Article

Hammill Institute on Disabilities

Journal of Learning Disabilities
2014, Vol. 47(6) 491–502

DOI: 10.1177/0022219412470518
journaloflearningdisabilities.sagepub.com
SAGE

It Depends: Conditional Correlation Between Frequency of Storybook Reading and Emergent Literacy Skills in Children With Language Impairments

Stephen A. Petrill, PhD[1] Jessica A. R. Logan, PhD[1] Brook E. Sawyer, PhD[2] and Laura M. Justice, PhD[1]

Abstract

The current study examined the association between frequency of storybook reading and emergent literacy in 212 children at risk for language impairment, assessed during the fall semester of kindergarten. Measures included parent-reported storybook reading, as well as direct assessments of print knowledge, letter awareness, and expressive vocabulary. Results suggested nonsignificant to moderate (r = .11 to .25) correlations between frequency of storybook reading and child emergent literacy across the entire range of environment and ability. Quantile regression results suggested that the association was highest at low frequency of storybook reading, particularly for print knowledge, approaching r = .50. Moreover, the association between frequency of storybook reading and emergent literacy was highest at higher levels of emergent literacy for print knowledge, but particularly for letter naming, approaching r = .80. These results suggest that in children with language difficulties, the relationship between aspects of the home environment and emergent literacy is conditional on the quality of the home environment as well as the child's proficiency in emergent literacy skills.

Keywords

language, precursors, dyslexia, literacy

Children with poor language skills (referred to throughout as having language impairments; LI) have significant deficits in the comprehension and/or expression of language, which includes but is not limited to difficulties in vocabulary, morphology, and syntax. LI occurs in roughly 7% to 13% of children entering kindergarten (Tomblin et al., 1997) and is associated with high risk for poor school performance, particularly reading difficulties (e.g., Catts, 2003; Dale, Price, Bishop, & Plomin, 2003; Snowling, Bishop, & Stothard, 2000; Whitehouse, Line, Watt, & Bishop, 2009). For instance, Catts, Fey, Tomblin, and Zhang (2002) found that more than one half (53%) of children with LI are diagnosed with reading disabilities in second grade, and Bishop and Adams (1990) found that preschool children with language problems are 6 times more likely to develop reading problems than children with typical language skills. In the present study, we examine the literacy skills of preschool children who have or who are at risk for developing LI.

Deficits in the environment may be one possible process through which LI leads to substantially higher risk for later academic difficulties, including reading problems (Bishop & Adams, 1990; Dale et al., 2003; Snowling et al., 2000; Whitehouse et al., 2009). The recent National Early Literacy Panel (U.S. National Early Literacy Panel, U.S. National Institute for Literacy, & U.S. National Center for Family Literacy, 2008) identified several aspects of the home and school environment that promote child precursor skills and later literacy development in unselected children, including a large literature exploring the relationship between shared reading and emergent literacy (e.g., Bennett, Weigel, & Martin, 2002; Bus, van IJzendoorn, & Pellegrini, 1995; Evans, Shaw, & Bell, 2000; Leseman & de Jong, 1998; Maclean, Bryant, & Bradley, 1987; Scarborough & Dobrich, 1994; Sénéchal, 2006), which is the focus of the current study. In general, these studies suggest that individual differences in shared reading are associated with variance in print knowledge and oral language skills, particularly in early literacy.

[1]Ohio State University, Columbus, OH, USA
[2]Lehigh University, Bethlehem, PA, USA

Corresponding Author:
Stephen A. Petrill, Ohio State University, 225 Psychology Building, 1835 Neil Avenue, Columbus, OH 43210, USA.
Email: petrill.2@osu.edu

Of particular interest in the current study is the relationship between shared reading and emergent literacy in children with LI. Children with LI are less likely to be exposed to literacy experiences in the home, including shared reading (e.g., Boudreau, 2005; Marvin & Wright, 1997; Skibbe, Justice, Zucker, & McGinty, 2008). Beyond the mean differences between children with and without LI, differences in shared reading experiences *within* groups of children with LI may also be associated with differences in emergent literacy skills. If so, promoting shared reading within children with LI may offer an important buffer against emergent reading difficulties. If not, LI may constitute a threshold that overrides the covariance between shared reading and emergent literacy.

Unfortunately, the literature addressing the relationship between shared reading and emergent literacy within children with LI is sparse. Skibbe et al. (2008) examined differences in parent literacy beliefs and practices, including shared reading, between children with and without LI. As expected, parents of children with LI demonstrated lower levels of literacy beliefs and practices, on average, than parents of children without LI. However, parent literacy beliefs and practices, including shared reading, were uncorrelated with children's emergent literacy performance within LI. Skibbe et al. (2008) concluded that genetic influences related to LI (e.g., DeThorne, Petrill, Hayiou-Thomas, & Plomin, 2005; Simmons et al., 2010) and/or the lack of requisite language skills (e.g., Justice, Sofka, & McGinty, 2007) attenuated the impact of the environment on emergent literacy outcomes within children with LI.

However, the Skibbe et al. (2008) results were based on correlations collapsed across the substantial range of environment and reading outcomes (even within children with LI). By doing so, this approach assumed that the relationship between environment and reading outcomes is consistent within LI. However, other literatures have suggested that the correlation between environment and outcome is conditional across levels of the environment and across skill level. For example, the behavioral genetic literature has suggested that environmental influences are most salient in children from lower-income environments for reading (Pennington et al., 2009; Rosenberg, Pennington, Willcutt, & Olson, 2012; Taylor & Schatschneider, 2010) and general cognitive ability (Rowe, Jacobson, & Van den Oord, 1999; Turkheimer & Waldron, 2000). Other studies have suggested that home literacy practices were associated with child-reading outcomes, but only in mothers who were lower-level readers (Johnson, Martin, Brooks-Gunn, & Petrill, 2008). Concerning the relationship between shared reading and reading outcomes within LI children, there may be a threshold of lower-frequency shared reading that yields or reflects poor emergent literacy skills. In this case, we would expect the correlation between shared reading and emergent literacy to be highest in lower amounts of shared reading. There may also be a threshold of higher-frequency shared reading that mitigates the impact of LI on early literacy outcomes. In this case, the correlation would be highest at higher levels of shared reading.

In addition to the potential moderating effects of the environment, the relationship between measures of the environment and reading skills may also be conditional on the type and level of reading skill being examined. As described in constrained skills theory (e.g., Paris, 2005; Paris & Wenshu Luo, 2010), reading-related skills may be divided into two types. First, unconstrained skills, such as vocabulary, are acquired over a lifetime, whereas constrained skills, such as letter naming, develop rapidly within a particular period of reading development. Thus, the correlation between measures of the environment and literacy skills may depend on whether the skill in question is constrained or unconstrained. In the case of constrained skills, the correlation between measure and environment may depend on whether the skill in question is in the process of being mastered. Considering the substantial heterogeneity of emergent literacy skills within children with LI, some children will have mastered constrained skills, whereas others will have not begun to master these skills. Other children will be in the process of learning these skills. Thus, we expect that the correlation between shared reading and emergent literacy may vary considerably as a function emergent literacy skills, particularly for those skills that are constrained. In particular, we expect that the correlation between shared reading and emergent literacy will be highest in constrained skills that are in the process of being mastered.

