



Published in final edited form as:

J Learn Disabil. 2010 ; 43(1): . doi:10.1177/0022219409345007.

Identifying Preschool Children At-Risk of Later Reading Difficulties: Evaluation of two Emergent Literacy Screening Tools

Shauna B. Wilson and Christopher J. Lonigan

Department of Psychology, Florida State University

Abstract

Emergent literacy skills are predictive of children's early reading success, and literacy achievement in early schooling declines more rapidly for children who are below-average readers. It is therefore important for teachers to identify accurately children at risk for later reading difficulty so children can be exposed to good emergent literacy interventions. In this study, 176 preschoolers were administered two screening tools, the Revised Get Ready to Read! (GRTR-R) and the Individual Growth and Development Indicators (IGDIs), and a diagnostic measure at two time points. Receiver operating characteristic curve analyses revealed that at optimal cut scores, GRTR-R provided more accurate classification of children's overall emergent literacy skills than did IGDIs. However, neither measure was particularly good at classifying specific emergent literacy skills.

Keywords

Emergent Literacy; Screening Tool; Receiver Operating Characteristic curves; *Revised Get Ready to Read!*; *Individual Growth and Development Indicators*

Early childhood education in the United States and elsewhere is increasingly viewed as a vital component of children's educational experiences and a key factor in some children's educational success. An increasing number of states have developed preschool standards for children's education-related achievement. These standards are often aligned with the states' kindergarten through 12th grade standards and include learning goals related to skills that are considered the precursors to reading achievement. Particularly with respect to children at risk for later literacy problems because of environmental (e.g., low SES) or cognitive risk factors, a goal of preschools is to identify and address children's development in key areas related to academic performance. Many states have funded universal or targeted state-funded preschool programs to accomplish this goal.

To address potential problems in preschool literacy development, children must first be identified correctly as at risk for later academic problems. Given that most preschools have limited funding and few preschools can afford to spend a significant amount of time assessing children (and thus spend less time in instruction or intervention), an assessment tool is needed that classifies accurately which children have developed age-appropriate literacy skills and which children are lagging behind in their literacy development. Such an assessment tool needs to involve minimal cost and time to administer. The gold standard for measuring children's early academic skill is diagnostic assessment, which provides accurate,

in-depth information regarding each child's unique set of strengths and weaknesses relative to his or her same-age peers. However, diagnostic assessment tools must be administered by trained personnel for scores to be valid, and, in general, teachers do not have the training to use these tools. This can be problematic for several reasons, including that immediate testing of a child may not be possible and that shy or anxious children may not cooperate with an unfamiliar adult. In addition, diagnostic assessments can be expensive. Although having specific and detailed data for every child may be optimal, it is not feasible for preschools with limited funds to assess every preschooler using a diagnostic assessment. Diagnostic assessment is not practical for the purpose of assessing all children for academic risk.

One viable alternative to diagnostic assessment is screening assessment. Screening tools, which are brief measures that allow snapshots of children's current academic skills, provide reliable and valid information regarding children's skills, and also meet financial and time constraints. Thus, using a screening tool to assess children's academic skills in preschool is a more practical way to meet the goals of identifying children who are at very high risk, are most in need of targeted instructional activities, or who have not responded to the basic classroom-wide curriculum. Although there have been a few studies concerning the psychometric characteristics of screening-type measures of early literacy skills, these studies have typically reported summaries of the measures' reliability, concurrent validity, or predictive validity. These metrics are useful for determining performance characteristics of the measures and for demonstrating that they provide measurement of specific domains, but the central question relevant to the value of screening measures relates to the ability to use the measures to make accurate classifications (e.g., correct classification of a child as either at risk or not at risk).

There is now a significant body of research concerning the development of literacy-related skills. Although the acquisition of reading skills was once thought to originate with the start of reading instruction in elementary school, research now supports the idea that learning to read is a continuous developmental process that emerges early in life (Lonigan, 2006; Snow, Burns, & Griffin, 1998; Whitehurst & Lonigan, 1998). Increasingly, research has focused on early literacy skill in an attempt to identify children who may be at risk for later reading difficulty to eliminate this potential risk before children begin elementary school (e.g., Scarborough, 1989; Whitehurst & Fischel, 2000; Whitehurst & Lonigan, 1998). The term "emergent literacy" refers to the skills, knowledge, and attitudes that children have about reading and writing before they are formally taught these skills (Sulzby & Teale, 1991; Teale & Sulzby, 1986; Whitehurst & Lonigan, 1998). Children's reading success throughout elementary school can be predicted from their emergent literacy skills (Lonigan, Burgess, & Anthony, 2000; Lonigan, Schatschneider, & Westburg, 2008a; Spira & Fischel, 2005; Storch & Whitehurst, 2002).

The three emergent literacy skills that are most predictive of reading ability are phonological awareness, print knowledge, and oral language (Lonigan, 2006; Lonigan et al. 2008a; Whitehurst & Lonigan, 1998). Phonological awareness refers to the ability to detect and manipulate the sounds of spoken language, independent of meaning (Lonigan, 2006; Wagner & Torgesen, 1987). This skill is strongly related to the acquisition of reading, even after accounting for other factors affecting reading ability, such as intelligence, receptive vocabulary, memory skills, and social class (Perfetti, Beck, Bell, & Hughes, 1987; Wagner, Torgesen, & Rashotte, 1994). Print knowledge refers to children's ability to comprehend how print is organized; skills include knowledge of the conventions of print as well as the letters of the alphabet. Knowledge of letter names before kindergarten is predictive of reading ability in late middle and early high school (Stevenson & Newman, 1986). Oral language refers to vocabulary as well as the ability to use words to understand and convey meaning. Relative to their same-age peers, children with larger vocabularies become more

proficient readers than children with smaller vocabularies (e.g., Bishop & Adams, 1990; Scarborough, 1989; Storch & Whitehurst, 2002).

Data from many sources indicate continuity between early language and literacy-related (emergent literacy) skills and later reading skills. In the absence of intervention, children's emergent literacy skills are stable over time (Lonigan et al., 2000; Wagner et al., 1994). For example, Lonigan et al. (2000) found that latent variables representing phonological awareness ($r = 1.00$) and letter knowledge ($r = .80$) were highly stable from late preschool to just before first grade. Stability of these domains for older children is also high; for example, the average correlation for letter knowledge in one study was .58 from grades K to 1 and 1 to 2 (Pettrill et al., 2007). For phonological awareness, stability coefficients in one study ranged from .83 to .95 for measurement from grades K to 1, 1 to 2, 2 to 3, and 3 to 4 (Wagner et al., 1997). With regard to vocabulary, Verhoeven and van Leeuwe (2008) found that stability was high from first to second grade ($r = .95$) as well as from grades 2 to 3, 3 to 4, and 4 to 5 ($r_s = .99$).

Results of studies examining the predictive relations between emergent literacy skills and later reading skills suggest that children's emergent literacy skills can provide an early indicator of their likely outcomes regarding the development of skilled versus problematic reading in the early elementary grades (e.g., Bishop & Adams, 1990; Perfetti et al., 1987; Scarborough, 1989; Stevenson & Newman, 1986; Storch & Whitehurst, 2002; Wagner et al., 1994). Therefore, it seems reasonable to examine the ability of screening measures to identify children who would be identified as having high risk for later reading problems on the basis of age-appropriate criterion measures of those skills that are both relatively stable over longer periods of time and that provide information about relative degree of risk for later reading problems. The purpose of this study was to determine the value of two emergent literacy screening measures, in terms of indices of classification accuracy with respect to children's emergent literacy skills. Within the logic of current models of early childhood education, identification of children with weak or slow development of these skills would allow a determination of children who are those most at risk of later reading problems.

Two currently available screening tools of children's emergent literacy skills are the *Get Ready to Read! Screening Tool* (GRTR; Whitehurst & Lonigan, 2001) and the *Individual Growth and Development Indicators* (IGDIs; McConnell, 2002). Teachers can administer either of these screening tools easily, and each usually takes less than 10 minutes to complete. The GRTR is a 20-item task that measures print knowledge and phonological awareness. To date, there have been four studies examining the psychometric and predictive characteristics of the GRTR. Whitehurst (2001) validated the GRTR on a sample of 342 preschool children and determined that the concurrent validity of the GRTR with a diagnostic measure of emergent literacy skills was high ($r = .69$). Molfese, Molfese, Modglin, Walker, and Neamon (2004) found that the correlations between the GRTR-R and measures of vocabulary, environmental print, phonological processing, and rhyming ranged from .12 to .51 (median $r = .46$) among a sample of 3-year-old children ($N = 73$) and .09 to .45 (median $r = .41$) among a sample of 4-year-old children ($N = 79$). Molfese et al. (2006) found that the one-year gains on a diagnostic measure of letter knowledge among low-income preschoolers were highly correlated with GRTR scores ($r = .48$). With regard to the predictive validity of the GRTR, Phillips, Lonigan, and Wyatt (2009) found that the GRTR was predictive of blending, elision, rhyming, letter knowledge, and word identification (r_s ranged from .25 to .40; median $r = .32$) at 20, 28, and 35 months after initial assessment using the GRTR.

