

An evaluation of curriculum, setting, and mentoring on the performance of children enrolled in pre-kindergarten

MICHAEL ANDREW ASSEL, SUSAN H. LANDRY, PAUL
R. SWANK, SUSAN GUNNEWIG

*Department of Pediatrics, University of Texas Health Science Center-Houston, 7000
Fannin, Ste. 2300, Houston, TX 77030, USA*

Abstract. An alarming number of American pre-school children lack sufficient language and literacy skills to succeed in kindergarten. The type of curriculum that is available within pre-kindergarten settings can impact children's academic readiness. This work presents results from an evaluation of two language and literacy curricula (i.e., Let's Begin with the Letter People and Doors to Discovery) from a random assignment study that occurred within three settings (i.e., Head Start, Title 1, and universal pre-kindergarten) and included a control group. The design included a mentoring and non-mentoring condition that was balanced across sites in either curriculum condition. A pre and post-test design was utilized in the analyses, with children ($n = 603$) tested before the intervention and at the end of the year. Multilevel growth curve modeling, where the child outcomes (dependent measures) are modeled as a function of the child's level of performance and rate of growth between pre and post-testing, was used for all analyses. Results indicated that in many key language/literacy areas, the skills of children in classrooms using either one of the target curricula grew at greater rates than children in control classrooms. This was especially true in the Head Start programs. The findings from this study indicate that at-risk children can benefit from a well-specified curriculum. Additionally, findings demonstrate that a well-detailed curriculum appeared to be less important for children from higher income families. The impact of mentoring was less clear and seemed dependent on the type of skill being measured and type of program.

Key words: Curriculum evaluation, Early literacy, Head Start, Pre-kindergarten

Introduction

In a recent report on the portrait of American children entering kindergarten, increasingly large numbers of children enter school without the cognitive readiness skills needed to succeed (Zill & West, 2001). This occurs to a large degree because too many children are attending early childcare programs of such low quality that effective learning cannot occur (Helburn, 1995). This problem is particularly striking for children

from low-income households (Smith, Blank, & Collins, 1992). As a large body of recent research documents the critical importance of early experiences for learning (Bradley, Caldwell, & Rock, 1988; Landry, Smith, Swank, Assel, & Vellet, 2001; Molfese, 1989; Neville et al., 1998), we are beginning to understand that quality early childhood education is an important factor in changing this picture.

One factor that has a significant impact on the educational environment is the type of curriculum materials that are available within pre-kindergarten (pre-K) settings. The impetus that has been placed on curriculum evaluation studies can be demonstrated by recent grant initiatives sponsored by the Institute of Educational Science's (IES) Pre-school Curriculum Evaluation Research program (PCER). In the spring of 2002, IES awarded seven grants to evaluate the effectiveness of pre-school curriculum within random assignment studies. This IES funded study will detail the results of one of the PCER grantee's efforts in the first year of the project. These results will be different from the IES Federal Evaluation of curriculum, as they represent a larger sample size, utilize different outcome measures, and seek to answer different questions considered complementary to the overall goals of the PCER project. Specifically, the study was designed to be able to systematically evaluate a number of factors that the literature suggested may be important for understanding the effectiveness of early childhood education programs. As curricula are often used in a broad range of program settings that vary in terms of program goals, teacher backgrounds, and characteristics of the children enrolled in the particular program, this study evaluated curricula across three different types of programs (i.e., Head Start, Title 1 pre-k, Universal pre-K). Additionally, with the design of the study systematically varying conditions (e.g., intensity of professional development provided to teachers) and examining curriculum effectiveness across a broad range of skills, this study has the potential to help us better understand the complex interplay of factors that influence curriculum efficacy.

A pre-kindergarten experience is thought to be important due to the fact that children from impoverished homes routinely have underdeveloped language competencies (e.g., Juel, Griffith, & Gough, 1986; Hart & Risley, 1995) and lack age appropriate letter and phonological awareness skills (e.g., Lonigan, Burgess, Anthony, & Barker, 1998; MacLean, Bryant, & Bradley, 1987). Pre-kindergarten is thought to provide the opportunity to have more focused learning experiences in language and literacy areas (e.g., Burchinal, Peisner-Feingberg, Bryant, & Clifford, 2000). However, a pre-school experience, in and of itself, is not sufficient for school readiness. Interventions that target the use of effective language and early reading