The purposes of this article are threefold. First, we examine the correlation between home literacy practices (focusing on the frequency of storybook reading) and emergent literacy skills in children at risk for LI. It is hypothesized that these relations will be weak, as found in Skibbe et al. (2008). Second, using quantile regression (described more fully in results), we examine whether the correlation between frequency of storybook reading and emergent literacy is conditional on the frequency of storybook reading. Given previous findings in typically developing children (e.g., Johnson et al., 2008; Petrill, Deater-Deckard, Schatschneider, & Davis, 2007; Taylor & Schatschneider, 2010; Turkheimer & Waldron, 2000), we hypothesize that environment–outcome relationships will be stronger at lower levels of the environment. In other words, below a certain threshold, poor literacy environments will have a negative impact on emergent literacy. We predict that variation in the literacy environment above this threshold will be uncorrelated with emergent literacy. Third, we examine whether the relationship between storybook reading and emergent literacy is conditional on emergent literacy skills. Given the findings of Paris (2005; Paris & Wenshu Luo, 2010), we hypothesize that these relationships will be strongest when examining constrained literacy skills (e.g., letter naming) that are in the process of being learned. In contrast, we hypothesize that the relationship between storybook reading and emergent literacy will be more uniform across the distribution of unconstrained literacy skills (e.g., vocabulary).

Method

Participants

This study is based on 212 children (72% male) drawn from the larger Sit Together and Read 2 (STAR 2) study. Briefly, STAR 2 is a pre–post randomized controlled trial examining the efficacy of a classroom-based book reading program to improve prereading skills in children with primary LI. Results for the current study are based on the first two cohorts of STAR 2 at the point of the fall preintervention assessment, prior to delivery of any intervention. Mean age for the children was 51.73 months (*SD* = 6.6, min = 36, max = 66) at the time of this fall assessment. The majority (71%) of parents reported that their children were "White/Caucasian," whereas 22% reported that their children were "Black/African American." The remaining children were multiethnic, Asian, Native American, or Filipino (6%). Maternal educational attainment ranged from "some high school but no diploma" (7%) to "doctoral degree" (6%). The modal maternal education attainment level was "some college but no degree" (29%). The median annual income was "\$50,001 to \$55,000" (range = "5,000 or less" to "85,000 or more").

All 212 children were enrolled in early child special education (ECSE) classrooms and were recruited for participation for this study via a questionnaire filled out by the teacher about each student's skills. Students who appeared to have specific language impairment (SLI; poor language skills in the absence of other cognitive delay or concurrent language-related diagnosis such as Down syndrome) were given priority for enrollment. Those with confirmed diagnosis of LI regardless of other diagnoses were given second priority, and those students who appeared to have LI (the teacher reported being concerned about the child's language skills) but were not diagnosed or being treated as such received third priority.

STAR 2 preintervention assessments indicated that children demonstrated lower than average standardized score (*M* = 75.59, *SD* = 17.95, min = 45, max = 106) on the Core Language Index of the *Clinical Evaluation of Language Fundamentals–Preschool–2* (CELF:P-2; Wiig, Semel, & Secord, 2004) as well as on the CELF:P-2 Expressive Index standardized score (*M* = 75.79, *SD* = 17.07, min = 45, max = 111) and CELF:P-2 Receptive Index standardized score (*M* = 77.26, *SD* = 16.33, min = 45, max = 109). Of the children, 15% scored below the 10th percentile on the CELF:P-2 Expressive Index, 9% scored below the 10th percentile on the CELF:P-2 Receptive Index, 45% scored below the 10th percentile on the Expressive and Receptive Indices, and 30% scored at or above the 10th percentile on the Expressive and Receptive Indices. Children were somewhat closer to average performance on the *Kaufman Brief Intelligence Test* (Kaufman & Kaufman, 1990), ranging from 53 to 124 (*M* = 85, *SD* = 18, *n* = 146).

Caregivers reported that 92% of children had individualized education programs (IEPs) and 91% were receiving language intervention services from a speech-language pathologist. The remaining children, although not receiving language intervention, were described by their teachers as exhibiting language skills uncharacteristic for their age, but they had not yet received an IEP. Caregivers also reported that 21% of these children (*n* = 46) also possessed other co-occurring disabilities, including autism (*n* = 24), cerebral palsy (*n* = 8), Down syndrome (*n* = 6), Stickler syndrome (*n* = 4), Tourette syndrome (*n* = 2), and apraxia (*n* = 2).

Procedure and Measures

The current study is based on four direct measures of child emergent literacy skills as well as one caregiver questionnaire examining frequency of storybook reading. Study examiners administered direct measures of emergent literacy skills within a 6-week window at the children's schools in the fall of the year. These measures were selected to assess constrained emergent literacy skills (print knowledge, lowercase letter naming, and uppercase letter naming) as well as a measure tapping an unconstrained skill associated with later reading (expressive vocabulary). Print knowledge was assessed using the *Preschool Word and Print Awareness Assessment* (PWPA; Justice & Ezell, 2001), which examines children's knowledge across 14 print concepts (e.g., book orientation, print directionality, print function, letters, words) as children are read a uniform picture book. A total of 17 points are possible on the PWPA. Item response theory used to validate the differential item functioning of the PWPA showed the tool to provide a valid representation of the construct it measures (see Justice, Bowles, & Skibbe, 2006). Interrater reliability was reported at .99 (Pearson product–moment correlation coefficient); internal consistency was reported at .84 (Cronbach's alpha) and .87 (Guttman split–half). Lowercase letter naming and uppercase letter naming were assessed via the *Phonological Awareness Literacy Screening for Preschool* (PALS–Pre-K; Invernizzi et al., 2004). A total of 26 points are possible on the Lower- and Upper-Case Alphabet Recognition tasks. Finally, we assessed expressive vocabulary using the *Test of Preschool Early Literacy* (TOPEL; Lonigan, Wagner, Torgesen, Rashotte, & PRO-ED, 2007), an age-normed measure with a mean of 100 and a standard deviation of 15.

Caregiver questionnaires were collected during one-on-one meetings with research staff also conducted in the fall of the year. As part of this assessment, caregivers were administered items adapted from Bennett et al. (2002),

Table 1. Descriptive Statistics.

Variable	*M*	*SD*	Min	Max	*n*
Frequency of storybook reading	14.05	6.46	1	24	204
Print knowledge	4.85	3.76	0	16	210
Lowercase letter naming	5.02	7.81	0	26	212
Uppercase letter naming	8.17	9.91	0	26	212
Expressive vocabulary	82.00	16.46	54	115	210

Table 2. Correlations Among Frequency of Storybook Reading, Print Knowledge, Lowercase Letter Naming, Uppercase Letter Naming, and Expressive Vocabulary.

Variable	1	2	3	4	5
1. Storybook reading	1.00				
2. Print knowledge	.25*	1.00			
3. Lowercase letter	.24*	.40*	1.00		
4. Uppercase letter	.20*	.43*	.84**	1.00	
5. Expressive vocabulary	.11	.59*	.36**	.38**	1.00

$^*p < .05$. $^{**}p < .01$.

assessing the frequency of literacy activities, either initiated by child or parent, assessed via 11 items. These items were scaled from 0 to 8 (i.e., "0 times" up to "8 times or more" per designated time period). Bennett et al. (2002) reported an alpha of .73, which has been supported by subsequent work in other samples (e.g., Skibbe et al., 2008).