The IGDIs contains a number of tasks designed to measure a diverse array of developmental domains from birth to approximately age eight. The tasks relevant to emergent literacy are Alliteration and Rhyming (measures of phonological awareness) as well as Picture Naming (a measure of oral language). Currently available psychometric data for the IGDIs demonstrate that it is a good measure of emergent literacy skills. According to the IGDIs Technical Report #8 (Missall & McConnell, 2004), 3-week test-retest reliability for Alliteration ranged from .46 to .80, Rhyming ranged from .83 to .89, and Picture Naming ranged from .44 to .78. With regard to concurrent validity (McConnell, Priest, Davis, & McEvoy, 2002; Missall, 2002, as cited by Missall & McConnell, 2004), correlations between all three IGDIs tasks and measures of print knowledge, phonological awareness, and vocabulary were moderate to large ($r_s = .32$ to $.79$; for McConnell et al., 2002, median $r = .58$). With regard to predictive validity (Missall et al. 2007), administration of these three IGDIs tasks in preschool was predictive of oral reading fluency at both the end of kindergarten ($r_s = .26$ to $.58$; median $r = .37$) and the end of first grade ($r_s = .26$ to $.50$, median $r = .42$). In addition to studies examining the GRTR and IGDIs individually, one study has examined the psychometrics of these two screening tools using the same criterion measure in the same population, allowing direct comparison of these measures (Wilson & Lonigan, 2008). In terms of reliability, concurrent validity, and predictive validity, Wilson and Lonigan found that the GRTR-R consistently performed better than or equal to the IGDIs.

The Phonological Awareness Literacy Screening (PALS) Pre-K version (PALS-PreK; Invernizzi, Sullivan, & Meier, 2001) also has been cited as a screening measure of emergent literacy skills. Results of a previous study suggested that the PALS-PreK was a comparable screening measure to the GRTR-R and IGDIs (e.g., Invernizzi, Cook, & Gellar, 2002–2003). However, whereas previous studies investigating the psychometric properties of the GRTR-R and IGDIs have used unrelated criterion measures, studies investigating the PALS-PreK have used another version of the PALS (e.g., PALS-Kindergarten version) as the criterion. This, in conjunction with the fact that the 121-item PALS-PreK takes much longer than either the GRTR-R or IGDIs to administer, suggests that the GRTR-R and IGDIs are better candidates for screening tools of emergent literacy skills, given the time and financial constraints of most preschools.

In this study, the GRTR-R and the IGDIs were compared using the *Test of Preschool Early Literacy* (TOPEL; Lonigan, Wagner, Torgesen, & Rashotte, 2007), a diagnostic measure of emergent literacy skills, as the predicted criterion measure of emergent literacy. Specifically, the GRTR-R and IGDIs were administered just prior to preschool entry and the TOPEL was administered just after preschool entry--a span of approximately three months. Optimal cut scores for the GRTR-R and IGDIs in predicting performance at or below the 25th percentile on the TOPEL were generated, and accuracy values for each screening tool were compared.

In the literature, several time intervals have been utilized for comparing screening tools to other measures, including 6 months (e.g., Molfese et al., 2002), 3 months (e.g., Missall & McConnell, 2004), 18 to 30 months (e.g., Missal et al., 2007), and 17 to 37 months (Phillips et al., 2009). The rationale in this study for choosing a three-month testing interval was threefold. First, in the pre-k context, one purpose of using a screening tool is to identify children with weak skills in an area and to provide additional assessment or intervention as needed. If effective intervention is applied, it is reasonable to assume that the prediction of a screening measure to an outcome measure will attenuate as children benefit from instruction. In this study, children's emergent literacy skills were screened using the GRTR-R and IGDIs just before children entered preschool--near the time when such a screening measure would be used in actual practice. After the children had acclimated to their new classrooms and teachers for about two months, their emergent literacy skills were assessed using a

diagnostic measure (the TOPEL). Had the TOPEL been administered later in the school year, results would no longer reflect just the screening capability of the GRTR-R and IGDIs for emergent literacy skills; they also would have represented variance in emergent literacy skills due to exposure to instructional activities. Given that the preschoolers in this study might have been engaging actively in some literacy activities and that the focus was on identifying children who would be classified as showing a high level of risk on a comprehensive measure of the precursor skills to the development of reading, the three-month testing interval from preschool entry to post-school entry was appropriate.

In addition to this practical use reason, the three-month interval used in this study was appropriate because the primary purpose of the study was to compare the level of accuracy in classification for the GRTR-R and IGDIs. Although the strength of the correlations between these screening measures and the TOPEL might have been attenuated over a testing interval longer than three months, the pattern of results would likely remain consistent given the stability of emergent literacy skills.

Finally, the three-month testing interval was appropriate because the primary question in evaluations of the benefit of screening is “screening to what?” In most studies of screening measures, the typical analysis reported is a correlation between a score on a screening measure and a score on a criterion measure, which does not address the question of the classification accuracy of a screening measure. The purpose of a measure like GRTR-R or IGDIs is to identify risk for later academic or reading problems. It is unlikely that even an extensive diagnostic battery would have high classification accuracy if the interval between administration of the diagnostic battery and the criterion measure were large. The goal of using a screening measure is to determine if there is value in conducting comprehensive assessment or providing an intervention.

In this study, receiver operating characteristic (ROC) curves were generated, and optimal cut scores were chosen for each screening tool’s prediction of criterion measures of print knowledge, phonological awareness, oral language, and overall emergent literacy skills. In this study, the minimum sensitivity allowed was set at .90 prior to data analysis because it was determined that correctly identifying children with poor emergent literacy skills (sensitivity) was more important than correctly identifying children with satisfactory emergent literacy skills (specificity). From optimal cut scores based on sensitivity values, indices of accuracy were calculated for each screening tool and compared. Setting sensitivity a priori to a certain level has implications for the number of children identified as weak in emergent literacy skills who in fact have satisfactory emergent literacy skills (i.e., false positives). Specifically, as sensitivity increases, the number of false positives increases as well, which has implications for the use of scarce resources for additional assessment or intervention, as discussed later. Based on previous findings (e.g., Missall et al., 2007; Molfese et al., 2006; Phillips et al., 2009; Wilson & Lonigan, 2008) and the item content of the measures, it was predicted that optimal cut scores for the GRTR-R and IGDIs would yield similar indices of accuracy for the criterion measures of phonological awareness and oral language but that indices of accuracy for the criterion measures of print knowledge and overall emergent literacy skills would be better for the GRTR-R than the IGDIs.

Method

Participants

Parents of 199 children from 21 preschools in north Florida signed consent forms allowing their children to participate in the study; only children remaining in preschool for the upcoming school year were included in the study. Children’s ages ranged from 42 to 55 months, with a mean age of 48.55 months ($SD = 3.69$ months). Child sex was divided

equally among boys (50%) and girls, and the majority of the children were Caucasian (70%; 19% African American; 11% other ethnicity). Although 199 children were assessed at Time 1 (July), 23 children were unavailable for assessment at Time 2 (October). These 23 children were mostly from ethnic minority groups (52% African American; 9% other ethnicity), boys (61%), and they obtained lower Time 1 scores on the GRTR-R, $F(1, 198) = 4.32, p = .04$; and IGDIs total score, $F(1, 198) = 5.28, p = .02$ (see Note 1). The 176 children remaining in the sample ranged in age from 42 to 55 months at Time 1, with a mean age of 48.49 months ($SD = 3.68$ months). Most of these children were boys (49%) and Caucasian (74%; 15% African American; 11% other ethnicity).

Preschools and Child-care Centers

Informal observations of the preschools and child-care centers were made to identify the general structure of these environments. Although materials available to children and activity structure varied between locations, the majority of centers did not engage in formal instruction; that is, children primarily spent their time in self-directed activities.

