Published in final edited form as:

Topics Early Child Spec Educ. 2013 May 1; 33(1): 48–64. doi:10.1177/0271121412455438.

Is a Response to Intervention (RTI) Approach to Preschool Language and Early Literacy Instruction Needed?

Charles R. Greenwood, PhD¹, Judith J. Carta, PhD¹, Jane Atwater, PhD¹, Howard Goldstein, PhD², Ruth Kaminski, PhD³, and Scott McConnell, PhD⁴

Charles R. Greenwood: greenwood@ku.edu; Judith J. Carta: carta@ku.edu; Jane Atwater: janea@ku.edu; Howard Goldstein: hgoldstein@usf.edu; Ruth Kaminski: rkamin@uoregon.edu; Scott McConnell: smcconne@umn.edu

¹University of Kansas, Kansas City, KS, USA

²The Ohio State University, Columbus, OH, USA

³Dynamic Measurement Group, Eugene, OR, USA

⁴University of Minnesota, Minneapolis, MN, USA

Abstract

Preschool experience plays a role in children's development. However, for programs with language and early literacy goals, the question remains whether or not preschool instructional experiences are sufficiently effective to achieve these goals for all children. In a multisite study, we conducted a process-product description of preschool instruction and children's growth and outcomes in typical programs (i.e., Pre-Kindergarten, Title 1, Head Start, Tuition-Based) using a Response to Intervention (RTI) perspective. Results indicated that (a) students in their preschool year prior to kindergarten made small gains, but students starting the year in lower Tier 2 and 3 performance levels did not close initial skills gaps, (b) variations were noted by program types with varying socio-demographics and instructional processes, and (c) the quality of instruction (Tier 1) received by all was low with room for improvement. Implications for future research on the application of the RTI approach and potential benefits are discussed.

Preschool experience plays a role in children's development of language and early literacy skills, including children at-risk for underachievement (e.g., Campbell, Pungello, Miller-Johnson, Burchinal, & Ramey, 2001; Reynolds, Temple, Robertson, & Mann, 2003) and sustained later life benefits (e.g., Reynolds, Temple, Ou, Arteaga, & White, 2011). We also have gained knowledge of what language and early literacy skills should be taught in preschool (e.g., Shanahan & Lonigan, 2008). However, for programs with stated language and early literacy outcome goals, the question remains whether the implementation of preschool instruction is sufficiently effective to reach these goals for all children served in the program?

Prior research has examined the effects of preschool instructional resources such as the length of the school day, days per year, number of teachers per classroom, teacher/pupil ratio, and teacher quality in terms of training and experience and annual professional

development. Much of this research reports weak to no relation to children's learning outcomes (Early et al., 2006; Pianta et al., 2005). Another line of preschool experimental research has compared the efficacy of alternative curriculum approaches and instruction models with rather disappointing results (Preschool Curriculum Evaluation Research Consortium, 2008). Only 2 of the 14 intervention curricula investigated in PCERS showed differential effects on student-level outcomes for the pre-kindergarten year. At the level of policy and practice, yet another approach was the funding of the Early Reading First (ERF) projects that required that funded programs use curricula, instructional practices, and measurement tools supported by evidence of effectiveness (Russell et al., 2007) and provide professional development to teachers in the use of these practice. Results indicated that professional development led to significant effects on teachers' classroom practices and on children's print and letter knowledge outcomes. However, neither phonological awareness nor oral language improved.

Reports describing the search for active ingredients in instruction processes at the child-level such as specific teacher-student interactions occurring in the classroom and their relation to children's learning outcomes are emerging. For example, evidence supporting the effects of factors Classroom Organization, Emotional Support, Instructional Support, Teacher-Student Interaction, Teacher's Focus on Literacy, Student Engagement, and other classroom instructional factors (e.g., feedback that extends learning) is substantial (Chien et al., 2010; Justice, 2006; Justice, Hamre, & Pianta, 2008; Mashburn, 2008; Mashburn et al., 2008; McGinty, Breit-Smith, Justice, Kaderavek, & Fan, 2011; McGinty, Justice, Piasta, & Kaderavek, 2011). However, we are still pondering how to make individual preschool instruction programs more responsive and effective for all children served. Despite evidence of gains in instructional effects, both ERF and PCERS provide sobering lessons indicating that much of what is available as curricula used by teachers are not always effective and positive results often are difficult to replicate. We need a vastly improved understanding of the processes involved in producing better child outcomes, and theory and practice based on that knowledge.

Children as a group bring wide variability in what they know and can do to in preschool. Many are at risk for school readiness and learning to read in K-3. These children do not close gaps in language and early literacy with exposure to most preschool instruction, and preschool instruction is rarely differentiated based on knowledge of children's risk level. In early childhood education, we are only just learning that differences in children's response to instruction predicts why some children readily learn literacy skills in school while some struggle (Connor, Morrison, & Petrella, 2004b; Connor et al., 2009). This observation lies at the heart of the Response to Intervention (RTI) approach in early childhood (Buysse & Peisner-Feinberg, 2009; Greenwood, Bradfield, et al., 2011), and suggests how we may achieve greater effectiveness teaching all children. RTI holds the promise of preventing early delays from becoming disabilities later by intervening sooner to meet children's needs. RTI also could produce improved growth by all children in a preschool program.

Response to Intervention (RTI) is an approach to early identification and differentiating instruction for children who lack language and early literacy experiences and for whom the current preschool curriculum is not promoting their progress. Its major premise lies at the

heart of modern learning theory; that is, instructional experiences should be adjusted based on knowledge of a student's measurable success or failure learning what is being taught (Fuchs & Deno, 1991). This big idea is new to some sectors of early childhood education and is just emerging in early education research and practice (Greenwood, Bradfield, et al., 2011).

As an approach, RTI typically includes the following components: (a) universal screening used to identify those children in a program not keeping up with peers; (b) ongoing measurement of the progress of children over time, conducted more frequently for those for whom additional intervention is provided; (c) multiple tiers of support linked to a (d) decision-making model so that children identified with weak and very weak skills receive more intensive instructional support in a timely fashion.

Like RTI in K-5 programs (Berkeley, Bender, Peaster, & Saunders, 2009), 3 tiers of greater intensity are often applied in preschool applications. Tier 1 support represents use of a high quality core curriculum. Tier 2 instruction typically supplements the Tier 1 curriculum and is provided in small groups or learning centers. Tier 3 instruction is more individualized to a child's particular needs and uses explicit and systematic strategies to teach fundamental skills and concepts (Greenwood et al., 2011). RTI espouses using evidence-based strategies in each and every tier of support, with decisions regarding movement between tiers (change in instruction) to greater or lesser levels of support based on each child's performance based on locally collected formative data.

However, the RTI approach is not widely adopted in early childhood programs (Justice, 2006; VanDerHeyden & Snyder, 2006) and its benefits have yet to be demonstrated in convincing research reports. In this investigation, we examined the fit of an RTI measurement approach to describing children's performances in a range of prekindergarten programs employing curricula with language and early literacy goals over a year's experience. We identified children at 3 tiers of performance (Tiers 1, 2, and 3) in the Fall and followed their progress and experiences they received over time in programs that were not guided by RTI principles nor were they using multiple tiers of intervention support to differentiate instruction. Study results and implications were seen as an initial step towards informing further experimental evaluations of RTI approaches, including tier interventions, and full models of tiered supports in future preschool research.

Our investigation included children served in classrooms in multiple program types (Pre-K, Head Start, Title 1, and Private-Tuition) and who qualified in 3 performance-level groups based on the *Get Ready To Read* early literacy screener (Whitehurst & Lonigan, 2001b). In this context, we addressed the following questions related to differences between Tier groups and program types:

1. What were the proportions of children populating the 3 *Get Ready to Read (GRTR)* screener performance-level groups where: Tier 1 = 9 to 20 (average and above); Tier 2 = 6 to 8 (weak skills); Tier 3 = 0 to 5 (very weak skills)) in the Fall at the beginning of the year?

2. Were there corresponding differences in the groups' performance on standardized diagnostic measures of language and early literacy skills at the beginning of the year? Did a year's exposure to preschool instruction close gaps between the groups on these measures by year's end?
3. Did a year's exposure to preschool instruction close gaps between groups' rates of growth on formative measures of language (i.e., picture naming) and early literacy skills (i.e., sound identification) (McConnell & Missall, 2008) by year's end and how large was annual growth?
4. What was the quality of instruction children received in terms of classroom instructional support, curriculum-quality, teacher-literacy focus, children's literacy engagement in these programs with language and early literacy goals?