Using exploratory factor analysis, we sought to derive select items related to frequency of storybook reading. One of the 11 items was dropped prior to factor analysis (concerning frequency with which children watched movies and videos) because it did not correlate significantly with the other 10 items. Analyses on the 10 remaining items suggested a three-factor solution, accounting for 60% of the variance. Frequency of storybook reading was measured by the three items that loaded on the first factor (Eigenvalue = 2.54): (a) "How many times did you (or another family member) read to your child last week?" (loading = .84), (b) "How many times did your child ask to be read to last week" (loading = .81), and (c) "How many times did your child look at books on his/her own last week" (loading = .67). As expected from these high factor loadings, Cronbach's alpha based on these three items was also high ($\alpha = .79$). The other two factors (Caregiver Literacy Teaching [Eigenvalue = 1.98] and Children's Print Interest [Eigenvalue = 1.63]) were not analyzed in the current study.

Results

Descriptive Statistics and Pearson Correlation

Descriptive statistics are presented in Table 1. Scores for the frequency of storybook reading averaged near the middle of the scale but ranged widely. Scores for print knowledge, lowercase letter naming, and uppercase naming were well below the midpoint of the scales for each measure, but also ranged across all possible responses. Mean expressive vocabulary ($M = 82.00$) was nearly 1 standard deviation below the unselected population mean of 100 but ranged widely ($SD = 16.46$, min = 54, max = 115).

Pearson correlations, presented in Table 2, suggest that frequency of storybook reading was moderately correlated with print knowledge ($r = .25$, $p < .05$), lowercase letter naming ($r = .24$, $p < .05$), and uppercase letter naming ($r = .20$, $p < .05$). In contrast, storybook reading was not significantly correlated with expressive vocabulary ($r = .11$, $p > .05$). Intercorrelations among preliteracy skills were moderate to large, ranging from $r = .36$ ($p < .01$) between lowercase letter naming and expressive vocabulary to $r = .84$ ($p < .01$) between lowercase and uppercase letter naming.

Quantile Regression

The primary goal of this study was to examine whether the relations between frequency of caregiver storybook reading and preliteracy skills was stronger or weaker depending on the amount of storybook reading that occurred and/or the level of child emergent literacy skills. These questions were addressed using quantile regression.

Quantile regression was developed by Koenker and Basset (1978) and expands on ordinary least squares (OLS) regression. Whereas OLS regression examines the relation of an independent variable with a dependent variable, quantile regression examines how the relation of an independent variable with a dependent variable changes based on the quantile (percentile) of the dependent variable. In this way, quantile regression is able to examine the relations between two variables at multiple different quantiles (percentiles) across a continuous distribution. Quantile regression is not equivalent to dividing a sample into multiple subgroups based on percentiles of the dependent variable and subsequently fitting an OLS regression to each subgroup. Instead, quantile regression uses all available data, estimating the relations between the two variables at multiple points across the distribution of scores through asymmetric weighting of the values across the distribution using bootstrapping, data resampling, and statistical inference. A full discussion of the technique is beyond the scope of the present article, but several excellent resources can inform the reader of the details of the technique (see Firpo, 2007; Koenker, 2005; Koenker & Hallock, 2001; Koenker & Machado, 1999).

Association between storybook reading and emergent literacy skills, conditional on frequency of storybook reading. First, we examined the relationship between frequency of storybook reading and emergent literacy skills (print knowledge,

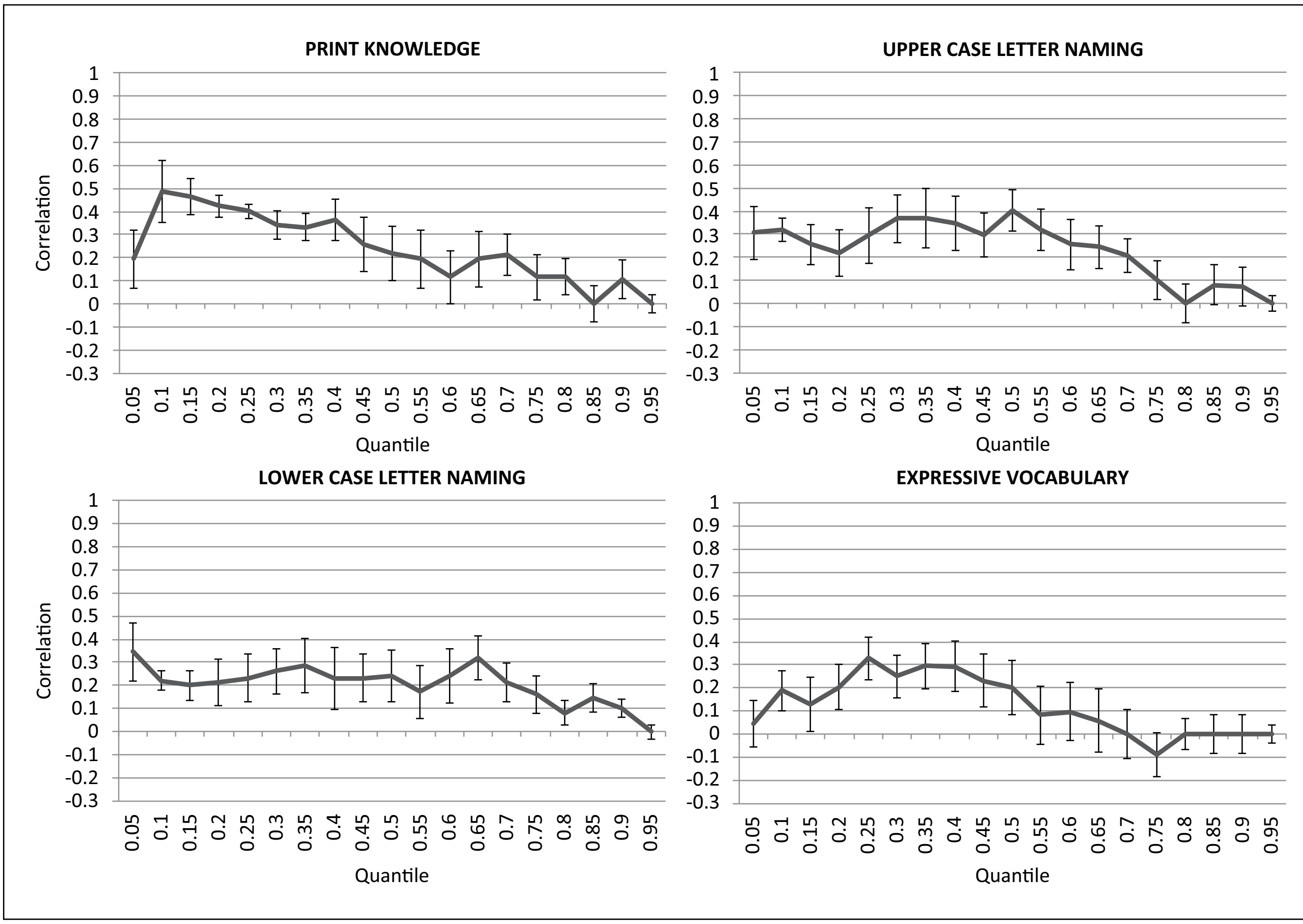

Figure 1. Correlation between frequency of storybook reading and preliteracy skills conditional on frequency of storybook reading.

lowercase letter naming, uppercase letter naming, and vocabulary), depending on the frequency of storybook reading that occurred (see Figure 1). The x-axes refer to the frequency of storybook reading (presented from the 5th to the 95th percentiles) and the y-axes to the correlation between frequency of storybook reading and each of the four direct measures (print knowledge, lowercase letter naming, uppercase letter naming, and vocabulary). The error bars around the estimates refer to the 90% confidence interval around each correlation at each quantile. Correlations with confidence intervals that do not cross zero on their respective y-axes are statistically significant.