Measures

Get Ready to Read! Screening Tool (GRTR! Whitehurst & Lonigan, 2001)—The GRTR was revised recently (*Revised Get Ready to Read!*; GRTR-R). The GRTR-R is a 25-item test that measures print knowledge and phonological awareness. For each item, the child is shown a page with four pictures. The test administrator reads the question at the top of each page aloud and the child answers by pointing to one of the four pictures. At the end of the GRTR-R, correct answers are summed into a single score encompassing both print knowledge and phonological awareness. Internal consistency reliability for the GRTR-R in the normative sample ($N = 866$ 3-, 4-, and 5-year-old children) was .88 (Lonigan & Wilson, 2008).

Individual Growth and Development Indicators (IGDIs; McConnell, 2002)—The IGDIs is a compilation of tests designed to describe young children's growth and development, including expressive communication, adaptive ability, motor control, social ability, and cognition. For this study, the three tasks related to emergent literacy skills were chosen: Alliteration and Rhyming (to evaluate phonological awareness), and Picture Naming (to evaluate oral language). There is no IGDIs test for print knowledge. For each of these three tasks, a set of flashcards is available as an item pool. The set of cards is shuffled between children so that each child is given a different set and order of cards. Scores on each task are the number of items completed correctly within a two-minute (for Alliteration and Rhyming) or one-minute (for Picture Naming) administration period. For both Alliteration and Rhyming tasks, if the child does not provide a correct answer for at least two of the first four cards shown, the remainder of the task is not administered. According to the IGDIs Technical Report #8 (Missall & McConnell, 2004), one-month test-retest reliability for these three tasks ranged from .44 to .78 (Picture Naming), .83 to .89 (Rhyming), and .46 to .80 (Alliteration). Internal consistency cannot be calculated for the IGDIs because items are not consistent across test administration.

Test of Preschool Early Literacy (TOPEL; Lonigan et al., 2007)—The development and testing of the TOPEL was based on the last decade of research concerning the development of emergent literacy, and the final version was normed on a sample of 842 children representative of the national population on several domains, including gender,

Note 1. Similar results were found for individual tasks: IGDIs Picture Naming, $F(1, 198), 5.45, p = .02$; TOPEL Print Knowledge, $F(1, 198), 4.54, p = .03$; and TOPEL Definitional Vocabulary, $F(1, 198), 6.39, p = .01$. However, neither IGDIs Phonological Awareness, $F(1, 198), 1.22, p = .27$; or TOPEL Phonological Awareness, $F(1, 198), .86, p = .36$; were different across these two groups.

ethnicity, family income, and highest level of parent education, all of which remained relatively consistent when stratified by children's ages (i.e., 3-, 4-, and 5-year-olds). The TOPEL includes three subtests: Print Knowledge, Definitional Vocabulary, and Phonological Awareness. Print Knowledge measures print concepts, letter discrimination, word discrimination, letter-name identification, and letter-sound identification. Definitional Vocabulary measures children's single word spoken vocabulary and their ability to formulate definitions for words. Phonological Awareness includes both multiple-choice and free-response items along the developmental continuum of phonological awareness from word awareness to phonemic awareness. According to the test manual (Lonigan et al., 2007), internal consistency reliability for these subtests ranges from .86 to .96 for 3- to 5-year-olds, and test-retest reliability over a one- to two-week period ranges from .81 to .89. In addition to these subtest scores, an Early Literacy Index (ELI) can be generated. The internal consistency reliability of the TOPEL ELI is .96.

The three TOPEL subtests and the TOPEL ELI have good evidence of convergent validity with other measures of similar constructs. According to the TOPEL Manual, the TOPEL is highly correlated with concurrent tests of alphabet knowledge (TERA-3, r with TOPEL Print Knowledge = .77), expressive vocabulary (EOWPVT, r with TOPEL Definitional Vocabulary = .71), phonological awareness (CTOPP, r s with TOPEL Phonological Awareness = .59 to .65), and overall early reading ability (TERA-3, r with TOPEL Early Literacy Index = .67). Data also support that the TOPEL is predictive of later reading skills. For example, scores on the TOPEL Print Knowledge and Phonological Awareness subtests administered in preschool were found to be significant correlates both of measures of phonological awareness (median r = .40), which is a strong concurrent correlate of decoding skills, and measures of reading skills (e.g., word identification, word attack; r s = .30 to .60) administered when children were in kindergarten and first grade (Sims & Lonigan, 2008; Lonigan & Farver, 2008). In fact, for the Print Knowledge subtest, the mean correlation with decoding skills measured near the end of first grade was .57 in one sample.

Procedure

Both written consent from parents and verbal assent from children were obtained prior to assessment. Both screening tools (GRTR-R and IGDIs) were administered just before the beginning of the preschool year. Approximately three months later, when children had been in their preschool classrooms for at least a month, the TOPEL was administered. To ensure that order of screener administration did not affect performance, the order of administration of GRTR-R and IGDIs was counterbalanced across children, determined randomly. Each testing session was conducted by trained examiners, all of whom were required to demonstrate proficiency with the assessments before testing any children.

Results

Preliminary Analysis and Descriptive Data

All variables were examined for accuracy of data entry, missing values, and fit between their distributions and the assumptions of multivariate analysis. The only variables containing missing values were the Alliteration and Rhyming tasks of the IGDIs. Because these missing values were due to children's inability to answer at least two of the four practice items correctly, all of these missing values were replaced with zeros.

The correlation between IGDIs Alliteration and IGDIs Rhyming was at least moderate (r = .32), and both measures are intended to assess phonological awareness. Therefore, these two IGDIs tasks were summed into a Phonological Awareness (PA) composite score. In addition to calculating an IGDIs PA composite, an IGDIs total score was calculated by summing raw

scores across IGDIs Alliteration, Rhyming, and Picture Naming. The correlation between IGDIs PA and IGDIs Picture Naming (PN) was weak ($r = .12$). To provide a comprehensive picture for how the IGDIs functions as an emergent literacy screener, the IGDIs composite score, IGDIs PA score, and IGDIs PN score were tested independently (see Note 2).

Descriptive statistics for scores on the screening measures at Time 1 and TOPEL scores at Time 2 are shown in Table 1. Order of test administration did not affect screening measure scores (GRTR-R score, $F[1,175] = .77, p = .38$; IGDIs score, $F[1, 175] = .01, p = .91$). Correlations between screening tool scores (i.e., GRTR-R, IGDIs composite score, IGDIs PA, and IGDIs PN) as well as between TOPEL scores (i.e., Print Knowledge, Definitional Vocabulary, Phonological Awareness, and Emergent Literacy Index) are shown in Tables 2 and 3, respectively.

Screening Accuracy Analysis

Diagnostic efficiency of the GRTR-R and IGDIs was tested by generating receiver operating characteristic (ROC) curves using TOPEL standard score cutoffs of 90 (26th percentile) for all three subtests as well as the Early Literacy Index. Like most other achievement-related individual differences, emergent literacy skill is a continuous variable. A child scoring just below the cut score (e.g., at the 24th percentile) is not qualitatively different than a child scoring just above the cut score (e.g., at the 27th percentile). The goal in choosing the 25th percentile as a cut score was to identify a group of children performing at the lower end of the distribution of emergent literacy skills and therefore children who were more likely candidate for additional assessment or need intervention than children scoring at higher percentiles.

Choice of optimal cut scores for the GRTR-R and IGDIs was determined by examining sensitivity values (i.e., proportion of children correctly identified as at risk by the screening tool). Cut scores were chosen such that sensitivity would be above .90; therefore, sensitivities are similar across all comparisons. Raw scores used to calculate indices of accuracy at optimal screening tool cut scores are shown in Table 4.

Effectiveness of the GRTR-R and IGDIs was determined by evaluating area under the curve (AUC), specificity (proportion of students correctly classified as not at risk by the screening tool), likelihood ratio, and overall correct classification. AUC is the proportion of the area falling below the ROC curve; values range from 0.5 to 1, with .5 denoting chance performance of the screening tool and 1 denoting perfect performance (Swets, 1988). Differences between AUCs were evaluated using the statistical procedure outlined by Hanley and McNeil (1983). The likelihood ratio is an indicator of the odds that a child identified as at risk by the screening tool will also be identified as at risk by the diagnostic measure. A value of 1 for this index of accuracy would mean that the screening tool was useless. A value of 2 would mean that failure on the screening tool was twice as likely for children at risk than children not at risk. Overall correct classification is the proportion of children who were either correctly identified as at-risk or not at-risk for later reading difficulties. These four indices were selected for evaluating the effectiveness of these screening tools because they are fixed properties of the test and thus not affected by base rate of diagnostic measure classification (Streiner, 2003).