Method

Sample

To address the research questions, a purposive, multisite sample was recruited and enrolled. Several criteria were used to select the sample of prekindergarten programs, teachers, and children participating at four major research sites (Kansas City, MO/KS, Columbus, OH, Eugene-Springfield, OR, and Minneapolis, MN). The goal was to represent typically available preschool programs in these locations. We selected only interested programs that reported teaching language and literacy goals and who were using an early literacy core curriculum with an identifiable scope and sequence. We selected programs where children attended for at least 12 hours per week, and the majority of early literacy instruction occurred in English. In these classrooms, we focused on 4 and 5 year-old children in their year prior to kindergarten. We selected children in programs where the students served communicated primarily only in English or Spanish, and that included children with disabilities (i.e., with Individual Education Plans). We did not include classrooms that were entirely or a majority composition of children with disabilities.

This resulted in a total of 65 classrooms (12-19 per site) in 23 programs/districts in during the 2009 - 2010 school year. The 65 classrooms reflected four program types in these sites: 31% were state-funded Pre-K, 30% were Title 1, 24% were Head Start, and 14% were Tuition-Based. The program types were unbalanced by site. The Pre-K classrooms consisted of students from KS, OH, and MN sites. Title 1 classrooms were contributed by KS and OH, Head Start classrooms only from OR and Private Tuition classrooms only from MN. The majority of classrooms were half day (63%) versus full day (37%) and also varied by site. Private-Tuition and Title 1 classrooms were more likely to be full day programs, while Head Start and Pre-K were more likely to be only half day.

Children and Parents—All 4-5 year-old children in their year prior to kindergarten and their parents in these classrooms were recruited for participation. Those enrolled in the study completed informed consent. The mean age of the children was 4.6 ($SD = .32$) months at the first assessment. Eighty-one percent were 4-year olds, 19% were 5 years old.

Children's and parent's socio-demographics varied by sites. Children were balance by gender, 50-50%. Forty-one percent of children were male in OH ranging to a high of 59% in MN. The total sample was 36% African American, 31% White, 20% Hispanic/Latino, 10% multi-race, 3% Asian, and .5% other, including Native American and Pacific Islanders. In OH, 80% of participating children were African American, whereas in KS a similarly large 80% proportion of children consisted of a relatively balanced mixture of African American, White, and Hispanic/Latino. In OR and MN, approximately 80% of children were a mixture of just White and Hispanic/Latino. The very smallest proportions of students in each site were Multi-race, Asian/Asian American, or other.

With respect to the numbers of young children learning two languages, the sample mean was 23%. Across sites, the percentages were 43% in KS, 20% in OR, 13% in MN, and 9% in OH. The mean percentage of children eligible for early childhood special education (with IEPs) was 11%, ranging from 19% in KS, 14% in OR, 6% in MN, and 4% in OH.

The mean educational attainment of the parent/caregiver was 22% reporting less than high school, 23% high school or GED, and 55% reporting education beyond high school. The MN sample was by far the most educated parents with 84% reporting education beyond high school as compared to 45%, 49%, and 53% in OR, OH, and KS. OR, OH, and KS were relative similar in proportions of parents with less than, and completing high school varying from 19 to 32% overall. Similar values for MN were much lower at 9% less than high school, and 7% with high school only.

Teachers—All teachers in these classrooms were recruited for participation and those enrolled provided informed consent. These 65 teachers reported 9.9 mean years of teaching experience. The majority of teachers reported having a 4-yr degree (47.7%) in early childhood, 7.4% had a 2-yr degree. The proportion of teachers having a graduate degree was 38.2%; 18.5% of these were early childhood degrees. Only 2.4% had a Child Development Associates (CDA) degree and 4.3% had no degree. By program type, State-funded Pre-Kindergarten (66.3%) and Tuition-Based (58.3%) had the highest percentage of teachers with a 4-year degree. Tuition-Based programs also had the highest percentage of teachers with no degree (26.4%). Title 1 program teachers reported the highest percentage with graduate degrees (62.7%). Head Start programs reported the highest number of teachers with 2-year degrees (33.1%) and CDA degrees (7.2%).

Design and Procedures

A descriptive, process-product design that included multiple measures assessed on as many as 3 occasions was used to address the research questions. Measures sampled 3 constructs that were (a) quantity and quality of classroom instruction assessed mid-year, (b) language and early literacy performance-level groups assessed in the Fall, (c) formative growth in picture naming and sound identification skills assessed Fall, Mid, and Spring, and (d) distal standardized language and literacy outcomes assessed in the Fall and Spring. Ancillary measures of family socio-demographics, and teacher preparation/experience were used to describe the sample. Parental consent was acquired for all children and family participants.

To maximize the sensitivity of the standardized measurement component across language and early literacy domains within study resources and reduce participant burden each child was randomly selected to be administered only one of the three standardized tests. As described below, the 3 test assignment groups were randomized within sites ($n \sim 219$ in each test group), one each receiving the *Peabody Picture Vocabulary Test* (PPVT-4) (Dunn & Dunn, 2007), or the *Comprehensive Evaluation of Language Fundamentals* (CELF-3) (Wiig, Secord, & Semel, 2004), or the *Test of Preschool Early Literacy* (TOPEL) (Lonigan, Wagner, & Torgesen, 2007) in each classroom. In the Spring children received the same test as in the Fall. Additionally, 5-6 children per classroom were randomly selected to receive classroom observations. Thus, the design included a planned missing data component.

A single measurement director planned and supervised implementation of the multisite data collection. This director worked with three designated cross-site coordinator who supervised and monitored the implementation of the measurement plan to control for site differences in measurement. Within this organization, staff members in local sites were trained in the use of the same measures that included calibration standards and procedural and measurement reliability.

Measurement

The measures used in the Fall were the Family and Teacher Surveys, the *Get Ready to Read* (GRTR) screener, the Picture Naming, and Sound Identification *Individual Indicators of Growth and Development* (IGDIs), and the PPVT-IV, CELF-3, and TOPEL Print Knowledge and Phonological Awareness sub-scales. At mid-year, the *Classroom Assessment Scoring System* (CLASS), *Code for Interactive Recording of Children's Learning Environments* (CIRCLE), the IGDIs, and *Preschool Curriculum Checklist* (PCC) were collected. And, in the spring, the IGDIs, and PPVT-IV, TOPEL, and CELF-3 measures were collected again. An average of 6 months separated administrations of the standardized tests.

Child and Family Characteristics—Socio-demographic characteristics of the children and their families were assessed using the 25-item parent survey. For the child, date of birth, age, gender, race/ethnicity, and disability status (Individual Education Program [IEP]) were collected. For the parent and family, the parents' marital status, attained level of education, and family income were collected. Information also was obtained on the primary and secondary languages spoken in the home to the child, and the child's language preference.

Teacher Preparation and Experience—Teachers also completed a 25-item survey developed by the research team. Items reported in this study, were each teacher's total years of teaching experience, and level of Early Childhood preparation including degrees/certificates held.

Process Measurement—The *CLASS Pre-K: Classroom Assessment Scoring System* (Pianta, La Paro & Hamre, 2008) was used. It is a widely used rating of preschool instructional quality. Composite scores are reported for three major dimensions of instruction: emotional support, classroom organization, and instructional support. Emotional

Support is comprised of subscales including Positive Climate, Negative Climate, Teacher Sensitivity, and Regard for Student Perspectives. Classroom Organization is comprised of Behavior Management, Productivity, and Instructional Learning Formats. Instructional Support is comprised of Concept Development, Quality of Feedback, and Language Modeling.

Classroom ratings, ranging from 1 (lowest) to 7 (highest), were based on 80 minutes of observation (four 20-minute cycles) in each classroom. Raters were trained by a CLASS-certified trainer and met rater-certification standards as required through the CLASS website. Reliability percentages averaged 90% across observers, with site averages ranging from 85% to 96%. In prior research, higher scores on CLASS dimensions of teacher-child interactions predicted growth in Pre-K children's achievement (Howes, 2008); gains in academic skills in kindergarten and first grade (Burchinal et al., 2008; Hamre, 2010, retrieved); and student engagement (Downer, Rimm-Kaufman, & Pianta, 2007).

Curriculum Quality Rating—*The Preschool Curriculum Checklist* (PCC) (Kaminski & Carta, 2010) provided ratings of each curricula in four areas of language and early literacy and was intended for estimating the quality that a curriculum adheres to evidence-based principles of curricular instructional design and practice. It was completed by a trained reviewer who evaluated the scope and sequence, materials, and procedures making up a preschools curriculum using a rubric. Along with each item to be evaluated in the rubric, a linked evidence source to be located and examined by the reviewer is provided. Reviewers are trained to base their ratings solely on the review of documents and not be influenced by any other knowledge of the curriculum's quality or implementation.

Scoring of a curriculum is based on a rubric of 10 criterion items each of which were rated as 0 (min), 1, or 2 (max). The obtained ratings summed across the 10 items produced a total summary score of 0 (min) to 30 (max). These scores are converted to a percentage ($\% = 100[\text{obtained}/\text{max score}]$). Reviewers provide these scores for (a) vocabulary and oral language, (b) alphabet knowledge, (c) phonological awareness, and (d) listening comprehension domains.