As shown in Figure 1, correlations between frequency of storybook reading and emergent literacy skills were more likely to be statistically significant at lower amounts of storybook reading. This pattern was strongest for print knowledge, where the correlation was approximately $r = .50$ ($p < .05$ as defined by the confidence intervals) at the 10th percentile of storybook reading, but it approached zero by the 80th percentile of storybook reading. A similar, but attenuated, pattern of results was found for lowercase letter naming, with the correlation ranging from $r = .30$ ($p < .05$) to 0, as well as uppercase letter naming ($r = .40$, $p < .05$, to 0) and expressive vocabulary ($r = .30$, $p < .05$, to 0). Specific values for the estimates represented in Figure 1 are presented in Table 3, which includes the values for the estimates at the 20th, 50th, and 80th percentiles for each outcome. Note that correlations were not significantly different from zero at the 80th percentile of frequency of storybook reading in all instances.

Though these descriptive patterns are interesting, we were also able to statistically compare the differences between the estimates at selected quantiles by adapting a method developed by Koenker (2006). This method allows for comparisons within one set of relations. Specifically, the correlation coefficient between X and Y at one given quantile can be compared to the correlation coefficient between X and Y at a different given quantile. Prior to analysis, three quantiles were selected for comparison between one another at the low end (20th percentile), middle (50th percentile), and high end (80th percentile) of the distribution of Y (following the example of Logan et al., 2012). Comparisons between percentiles were conducted in the R statistical software package (R Development Core Team, 2011).

Table 3. Results of Quantile Regression of Frequency of Storybook Reading Predicting Preliteracy Skills (conditional on storybook reading).

Quantile	Correlation	*SE*	LB	UB	*t*	*p*
Print knowledge						
QR-20	.42	0.05	0.33	0.51	9.18	< .0001
QR-50	.22	0.12	−0.01	0.45	1.85	.07
QR-80	.12	0.08	−0.04	0.27	1.50	.13
Lowercase letter naming						
QR-20	.21	0.10	0.01	0.41	2.10	.04
QR-50	.24	0.11	0.03	0.46	2.20	.03
QR-80	.08	0.05	−0.03	0.19	1.49	.14
Uppercase letter naming						
QR-20	.22	0.10	0.02	0.42	2.19	.03
QR-50	.40	0.09	0.22	0.58	4.44	< .0001
QR-80	.00	0.08	−0.17	0.17	0.00	1.00
Expressive vocabulary						
QR-20	.20	0.10	0.01	0.40	2.05	.04
QR-50	.20	0.12	−0.03	0.43	1.75	.08
QR-80	.00	0.07	−0.14	0.14	0.00	1.00

Note: LB = lower bound; QR-20 = quantile regression at the 20th quantile, etc.; UB = upper bound. This table corresponds to Figure 1. Statistical tests comparing differences across quantiles are presented in Table 4.

Table 4. Significance Tests for Differences Between Quantiles Conditional on Frequency of Storybook Reading.

Quantile Comparison	*F*	*p*
Print knowledge		
20 vs. 50	10.19	.001
50 vs. 80	11.24	.001
20 vs. 80	25.68	< .001
Lowercase letter naming		
20 vs. 50	0.07	.78
50 vs. 80	2.76	.09
20 vs. 80	1.60	.20
Uppercase letter naming		
20 vs. 50	3.75	.05
50 vs. 80	21.01	< .001
20 vs. 80	3.80	.05
Expressive vocabulary		
20 vs. 50	N/A	N/A
50 vs. 80	3.95	.05
20 vs. 80	3.67	.05

Note: N/A = comparison not analyzed because the difference in correlation between quantiles was zero.

The results of the statistical comparisons between these selected percentiles are presented in Table 4. For print knowledge, the general shape of the distribution suggests that the correlation between frequency of storybook reading and print knowledge declined as frequency of storybook reading increased. Tables 3 and 4 demonstrate that the estimated relation decreased from 0.42 at the 20th percentile to 0.22 at the 50th percentile, to 0.12 at the 80th percentile. The significance tests confirmed that these comparisons were all significant, all Fs > 10, ps ≤ .001 (see Table 4). In the case of lowercase letter naming, no significant differences between the selected percentiles were observed. For uppercase letter naming, the decrease in the relation between the 50th and 80th percentile was found to be statistically significant, $F = 21.01$, $p < .001$ (see Table 4). For expressive vocabulary, as noted in Table 3, the correlation decreased from $r = .20$ to $r = .00$ across the range of frequency of storybook reading. However, the magnitude of this decrease was not statistically significant ($p = .05$; see Table 4).

To provide additional visual representation of the conditional relations between print knowledge and storybook reading, Figure 2 presents the scatterplot for print knowledge (x-axis) and frequency of storybook reading (y-axis). Regression lines were fit at the 10th, 20th, 50th, 80th, and 90th percentiles of frequency of storybook reading. Consistent with the results in Tables 3 and 4, lower frequency of storybook reading was associated with lower levels of print knowledge (e.g., regression line plotted at the 20th percentile), but high frequency of storybook reading was neither necessary nor sufficient for high levels of print knowledge (e.g., regression line plotted at the 80th percentile).

Association between storybook reading and emergent literacy skills, conditional on levels of emergent literacy skills. We also examined whether the relationship between frequency of storybook reading and emergent literacy skills was stronger or weaker depending on the level of emergent constrained versus unconstrained literacy skills (see Figure 3). Similar to the previous analyses, the quantiles were compared at each of three selected quantiles. The point estimates of each of the three selected quantiles are presented in Table 5 and the significance tests for comparisons between each of these selected quantiles is presented in Table 6.

The correlation between frequency of storybook reading and print knowledge was at or close to $r = .00$ at lower levels of print knowledge, but more likely to be statistically significant at higher levels of print knowledge. The increase was gradual enough that the differences between the 20th and 50th quantiles, $F = 3.14$, $p = .07$ (see Table 6), and the 50th and 80th quantiles, $F = 0.95$, $p = .33$ (see Table 6), were nonsignificant. Only the difference between the 20th and 80th quantiles was large enough to be considered a statistically significant difference, $F = 15.45$, $p < .001$ (see Table 6). These results suggest that the correlation between frequency of storybook reading and print knowledge increased as levels of print knowledge increased.