Whereas sensitivity, specificity, and likelihood ratio values are fixed properties of the test and refer to how good the screening tool is at identifying at risk and not at risk status, predictive power is highly dependent on the prevalence of weak emergent literacy skills in

Note 2. In addition to using raw screening tool scores to calculate ROC curves, age-standardized scores were used to calculate an alternative set of ROC curves. Results were similar to those presented.

the sample (Meehl & Rosen, 1955) and refers to the probability that a child has been correctly identified as either at risk or not at risk by the screening tool. Positive predictive power is the proportion of children with poor emergent literacy skills who are correctly classified by the screening tool, and it is correlated with sensitivity. In this study, positive predictive power was comparable across calculations. Negative predictive power is the proportion of children with good emergent literacy skills who are correctly classified by the screener, and it is related to specificity.

A final piece of information taken into consideration was the percentage of children falling below the optimal cut scores. From a practical perspective, a screening tool that identifies 100% of children as at risk among a population of students with a wide range of emergent literacy abilities is not helpful, as its use would provide no discrimination between children. The percentage of children falling below the optimal cut score on a screening tool should be approximately equal to the percentage of children truly at risk according to the diagnostic measure. Note that percentages of children falling below optimal cut scores are limited in their generalizability, as they are affected by the base rate of the presented condition in the sample. If there are more children at risk in a sample than children not at risk, then the percentage of children falling below the cut score will be inflated. In this study, direct comparisons can be made between percentages of children falling below the cut score, as the same children were tested using both the GRTR-R and the IGDIs.

The ROC curves for the GRTR-R and IGDIs total score in predicting the TOPEL Early Literacy Index are shown in Figure 1. Indices of accuracy for the GRTR-R and IGDIs ROC curves are reported in Table 5. With regard to the TOPEL Early Literacy Index, the AUC values for both screening tools were at or above .70, indicating good accuracy for both screening tools. However, the GRTR-R (AUC = .86) had higher accuracy than the IGDIs total score (AUC = .73; $z = 2.57, p < .01$), the IGDIs PA score (AUC = .68; $z = 2.71, p < .01$), and the IGDIs PN score (AUC = .67, $z = 2.81, p < .01$). The GRTR-R showed higher specificity and negative predictive power than any of the IGDIs scores, meaning that scores on the GRTR-R correctly identified children not at risk for later reading difficulties more often than did the IGDIs. Scores from the IGDIs identified children with satisfactory emergent literacy skills as at risk more often than did the GRTR-R. Additionally, the likelihood ratio was higher for the GRTR-R than all three IGDIs scores. For predicting to the TOPEL Early Literacy Index, the GRTR-R had higher overall correct classification than the IGDIs. Finally, the GRTR-R identified 41% of the current sample as at risk for later reading difficulty. In contrast, the IGDIs total, PA, and PN scores identified 67%, 74%, and 77%, respectively, of the current sample as at risk for later reading difficulty. Thus, use of the GRTR-R resulted in identification of a smaller group of children potentially in need of intervention.

In addition to computing ROC curves for scores on the TOPEL Early Literacy Index, ROC curves for TOPEL subtest scores were computed to determine how the GRTR-R and IGDIs related to the three separate domains of emergent literacy measured by the TOPEL. With regard to TOPEL Print Knowledge, the GRTR-R (AUC = .84) had significantly higher accuracy than all three IGDIs scores in the domains of area under the curve (IGDIs total score, AUC = .76, $z = 1.87, p = .03$; IGDIs PA, AUC = .69, $z = 2.78, p < .01$; IGDIs PN, AUC = .70, $z = 2.71, p < .01$), specificity, negative predictive power, likelihood ratio, overall correct classification, and percentage of children falling below the optimal cut score.

For TOPEL Definitional Vocabulary, the AUC for the GRTR-R (.75) was significantly higher than the AUCs for IGDIs PA (AUC = .60, $z = 1.76, p = .04$) and IGDIs PN (AUC = .71, $z = 1.85, p = .03$); the AUC for the GRTR-R was statistically equivalent to the AUC for the IGDIs total score (AUC = .71, $z = .55, p = .29$). However, at the optimal cut scores for

these screening tools, IGDIs PN was more accurate than the GRTR-R as well as IGDIs total score and IGDIs PA with regard to specificity, negative predictive power, likelihood ratio, overall correct classification, and percentage of children scoring below the cut score.

For TOPEL Phonological Awareness, the AUC for the GRTR-R (.68) was statistically equivalent to the AUCs for IGDIs total score (AUC = .64, $z = .94$, $p = .17$) but marginally higher than for IGDIs PA (AUC = .60, $z = 1.52$, $p = .06$) and IGDIs PN (AUC = .62, $z = 1.54$, $p = .06$). At optimal cut scores for these screening tools, results were similar: the GRTR-R was slightly more accurate than all three IGDIs scores with regard to specificity, negative predictive power, likelihood ratio, overall correct classification, and percentage of children scoring below the cut score.

Discussion

Overall, these findings indicate that it is possible to effectively screen preschool children with less well developed emergent literacy skills, who are at higher risk of later reading problems than children with more well developed emergent literacy skills. In general, the results indicated that use of the GRTR-R yielded more accurate classification of children into at-risk or not at-risk groups with regard to their overall emergent literacy skills than did the IGDIs. In terms of classification accuracy for overall early literacy skills, the GRTR-R did about as well as established “screening tools” that are used with kindergarten and early elementary school age children (e.g., Invernizzi et al., 2002–2003). With regard to specific domains of emergent literacy skills, the GRTR-R and IGDIs were more accurate in predicting children’s print knowledge than they were in predicting children’s oral language and phonological awareness. The results of this study demonstrate that screening measures could be used to effectively screen preschool children who may be in need of more in-depth assessment or to identify preschool children who are most in need of additional or more intensive exposure to instructional activities to promote the development of early literacy skills. The results of this study also demonstrate that the use of these measures do not extend to the identification of specific weaknesses or strengths in specific emergent literacy domains.

Whereas previous studies have examined the psychometrics of these screening tools (e.g., Missall et al., 2007; Missall & McConnell, 2004; Molfese et al., 2006; Phillips et al., 2009; Whitehurst, 2001), this study expands on earlier findings by addressing the question of how well these assessments work as “screening tools” in the traditional sense. That is, whereas prior studies have established acceptable levels of internal consistency, cross-time stability, or validity for these measures, this study established the utility of the measures for screening to an established level of performance on a nationally standardized measure of emergent literacy skills that reflected performance below the expected level given a child’s age. Given that the intended goal of screening is identification, the question of how well a measure works (i.e., its accuracy) for identification of children is the one that provides educators with the necessary information for measure selection. Whereas traditional psychometric indices are factors that likely contribute to a measure’s utility as a screening tool, acceptable psychometrics in terms of reliability and validity do not insure classification accuracy. One previous study comparing the psychometric properties of the GRTR-R and IGDIs found that the GRTR-R had equivalent or higher indices of reliability and validity than did the IGDIs (Wilson & Lonigan, 2008). However, results indicated that although the GRTR-R was more accurate than the IGDIs in predicting to overall emergent literacy skills, print knowledge, and phonological awareness, the IGDIs Picture Naming task was a more accurate predictor of risk status in oral language. Thus, it is important to examine classification accuracy in addition to traditional psychometric properties, as these analyses reveal different results.

Benefits of Examining ROC Curves

Analysis using ROC curves allows two ways of examining data. First, the AUC is an indicator of a screening tool's overall ability to differentiate between children with lower than average emergent literacy skills and children with average or above emergent literacy skills, and it is calculated at all possible cut-scores. All other ROC statistics reported in this study were based on cut-scores set where sensitivity was at least .90 (i.e., where 90% of children at risk for later reading difficulty were correctly identified) and allowed examination of the properties of these screening tools at these optimal cut scores. Whereas examination of the AUC allows a broad determination of the overall effectiveness of the screening tool at all possible cut scores, using optimal cut-score statistics allows examination of the utility of the screening tool under the circumstances in which it would typically be used. It is unlikely that an educator would be interested in a cut score with a sensitivity of, for example, .20, as this would result in correct identification of only 20% of students in the risk category on the criterion measure. Conversely, a cut score with a sensitivity of .99 would most likely result in a correspondingly low specificity, resulting in identification of almost all children in the risk category on the criterion measure.

