n^a	M	SD	Range
Preschool uppercase letter-naming ability	369	17.60	8.99	0–26
Preschool lowercase letter-naming ability	371	14.87	9.08	0–26
First-grade Letter-Word Identification	371	108.85	12.83	51–139
First-grade Spelling	369	107.66	13.65	61–148
First-grade Passage Comprehension	370	99.61	12.79	43–134

Note. Raw numbers of letter names known are presented for letter-naming abilities; standard scores ($M = 100$, $SD = 15$) are presented for first-grade literacy achievement measures, representing, respectively, the Letter– Word Identification, Spelling, and Passage Comprehension subtests of the Woodcock–Johnson Tests of Achievement (3rd ed.; Woodcock, McGrew, & Mather, 2001).

[a]Sample sizes ranged because two children did not complete the uppercase letter-naming and Spelling subtests, and one child did not to complete the Passage Comprehension subtest.

Table 3

Correlations Among Primary Study Measures

Variable	Lowercase letters	Uppercase letters	Letter-Word Identification	Spelling	Passage Comprehension
Lowercase letters	—				
Uppercase letters	.93	—			
Letter-Word Identification	.45	.45	—		
Spelling	.42	.41	.85	—	
Passage Comprehension	.41	.42	.86	.78	—
Lowercase letter benchmark - 15			.35	.33	.32
Uppercase letter benchmark - 18			.38	.35	.35

Note. Lowercase letter benchmark–15 and uppercase letter benchmark–18 represent dichotomous variables of the two optimal benchmarks (15 letters for lowercase and 18 letters for uppercase). The correlations represent the associations between these dichotomous variables and the three subtests of the Woodcock–Johnson Tests of Achievement (3rd ed.; Woodcock, McGrew, & Mather, 2001).

Table 4

Diagnostic Efficiency of Current Benchmarks for Uppercase Letters

First-grade outcome and letter-naming benchmark	SE	SP	NPP	PPP	φ
Letter-Word Identification					
10	.60	.80	.92	.33	.63
13	.63	.74	.92	.28	.49
16	.67	.70	.93	.27	.45
18	.73	.64	.94	.25	.42
19	.73	.62	.93	.24	.38
20	.79	.61	.95	.25	.43
26	.90	.32	.95	.18	.21
Spelling					
10	.50	.80	.89	.34	.47
13	.55	.74	.89	.30	.39
16	.61	.70	.90	.29	.40
18	.65	.64	.90	.27	.34
19	.66	.62	.90	.26	.32
20	.69	.61	.91	.26	.34
26	.87	.33	.93	.21	.20
Passage Comprehension					
10	.42	.83	.74	.56	.41
13	.50	.78	.76	.53	.42
16	.58	.76	.78	.54	.48
18	.63	.70	.79	.52	.46
19	.64	.67	.79	.50	.42
20	.67	.67	.80	.50	.46
26	.87	.38	.85	.41	.31

Note. SE = sensitivity [TP/(TP + FN)]; SP = specificity [TN/(FP + TN)]; NPP = negative predictive power [TN/(TN + FN)]; PPP = positive predictive power [TP/(TP + FP)]. TP = true positive; FN = false negative; TN = true negative; FP = false positive.

Table 5

Diagnostic Efficiency of Current Benchmarks for Lowercase Letters

First-grade outcome and letter-naming benchmark	SE	SP	NPP	PPP	φ
Letter-Word Identification					
10	.64	.72	.92	.28	.46
13	.68	.66	.93	.25	.39
16	.77	.59	.94	.24	.38
18	.79	.54	.94	.22	.33
19	.79	.51	.94	.21	.30
20	.79	.48	.93	.20	.26
26	.98	.09	.97	.15	.10
Spelling					
10	.59	.73	.90	.31	.42
13	.60	.66	.89	.27	.32
16	.68	.58	.90	.25	.29
18	.73	.54	.91	.25	.29
19	.73	.51	.90	.24	.25
20	.73	.48	.90	.22	.21
26	.98	.09	.97	.18	.11
Passage Comprehension					
10	.54	.77	.77	.54	.45
13	.57	.71	.76	.50	.37
16	.67	.64	.79	.49	.41
18	.71	.60	.80	.47	.39
19	.72	.56	.80	.46	.35
20	.74	.53	.80	.44	.33
26	.98	.11	.93	.36	.18

Note. SE = sensitivity [TP/(TP + FN)]; SP = specificity [TN/(FP + TN)]; NPP = negative predictive power [TN/(TN + FN)]; PPP = positive predictive power [TP/(TP + FP)]. TP = true positive; FN = false negative; TN = true negative; FP = false positive.