Similarly, the correlation of lowercase letter naming and uppercase letter naming with frequency of storybook reading (Figure 3) approached zero at and below the 30th percentile of letter naming, but at higher levels of skill it approached $r = .70$ ($p < .05$) for lowercase letter naming and $r = .60$ ($p < .05$)

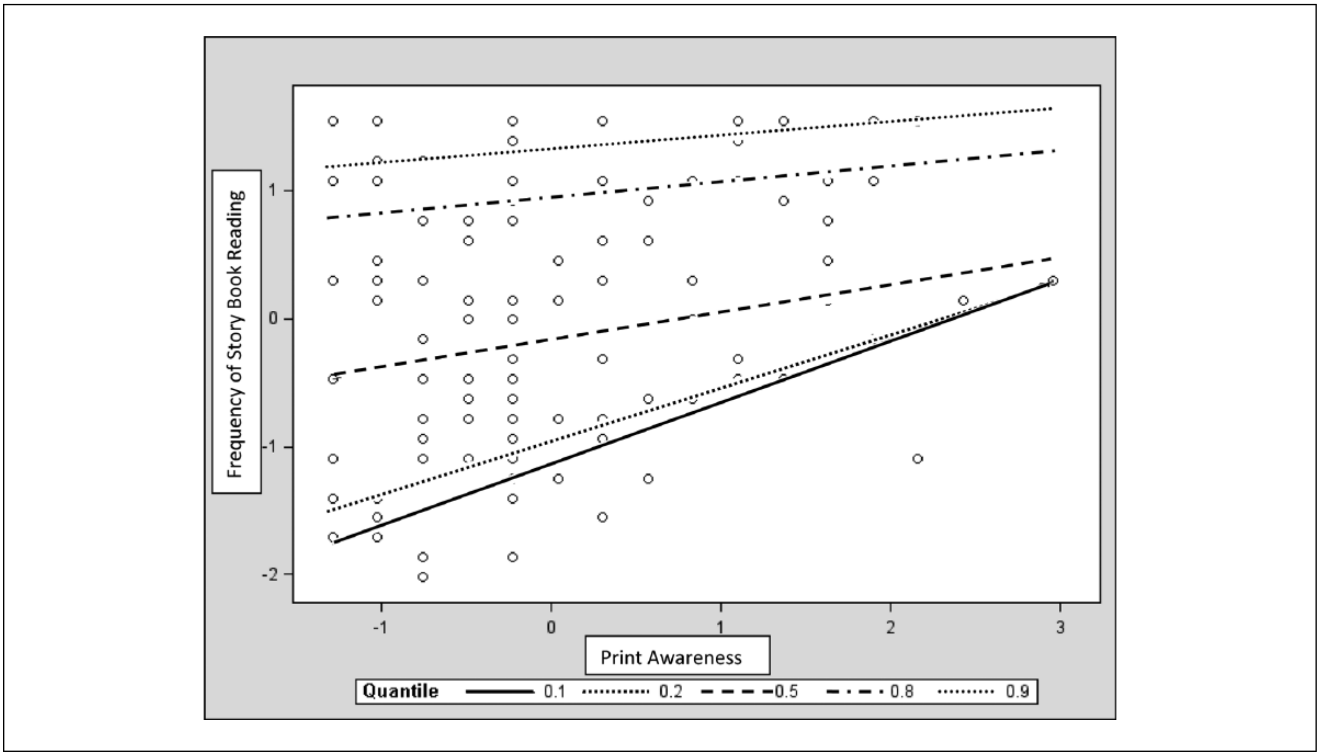

Figure 2. Scatterplot of print knowledge by frequency of storybook reading at the 10th, 20th, 50th, 80th, and 90th percentiles of storybook reading.

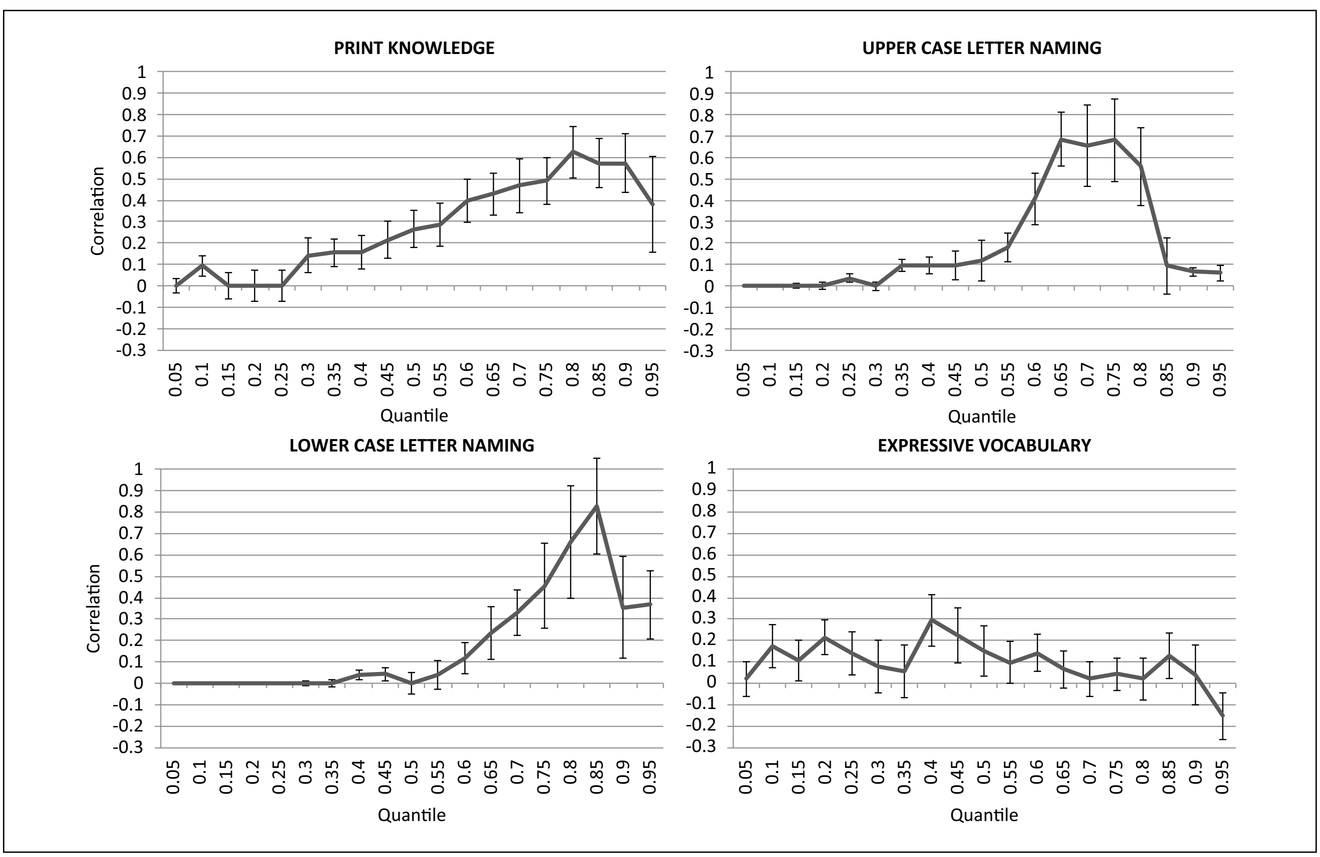

Figure 3. Correlation between frequency of storybook reading and preliteracy skills, conditional on preliteracy skill.