Discrepancies between the AUC and other ROC statistics (e.g., when the AUC favored one screening tool and other indices of accuracy favored the other screening tool) were due to the screening tools performing differentially at different cut scores. For example, as shown in Figure 1, the GRTR-R was more accurate than the IGDIs total score at very low levels of sensitivity, as evidenced by the higher specificity of the GRTR-R than the IGDIs total score. However, at very high levels of sensitivity (i.e., above .98), the GRTR-R and IGDIs total score were equal in terms of screening to overall emergent literacy skills. Although AUC values for the GRTR-R were generally higher than AUC values for the IGDIs, at optimal cut-scores, alternative indices of accuracy differentially supported these two screening tools depending on the domain of emergent literacy measured (i.e., print knowledge, oral language, and phonological awareness).

Results of this study allowed determination of cut-scores on the IGDIs and GRTR-R for predicting later risk status, defined as scoring below average on a diagnostic measure of early literacy skills. Given the significant predictive relations between emergent literacy skills and reading and writing abilities measured when children are in kindergarten and early elementary school grades (e.g., Lonigan et al., 2000; Lonigan et al., 2008a; Storch & Whitehurst, 2002), teachers or other early childhood professionals can use information concerning whether a child scored at, above, or below the cut-score on the screening tool to determine if the child is likely to be at risk for later reading difficulty or not. Unlike predictive correlations, which provide information regarding a measure's overall ability to predict later scores, ROC curve analyses provide specific information about the dichotomous outcome of whether a child will or will not fall into a particular category--in this case scoring at or above versus below the average range on a diagnostic measure of emergent literacy skills. Accurate classification of children into at-risk and not at-risk categories is salient and useful information for educators attempting to ready children for formal schooling. Children scoring at or below the cut score either can be referred for additional and more in-depth assessment to identify their patterns of strengths and weaknesses or can be provided with more frequent or more intensive instructional activities to enhance development of their early literacy skills.

The GRTR-R score was a more accurate predictor of children's overall emergent literacy skills than the IDGIs total score, IGDIs phonological awareness composite, and IGDIs Picture Naming score. In this study, not only was the GRTR-R superior in overall classification at all cut scores (i.e., AUC), it was more accurate than were the IGDIs under the circumstances in which a screening tool would typically be used (i.e., sensitivity, or

correct classification of 90% of at-risk children). These results align with a previous study by Wilson & Lonigan (2008) in which the psychometric properties of the GRTR-R and IGDIs were compared directly; the GRTR-R ($r = .65$) was a more accurate predictor of children's emergent literacy skills than were the IGDIs ($r = .43$). Additional studies have examined the GRTR and IGDIs separately and over a longer timeframe and found comparable data across these two screening tools. For example, Phillips et al. (2009) reported that correlations between the GRTR and measures of phonological awareness, print knowledge, and decoding were moderate nearly three years after initial preschool assessment (r s ranged from .25 to .40; median $r = .32$). Missall et al. (2007) also found moderate correlations between preschool IGDIs scores and oral reading fluency at the end of first grade (r s ranged from .26 to .50; median $r = .42$). Future studies should examine the long-term predictive validity of these two screening tools using the same sample of children to further clarify the relation of the GRTR-R and IGDIs to diagnostic measures of emergent literacy and which is the more accurate screening tool over a period of time longer than three months.

Comparison of the GRTR-R and IGDIs with Other Screening Measures

Screening or monitoring children's literacy skills has become increasingly common in kindergarten and early elementary school. For example, the Dynamic Indicators of Basic Early Literacy Skills (DIBELS; Good & Kaminski, 2002) and the Texas Primary Reading Inventory (TPRI; Texas Education Agency, 1999) are nationally recognized screening-type measures of literacy-related skills performance. Whereas most of the reported data for available literacy assessments only include psychometric properties such as reliability and validity, published reports for the DIBELS and TPRI include classification accuracy data. Results from this study indicate that the GRTR-R was about as accurate or more accurate than these two nationally recognized measures, as evidenced by GRTR-R classification accuracy data in the current study (sensitivity = .69, specificity = .69) approximating those for the DIBELS and TPRI. Specifically, the sensitivity and specificity of the DIBELS administered at the end of kindergarten in predicting academic outcomes at the end of first grade ranged from .80 to 1.00 and .23 to .39, respectively ($N = 86$; Good et al. in preparation, as cited by Glover & Albers, 2007). For the TPRI, sensitivity in predicting first grade outcome from kindergarten was .90; specificity was .62 ($N = 421$; Foorman et al. 1998).

Neither the DIBELS nor the TPRI measures are appropriate for assessing preschool children. In fact, these screening tools are not to be administered before the middle of kindergarten. We are aware of only one previous study in which the classification accuracy of a "screening tool" for use with preschool children was examined. Invernizzi et al. (2002–2003) reported that the Phonological Awareness Literacy Screening (PALS) Pre-K version (PALS-PreK; Invernizzi et al, 2001) had an overall classification accuracy of .88 to the kindergarten version of PALS (PALS-K; Invernizzi, Meier, Swank, & Juel, 1999) in a sample of 198 preschool children assessed in the spring of their preschool year using the PALS-PreK and in the spring of their kindergarten year using the PALS-K. Although the level of classification accuracy reported for the PALS-PreK was higher than that achieved by the GRTR-R (.73), the IGDIs total score (.47), the IGDIs Phonological Awareness tasks (.39), or the IGDIs Picture Naming task (.38), there are a number of critical issues about the PALS-PreK study to address, including the measures and the sample used. First, whereas classification accuracies of the GRTR-R and IGDIs were determined in reference to a completely different and nationally standardized diagnostic measure, classification accuracy of the PALS-PreK was determined in reference to the kindergarten version of the same measure (i.e., the PALS-K). Second, the PALS-PreK contains 121 items and takes at least 45 min to administer. In contrast, the GRTR-R has only 25 items, and the relevant IGDIs

measures consist of three tasks that take a total of 5 minutes to administer. Finally, the sensitivity of the PALS-PreK for predicting risk status on the PALS-K was only .15, meaning that the measure only correctly identified 15% of the children who were later found to be at actual risk based on their PALS-K scores, as compared with at least 90% for the GRTR-R and IGDIs. Taken together, the above facts suggest that the PALS-PreK is not a practical screening tool for emergent literacy skills. The GRTR-R and IGDIs, by virtue of their superior sensitivity as well as their speed and ease of administration, are more effective screening tools of emergent literacy skills than the PALS-PreK.

Another issue that deserves consideration is that although it would be optimal if every child could receive intensive or individualized instruction, this is not currently practical in most early childhood environments. Use of a screening tool like the GRTR-R or IGDIs provides a valid and effective means of distributing scarce resources (i.e., teacher time) to the children most in need of intervention. Studies have shown that children's rate of growth in literacy achievement declines throughout schooling and that below-average readers' growth declines earlier and more steeply than for above-average readers (Chall, Jacobs, & Baldwin, 1990). Work toward closing the achievement gap will require that time is spent ensuring that young children with weak emergent literacy skills are exposed to effective interventions as early as possible. To maximize teachers' time spent instructing students, screening tools must be easy to administer; the longer a teacher spends screening students, the less time he or she has available to work with students. The GRTR-R is a good choice as a screening tool for emergent literacy skills because it is accurate at identifying children in need of individualized instruction without taking too much of the teacher's time for administration.

Limitations in Use of Screening Tools

Although the results of this study supported the use of the GRTR-R and to some extent the IGDIs as screening tools for identifying children who are at-risk of later reading difficulties because of weaknesses in their overall early literacy skills, the results of this study do not support the use of these measures for classification of the specific components of emergent literacy skills. In comparison to the classification accuracy to overall early literacy skills, classification accuracies to print knowledge, oral language, and phonological awareness skills were substantially lower, and with the exception of screening to print knowledge, neither screening tool evidenced clear superiority. For the specific emergent literacy domains overall and oral language in particular, rates of false positives (i.e., children classified as at risk who did not score below average levels on the relevant TOPEL subscales) were very high. If a screening tool is correlated highly with the diagnostic measure, the false positive rate will be lower than if a screening tool is correlated weakly with the diagnostic measure. In this study, the overall correct classification to specific emergent literacy domains, particularly oral language, was low; hence, false positive rates were higher for these domains than for classification to overall emergent literacy skills.

In this study, focus was placed on increasing the sensitivities of the screening tools (i.e., correct classification of children with weak emergent literacy skills), allowing the false positive rate free to vary. The practice of setting a minimum sensitivity (in this study .90) will often result in correspondingly high false positive rates. For example, in this study, the best screening result was from GRTR-R to TOPEL Emergent Literacy Index. Practical use of these screening data in a school would result in 27 children truly needing intervention (true positives, 15%) and 45 children not needing intervention (false positives, 26%) being labeled as in need of intervention. In a school with few resources, it is important to provide intervention only to those children who truly have low scores on the criterion. Overall, the GRTR-R resulted in a smaller number of false positives than the IGDIs, particularly for screening to overall emergent literacy skills and to print knowledge skills.