activities show gains in these skills that carry into kindergarten (Whitehurst et al., 1994; Whitehurst & Lonigan, 1998; Zevenbergen et al., 1997). However, publicly funded programs have not placed a strong emphasis on the use of quality language and early literacy activities. Reports on children attending programs such as Head Start, where cognitive readiness has been less of a priority, demonstrate that children exit these programs at much lower levels than their non-poverty, beginning kindergarten peers (Lee, Brooks-Gunn, & Schnur, 1988). More recent evaluations of such programs show that when observations are made of teaching strategies in these types of classrooms, teachers give low priority to reading and writing goals and to the way children are involved in reading activities (Legislative Office of Education, 1998). Thus, lack of attention to stimulating language and literacy activities, in large part, explain low scores on language and emergent literacy tests for pre-school children.

Research has determined that pre-school instructional activities should be designed to help children develop language (e.g., expression and listening skills, vocabulary, phonological awareness) and early reading skills (e.g., read alouds, print concepts, letter knowledge, written expression, and sound sequencing) (Dickinson & Smith, 1994; Frede, 1995; Whitehurst et al., 1994; Whitehurst & Lonigan, 1998; Zevenbergen, et al., 1997) in order to ensure school readiness. When early childhood programs do attempt to include language and literacy activities, they often use poorly specified teaching approaches (Barnett, 1995). This is due, in part, to a lack of understanding of the need for specific instructional goals that can be supported by the use of well-detailed curriculum. With the national focus and available research on this important issue (Arnold, Lonigan, Whitehurst, & Epstein, 1994; Dickinson & Smith, 1994; Reynolds, 1995; Whitehurst et al., 1994), a number of research based curricula that target language and early literacy skills for pre-kindergarten age children have recently been developed. However, scientific evidence is lacking to determine whether these curricula result in educationally meaningful differences. Rigorous evaluations, such as the PCER program, are needed to allow early childhood programs to make informed choices concerning classroom curricula that have been developed to enhance language and literacy skills for this age group.

Pre-kindergarten programs typically occur within a variety of settings (e.g., state funded school districts with universal pre-K enrollment, federally funded Head Start Programs, Title 1 programs, or privately funded child care settings). Secondary to these different funding mechanisms, there appears to be a discrepancy in the types of curriculum materials that are available within settings. Additionally, teachers within these types of settings vary greatly in their level of education and

experiences. This is important, as recent research demonstrates the importance of the teacher as a critical agent of learning for children with well-planned, intentional instructional goals (International Reading Association, & National Association for the Education of Young Children, 1998) and avoidance of intensive drill and practice on isolated skills.

Another factor that is likely to influence whether or not a particular curriculum can demonstrate a significant impact on the academic readiness of children is the quality of the professional development activities that support curriculum implementation. It is generally agreed that a substantial investment in the training of those who work with children is required for programs to be successful (Bowman, Donovan, & Burns, 2001). Other research (Howes, Phillips, & Whitebook, 1992; Kontos, Howes, & Galinsky, 1997) has also reported that the quality of professional development provided to teachers is related to program success. One method that may be a particularly effective way to enhance teachers' skills is ongoing coaching, or mentoring (Spodek, 1996). In other words, a strong curriculum might not prove to be effective unless teachers have the ongoing support within the classroom to ensure effective implementation. This might be especially true for teachers with lower levels of education, training, and professional development. Research that has evaluated teacher training and professional development has indicated that teachers benefit from opportunities to try new approaches in supportive environments (Eisenhower National Clearinghouse for Mathematics and Science Education, 1998; Spodek, 1996). Mentoring/coaching is one way to achieve this supportive environment and may be important when implementing new curricula, as this requires understanding new instructional approaches and use of new sets of materials.

Two pre-kindergarten curricula that have been developed using current research and best practice strategies are *Let's Begin with the Letter People* (Abrams & Company, 2000), and *Doors to Discovery* (Wright Group, McGraw-Hill, 2001). Both curricula target those skills (e.g., print knowledge, phonological awareness activities, and language) that are most highly predictive of academic success in the early years of elementary school. These curricula were developed with sensitivity to teaching practices that are appropriate to the special learning needs of pre-school children based on the position presented by the International Reading Association and the National Association for the Education of Young Children (1998).