Table 5. Results of Quantile Regression of Frequency of Storybook Reading Predicting Preliteracy Skills (conditional on emergent literacy skill).

Quantile	Correlation	*SE*	LB	UB	*t*	*p*
Print knowledge						
QR-20	.00	0.07	−0.15	0.15	0.00	1.00
QR-50	.26	0.09	0.09	0.44	3.03	.00
QR-80	.62	0.12	0.39	0.86	5.31	< .0001
Lowercase letter naming						
QR-20	.00	—	—	—	—	—
QR-50	.00	0.05	−0.10	0.10	0.00	1.00
QR-80	.66	0.26	0.14	1.18	2.52	.01
Uppercase letter naming						
QR-20	.00	0.01	−0.03	0.03	0.00	1.00
QR-50	.12	0.10	−0.07	0.31	1.23	.22
QR-80	.56	0.18	0.20	0.92	3.08	.00
Expressive vocabulary						
QR-20	.21	0.08	0.05	0.37	2.65	.01
QR-50	.15	0.12	−0.08	0.38	1.31	.19
QR-80	.02	0.10	−0.17	0.22	0.21	.83

Note: LB = lower bound; QR-20 = quantile regression at the 20th quantile, etc.; UB = upper bound. Cells with dash = could not be estimated. This table corresponds to Figure 3. Statistical tests comparing differences across quantiles are presented in Table 6.

Table 6. Significance Tests for Differences Between Quantiles, Conditional on Preliteracy Skills.

Quantile Comparison	*F*	*p*
Print knowledge		
20 vs. 50	3.14	.07
50 vs. 80	0.95	.33
20 vs. 80	15.45	< .001
Lowercase letter naming		
20 vs. 50	N/A	N/A
50 vs. 80	7.13	.007
20 vs. 80	6.30	.01
Uppercase letter naming		
20 vs. 50	1.67	.19
50 vs. 80	5.57	.018
20 vs. 80	9.67	.002
Expressive vocabulary		
20 vs. 50	0.37	.54
50 vs. 80	1.44	.23
20 vs. 80	3.01	.08

Note: N/A = comparison not analyzed because the difference in correlation between quantiles was zero.

for uppercase letter naming. The statistical comparisons were also significant when comparing the 50th and the 80th quantiles, $F = 7.13$, $p = .007$ (see Table 6), and when comparing the 20th to the 80th quantiles, $F = 6.30$, $p = .01$ (see Table 6) for lowercase letter naming. Similar results were obtained for uppercase letter naming, $F = 5.57$, $p = .018$ and $F = 9.67$, $p = .002$ (see Table 6). Finally, in the case of expressive vocabulary, the results in Figure 3 and Table 5 suggest that the relations were significant at the 20th percentile of expressive vocabulary (e.g., $r = .21$, $p = .00$; see Table 5) but not statistically significant at the 50th ($r = .15$, $p = .19$; see Table 5) or 80th percentiles ($r = .02$, $p = .83$; see Table 5). However, as shown in Table 6, the change in the correlation between expressive vocabulary and storybook reading across levels of expressive vocabulary was not statistically significant; the largest difference was between the 20th and 80th percentiles, $F = 3.01$, $p = .08$.

Figure 4 provides further visual representation of these results, which presents scatterplot results for uppercase letter naming with regression lines representing the 10th, 20th, 50th, 80th, and 90th percentiles for letter naming. Very low levels (10th, 20th percentiles) and very high levels (90th percentile) of letter naming were uncorrelated with frequency of storybook reading (represented by flat lines). However, there was a strong association between letter naming and home literacy environment at the moderate to high quantiles of letter naming, presumably when children are actively learning their letters.

Discussion

As expected, children who are at risk for developing LI come to school with lower than average levels of vocabulary knowledge. Also as expected, the current study suggests nonsignificant to modest correlations ($r = .11$ to $r = .25$) between frequency of storybook reading and emergent literacy skills when looking across the wide range of reported storybook reading and emergent literacy skills. Descriptively, the effect sizes in the current study were somewhat larger than those found in Skibbe et al.'s (2008) study of language-impaired children with LI, but smaller than those in studies examining children without LI (e.g., Bus et al., 1995; Scarborough & Dobrich, 1994; Sénéchal, 2006). It is notable that the Pearson correlation of shared book reading and expressive vocabulary was not significant. However, as described in the results, the association was statistically significant at low levels of storybook reading and at low levels of expressive vocabulary.

We employed quantile regression to consider the association between frequency of storybook reading and emergent literacy across different points in the distribution of storybook reading. Results suggested that the correlation between storybook reading and preliteracy skills was highest at a lower frequency of storybook reading, particularly for print knowledge, which approached $r = .50$. Low frequency of storybook reading was generally associated with low levels of print knowledge. However, for some children with LI, the presence of a high frequency of storybook reading was not sufficient for success in print knowledge skills. A similar, but attenuated, pattern of results was found for

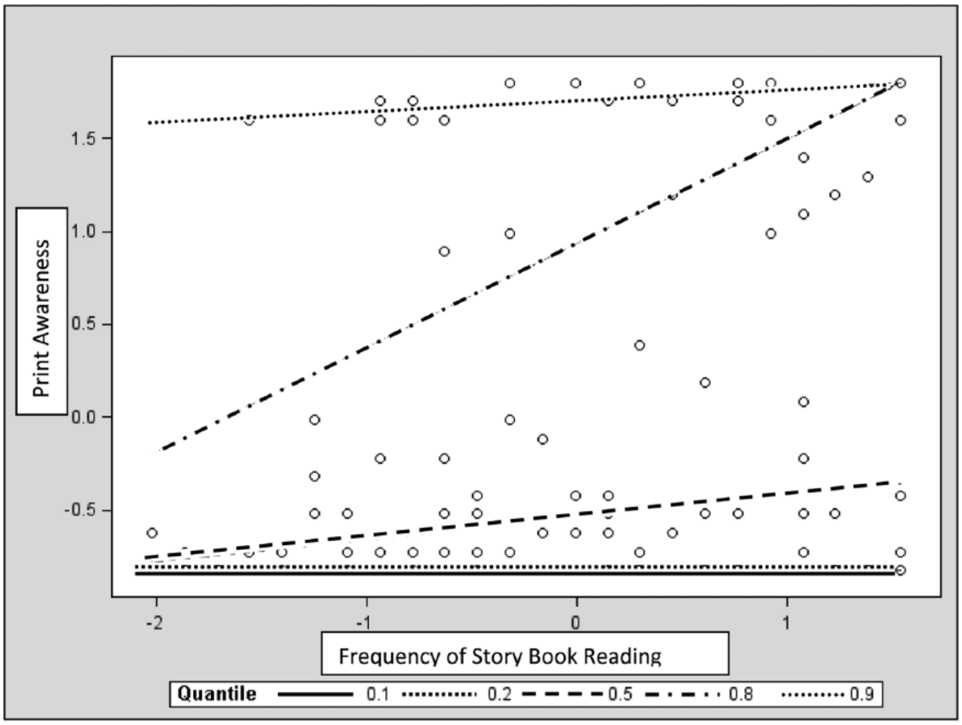

Figure 4. Scatterplot of uppercase letter naming by frequency of storybook reading at the 10th, 20th, 50th, 80th, and 90th percentiles of uppercase letter naming.

letter naming. Finally, a moderate association was found between frequency of storybook reading and expressive vocabulary at lower frequency of storybook reading.