The results of this study indicate that the GRTR-R was a more accurate screening tool for emergent literacy skills than the IGDIs. Future research focused on developing more accurate screening tools for the specific domains of emergent literacy, particularly phonological awareness and oral language, may be valuable. Alternatively, it may be that screening to overall emergent literacy skills alone is sufficient in the preschool years given that the specific skills are correlated (e.g., in this sample, the three TOPEL subtests were moderately correlated) and teacher time is limited. Children with weaker overall emergent literacy skills can be grouped together for instruction in all three domains rather than singled out for individualized instruction in particular domains. Ultimately, validated assessment-instruction rubrics that optimize outcomes and help reduce children's risk for later academic difficulties are needed.

Potential Limitations and Future Directions

Despite the strengths of this study and the valuable information provided, there are a few potential limitations of the study that need to be recognized. One potential limitation was that the current sample over-represented children not at risk for later reading difficulties (16% were considered "at risk" or scoring below the 25th percentile on the TOPEL at Time 2) versus the expected 25% based on a population distribution. In looking at this issue practically, the utility of a screening tool's ability to identify at-risk status in a sample of mostly at-risk children is debatable; that is, if most children at a school are at risk for later reading difficulties, then screening is not necessary. Further, Streiner (2003) concluded that at high base rates, tests should be used to rule a condition in but not to rule it out--meaning that in a high-risk sample, administration of a screening tool should not be used to rule out at-risk status. It would be more beneficial to use a screening tool in a heterogeneous population of at-risk children. In this latter case, teachers would be able to differentiate between children most in need of additional assessment or intervention and provide additional services to those children. Therefore, the percentage of at-risk children in the current sample is desirable in determining indices of accuracy for these screening tools.

A second potential limitation of this study is that the cut scores presented are based on a single sample of children. Future studies should be conducted to determine if these cut scores generalize to other samples of children. This can be accomplished by cross-validating the cut scores from this study in new samples using the TOPEL as well as other criterion measures of emergent literacy. Whereas these new studies will likely recover different predictive power and percentages of children falling below the cut score (because these indices of accuracy are dependent on the base rate of weak emergent literacy in the sample), we expect that at a sensitivity of approximately .90, specificity, likelihood ratio, and area under the curve will remain consistent with those provided in this study. It is important to note that the sample used in this study were prekindergarten children assessed at the beginning and middle of their preschool year. Future studies should examine the predictive utility of the GRTR-R and IGDIs later in the preschool year (while still keeping consistent the three-month testing interval). For example, Missall et al. (2007) found that although IGDIs Rhyming was relatively consistent in predicting kindergarten skills, whether measured in the fall, winter, or spring of the preschool year, IGDIs Alliteration demonstrated higher correlations with criterion measures when administered later in the preschool year compared to earlier administration. Therefore, using the IGDIs to screen a child's phonological awareness at the beginning of preschool might not provide much useful information; alternatively, using the IGDIs to screen a child's phonological awareness at the end of the preschool year might be more informative. Future studies comparing the performance of the IGDIs and the GRTR-R at different times throughout the preschool year would allow determination how these screening tools function with older children and children with more developed emergent literacy skills.

Summary and Conclusions

In recent years, there has been an increasing emphasis placed on early childhood education as one piece of a system of education designed to increase academic outcomes for children. Increasingly, states are adopting preschool learning standards that address critical components of the precursors to later formal academic skills, like reading and mathematics. Use of screening tools that can accurately identify children who are at-risk of later reading difficulties because of below average development of emergent literacy skills is one way to reduce the likelihood that children will later receive a learning disabled classification or experience significant academic difficulties. Identification and deployment of effective preschool interventions to promote the development of the relevant skills that increase the probability of success in school is also required. High quality, causally interpretable research has identified a number of interventions capable of boosting the emergent literacy skills of children identified as needing additional instruction (e.g., see Lonigan, Schatschneider, Westberg, 2008b; Lonigan, Shannahan, & Cunningham, 2008; Lonigan & Phillips, 2007; What Works Clearinghouse [<http://ies.ed.gov/ncee/wwc/>] for summaries). However, before successful interventions can be used with children who are at-risk, educators must be able to identify accurately those children who indeed have below average skills in critical domains. Brief, but accurate, screening tools are an excellent way for educators to obtain a snapshot of children's emergent literacy skills. Although there are several available measures of emergent literacy skills, very few of these measures are as simple and quick to administer as the GRTR-R and IGDIs.

The results of this study indicate that the GRTR-R was a more accurate screening tool for children's overall emergent literacy skills than was the IGDIs. Its ability to accurately classify children into at-risk and not at-risk categories was relatively high, with most incorrect classifications resulting from false positives. The results did not support the use of either the GRTR-R or the IGDIs for identifying children with at-risk status in a specific domain of emergent literacy (i.e., oral language, phonological awareness, print knowledge). Consequently, at this time, these screening measures should only be used to identify a broad classification of risk status, and additional in-depth measures would need to be employed to identify children's patterns of strengths and weaknesses in specific emergent literacy skills. Although this study identified cut scores for the measures that could be used to classify children into at-risk and not at-risk groups, additional studies are needed to cross-validate these cut scores with additional samples of children that vary in the expected proportion of children at risk, with additional criterion measures of true risk status, and with varying time frames of assessment (e.g., beginning, middle, end of preschool year).

Acknowledgments

The research reported here was supported, in part, by grants from the Institute of Education Sciences, U.S. Department of Education (R305B04074, R324E06086) and the National Institute of Child Health and Human Development (HD052120) to Florida State University. The opinions expressed are those of the authors and do not represent views of granting agencies.

References

- Bishop DVM, Adams C. A prospective study of the relationship between specific language impairment, phonological disorders and reading retardation. *Journal of Child Psychology and Psychiatry and Allied Disciplines*. 1990; 31:1027–1050.
- Chall, JS.; Jacobs, VA.; Baldwin, LE. *The reading crisis: Why poor children fall behind*. Cambridge, Massachusetts: Harvard University Press; 1990.
- Foorman, BR.; Fletcher, JM.; Frances, DJ.; Carlson, CD.; Chen, D.; Mouzaki, A., et al. *Technical report for the Texas Primary Reading Inventory*. 1998 Ed.. Houston, TX: Center for Academic and Reading Skills and the University of Houston; 1998.