While school districts and Head Start programs across the country are routinely adopting curricula in pre-kindergarten programs, decisions surrounding which curriculum to adopt are typically being made without

evidence of a curriculum's efficacy from systematic investigations. Additionally, when effectiveness of pre-kindergarten programs has been evaluated, randomized clinical trial approaches have not been used, particularly with regard to effectiveness of specific language/literacy curricula. Randomization is critical in order to assure that results are not due to systematic differences that exist across classrooms and programs independent of the targeted curricula.

The primary objective of this study was to evaluate the effectiveness of two curricula (*Let's Begin with the Letter People* and *Doors to Discovery*) within multiple settings (Head Start, Title 1, and Universal pre-K classrooms). Both curricula were expected to result in greater gains in child outcome measures evaluating language and literacy skills when compared to the control classrooms. Additionally, as each of these curricula differs in their emphasis on language versus literacy skills, we expected differences on the impact that each would have on child skill areas. As *Let's Begin with the Letter People* has a particularly strong emphasis on letter knowledge and phonological awareness, it was expected to result in stronger gains in these skills when compared to classrooms using *Doors to Discovery*, or the control condition. Given *Doors to Discovery's* strong emphasis on language, it was expected to result in greater gains in vocabulary and language comprehension compared to *Let's Begin with the Letter People*, or the control condition.

Another central question of interest to this study was the relation between the particular site (Head Start versus classrooms within a public school district) and the curricula's effectiveness. As described above, Head Start differs from independent school districts in the educational backgrounds of teachers (i.e., 4 year degrees not required of teachers) and serves families and children from a deeper poverty level. As teachers with less education and training were expected to benefit the most from a systematic set of instructional goals and materials, differential effects on the rates of growth on child outcomes were expected. Specifically, greater gains were expected in Head Start classrooms using either language and literacy curriculum examined in this study compared to either the Title 1 or universal pre-kindergarten classrooms within school districts. In relation to the impact of ongoing professional development for teachers, mentored classrooms were expected to produce greater gains in child outcomes than non-mentored classrooms. Finally, the impact of mentoring on child outcomes was expected to be greater in Head Start classrooms, where teachers typically had lower levels of educational attainment and were expected to benefit to a greater extent from individualized support.

Method

Participants

The participants were 603 typically developing pre-kindergarten children (308 boys and 295 girls) enrolled in full day early childhood programs where English was the language of instruction. Children and their teachers were recruited from two primary organizations. The average age of children enrolled in the study was 4.6 years of age at the midpoint of the school year. Title 1 and universal pre-K classes were selected from a fairly large school district in greater Houston, Texas. At the time of recruitment, the school district contained 40 schools and had 31,599 enrolled children and was selected as it had naturally occurring groups of children of interest to this study. The district is economically diverse, supporting a large Title 1 pre-kindergarten population (26 classrooms), as well as offering universal pre-kindergarten for all of the 4-year-old children, irrespective of income (19 classrooms). The classrooms are distributed across 22 ethnically diverse elementary schools. The free-standing Head Start Centers (31 classrooms) in the study were recruited from a large non-profit agency that provides a broad range of services to the community of Houston and includes 10 Head Start school sites, each with multiple classrooms. As the goal of this investigation was to determine whether the two curricula were effective in different types of early childhood programs serving different populations of children, the differences across program types were not considered a confound to the study but rather necessary to answer our research questions.

Both programs allowed researchers to randomly assign treatment conditions. Across all three program sites, school sites were randomly assigned into one of three conditions (i.e., *Let's Begin with the Letter People*, *Doors to Discovery*, or Control). Figure 1 illustrates the experimental design of the study. After random assignment into curriculum condition, those school sites receiving one of the two curricula were randomly assigned to receive mentoring versus no mentoring. Random selection was done by school site, rather than individual classroom, in order to assure that teachers across the three curricula conditions did not share information.