Development of the PCC went through several iterations of definition, consensus, trial, and revision that included multiple paired evaluations of interrater agreement on use of the rubric covering 6 different curriculum rated by 7 raters for a total of 58 paired evaluations. Mean Pearson r 's across raters' scores and curricula were .70 (Phonological Awareness), .69 (Alphabet Knowledge), .72 Vocabulary and Oral Language, and .67 (Listening Comprehension).

Teacher literacy focus and student literacy engagement—*The Classroom CIRCLE: Code for Interactive Recording of Children's Learning Environments* (Atwater, Lee, Montagna, Reynolds & Tapia, 2009), was used to quantify the literacy experiences provided the children and their engagement in literacy behaviors. Classroom CIRCLE quantifies (a) classroom contexts (e.g., small groups, individual), (b) teacher's behavior (e.g., focus of instruction, follows lead), and (c) target child's response (e.g., academic engagement, other engagement, inappropriate) using a momentary time sample method.

Trained observers collected the data using collection software running on a palm computer (Atwater, Lee, Montagna, Reynolds, & Tapia, 2009).

Thirty minutes of data were collected for each of 360 children randomly sampled during a range of classroom activities in which literacy was likely to be embedded (e.g., dramatic play, centers, art, story reading, nature, music, etc.). Scores produced were the percentage of intervals observed for: a Teacher Literacy Focus composite and a target child's Literacy Engagement composite. Observers were trained locally by a local coordinator, who was trained at the Kansas City site and certified by the CIRCLES developer. The local coordinators trained and certified local observers using the same procedures. Indices of inter-observer agreement were high and were similar across sites. Overall percentage agreement ranged from 84.6% to 97.5%; Kappa ranged from 0.70 to .88 per site.

Language and early literacy screener—The *Get Ready to Read* (GRTR) screener and its cut points for identifying children with weak and very weak early literacy skills were used to identify children (Whitehurst & Lonigan, 2001a, 2001b). The GRTR is a brief, widely used, 20-item screener that taps print knowledge, emergent writing, and phonological awareness. Alpha coefficient reliability of .78, and split-half reliability of .80 are reported with validity ranging from .58 - .69 (Phillips, Lonigan, & Wyatt, 2009). Phillips, Lonigan, and Wyatt (2009) also reported longer-term predictive correlations showing that the early literacy screener was correlated with some later reading-related measures. And, that a cut point on the GRTR produced 68% to 86% classification accuracy (predictive utility) with the three domains of *Test of Preschool Early Literacy* (TOPEL) (Wilson & Lonigan, 2010). Estimates of the predictive validity in the current study were correlations between the Fall GRTR total score and the Spring standard scores on the PPVT-IV, CELF-3, TOPEL-Phonological Awareness, and TOPEL-Print Knowledge that were $r = .65, .59, .46, \text{ and } .62$, respectively. Based on the GRTR total score and recommended cut point ranges (Whitehurst & Lonigan, 2001b), we formed three performance-level groups: Tier 1 = 9 to 20 (average and above), Tier 2 = 6 to 8 (weak skills), and Tier 3 = 0 to 5 (very weak skills) for use in later analyses.

Formative Growth and Development Measures—Two Individual Growth and Development Indicators (IGDIs, Version 2.0) developed by the research team were used (McConnell & Greenwood, In press). These were Picture Naming and Sound Identification (Greenwood, Carta, & McConnell, 2011). The Picture Naming IGDI is an individually administered untimed task in which the student verbally identifies a single picture presented to them (Bradfield, Besner, Wackerle-Hollman, Rodriguez, & McConnell, in preparation). Picture Naming administration consisted of four sample cards and 40 test cards, each depicting a familiar object. The child was asked to name each card. The score on this untimed test was the number of pictures correctly named out of 40 that was converted to a Rasch scale score and a correct card equivalent score. Picture Naming has a reported person-level (as opposed to item level) reliability score of .81 and criterion validity correlation coefficients of .69 with the *Clinical Evaluation of Language Fundamentals 3* (CELF-3) expressive vocabulary subtest and .62 with the PPVT-4. A cut score of 28 out of 40 was associated with a probability of .70 of being in Tier 1. This score was associated with a

balance in classification accuracy of 70% between correctly needing Tier 2/3 (sensitivity) versus falsely needing Tier 2/3 (selectivity). The criterion used for classification accuracy was teachers' ratings as defined by a 3-tier rubric classification of students' skills developed by the authors. Estimates of the predictive validity in the current study between the Fall Picture Naming score and the Spring standard scores on the PPVT-IV and CELF-3 were $r = .74$ and $.66$, respectively.

The Sound Identification IGDIs also is an individually administered untimed task in which the student identifies from three choices the letter that makes a particular letter sound (Bradfield, Clayton, et al., in preparation). Sound Identification also consists of card items; however, each card depicts 3 letters (upper and lower case). The child is asked to point to the letter that makes the sound modeled by the test administrator. The score on this untimed test is the number of letters correctly identified and it also was converted to a Rasch scale score and a correct card count equivalent score. The Rasch person reliability was $.60$. Sound Identification has a criterion validity correlation coefficient of $.54$ with the TOPEL Print Knowledge subtest. A cut score of 10 out of 20 possible correct was associated with a probability of $.74$ of being in Tier 1. This score was associated with a balance in classification accuracy of 70% for Tier 2/3 truly needing Tier 2 (sensitivity) versus falsely needing Tier 2/3 (selectivity). Estimates of the predictive validity in the current study between the Fall SID score and the Spring standard scores on the TOPEL- PA, and TOPEL- PK were $r = .42$, and $.44$, respectively.

Assessors learning to administer the IGDIs qualified based on meeting a protocol assessing their administration fidelity. After receiving training from the lead child assessor, all child assessors practiced using all IGDIs with at least two adults. Next, child assessors practiced administration with one preschool aged child while the lead assessor recorded the number of accurate administration steps using the protocol. Child assessors unable to complete administration with a preschool aged child accurately practiced with an additional two adults and then tried again with a preschool aged child. Once initial administration reliability was attained, child assessors were observed again by the lead assessor once per week during heavy data collection periods to ensure that standardized administrations were continuing. They provided corrective feedback as needed. The picture naming and sound identification measures were assessed 3 times, in the Fall, mid-year, and in Spring.

Fall and Spring Preschool Language and Early Literacy—Receptive vocabulary was measured using the *Peabody Picture Vocabulary Test-4th Edition* (PPVT- 4) (Dunn & Dunn, 2007). Reliability for the PPVT-IV is reported to be $.92$ -. $.98$ (Alpha) and $.87$ -. $.97$ (split-half). Validity was $.41$ -. $.84$. Expressive vocabulary, word structure, sentence structure, and core language skills were measured by the *Clinical Evaluation of Language Fundamentals Preschool* (CELF-P2). Reliability is reported to range from $.77$ -. $.95$ for alpha, $.80$ -. $.97$ for split half. Validity is reported to range from $.57$ -. $.84$. Phonological awareness and print knowledge were measured using the *Test of Preschool Early Literacy* (TOPEL: Lonigan, Wagner, & Torgesen, 2007). It provides raw, scale, and standard scores. Split half reliability is reported to range from $.87$ -. $.96$. Criterion validity ranges from $.59$ -. $.77$. Estimates of the concurrent validity in the current study were the pattern of correlations between the Fall standard scores on the PPVT-IV vs. CELF-P2, the TOPEL- PA, and the

TOPEL- PK that were $r = .70, .65, .51$, respectively, and between the CELF-P2 vs. the TOPEL- PA and TOPEL- PK that were $r = .61, .45$, respectively, and between the TOPEL- PA vs. TOPEL- PK at $r = .43$.

Statistical Analysis

The pre-imputation child-level descriptive statistics for all measures are reported in Table 1 by Fall, Mid-year and Spring occasions. The number of children with specific data ranged from $n = 644$ in the Fall on the GRTR to $n = 189$ for the CELF-P2 administered in the Spring. Evaluation of the patterns of incomplete data indicated that all variables in the data set had at least one missing value on a case and all cases had at least one missing value on a variable. While 2,889 of the 8,400 values (cases $\times$ variables) were missing for a total of about 34% missing data, only about 1,156 of the 8,400 values were unplanned missing data for about 14% missing data. Because planned missing data designs produce missing completely at random (i.e., the missing information is unrelated to the study variables), the planned missing data were completely recoverable and cannot introduce bias into the parameter estimates (see Enders, 2010).