Quantile regression also examined whether the correlation between frequency of storybook reading and emergent literacy varied as a function of emergent literacy. Results suggested that associations were highest at higher levels of constrained emergent literacy skills, particularly uppercase letter naming and lowercase letter naming. The magnitude of these associations also changed most dramatically across levels of constrained as compared to unconstrained skills. Correlations were zero at low levels of letter naming, large in magnitude in the middle of the distribution of letter naming (approaching r = .80), then attenuated to zero at every high levels of letter naming, once children attained mastery of letter knowledge. These results were reflected in the secondary analysis presented in Figure 4. The lowest and highest quantiles of letter naming showed no correlation with frequency of storybook reading.

These findings are relevant for the debate concerning Paris's (2005) constrained skills theory (CST). CST makes two assertions: that nearly all readers learn to master some skills completely (e.g., letter knowledge) and that constrained skills, although important gateways to reading, are poor indices of later reading success. Our findings are consistent with CST in that the association between the environment and letter knowledge was zero when letter knowledge was at the floor, as well as when the skill had been mastered. However, we also assert that mastery (or lack of mastery) of these skills, in the context of the presence or absence of storybook reading, may be an important time-limited marker of performance in children with LI.

Study Limitations

When considering these results, several limitations must be noted. First, given the scope of the statistical analyses, we elected to focus on the specific relationship between frequency of storybook reading and emergent literacy using a measure of storybook reading that was limited to a single questionnaire. It is therefore possible that results may vary when examining different aspects of the home environment, such as parent literacy practices and attitudes, as well as in other measures of storybook reading. Moreover, the results of this study were based on a select group of children (who are at risk for LI) at a particular developmental window (fall of prekindergarten). Thus, the results of this study should be interpreted as a snapshot of an important moment in an ongoing developmental process involving growth of constrained and unconstrained language and literacy skills in the context of home and school environments, set against the backdrop that all children were at risk for developing language difficulties. The results also assume that the reliability of measurement of frequency of storybook reading and emergent literacy is consistent across the range of ability.

Finally, there is considerable heterogeneity in the composition of children at risk for LI in ECSE classrooms. This was evident in the multilevel selection strategy we employed to identify children for the STAR 2 study. This obfuscates the causal pathways between storybook reading and emergent literacy. For example, our data are consistent with the hypothesis that lack of shared book reading puts children at a disadvantage for emergent literacy development compared to their peers. Our data are also consistent with the hypothesis that the relationship between shared book reading and emergent literacy is conditional on levels of emergent literacy, particularly for constrained skills. However, we cannot rule out that the results of the study are the result of differential characteristics related to selection into ECSE classrooms, profiles of child performance, or profiles of parent performances. We cannot empirically address these possibilities, but we intend to do so in future studies.

Implications

Despite these limitations, study results have important educational implications. Our study suggests that the correlation between storybook reading and emergent literacy is particularly strong at low levels of storybook reading. However, our results also suggest that children with LI come from homes displaying a wide range of storybook reading. In other words, not all children with LI come from homes with impoverished reading environments. We argue that systematically examining the conditions under which storybook reading, and other aspects of the home literacy environment, correlate and do not correlate with emergent literacy may constitute an important untapped marker for understanding the nature, causal mechanisms, and development of emergent literacy problems in children with LI. In particular, LI and related emergent literacy deficits may be a function of familial risk, whereas in other cases LI and related emergent literacy may be emerging despite a supportive home environment.

This possibility has important implications for intervention. Children with high storybook reading in the home and low emergent literacy may be less likely to benefit from additional storybook-based instruction, relative to children who come from homes with lower levels of storybook reading. In addition, the association between storybook based interventions may be heavily time and/or skill dependent for constrained skills such as letter naming. Put another way, parents are correctly thought of as untapped resources for improving outcomes in children at risk for LI (Reese, Sparks, & Leyva, 2010). However, implicit in this argument is that children with LI come from homes where more could be done to improve the literacy environment. Our results suggest that this is not always the case: Some children with LI come from impoverished home literacy environments that are correlated with their emergent literacy, whereas other children with LI come from rich home literacy environments. As a result, rather than merely statistically or experimentally controlling for person- and/or home-based individual differences, including home- and child-based markers may elucidate the conditions under which home-based interventions may be more or less effective.

Acknowledgments

We would like to thank all of the teachers, families, and children who participated in this study as well as Joan Kaderavek, Jaclyn Dynia, Rashaun Geter, Anita Hunt, Kim Thompson, Charla Forsty, Jason Flowers, and many additional members of our research team for their assistance with this project.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: We gratefully acknowledge the support provided by Grant R324A080037 from the National Center for Special Education Research, Institute of Education Sciences.

References

Bennett, K. K., Weigel, D. J., & Martin, S. S. (2002). Children's acquisition of early literacy skills: Examining family contributions. *Early Childhood Research Quarterly*, *17*(3), 295–317.

Bishop, D. V., & Adams, C. (1990). A prospective study of the relationship between specific language impairment, phonological disorders and reading retardation. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *31*(7), 1027–1050.

Boudreau, D. (2005). Use of a parent questionnaire in emergent and early literacy assessment of preschool children. *Language, Speech, and Hearing Services in Schools*, *36*(1), 33–47.

Bus, A. G., van IJzendoorn, M. H., & Pellegrini, A. D. (1995). Joint book reading makes for success in learning to read: A meta-analysis on intergenerational transmission of literacy. *Review of Educational Research*, *65*(1), 1–21.

Catts, H. W. (2003). Language basis of reading disabilities and implications for early identification and remediation. *Reading Psychology*, *24*(3), 223–246.

Catts, H. W., Fey, M. E., Tomblin, J. B., & Zhang, X. (2002). A longitudinal investigation of reading outcomes in children with language impairments. *Journal of Speech, Language, and Hearing Research*, *45*, 1142–1157.

Dale, P. S., Price, T. S., Bishop, D. V. M., & Plomin, R. (2003). Outcomes of early language delay: I. Predicting persistent and transient language difficulties at 3 and 4 years. *Journal of Speech, Language, and Hearing Research*, *46*(3), 544–560.

DeThorne, L. S., Petrill, S. A., Hayiou-Thomas, M. E., & Plomin, R. (2005). Low expressive vocabulary: Higher heritability as a function of more severe cases. *Journal of Speech, Language, and Hearing Research*, *48*(4), 792–804.

Evans, M. A., Shaw, D., & Bell, M. (2000). Home literacy activities and their influence on early literacy skills. *Canadian Journal of Experimental Psychology/Revue Canadienne De Psychologie Expérimentale*, *54*(2), 65–75.

Firpo, S. (2007). Efficient semiparametric estimation of quantile treatment effects. *Econometrica*, *75*(1), 259–276.