- Glover TA, Albers CA. Considerations for evaluating universal screening assessments. *Journal of School Psychology*. 2007; 45:117–135.
- Good, RH.; Kaminiski, RA., editors. *Dynamic Indicators of Basic Early Literacy Skills (DIBELS)*. 6th Ed.. Eugene, OR: Institute for the Development of Educational Achievement; 2002.
- Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology*. 1983; 148:839–843. [PubMed: 6878708]
- Invernizzi, M.; Cook, A.; Geller, K. *PALS-PreK: Phonological Awareness Literacy Screening for Preschool (Technical Report)*. Charlottesville, VA: University of Virginia; 2002–2003.
- Invernizzi, M.; Meier, J.; Swank, L.; Juel, C. *Phonological Awareness Literacy Screening-Kindergarten*. Charlottesville, VA: University of Virginia; 1999.
- Invernizzi, M.; Sullivan, A.; Meier, J. *Phonological Awareness Literacy Screening-Pre-Kindergarten*. Charlottesville, VA: University of Virginia; 2001.
- Lonigan CJ. Development, assessment, and promotion of pre-literacy skills. *Early Education and Development*. 2006; 17:91–114.
- Lonigan CJ, Burgess SR, Anthony JL. Development of emergent literacy and early reading skills in preschool children: Evidence from a latent-variable longitudinal study. *Developmental Psychology*. 2000; 36:596–613. [PubMed: 10976600]
- Lonigan, CJ.; Farver, JM. Development of reading and reading-related skills in preschoolers who are Spanish-speaking English language learners. Asheville, NC: Presented at the annual meeting of the Society for the Scientific Study of Reading; Jul. 2008
- Lonigan, CJ.; Phillips, BM. *Encyclopedia of Language and Literacy Development*. Canadian Language and Literacy Research Network; 2007. Research-based instructional strategies for promoting children's early literacy skills. <http://www.literacyencyclopedia.ca/index.php?fa=items.show&topicId=224>
- Lonigan, CJ.; Schatschneider, C.; Westberg, L. *Developing Early Literacy: Report of the National Early Literacy Panel*. Washington, DC: National Institute for Literacy; 2008a. Identification of children's skills and abilities linked to later outcomes in reading, writing, and spelling; p. 55-106.
- Lonigan, CJ.; Schatschneider, C.; Westberg, L. *Developing Early Literacy: Report of the National Early Literacy Panel*. Washington, DC: National Institute for Literacy; 2008b. Impact of code-focused interventions on young children's early literacy skills; p. 107-151.
- Lonigan, CJ.; Shannahan, T.; Cunningham, A. *Developing Early Literacy: Report of the National Early Literacy Panel*. Washington, DC: National Institute for Literacy; 2008. Impact of shared-reading interventions on young children's early literacy skills; p. 153-171.
- Lonigan, CJ.; Wagner, RK.; Torgesen, JK.; Rashotte, CA. *Test of Preschool Early Literacy*. Austin, TX: Pro-Ed; 2007.
- Lonigan, CJ.; Wilson, SB. *Report on the Revised Get Ready to Read! Screening Tool: Psychometrics and normative information*. Technical Report: National Center for Learning Disabilities; 2008.
- McConnell, SR. *Individual Growth and Development Indicators*. Minneapolis, Minnesota: University of Minnesota; 2002.
- McConnell, SR.; Priest, JS.; Davis, SD.; McEvoy, MA. Best practices in measuring growth and development in preschool children. In: Thomas, A.; Grimes, J., editors. *Best Practices in School Psychology*. 4th ed.. Vol. Vol 2. Washington, DC: National Association of School Psychologists; 2002. p. 1231-1246.
- Meehl PE, Rosen A. Antecedent probability and the efficiency of psychometric signs, patterns, and cutting scores. *Psychological Bulletin*. 1955; 52:194–216. [PubMed: 14371890]
- Missall, KN.; McConnell, SR. *Psychometric characteristics of Individual Growth and Development Indicators: Picture naming, rhyme, and alliteration (Tech. Rep. No. 8)*. Minneapolis, MN: University of Minnesota, Center for Early Education and Development; 2004.
- Missall K, Reschly A, Betts J, McConnell S, Heistad D, Pickart M, et al. Examination of the predictive validity of preschool early literacy skills. *School Psychology Review*. 2007; 36:433–452.
- Molfese VJ, Modglin AA, Beswick JL, Neamon JD, Berg SA, Berg CJ, et al. Letter knowledge, phonological processing, and print knowledge: Skill development in nonreading preschool children. *Journal of Learning Disabilities*. 2006; 39:296–305. [PubMed: 16895155]

- Molfese VJ, Molfese DL, Modglin AT, Walker J, Neamon J. Screening early reading skills in preschool children: Get Ready to Read. *Journal of Psychoeducational Assessment*. 2004; 22:136–150.
- Perfetti CA, Beck I, Bell LC, Hughes C. Phonemic knowledge and learning to read are reciprocal: A longitudinal study of first grade children. *Merrill-Palmer Quarterly*. 1987; 33:283–319.
- Petrill SA, Deater-Deckard K, Thompson LA, Schatschneider C, Dethorne LS, Vandenberg DJ. Longitudinal genetic analysis of early reading: The Western Reserve Reading Project. *Reading and Writing*. 2007; 20:127–146. [PubMed: 19829751]
- Phillips BM, Lonigan CJ, Wyatt M. Predictive Validity of the Get Ready to Read! Screener: Concurrent and Long-term Relations with Reading-related Skills. *Journal of Learning Disabilities*. 2009; 42:133–147. [PubMed: 19074622]
- Scarborough HS. Prediction of reading dysfunction from familial and individual differences. *Journal of Educational Psychology*. 1989; 81:101–108.
- Sims, DM.; Lonigan, CJ. The predictive utility of phonological and print awareness for later reading skills. Washington, DC: Presented at the annual meeting of the Institute of Education Sciences; Jun. 2008
- Snow, CE.; Burns, MS.; Griffin, P. Preventing reading difficulties in young children. Washington, D.C: National Academy Press; 1998.
- Spira EG, Fischel JE. The impact of preschool inattention, hyperactivity, and impulsivity on social and academic development: A review. *Journal of Child Psychology and Psychiatry*. 2005; 46:755–773. [PubMed: 15972069]
- Stevenson HW, Newman RS. Long-term prediction of achievement and attitudes in mathematics and reading. *Child Development*. 1986; 57:646–659. [PubMed: 3720396]
- Storch SA, Whitehurst GJ. Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology*. 2002; 38:934–947. [PubMed: 12428705]
- Streiner DL. Diagnostic tests: Using and misusing diagnostic and screening tests. *Journal of Personality Assessment*. 2003; 81:209–219. [PubMed: 14638445]
- Sulzby, E.; Teale, W. Emergent literacy. In: Barr, R.; Kamil, M.; Mosenthal, P.; Pearson, PD., editors. *Handbook of reading research*. Vol. Vol. 2. New York: Longman; 1991. p. 727-758.
- Swets JA. Measuring the accuracy of diagnostic systems. *Science*. 1988; 240:1285–1293. [PubMed: 3287615]
- Teale, WH.; Sulzby, E., editors. *Emergent literacy: Writing and reading*. Norwood, New Jersey: Ablex; 1986.
- Texas Education Agency. *Texas Primary Reading Inventory*. Austin, TX: Author; 1999.
- Verhoeven L, van Leeuwe J. Prediction of the development of reading comprehension: A longitudinal study. *Applied Cognitive Psychology*. 2008; 22:407–423.
- Wagner RK, Torgesen JK, Rashotte CA, Hecht SA, Barker TA, Burgess SR, et al. Changing relations between phonological processing abilities and word-level reading as children develop from beginning to skilled readers: A 5-year longitudinal study. *Developmental Psychology*. 1997; 33:468–479. [PubMed: 9149925]
- Wagner RK, Torgesen JK. The nature of phonological processing and its casual role in the acquisition of reading skills. *Psychological Bulletin*. 1987; 101:192–212.
- Wagner RK, Torgesen JK, Rashotte CA. Development of reading-related phonological processing abilities: New evidence of bidirectional causality from a latent variable longitudinal study. *Developmental Psychology*. 1994; 30:73–87.
- Whitehurst, GJ. *The NCLD Get Ready to Read! screening tool technical report (Tech. Rep.)*. Stony Brook, NY: The State University of New York at Stony Brook, National Center for Learning Disabilities; 2001.
- Whitehurst, GJ.; Fischel, JE. Reading and language impairments in conditions of poverty. In: Bishop, DVM.; Leonard, LB., editors. *Speech and language impairments in children: Causes, characteristics, intervention and outcome*. New York: Psychology Press; 2000. p. 53-71.
- Whitehurst GJ, Lonigan CJ. Child development and emergent literacy. *Child Development*. 1998; 69:848–872. [PubMed: 9680688]

Whitehurst, GJ.; Lonigan, CJ. Get Ready to Read! Screening Tool. New York: National Center for Learning Disabilities; 2001.

Wilson, SB.; Lonigan, CJ. An evaluation of two emergent literacy screening tools for preschool children. Manuscript submitted for publication; 2008.