Across sites, classroom size ranged between 15 and 20 students per class. There were 245 children enrolled in the Head Start program, 213 Title 1 children, and 145 children in the universal pre-K program. Ethnicity information for children is presented in Table 1. There were differences in terms of the ethnic backgrounds of the children enrolled in each program. For instance, the children enrolled in universal pre-K were

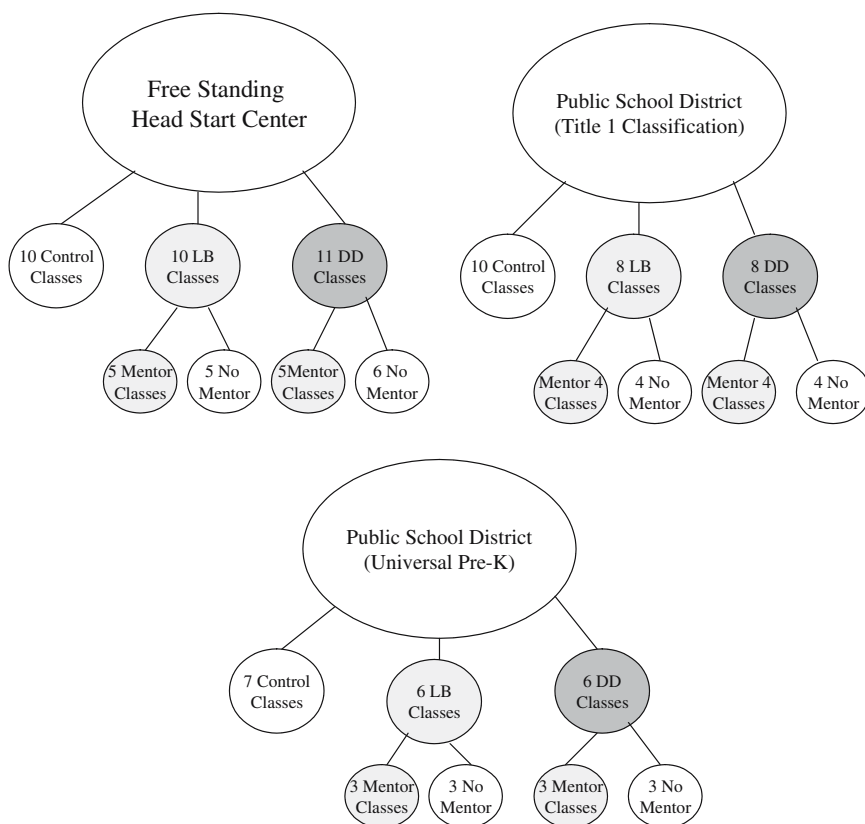


Figure 1. Experimental design.

primarily Caucasian, as compared to only approximately one fourth of the children enrolled in Title 1 programs, and 7% of Head Start children. Hispanic children were equally prevalent in the Head Start program (51%) and Title 1 program (52%) but were a much smaller proportion of

Table 1. Ethnicity of the sample by site and associated percentages.

Site	Child ethnicity				Total
	AA	His.	Cau.	Other	
Head Start	93 (38)	127 (51)	16 (7)	9 (4)	245
Title 1	25 (12)	112 (53)	58 (27)	18 (8)	213
Universal Pre-K	6 (4)	15 (10)	103 (71)	21 (15)	145
<i>n</i>					603

Note: AA, African American; His., Hispanic; Cau, Caucasian. Numbers in () are percentages.

children enrolled in universal pre-K programs (10%). African American children comprised 37% of the children enrolled in Head Start, 11% of children enrolled in Title 1 programs, and only 4% of the children enrolled in universal pre-K. Additional similarities between the programs included the fact that both programs (Head Start & Independent School District) had a teacher and aide in each classroom.

One difference between sites in the project was their approach to children identified as English Language Learners (ELLs). In the school district, monolingual Spanish speaking children are typically enrolled in what the district classifies as “bilingual classrooms.” No bilingual classrooms were recruited for this study. Thus, children in the Title 1 and Universal pre-K classrooms within the school district who were of Spanish ethnicity were considered ELLs or English speakers and enrolled in classrooms where English was the language of instruction. In the Head Start site, children from Spanish speaking homes were instructed in English by the Head Start staff and in Spanish by a bilingual partnering teacher from local school districts. Therefore, the children in the study from the Head Start site had language and literacy concepts presented in both languages.

Teacher ethnicity, education, and certification information is provided in Table 2. Teachers across all sites were predominantly female but their ethnicity varied. The majority of teachers within the Head Start site were African American and those in the other two programs, Caucasian. Teacher education varied with almost all of the teachers in the independent school district having 4-year college degrees. In contrast, less than half of the Head Start Teachers had a 4-year college degree or a Child Development Associate’s degree. Two of the teachers within the Head Start site had only a high school degree. As expected, teachers within public school settings had more teaching certificates. Additionally, almost all of the teachers in the public school district’s Title 1 programs had an additional certification in the area of English as a Second Language (ESL), while only a majority of the teachers in the universal pre-K held the ESL certification. Finally, more than half of Head Start Teachers reported that they had no certifications. In summary, it is clear that the teachers within the public school settings (Title 1 and universal pre-K) had more education, were predominantly Caucasian, and had more certifications than teachers in Head Start.