To appropriately manage incomplete data, we used the multiple imputation technique. Research suggests that the multiple imputation technique improves accuracy and power relative to more traditional approaches of addressing missing data (Enders, 2010; Graham, 2009; Graham, Cumsille, & Elek-Fisk, 2003; Graham, Taylor, Olchowski, & Cumsille, 2006). We used the SAS 9.2 MI procedure (PROC MI) with the Markov Chain Monte Carlo (MCMC) estimation to create 100 imputations (i.e., datasets) to retain optimal power (Collins, Schafer, & Kam, 2001; Graham, Olchowski, & Gilreath, 2007) (Schafer & Graham, 2002). We determined that the imputed data sets should be separated by at least 60 iterations, so each data set was saved after a much more conservative 200 iterations.

Descriptive statistics were computed for each imputation and pooled over all imputations to form overall estimates using the SPSS Version 18 missing value analysis module. Rubin's Rules were used to provide pooled descriptives (Rubin, 1987). For categorical variables, simple percentages were reported. For continuous variables, the mean and standard error (*SE*) rather than the standard deviation was reported because the pooled statistics were parameter estimates. M-Plus was used to compute linear slope and intercepts for the two IGDIs. Fall and Spring mean intercepts were computed reflecting start and year-end performance. Slope estimates reflected the linear gain in correct card count per mean month of age. To address product and process differences in children's processes and products the main effect and interaction effect means Tier (3) $\times$ Program Types (4) were computed and reported in tables or graphical displays.

Results

Research Question 1. What were the proportions of children populating the 3 performance-level Tier groups?

The Fall pooled mean GRTR screener total score was 11.2 ($SE = .17$), ranging from 1 to 20 indicating that children's initial early literacy skill levels varied across the entire spectrum of

the GRTR scale of 1 to 20. The proportional breakdown of children overall by GRTR tier levels was 70%, 20%, and 10% for Tiers 1, 2, and 3, respectively. The pooled mean GRTR tier group scores were 13.4 ($SE = .15$) vs. 7.2 ($SE = .07$), vs. 4.2 ($SE = .14$) for Tier 1, 2, and 3, respectively. Pooled GRTR means by program type from highest to lowest were 14.4 ($SE = .26$) for Tuition-Based, 11.1 ($SE = .30$) for Title 1, 11.0 ($SE = .26$) for Pre-K, and 10.3 ($SE = .37$), for Head Start programs. The proportional breakdown of program types (4) by tier levels (3) revealed additional variability in tier group proportions. As shown in Figure 1, these data confirmed that children in the three low-income eligibility programs (i.e., Pre-K, Head Start, and Title 1) had greater proportions of lower performing children in the Tier 2 (ranging from 19 to 23%) and Tier 3 (ranging from 8 to 13%) groups compared to the Tuition-Based program. The Tuition-Based program had the greatest proportion of higher performing children in Tier 1 (92%) and the smallest proportions of children at Tier 2 (6%) and Tier 3 (3%) levels.

2. Were there corresponding differences in the 3 groups' performance on standardized measures of language and early literacy skills at the beginning of the year and did a year's exposure to preschool instruction close gaps observed between the groups by year's end?

GRTR tier groups were associated with corresponding gaps in the four Fall language and early literacy outcome measures with minor exceptions (see Figure 2 and Table 2). Children in the Tier 1 group generally performed at the normative mean level. Gaps between Tier 1 versus Tier 2 were sometimes as large as $-1.0 SD$, and versus Tier 3 as large as $-2.0 SD$. In language, receptive vocabulary gaps on the PPVT-IV were: Tier 1 $M = 95.9$ ($SE = .8$), versus Tier 2 $M = 80.4$ ($SE = 1.6$), versus Tier 3 $M = 66.7$ ($SE = 2.2$) while total language gaps on the CELF-P2 were Tier 1 $M = 92.1$ ($SE = .8$), versus Tier 2 $M = 75.6$ ($SE = 1.6$), versus Tier 3 $M = 66.0$ ($SE = 2.2$). Similar gaps for early literacy skills were observed for print knowledge, Tier 1 $M = 100.5$ ($SE = 0.7$) versus, Tier 2 $M = 88.4$ ($SE = 1.9$) versus Tier 3 $M = 82.9$ ($SE = 2.7$). Phonological Awareness also had sizable gaps between the tier groups: Tier 1 $M = 93.3$ ($SE = .8$), versus Tier 2 $M = 84.0$ ($SE = 2.3$), versus Tier 3 $M = 74.7$ ($SE = 3.2$).

Program types also accounted for similar gaps between Tier groups with the exception of the Tuition-Based program that contained only 6 of their 72 children in the Tier 2 and 3 groups (see Table 2). Also, the three low-income eligibility programs (i.e., Pre-K, Head Start, and Title 1) were consistently lower than the Tuition-Based children at start in the Fall.

After a year of preschool in the Spring, the good news was that tier groups overall made standard score gains ranging from 4.6 to 11.1 standard score points (see Figure 2 and Table 2). Gains also varied by tier groups within program types ranging from a low of 0.1 (Head Start Tier 3 print knowledge) to 11.7 (Head Start Tier 3 phonological awareness). Only in the case of print knowledge did Tier 2 and 3 Head Start children make very little growth Fall to Spring. Even so, however, Tier 2 and 3 groups maintained their relative rank order in status from Fall to Spring even though some gains were made by each group. Two exceptions, Head Start phonological awareness and Pre-K and Title 1 print knowledge, children in the 3 low-income programs reduced gaps between Tiers.

3. Did a year's exposure to preschool instruction close gaps on formative measures of picture naming and sound identification between groups by year's end, and how large were children's rates of growth?

The large gaps between tier group levels initial status and the language and early literacy outcome measures also were reflected in the children's initial picture naming and sound identification status in the Fall (see Table 3). Mean picture naming intercepts by tier groups were 29.5 (Tier 1), versus 22.7 (Tier 2), versus 17.2 (Tier 3) in the Fall, respectively. These differences between Tier groups were 6.6 and 5.5 cards named correctly in the Fall respectively between Tier 1 versus 2, and Tier 2 versus 3. Spring mean picture naming intercepts were increased to 33.3 (Tier 1) versus 28.6 (Tier 2), versus 23.8 (Tier 3), with differences of 4.8 and 4.7 correct cards remaining between adjacent tier groups.

Fall mean intercepts values for sound identification were 11.8 versus 7.4, versus 5.9 respectively by tier level. Gaps between adjacent groups were 4.4 and 1.5 cards correct. Spring mean sound identification intercepts also had increased to 14.86 (Tier 1), versus 9.3 (Tier 2), versus 8.0 (Tier 3). However, the gaps remaining between adjacent tier groups were 5.6 and 1.3 cards correct.

Annual overall growth patterns also differed by formative skill. In the case of picture naming, lower performing children were growing faster in linear slope (Tier 2 $M_{\text{slope}} = 2.94$) versus (Tier 3 $M_{\text{slope}} = 3.26$) than were the children in Tier 1 ($M_{\text{slope}} = 1.87$) over the preschool year. This was not the case for sound identification where the children in Tier 1 were growing moderately faster ($M_{\text{slope}} = 1.49$) than the low performance groups (Tier 2 $M_{\text{slope}} = .95$, versus Tier 3 $M_{\text{slope}} = 1.06$) (see Table 3).

Additional variation was seen for program types. All programs and Tier groups showed growth over time with the exception of Head Start where Tier 2 and 3 children actually showed negative slope values in sound identification compared to all others that were positive (see Table 3 and Figures 3 and 4). With respect to picture naming there was a tendency for high performing children in the Fall to grow slower than lower performing children at start of prekindergarten. For example, children in the Head Start and Tuition-Based program Tier 1 groups started and ended the year with the greatest picture naming and the lowest rates of growth ($M_{\text{slope}} = .82$ and 1.52 in order) compared to the other Pre-K and Title 1 Tier 1 groups ($M_{\text{slope}} = 2.21$ and 2.03 , respectively). Head Start and Tuition-Based Tier 1 groups also made the least growth in sound identification compared to children in Pre-K and Title 1 programs (see Table 3).

4. What was the quality of instruction children received in terms of classroom instructional support, curriculum-quality, teacher-literacy focus, and children's literacy engagement in these programs with language and early literacy goals?

Overall, the instructional support provided by the teacher at the classroom level of analysis was low as measured by the CLASS, $M = 2.6$ ($SD = .80$), whereas CLASS ratings were twice as high for emotional support ($M = 5.4$, $SD = .90$) and classroom organization ($M = 5.0$, $SD = 1.00$) (see Table 4). Emotional support ranged from 6.0 in Head Start to 5.2 in Title 1 programs. Similar values for Classroom organization ranged from 5.6 also in Head

Start to 4.8 in Title 1 programs (see Table 4). Instructional support ratings were nearly twice as low, ranging from 2.2 in Title 1 to a high of 3.3 in Head Start programs. The three CLASS scales were the very lowest in the Tuition-Based program as compared to the other three programs (i.e., emotional support [4.5], classroom organization [4.1] and instructional support [2.0]).