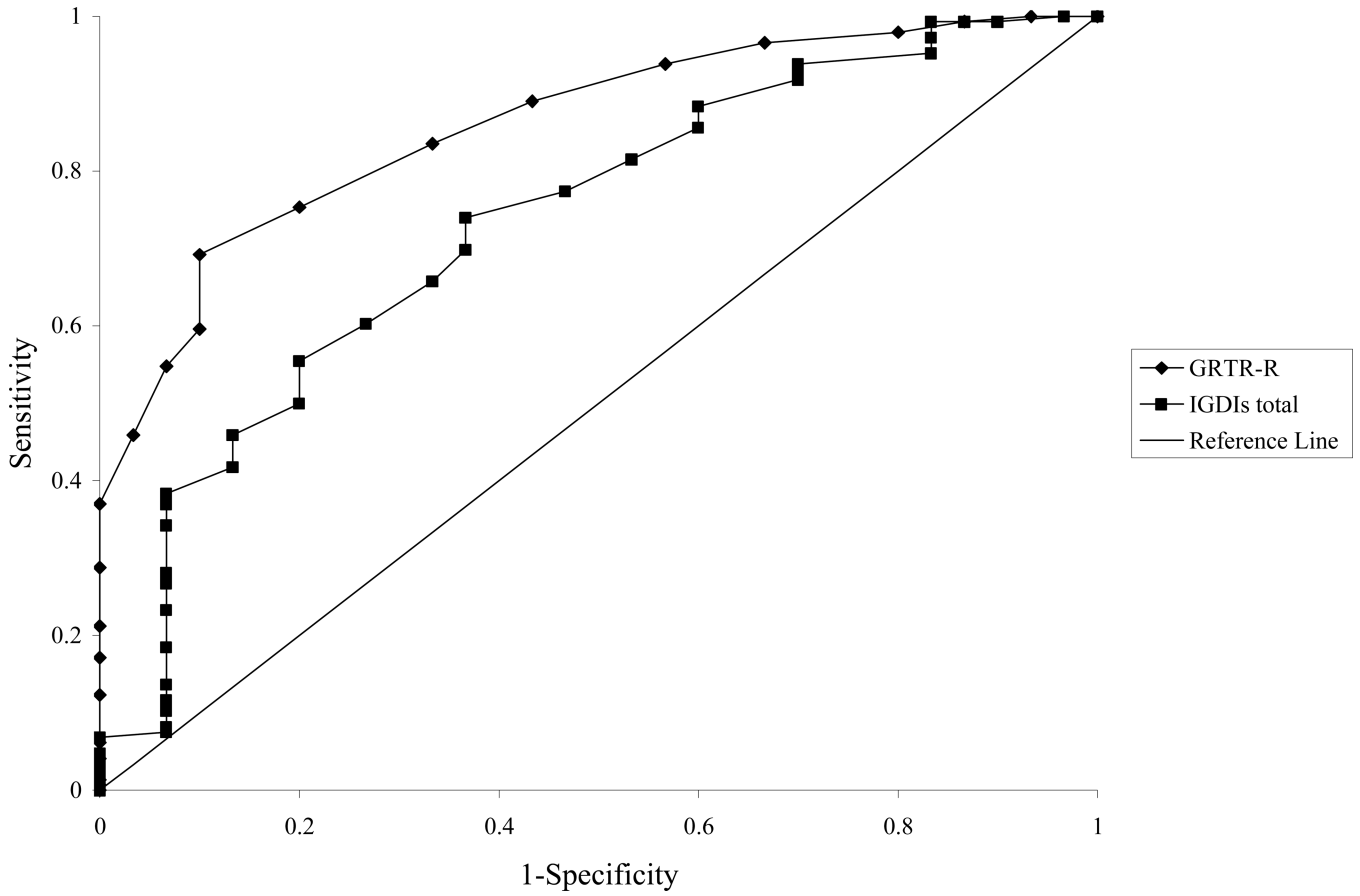


Figure 1. Receiver operating characteristic curves for the *Test of Preschool Early Literacy* Early Literacy Index with the *Revised Get Ready to Read!* score and the *Individual Growth and Development Indicators* total score; the reference line represents where accuracy = .50.

Table 1

Descriptive Statistics for Time 1 and Time 2 Emergent Literacy Skills

Measures	Mean	(SD)	Range
<i>Time 1 Screening Tool Raw Scores</i>			
GRTR-R Total Score	12.95	(4.70)	3–25
IGDIs Composite Score	24.42	(9.52)	3–64
IGDIs Picture Naming	18.46	(6.48)	2–34
IGDIs Phonological Awareness	5.96	(6.27)	0–31
<i>Time 2 TOPEL Standard Scores</i>			
Early Literacy Index	103.22	(14.10)	68–142
Print Knowledge	106.20	(15.65)	77–142
Definitional Vocabulary	104.52	(12.49)	51–127
Phonological Awareness	97.71	(15.48)	57–134

Note. N = 176. TOPEL = Test of Preschool Early Literacy; GRTR-R = Revised Get Ready to Read!; IGDIs = Individual Growth and Development Indicators.

Table 2

Zero Order Correlations within Time 1 Screening Measure Scores

	1	2	3
1. GRTR-R Total Score			
2. IGDIs Composite Score	.57***		
3. IGDIs Picture Naming	.41***	.76***	
4. IGDIs Phonological Awareness	.44***	.74***	.12

Note. $N = 176$. GRTR-R = *Revised Get Ready to Read!*; IGDIs = *Individual Growth and Development Indicators*.

 $p < .005$.

Table 3

Zero Order Correlations within Time 2 Test of Preschool Early Literacy Scores

TOPEL Score	1	2	3
1. Early Literacy Index			
2. Print Knowledge	.76***		
3. Definitional Vocabulary	.74***	.33***	
4. Phonological Awareness	.82***	.40***	.49***

Note. $N = 176$. TOPEL = *Test of Preschool Early Literacy*.

 $p < .005$.

Contingency Table Values Based on Optimal Cut Scores for the Revised Get Ready to Read¹ and the Individual Growth and Development Indicators

Table 4

Comparison	OCC	Risk Classification Based on TOPEL Scores			
		Truly At Risk		Truly Not At Risk	
		True Positive S ⁺ C ⁺	False Negative S ⁻ C ⁺	True Negative S ⁻ C ⁻	False Positive S ⁺ C ⁻
TOPEL Early Literacy Index					
GRTR-R	.73	27	3	101	45
IGDIs Score	.48	28	2	56	90
IGDIs PA	.39	27	3	42	104
IGDIs PN	.38	28	2	39	107
TOPEL Print Knowledge					
GRTR-R	.64	33	3	79	61
IGDIs Score	.51	34	2	56	84
IGDIs PA	.33	34	2	24	116
IGDIs PN	.46	33	3	48	92
TOPEL Definitional Vocabulary					
GRTR-R	.24	19	1	24	132
IGDIs Score	.16	19	1	9	147
IGDIs PA	.18	18	2	13	143
IGDIs PN	.38	18	2	49	107
TOPEL Phonological Awareness					
GRTR-R	.45	52	4	27	93
IGDIs Score	.38	52	4	15	105
IGDIs PA	.41	51	5	21	99
IGDIs PN	.43	52	4	23	97

Note. $N = 176$. TOPEL = Test of Preschool Early Literacy; GRTR-R = Revised Get Ready To Read¹; IGDIs = Individual Growth and Development Indicators; PA = Phonological Awareness; PN = Picture Naming; OCC = Overall Correct Classification; S⁺ = at risk according to screening tool; S⁻ = not at risk according to screening tool; C⁺ = at risk according to criterion measure; C⁻ = not at risk according to criterion measure.

Table 5

Summary of ROC Curve Statistic Comparisons of the Revised Get Ready To Read! and the Individual Growth and Development Indicators with the Test of Preschool Early Literacy

Comparison	Indices of Accuracy							% Children Below Cut Score
	Area Under Curve	Sensitivity	Specificity	Negative Predictive Power	Positive Predictive Power	Likelihood Ratio	Optimal Cut Score	
TOPEL Early Literacy Index								
GRTR-R	.86***	.90	.69	.97	.38	2.90	12	41%
IGDIs Score	.73***	.93	.38	.97	.24	1.50	28	67%
IGDIs PA	.68**	.90	.29	.93	.21	1.27	10	74%
IGDIs PN	.67**	.93	.27	.95	.21	1.27	24	77%
TOPEL Print Knowledge								
GRTR-R	.84***	.92	.56	.96	.35	2.09	14	53%
IGDIs Score	.76***	.94	.40	.97	.29	1.57	28	67%
IGDIs PA	.69**	.94	.17	.92	.23	1.13	13	85%
IGDIs PN	.70***	.92	.34	.94	.26	1.39	23	71%
TOPEL Definitional Vocabulary								
GRTR-R	.75***	.95	.15	.96	.13	1.12	19	86%
IGDIs Score	.71**	.95	.06	.90	.11	1.01	40	94%
IGDIs PA	.60	.90	.08	.87	.11	.98	15	92%
IGDIs PN	.71**	.90	.31	.96	.14	1.30	23	71%
TOPEL Phonological Awareness								
GRTR-R	.68***	.93	.23	.87	.36	1.21	18	82%
IGDIs Score	.64**	.93	.13	.79	.33	1.07	36	89%
IGDIs PA	.60*	.91	.18	.81	.34	1.11	13	85%
IGDIs PN	.62**	.93	.19	.85	.35	1.15	26	85%

Note. $N = 176$. TOPEL = Test of Preschool Early Literacy; GRTR-R = Revised Get Ready To Read!; IGDIs = Individual Growth and Development Indicators; PA = Phonological Awareness; PN = Picture Naming.

* $p < .05$.

**

 $p < .01$.
 $p < .001$.

Note. In the published version of this table, the column labels for Negative Predictive Power and Positive Predictive Power are reversed.