Table 6

Diagnostic Efficiency of Optimal Benchmarks for Uppercase and Lowercase Letters

First-grade outcome and letter-naming benchmark	SE	SP	NPP	PPP	φ
Uppercase letters					
Letter-Word Identification					
17	.69	.67	.93	.26	.42
18	.73	.64	.94	.25	.42
20	.79	.61	.95	.25	.43
Spelling					
17	.63	.67	.90	.28	.37
18	.65	.64	.90	.27	.34
20	.69	.61	.91	.26	.34
Passage Comprehension					
17	.61	.73	.79	.53	.48
18	.63	.70	.79	.52	.46
20	.67	.67	.80	.50	.46
Lowercase letters					
Letter-Word Identification					
13	.68	.66	.93	.25	.39
15	.70	.61	.92	.23	.33
16	.77	.59	.94	.24	.33
Spelling					
13	.60	.66	.89	.27	.32
15	.67	.64	.79	.49	.26
16	.68	.58	.90	.25	.29
Passage Comprehension					
13	.57	.71	.76	.50	.37
15	.62	.67	.78	.49	.38
16	.67	.64	.79	.49	.41

Note. SE = sensitivity [TP/(TP + FN)]; SP = specificity [TN/(FP + TN)]; NPP = negative predictive power [TN/(TN + FN)]; PPP = positive predictive power [TP/(TP + FP)]. TP = true positive; FN = false negative; TN = true negative; FP = false positive.

Table 7

Test of Differential Risk Identification by Benchmarks and Demographic Characteristics

First-grade outcome and variable	df	Estimate	SE	Wald χ^2	p
Uppercase					
Letter-Word Identification					
Letter Risk × Black	1	−0.18	0.69	0.07	.80
Letter Risk × Hispanic	1	−0.35	1.30	0.07	.78
Letter Risk × Age	1	0.00	0.09	0.00	.99
Letter Risk × Attendance	1	−0.01	0.01	0.70	.40
Spelling					
Letter Risk × Black	1	0.34	0.51	0.43	.51
Letter Risk × Hispanic	1	−0.57	1.02	0.31	.58
Letter Risk × Age	1	−0.13	0.07	3.39	.07
Letter Risk × Attendance	1	0.00	0.00	0.03	.87
Passage Comprehension					
Letter Risk × Black	1	−0.15	0.60	0.06	.80
Letter Risk × Hispanic	1	−2.71	1.38	3.89	.05
Letter Risk × Age	1	0.06	0.09	0.49	.48
Lowercase					
Letter-Word Identification					
Letter Risk × Black	1	−0.26	0.66	0.15	.69
Letter Risk × Hispanic	1	−0.12	1.28	0.01	.93
Letter Risk × Age	1	−0.06	0.09	0.40	.53
Letter Risk × Attendance	1	−0.27	0.35	0.60	.44
Spelling					
Letter Risk × Black	1	−0.05	0.49	0.01	.91
Letter Risk × Hispanic	1	−0.73	1.02	0.50	.48
Letter Risk × Age	1	−0.09	0.06	2.06	.15
Letter Risk × Attendance	1	0.01	0.01	0.92	.34
Passage Comprehension					
Letter Risk × Black	1	−0.22	0.59	0.14	.71

First-grade outcome and variable	df	Estimate	SE	Wald χ^2	p
Letter Risk × Hispanic	1	−2.06	1.20	2.97	.08
Letter Risk × Age	1	0.01	0.09	0.01	.92
Letter Risk × Attendance	1	0.01	0.01	0.48	.49