Also at the classroom level of analysis, 12 curricula were used in part or in combination to guide the core language and early literacy instruction. The mostly frequently reported curriculum was used in 33% of programs with four others used in only 2% of programs according to teachers. Curriculum quality ratings for vocabulary/oral language, phonological awareness, and listening comprehension overall were in the low 60% range, ranging from 60% (listening comprehension to 63.6% (vocabulary/oral language). With the exception of vocabulary/oral language, the Tuition-Based program was lowest in quality related to alphabet knowledge; phonological awareness, and the very least in listening comprehension (see Table 4).

At the student level of analysis, the amount of teacher literacy focus they experienced was low overall at 16.5% and did not differ by children's tier groups at 16.0%, 16.2%, and 17.7% for Tier 1, 2, and 3, respectively. Student literacy engagement was correspondingly low with means of 22.7%, 21.8%, and 24.2% occurrence in order by Tier group (see Figure 5).

Program type difference, however, were observed. The Head Start program was lowest in teacher literacy focus $M = 5.9\%$ ($SE = .98$) compared to the other three programs, while the Pre-K program was highest $M = 24.6\%$ ($SE = 1.06$) compared to the other three programs. Similar values of teacher literacy focus in the Title 1 and Tuition-Based programs were 14.2% ($SE = 1.17$) and 12.19% ($SE = 1.90$), respectively. Breaking these down by foci, teachers in Head Start and Tuition-Based programs were least engaged in phonological focus, story comprehension, and vocabulary in favor of reading to the children. In contrast, Pre-K instruction provided the children with greater vocabulary focus, story and other forms of comprehension, alphabet knowledge, and phonological awareness.

Observed values for student literacy engagement indicated that the Pre-K program with the highest teacher literacy engagement also had the highest levels of student literacy engagement $M = 28.0\%$ ($SE = 1.24$). Children in Pre-K were much more engaged in writing, academic manipulation, academic verbalization, and academic attention. Head Start also had a much lower value at 17.9% ($SE = 1.47$). Similar levels for Title 1 and Tuition-Based were $M = 21.5\%$ ($SE = 1.55$) and $M = 16.0\%$ ($SE = 1.98$), respectively. Children in Head Start were most engaged in academic manipulation least in academic attention. Behavior problems were low overall and not different by Tier group or program types.

Discussion

We applied an RTI measurement approach to describing children's performances in a multi-site sample of prekindergarten programs employing with language and early curricula to reach they literacy goals. Consistent with the RTI approach, we identified children at 3 tiers of performance (Tiers 1, 2, and 3) in the Fall and followed their progress and experiences

received through Spring. These programs were not using an RTI approach to track progress of their lowest performing children or to provide multiple tiers of support. In analyses, we reported educationally rather than statistically significant differences between Tier levels and Program Types as would likely be the case in a real field evaluation or RTI implementation.

Overall, findings supported the feasibility and validity of a preschool RTI measurement approach. Findings were also consistent with prior concerns reported in the preschool literature regarding the low quality of instruction and the lack of differentiated instruction for those children most at risk for language and early literacy delays. With respect to feasibility, it was possible to universally screen and identify groups of children falling below benchmark cut points (i.e., Tiers 1, 2, and 3 groups) in the multi-site sample. Results confirmed that children identified to 3 performance groups in Fall were highly divergent on norm-referenced, standardized and formative progress monitoring language and literacy skills/abilities. Fall results also confirmed SES differences in children's entry performance, wherein the three low-income programs consistently performed lower than the higher SES Tuition-Based group. Fall mean performance differences in tier groups were often as large as -1 to -2 standard deviations below the normative mean.

By Spring testing, children overall, by tier groups and by program types made standard score gains in outcomes. However, tier groups in the 3 low-income programs generally, had not closed the gaps identified in the Fall. Because of low numbers of Tuition-Based children in these tier groups, these children were not analyzed. Associated with these findings, results also indicated low and the wide variability in the quality of the curriculum and instructional experiences provided in the 4 program type contexts with no evidence of differentiation based on skill level in terms of Tier group status.

Collectively, these programs with language and early literacy outcome goals were not able to reduce initial skill gaps in some 30 percent of students compared to normative standards, or even compared to children in the study who were initially highest performing in Tier 1 (i.e., the Tuition-Based program). Nor did they produce sizeable improvements within children over time. These findings are consistent with other similar reports in the literature discussed earlier. The process findings in terms of instructional support, curriculum quality, teacher literacy focus, and student literacy engagement while complex converged in many ways supporting the conclusion that basic classroom instruction quality remains low and variable.

Limitations

Several limitations are noted. This was a multivariate, longitudinal descriptive study over one preschool year. Consequently, causal relationships between processes and products were not possible to identify. Programs/classrooms in the sample were drawn from regional locations that were accessible by the research teams in each site. They represented typical early childhood programs serving children in those locations. Because of site differences in socio-demographics and lack of balance in program representation within sites, it cannot be assumed that the program type differences reported in this study were representative of such programs nationally.

Implications for Research and Practice

Considering these limitations, the RTI measurement model did reveal theoretically interesting findings. Overall, it appeared that preschool instructional quality remains a universal preschool problem and the question of how to improve the instructional experience and improve the outcomes of all students still remains to be addressed. So, is the RTI approach a possible alternative?

An assumption of RTI is the use of evidence-based practice without which it is not possible to achieve greater effectiveness than just *business as usual*. Thus, the wisdom of overlaying Tier 2 and Tier 3 interventions on low-quality Tier 1 arises. Most experienced proponents of RTI note that installing an RTI model in a program, school, or district, takes multiple years to accomplish with implementation including local measurement tools and Tier 1 curriculum improvements given first priority (Greenwood, Horner, Kratochwill, & Clements, 2008). Tier 1 improvement is an early priority in such efforts because without it, Tier 1 remains a continuing source of larger numbers of children needing Tier 2 and 3 interventions leading to lower outcomes and higher costs (Chard et al., 2008).

One of the most intriguing findings was the low and wide variation in the instructional quality profile of measures: the curriculum ratings, the CLASS, the teacher literacy focus, and student literacy engagement scores in the various program contexts. For example, the Tuition-based programs had the lowest CLASS and curriculum quality ratings (with the exception of vocabulary and oral language), and the second lowest teacher literacy focus and student literacy engagement outcomes. However, they also had very few students in Tier 2 or 3 in the Fall (only a combined 9% of all children). By Spring, students in the Tuition-Based program had made the least overall gains on the four test measures and only small to moderate rates of growth in picture naming and sound identification. This evokes several questions. One is the question of whether or not an RTI approach is appropriate or even needed in such a low-risk program-level population?; and two, as these findings suggest, might not the Tuition-Based students also benefit from a higher quality Tier 1 experience?

Because RTI is an approach that embraces all children, small numbers at risk does not seem to obviate RTI because even small numbers potentially benefiting can be identified and served at Tier 2 or 3. An assumption of RTI is that individual children vary in cumulative history of exposure to early literacy experiences (e.g., Hart & Risley, 1995) and response to instructional intervention, even those from higher SES families. This seems particularly true in preschool, where many children are just beginning their schooling, arriving with the early literacy experiences and skills received earlier at home. The result is that there will also be some number of children in a program potentially benefitting from receipt of more than just Tier 1, core-level instructional support.

And, yes, when one examines the gains made by the Tuition-based children after exposure to its low-quality program, and children's small gains, by these children who entered preschool with average to high performance levels, these children we argue, would highly likely benefit from a higher quality Tier 1 experience.

However, improvement for these children or those lower performing at entry cannot be expected to happen just because children are placed in 3 diagnostic groups in the absence of high quality instructional experiences associated with their particular needs. RTI's 3 tiers of support after all are conceived to be intervention/prevention groups, not simply diagnostic groups in the absence of specific, standard instructional practices to be received (Fuchs & Fuchs, 2007).

Head Start had the highest curriculum ratings and CLASS scores but was lowest teacher literacy focus and student literacy engagement. Yet, Head Start children closed the Tier 3 gap in phonological awareness and gained in other outcomes, but made the very least progress in picture naming, sound identifications, and print knowledge. These findings suggest complex relationships at work here and not all simply convergent.

Pre-K programs had the highest observed teacher literacy focus and student literacy engagement, and the Pre-K students did well in a number of ways by making progress on all four outcomes – closing the Tier 2 gap in phonological awareness, and showing the best growth in picture naming and sound identification skills, even though they served a high proportion of young children learning two languages.