Invernizzi, M., Sullivan, A., Meier, J., & Swank, L., Virginia Department of Education, & Curry School of Education. (2004). *PALS PreK assessment materials*. Richmond, VA: Virginia State Department of Education, University of Virginia, Curry School of Education.

Johnson, A. D., Martin, A., Brooks-Gunn, J., & Petrill, S. A. (2008). Order in the house! Associations among household chaos, the home literacy environment, maternal reading ability, and children's early reading. *Merrill Palmer Quarterly*, *54*(4), 445–472.

Justice, L. M., Bowles, R. P., & Skibbe, L. E. (2006). Measuring preschool attainment of print-concept knowledge: A study of typical and at-risk 3- to 5-year-old children using item response theory. *Language, Speech, and Hearing Services in Schools*, *37*(3), 224–235.

Justice, L. M., & Ezell, H. (2001). Written language awareness in preschool children from low-income households. *Communication Disorders Quarterly*, *22*(3), 123–134.

Justice, L. M., Sofka, A. E., & McGinty, A. (2007). Targets, techniques, and treatment contexts in emergent literacy intervention. Seminars in Speech and Language, 28(1), 14–24.

Kaufman, A. S., & Kaufman, N. L. (1990). *K-BIT: Kaufman Brief Intelligence Test: Manual*. Circle Pines, MN: American Guidance Service.

Koenker, R. (2005). *Quantile regression*. New York, NY: Cambridge University Press.

Koenker, R. (2006). *Quantreg: Quantile regression R Package Version 3.90*. Retrieved from http://www.cran.r-project.org

Koenker, R., & Bassett, G. (1978). Regression quantiles. *Econometrica*, *46*(1), 33–50.

Koenker, R., & Hallock, K. F. (2001). Quantile regression. *Journal of Economic Perspectives*, *15*(4), 143–156.

Koenker, R., & Machado, J. A. F. (1999). Goodness of fit and related inference processes for quantile regression. *Journal of the American Statistical Association*, *94*(448), 1296–1310.

Leseman, P. P. M., & de Jong, P. F. (1998). Home literacy: Opportunity, instruction, cooperation and social-emotional quality predicting early reading achievement. *Reading Research Quarterly*, *33*(3), 294–318.

Logan, J. A. R., Petrill, S., Hart, S. A., Thompson, L. A., Deater-Deckard, K., DeThorne, L. S., & Schatschneider, C. (2012). Heritability across the distribution: An application of quantile regression. *Behavior Genetics*, *42*(2), 256–267.

Lonigan, C. J., Wagner, R. K., Torgesen, J. K., & Rashotte, C. A., & PRO-ED (Firm). (2007). *TOPEL Test of Preschool Early Literacy*. Austin, TX: PRO-ED.

Maclean, M., Bryant, P., & Bradley, L. (1987). Rhymes, nursery rhymes, and reading in early childhood. *Merrill-Palmer Quarterly: Journal of Developmental Psychology*, *33*(3), 255–281.

Marvin, C. A., & Wright, D. (1997). Literacy socialization in the homes of preschool children. *Language, Speech, and Hearing Services in Schools*, *28*(2), 154–163.

Paris, S. G. (2005). Reinterpreting the development of reading skills. *Reading Research Quarterly*, *40*(2), 184–202.

Paris, S., & Wenshu Luo, S. (2010). Confounded statistical analyses hinder interpretation of the NELP report. *Educational Researcher*, *39*(4), 316–322.

Pennington, B. F., McGrath, L. M., Rosenberg, J., Barnard, H., Smith, S. D., Willcutt, E. G., & Olson, R. K. (2009). Gene x environment interactions in reading disability and attention-deficit/hyperactivity disorder. *Developmental Psychology*, *45*, 77–89.

Petrill, S. A., Deater-Deckard, K., Schatschneider, C., & Davis, C. (2007). Environmental influences on reading-related outcomes: An adoption study. *Infant and Child Development*, *16*(2), 171–191.

Petscher, Y., & Kim, Y. S. (2011). The utility and accuracy of oral reading fluency score types in predicting reading comprehension. *Journal of School Psychology*, *49*(1), 107–129.

R Development Core Team. (2011). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org

Reese, E., Sparks, A., & Leyva, D. (2010). A review of parent interventions for preschool children's language and emergent literacy. *Journal of Early Childhood Literacy*, *10*(1), 97–117. doi:10.1177/1468798409356987

Rosenberg, J., Pennington, B. F., Willcutt, E. G., & Olson, R. K. (2012). Gene by environment interactions influencing reading disability and the inattentive symptom dimension of attention deficit/hyperactivity disorder. *Journal of Child Psychology and Psychiatry*, *53*, 243–251.

Rowe, D. C., Jacobson, K. C., & Van den Oord, E. J. (1999). Genetic and environmental influences on vocabulary IQ: Parental education level as a moderator. *Child Development*, *70*(5), 1151–1162.

Scarborough, H. S., & Dobrich, W. (1994). On the efficacy of reading to preschoolers. *Developmental Review*, *14*(3), 245–302.

Sénéchal, M. (2006). Testing the home literacy model: Parent involvement in kindergarten is differentially related to grade 4 reading comprehension, fluency, spelling, and reading for pleasure. *Scientific Studies of Reading Scientific Studies of Reading*, *10*(1), 59–87.

Simmons, T. R., Flax, J. F., Azaro, M. A., Hayter, J. E., Justice, L. M., Petrill, S. A., & Bartlett, C. W. (2010). Increasing genotype-phenotype model determinism: Application to bivariate reading/language traits and epistatic interactions in language-impaired families. *Human Heredity*, *70*(4), 232–244.

Skibbe, L. E., Justice, L. M., Zucker, T. A., & McGinty, A. S. (2008). Relations among maternal literacy beliefs, home literacy practices, and the emergent literacy skills of preschoolers with specific language impairment. *Early Education and Development*, *19*(1), 68–88.

Snowling, M., Bishop, D. V. M., & Stothard, S. E. (2000). Is preschool language impairment a risk factor for dyslexia in adolescence? *Journal of Child Psychology and Psychiatry*, *41*(5), 587–600.

Taylor, J., & Schatschneider, C. (2010). Genetic influence on literacy constructs in kindergarten and first grade: Evidence from a diverse twin sample. *Behavior Genetics*, *40*, 591–602.

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research*, *40*(6), 1245–1260.

Turkheimer, E., & Waldron, M. (2000). Nonshared environment: A theoretical, methodological, and quantitative review. *Psychological Bulletin*, *126*(1), 78–108.

U.S. National Early Literacy Panel, U.S. National Institute for Literacy, & U.S. National Center for Family Literacy. (2008). *Developing early literacy: Report of the National Early Literacy Panel: A scientific synthesis of early literacy development and implications of intervention*. Washington, DC: National Institute for Literacy.

Whitehouse, A. J., Line, E. A., Watt, H. J., & Bishop, D. V. (2009). Qualitative aspects of developmental language impairment relate to language and literacy outcome in adulthood. *International Journal of Language & Communication Disorders/Royal College of Speech & Language Therapists*, *44*(4), 489–510.

Wiig, E. H., Semel, E. M., & Secord, W. (2004). *CELF-Preschool: Clinical Evaluation of Language Fundamentals–Preschool*. San Antonio, TX: Psychological Corp.