Pre-K classrooms were highest in observed teacher literacy focus and student literacy engagement and had second highest curriculum ratings and CLASS scores. A particularly bright spot were the findings for the children learning two languages (ELL), many of whom were identified needing Tier 2 and 3 supports in the Fall. These children also had greater picture naming (vocabulary) growth rates of children perhaps due to this more intensive exposure. Alternately, teacher's literacy focus was least large in Head Start, and children were not growing in picture naming and sound identification with some groups actually declining over time.

Collectively, we interpret these results to mean that the quality of what occurs in the classroom as instruction does matter in terms of what students are learning. Untangling these relationships within and among these instructional variables under teachers control was not possible in this report, however, additional analyses examining moderating and mediating roles in students' outcomes independently of program type is currently underway for a future report. Just one hypothesis being investigated is that the path between the instructional quality process variables and children's long-term language and early literacy outcomes is through their growth in short-term literacy skills outcomes like picture naming and sound identification? While informative empirical findings may result from such analyses they will not be convincing causal information.

Thus, experimental studies using RTI measurement models like that in this study either for purposes of development and validation of Tier interventions (Goldstein, Spencer, & Greenwood, 2012, March) or for the purpose of testing the effects of full RTI models will offer greater insight into the causal relationships. Beyond new knowledge, such work will likely widen the choices of specific interventions for use by early childhood practitioners in multi-tier models and also provide efficacy of RTI models when implemented over time (e.g., Bailet, Kepper, Piasta, & Murphy, 2009; Buzhardt et al., 2011; Gettinger & Stoiber,

2007). Experimental, developmental work is needed to create the effective tools and practices the preschool teachers can use for structuring and monitoring successfully implemented RTI programs in ways that supersede just initial professional development and training. For example, intervention development work is needed to provide not only the evidence, but also the measures, curricula, media, websites, TA, and management practices needed to support and sustain the implementation of RTI programs and bring them to scale (Linas, Carta, & Greenwood, 2010, June). Unique characteristics of the RTI approach that could or should make this infusion of new tools and practices easier may be its focus on inclusion of all children, prevention, evidence-based practice, intentional teaching, data-based decision making, and fidelity of intervention. Like the development of progress monitoring measurement for early childhood (Greenwood, Carta, et al., 2011), the feasibility and efficacy of the RTI approach in preschool language and literacy instruction remains to be demonstrated.

References

- Atwater, JB.; Lee, Y.; Montagna, D.; Reynolds, LH.; Tapia, Y. Classroom CIRCLE: Classroom Code for Interactive Recording of Children's Learning Environments. In: Utley, C.; Greenwood, CR.; Reynolds, SH.; Douglas, K.; Bannister, H.; Kim, JM., editors. EBASS-Mobile: Ecobehavioral Assessment System Software. Kansas City, KS: Juniper Gardens Children's Project, Institute for Life Span Studies, University of Kansas; 2009.
- Bailet LL, Kepper KK, Piasta SB, Murphy SP. Emergent literacy intervention for prekindergarteners at risk for reading failure. *Journal of Learning Disabilities*. 2009; 42(4):336–155. [PubMed: 19398614]
- Berkeley S, Bender WN, Peaster LG, Saunders L. Implementation of response to intervention: A snapshot of progress. *Journal of Learning Disabilities*. 2009; 42(1):85–95. [PubMed: 19103800]
- Bradfield, T.; Besner, A.; Wackerle-Hollman, A.; Rodriguez, M.; McConnell, S. Research and development of new individual growth and development indicators of oral language: Picture Naming. University of Minnesota Minneapolis: Center for Response to Intervention in Early Childhood; in preparation
- Bradfield, T.; Clayton, C.; Wackerle-Hollman, A.; Hays, A.; Rodriguez, M.; McConnell, S. Research and development of new individual growth and development indicators of alphabetic principle: Sound Identification. University of Minnesota Minneapolis: Center for Response to Intervention in Early Childhood; in preparation
- Burchinal M, Howes C, Pianta R, Bryant D, Early D, Clifford R, et al. Predicting child outcomes at the end of kindergarten from the quality of pre-kindergarten teacher-child interactions and instruction. *Applied Development Science*. 2008; 12:140–153.
- Buyse V, Peisner-Feinberg E. Recognition and Response. *Early Childhood RTI Roadmap*. 2009:7–8.
- Buzhardt J, Greenwood CR, Walker D, Anderson R, Howard WJ, Carta JJ. Effects of web-based support on Early Head Start home visitors' use of evidence-based intervention decision making and growth in children's expressive communication. *NHSA Dialog: A Research-to-Practice Journal for the Early Childhood Field*. 2011; 14(3):121–146.
- Campbell R, Pungello EP, Miller-Johnson S, Burchinal M, Ramey CT. The development of cognitive and academic abilities: Growth curves from an early childhood educational experiment. *Developmental Psychology*. 2001; 27:231–242. [PubMed: 11269391]
- Chard, DJ.; Harn, BA.; Sugai, G.; Horner, RH.; Simmons, DC.; Kame'uei, EJ. Core features of multi-tiered systems of reading and behavior support. In: Greenwood, CR.; Kratochwill, T.; Clements, M., editors. *School-wide prevention models: Lessons learned in elementary schools*. New York: Guilford; 2008. p. 31–58.
- Chien NC, Howes C, Burchinal M, Pianta RC, Ritchie S, Bryant DM, Barbarin OA. Children's classroom engagement and school readiness gains in prekindergarten. *Child Development*. 2010; 81(5):1534–1549. [PubMed: 20840239]

- Connor CM, Morrison FJ, Petrella JN. Effective reading comprehension instruction: Examining child by instruction interactions. *Journal of Educational Psychology*. 2004b; 96:682–698.
- Connor CM, Piasta SB, Fishman B, Glasney S, Schatschneider C, Crowe E, Morrison FJ. Individualizing Student Instruction Precisely: Effects of Child x Instruction Interactions on First Graders' Literacy Development. *Child Development*. 2009; 80(1):77–100. [PubMed: 19236394]
- Downer J, Rimm-Kaufman S, Pianta R. How do classroom conditions and children's risk for school problems contribute to children's engagement in learning? *School Psychology Review*. 2007; 36(3):413–433.
- Dunn, LM.; Dunn, DM. *Peabody Picture Vocabulary Test - Fourth Edition*. Bloomington, MN: Pearson Assessments; 2007.
- Early D, Bryant D, Pianta RC, Clifford D, Burchinal M, Ritchie S, Barbarin O. Are teachers' education, major, and credentials related to classroom quality and children's academic gains in pre-kindergarten? *Early Childhood Research Quarterly*. 2006; 21(2):174–195.
- Fuchs LS, Deno SL. Paradigmatic distinctions between instructionally relevant measurement models. *Exceptional Children*. 1991; 57:488–500.
- Fuchs, LS.; Fuchs, D. The role of assessment in the three-tier approach to reading instruction. In: Haager, D.; Klingner, J.; Vaughn, S., editors. *Evidence-based reading practices for response to intervention*. Baltimore, MD: Brookes; 2007. p. 29–44.
- Gettinger M, Stoiber K. Applying a response to intervention process for early literacy development in low-income children. *Topics in Early Childhood Special Education*. 2007; 27:198–213.
- Goldstein, H.; Spencer, E.; Greenwood, CR. Applying RTI to language and early literacy in pre-k settings. Paper presented at the Conference on Research Innovations in Early Intervention; San Diego. 2012 Mar.
- Greenwood CR, Bradfield T, Kaminski R, Linas M, Carta JJ, Nylander D. The Response to Intervention (RTI) approach in early childhood. *Focus on Exceptional Children*. 2011; 43(9):1–22.
- Greenwood CR, Carta JJ, McConnell S. Advances in measurement for universal screening and individual progress monitoring of young children. *Journal of Early Intervention*. 2011; 33(4):254–267.
- Greenwood, CR.; Horner, R.; Kratochwill, T.; Clements, M. Introduction. In: Greenwood, CR.; Kratochwill, T.; Clements, M., editors. *School-wide prevention models: Lessons learned in elementary schools*. New York: Guilford; 2008. p. 3–30.
- Hamre, B. K-3 CLASS: Technical Appendix. Baltimore, MD: Brookes; 2010. retrieved online at http://www.brookespublishing.com/class2007/CLASS_K-3.pdf
- Hart, B.; Risley, TR. *Meaningful differences in the everyday experience of young American children*. Baltimore: Brookes; 1995.
- Howes C, Burchinal M, Pianta R, Bryant D, Early D, Clifford R, et al. Ready to learn? Children's pre-academic achievement in pre-kindergarten programs. *Early Childhood Research Quarterly*. 2008; 23:27–50.
- Justice LM. Evidence-based practice, response-to-intervention, and prevention of reading difficulties. *Language, Speech, and Hearing Services in Schools*. 2006; 37:1–14.
- Justice LM, Hamre B, Pianta R. Quality of language and literacy instruction in preschool classrooms serving at-risk pupils. *Early Childhood Research Quarterly*. 2008; 23:51–68. [PubMed: 22773887]
- Kaminski, R.; Carta, JJ. *Preschool Curriculum Quality Checklist (PCC)*. Kansas City, KS: Center for Response to Intervention in Early Childhood, University of Kansas; 2010.
- Linas, MW.; Carta, JJ.; Greenwood, CR. Taking a Snapshot of Early Childhood Response to Intervention Across the United States: 2009 and 2010. Washington DC: Head Start Research Conference; 2010 Jun.
- Lonigan, CJ.; Wagner, DA.; Torgesen, JK. *Test of Preschool Early Literacy (TOPEL)*. Austin, TX: Pro-Ed; 2007.
- Mashburn, A. Evidence for creating, expanding, designing, and improving high-quality preschool programs. In: Justice, LM.; Vukelich, C., editors. *Achieving excellence in preschool literacy instruction*. New York: Guilford; 2008. p. 5–24.

- Mashburn A, Pianta R, Hamre BK, Downer JT, Barbarin O, Bryant D, Howes C. Measures of Classroom Quality in Pre-Kindergarten and Children's Development of Academic, Language and Social Skills. *Child Development*. 2008; 79:732–749. [PubMed: 18489424]
- McConnell, SR.; Greenwood, CR. In press. General outcome measures in early childhood and the Individual Growth and Development Indicators (IGDIs). In: Buysee, V.; Peisner-Feinberg, E., editors. *Handbook of Response to Intervention (RTI) in Early Childhood*. Baltimore: Brookes;
- McConnell, SR.; Missall, KN. Best practices in monitoring progress for preschool children. In: Thomas, A.; Grimes, J., editors. *Best practices in school psychology*. 5th. Washington, D.C.: National Association of School Psychologists; 2008. p. 561-573.
- McGinty AS, Breit-Smith A, Justice LM, Kaderavek J, Fan X. Does intensity matter? Preschoolers' print knowledge development within a classroom-based intervention. *Early Childhood Research Quarterly*. 2011; 26:255–267.
- McGinty AS, Justice LM, Piasta SB, Kaderavek JN. Does context matter? Explicit print instruction during reading varies in its influence by child and classroom factors. *Early Childhood Research Quarterly*. 2011; 27:77–89.
- Phillips BM, Lonigan CJ, Wyatt MA. Predictive validity of the Get Ready to Read! Screener: Concurrent and long-term relations with reading-related skills. *Journal of Learning Disabilities*. 2009; 42(2):133–147. [PubMed: 19074622]
- Pianta R, Howes C, Burchinal M, Bryant D, Clifford R, Early D, Barbarin O. Features of pre-kindergarten programs, classrooms, and teachers: Do they predict observed classroom quality and child-teacher interactions? *Applied Developmental Science*. 2005; 9(3):144–159.
- Preschool Curriculum Evaluation Research Consortium. Effects of preschool curriculum programs on school readiness: Report from the Preschool Curriculum Evaluation Research Initiative. 2008. Retrieved May 2009 online at <http://ies.ed.gov/pubsearch/pubsinfo.asp?pubid=NCER20082009rev>
- Reynolds AJ, Temple JA, Ou SR, Arteaga IA, White BAB. School-Based Early Childhood Education and Age-28 Well-Being: Effects by Timing, Dosage, and Subgroups. *Science*. 2011; 333(6040): 360–364. [PubMed: 21659565]
- Reynolds AJ, Temple JA, Robertson DL, Mann EA. Age 21 cost-benefit analysis of the Title I Chicago child-parent centers. *Educational Evaluation and Policy Analysis*. 2003; 24:267–303.
- Rubin, DB. Multiple imputation for nonresponse in surveys. New York: J. Wiley & Sons; 1987.
- Russell, J.; McCoy, A.; Pistorino, C.; Wilkinson, A.; Burghardt, J.; Clark, M.; Swank, P. National Evaluation of Early Reading First: Final Report. Washington, DC: U.S. Government Printing Office; 2007. online at <http://ies.ed.gov/ncee/pdf/20074007.pdf> U.S Department of Education, Institute of Education Sciences
- Schafer JL, Graham JW. Missing data: Our view of the state of the art. *Psychological Methods*. 2002; 7:147–177. [PubMed: 12090408]
- Shanahan, T.; Lonigan, CJ. Developing Early Literacy: A Report of the National Early Literacy Panel. Washington DC: National Early Literacy Panel; 2008. online at <http://www.nifl.gov/publications/pdf/NELPReport09.pdf>
- VanDerHeyden AM, Snyder P. Integrating frameworks from early childhood intervention and school psychology to accelerate growth for all young children. *School Psychology Review*. 2006; 35(4): 519–534.
- Whitehurst, GJ.; Lonigan, CJ. Emergent literacy: Development from prereaders to readers. In: Neuman, SB.; Dickinson, DK., editors. *Handbook of early literacy research*. New York: Guilford Press; 2001a. p. 11-29.
- Whitehurst, GJ.; Lonigan, CJ. *Get Ready to Read*. Columbus, OH: Pearson; 2001b.
- Wiig, E.; Secord, W.; Semel, E. *Clinical Evaluation of Language Fundamentals - Preschool Second Edition*. San Antonio, TX: Harcourt Assessment; 2004.
- Wilson SB, Lonigan CJ. Identifying preschool children at risk of later reading difficulties: Evaluation of two emergent literacy screening tools. *Journal of Learning Disabilities*. 2010; 43(1):62–76. [PubMed: 19822699]

Figure 1.
Proportions of children identified at GRTR Tier group performance levels by program types.

Figure 2.

Child language and early literacy outcomes by program type, and GRTR Tier groups Fall to Spring. (Note that due to small n in the Tuition-Based Tier 2 and 3 groups, means were not plotted)

Figure 3.

Growth in picture naming performance by program types and GRTR Tier groups over time.
(Note that due to small n in the Tuition-Based Tier 2 and 3 groups, means were not plotted)

Figure 4.

Growth in sound identification performance by program types and GRTR Tier groups over time. (Note that due to small n in the Tuition-Based Tier 2 and 3 groups, means were not plotted)

Figure 5.
Mid-year teacher literacy focus and student literacy engagement by program types and GRTR Tier groups.

Table 1
Descriptive Child-level Risk and Language and Early Literacy Status Statistics Pre-Imputation

Variable	Fall			Mid-Year			Spring		
	N	Mean	SD	N	Mean	SD	N	Mean	SD
GRTR: Total Score	644	11.3	4.3						
IGDI: Sound ID (True)	612	10.1	5.7	590	12.9	5.7	581	13.8	5.5
IGDI: Picture Naming (True)	624	27.0	8.7	583	29.2	7.5	576	30.0	6.9
CLASS: Emotional Support (Classroom level)				66	5.3	0.9			
CLASS: Classroom Organization				66	5.0	1.0			
CLASS: Instructional Support				66	2.6	0.8			
CIRCLE: Teacher Literacy Focus				353	15.7	15.3			
CIRCLE: Student Literacy Engagement				353	22.8	18.0			
PCC: Vocabulary/Oral Language (Classroom level)				67	63.9	18.3			
PCC: Alphabet Knowledge				67	62.5	18.6			
PCC: Phonological Awareness				67	63.0	17.9			
PCC: Listening Comprehension				67	60.2	22.0			
CELF-P2: Core Skills Standard Score	193	86.6	18.6				189	91.4	17.7
PPVT-IV: Vocabulary Standard Score	205	89.5	19.6				194	95.5	17.3
TOPEL: Print Knowledge Standard Score	198	94.8	14.3				197	102.5	14.4
TOPEL: Phonological Awareness Standard Score	198	89.0	15.6				196	93.2	15.9

Note. Abbreviations are as follows: GRTR = Get Ready to Read, IGDI = Individual Growth and Development Indicator, CLASS = Classroom Assessment Scoring System, CIRCLE = Code for Interactive Recording of Children's Learning Environments, PCC = Preschool Curriculum Checklist, CELF-P2 = Clinical Evaluation of Language Fundamentals (Preschool), PPVT-IV = Peabody Picture Vocabulary Test, TOPEL = Test of Preschool Early Literacy.

Table 2
Pooled Fall-Spring Outcomes and Gains Overall, by Program Type, and Fall GRTR Risk Level.

Outcome	Program	Tier 1				Tier 2				Tier 3						
		Fall		Spring		Fall		Spring		Fall		Spring				
		M	SE	M	SE	M	SE	M	SE	M	SE	M	SE			
PPVT-IV	Overall	95.9	0.8	100.50	0.7	4.6	80.4	1.6	85.5	1.4	5.1	66.7	2.2	73.8	1.9	7.1
	Pre-K	93.1	1.3	97.3	1.2	4.2	74.8	2.5	80.4	2.1	5.6	61.5	3.4	68.0	2.8	6.5
	Head Start	95.7	1.9	102.9	1.6	7.2	82.3	2.9	88.4	2.6	6.1	72.9	3.9	82.8	3.4	9.9
	Title 1	94.9	1.5	98.3	1.3	3.4	83.3	2.8	87.8	2.4	4.5	67.1	4.4	70.7	3.7	3.6
	Tuition-Based	105.5	2.1	110.8	1.9	5.3	--	--	--	--	--	--	--	--	--	--
CELF-P2	Overall	92.1	0.8	96.8	0.7	4.7	75.6	1.6	83.0	1.4	7.4	66.0	2.2	71.6	2.0	5.6
	Pre-K	87.2	1.3	92.6	1.2	5.4	67.8	2.4	75.8	2.2	8.0	59.9	3.2	65.1	2.8	5.2
	Head Start	90.3	1.9	95.5	1.7	5.2	78.2	3.0	86.1	2.6	7.9	71.3	3.7	76.2	3.4	4.9
	Title 1	93.9	1.5	98.8	1.3	4.9	82.3	2.7	87.9	2.5	5.6	67.7	4.3	74.3	3.9	6.6
	Tuition-Based	103.4	2.1	105.7	1.9	2.3	--	--	--	--	--	--	--	--	--	--
TOPEL-PK	Overall	100.5	0.7	106.6	0.7	6.1	88.4	1.9	92.3	1.8	3.9	82.9	2.7	83.7	2.6	0.8
	Pre-K	99.6	1.1	109.4	1.0	9.8	83.2	2.0	94.1	1.8	10.9	83.4	2.5	87.6	2.4	4.2
	Head Start	94.8	1.5	98.3	1.4	3.5	84.8	2.3	85.4	2.2	0.6	82.0	3.1	82.1	2.9	0.1
	Title 1	99.5	1.2	108.3	1.1	8.8	86.2	2.2	95.0	2.0	8.8	78.9	3.5	84.6	3.2	5.7
	Tuition-Based	108.0	1.7	110.5	1.6	2.5	--	--	--	--	--	--	--	--	--	--
TOPEL-PA	Overall	93.3	0.8	99.2	0.9	5.9	84.0	2.3	90.6	2.6	6.6	74.7	3.2	85.8	3.6	11.1
	Pre-K	89.3	1.2	95.3	1.4	6.0	76.8	2.3	86.8	2.6	10.0	69.9	3.1	72.9	3.3	3.0
	Head Start	94.8	1.7	103.1	2.0	8.3	88.6	2.7	96.6	3.1	8.0	78.7	3.8	91.5	4.1	12.8
	Title 1	90.8	1.4	96.5	1.6	5.7	81.4	2.5	86.0	2.8	4.6	70.9	4.2	82.6	4.6	11.7
	Tuition-Based	98.3	2.0	102.0	2.2	3.7	--	--	--	--	--	--	--	--	--	--

Note. Abbreviations are as follows: PPVT = Peabody Picture Vocabulary Tests, CELF-P2 = Preschool Clinical Evaluation of Language Fundamentals, TOPEL = Test of Preschool Early Literacy, PA = Phonological Awareness, PK = Print Knowledge; -- = missing due to small n

Table 3
Growth in Picture Naming and Sound Identification Overall, by Program type and GRTR Risk Group.

Parameter	Tier 1			Tier 2			Tier 3		
	Est.	SE	p	Est.	SE	p	Est.	SE	p
<i>Overall</i>									
<i>IGDI Picture Naming</i>									
Mean intercept (initial status)	29.54	0.29	< .001	22.69	0.81	< .001	17.22	1.20	< .001
Mean slope	1.87	0.15	< .001	2.94	0.37	< .001	3.26	0.52	< .001
Mean intercept (final status)	33.29	0.27	< .001	28.58	0.66	< .001	23.75	1.21	< .001
<i>IGDI Sound ID</i>									
Mean intercept (initial status)	11.88	0.26	< .001	7.40	0.36	< .001	5.88	0.39	< .001
Mean slope	1.49	0.14	< .001	0.95	0.30	< .01	1.06	0.42	< .05
Mean intercept (final status)	14.86	0.24	< .001	9.29	0.48	< .001	8.00	0.68	< .001
<i>Pre-K</i>									
<i>IGDI Picture Naming</i>									
Mean intercept (initial status)	27.74	0.59	< .001	18.13	1.46	< .001	14.10	1.75	< .001
Mean slope	2.21	0.25	< .001	4.16	0.70	< .001	3.42	0.82	< .001
Mean intercept (final status)	32.16	0.53	< .001	26.46	1.19	< .001	20.93	1.72	< .001
<i>IGDI Sound ID</i>									
Mean intercept (initial status)	12.67	0.44	< .001	6.75	0.60	< .001	6.54	0.65	< .001
Mean slope	1.67	0.23	< .001	2.54	0.44	< .001	1.51	0.66	< .05
Mean intercept (final status)	16.01	0.38	< .001	11.82	0.72	< .001	9.55	1.01	< .001
<i>Head Start</i>									
<i>IGDI Picture Naming</i>									
Mean intercept (initial status)	32.00	0.49	< .001	24.07	1.63	< .001	22.60	2.31	< .001
Mean slope	0.82	0.35	< .05	2.26	0.65	< .01	2.25	1.13	< .05
Mean intercept (final status)	33.64	0.63	< .001	28.59	1.36	< .001	27.10	2.23	< .001
<i>IGDI Sound ID</i>									

<i>Parameter</i>	Tier 1			Tier 2			Tier 3		
	<i>Est.</i>	<i>SE</i>	<i>p</i>	<i>Est.</i>	<i>SE</i>	<i>p</i>	<i>Est.</i>	<i>SE</i>	<i>p</i>
Mean intercept (initial status)	10.39	0.57	< .001	7.91	0.69	< .001	5.31	0.67	< .001
Mean slope	0.84	0.34	< .05	-1.32	0.52	< .05	-0.22	0.58	> .05
Mean intercept (final status)	12.07	0.55	< .001	5.27	0.72	< .001	4.87	0.88	< .001
<i>Title I</i>									
<i>IGDI Picture Naming</i>									
Mean intercept (initial status)	29.39	0.39	< .001	26.11	0.79	< .001	17.87	1.41	< .001
Mean slope	2.03	0.25	< .001	2.08	0.53	< .001	3.10	0.63	< .001
Mean intercept (final status)	33.44	0.41	< .001	30.27	0.84	< .001	22.80	2.38	< .001
<i>IGDI Sound ID</i>									
Mean intercept (initial status)	12.00	0.44	< .001	7.53	0.60	< .001	6.03	0.48	< .001
Mean slope	1.63	0.27	< .001	1.09	0.46	< .05	0.81	0.48	> .05
Mean intercept (final status)	15.25	0.50	< .001	9.70	0.81	< .001	8.94	1.59	< .001
<i>Tuition-Based</i>									
<i>IGDI Picture Naming</i>									
Mean intercept (initial status)	32.57	0.40	< .001	-	-	-	-	-	-
Mean slope	1.52	0.32	< .001	-	-	-	-	-	-
Mean intercept (final status)	35.62	0.48	< .001	-	-	-	-	-	-
<i>IGDI Sound ID</i>									
Mean intercept (initial status)	13.91	0.69	< .001	-	-	-	-	-	-
Mean slope	0.82	0.30	< .001	-	-	-	-	-	-
Mean intercept (final status)	15.55	0.52	< .001	-	-	-	-	-	-

Note. Est. = parameter estimate; SE = standard error, - = not estimated due to small n in the Tuition-Based Tier 2 and 3 groups

Table 4

Classroom-Level (N = 65) Mid-Year Quality Ratings by Program Types.

Measure	Program Types											
	Pre-K			Head Start			Title 1			Tuition-Based		
	M	SD		M	SD		M	SD		M	SD	
CLASS: Emotional Support	5.3	0.5		6.0	0.6		5.2	0.6		4.5	1.4	
CLASS: Classroom Organization	5.2	0.9		5.6	0.7		4.8	0.9		4.1	1.3	
CLASS: Instructional Support	2.6	0.7		3.3	0.5		2.2	0.6		2.0	0.5	
PCC: Vocabulary/Oral Language	66.9	18.4		75.0	0.0		51.5	24.4		54.0	0.0	
PCC: Alphabet Knowledge	63.0	23.9		75.0	0.0		59.0	13.2		44.0	0.0	
PCC: Phonological Awareness	64.8	22.0		75.0	0.0		58.7	13.7		44.0	0.0	
PCC: Listening Comprehension	66.2	20.4		75.0	0.0		50.5	24.2		33.0	0.0	

Note. Abbreviations are: CLASS = Classroom Assessment Scoring System, PCC = Preschool Curriculum Checklist