Incentives

Sites that participated in the study were provided some incentives for their participation. Specifically, teachers received 4 days of teacher training

Table 2. Treatment and control teachers' ethnicity, educational level, and certifications in percentages.

	Head Start	Title 1	Universal pre-K
<i>Ethnicity</i>			
African American	71	0	0
Hispanic	13	0	11
Caucasian	6	100	84
Other	10	0	5
<i>Education</i>			
High school	6	0	0
CDA	39	0	0
2 Year	10	0	5
4 Year	39	81	79
Graduate	6	19	16
<i>Certification^a</i>			
None	58	0	0
Teaching certificate	13	92	84
SPED	3	15	10
ESL	3	88	58
<i>n</i>	31	26	19

Note: CDA, Child Development Associate's degree; SPED, Special Education Certification; ESL, English as a Second Language Certification.

^aAs teachers can have more than one type of certification, the percentages in the table do not total 100%.

specific to the curricula and were provided a complete set of materials for either *Let's Begin with the Letter People* or *Doors to Discovery*. Finally, administrators at both sites received a comprehensive summary report surrounding the language and literacy skills of the children enrolled in their programs (i.e., classroom based data on pre-test and post-test scores of the children enrolled in the study). In order to avoid potential confounds, teachers were not provided data surrounding how individual children scored on assessments, nor were they provided with information surrounding how the students in their classroom compared to other classrooms in the project.

Prior to the start of the project, extensive communication with program directors (e.g., Assistant Superintendent of Early Childhood Programs and the Head Start Program Director) and principals from each school were briefed regarding classroom identification, randomization, and planning for program implementation. Because previous experience has demonstrated that "buy in" from all levels of administration is critical and can be

best accomplished when there is a sense of ownership of the program, the senior research team met monthly across the year with head administrators of each program. Each major site also assigned a project coordinator to work with researchers to understand the design/scope of the project.

Curriculum

There were several similarities across both curricula including activities that focused on building skills in letter knowledge, phonological awareness, as well as language and motivation to read. For both curricula, skill building in these areas occurred across multiple contexts (e.g., circle time, center activities, small and large group activities, and read alouds) with a set of materials (e.g., small and big books, fiction and non-fiction books, family readers, and comprehensive teacher manuals). Differences across the two curricula included a greater focus in *Let's Begin with the Letter People* (abbreviated as *Let's Begin* in the rest of this report) on letter knowledge and early skills across the phonological continuum (e.g., rhyming). Both curricula also utilize similar instructional approaches (i.e., print rich materials, book reads that introduce new vocabulary, guided instruction within small group activities where children learn phonological awareness skills at an age-appropriate level, and language building within thematic units) that have been shown to be an effective method for early childhood education programs (e.g., Biemiller, 2003; Senechal, 1997; Snow, Burns, & Griffin, 1998; Whitehurst et al., 1994). The expectation that both curricula will support enhanced print concepts is supported by research documenting that teachers' encouragement of children's involvement with books, particularly predictable books, results in stronger print concept skills than a comparison group (Box & Aldridge, 1993). Both curriculum avoided didactic highly structured approaches, but rather emphasized children learning in meaningful contexts in highly engaging "hands on" activities (e.g., use of story extenders, recognizing and using rhyming words, tracking print from left to right, recognizing and naming letters, and predicting story events). Both curricula include home components that contain materials for parents describing the activities occurring in the classroom and reading materials.

In *Let's Begin*, letter knowledge activities were encouraged through the use of the Letter People "Huggables." Additionally, story books about each "Huggable" were presented to children within circle time activities. Throughout the *Let's Begin with the Letter People* (Abrams & Company, 2000) the curriculum focuses on development of letter knowledge in multiple contexts and a set of research based activities that support

children's development of phonological awareness. The *Let's Begin* curriculum has 26 thematic units and teachers typically devote at least one week to each unit. Each of the 26 units has activities divided across seven domains: (1) Oral Language/Listening; (2) Alphabetic/Story Knowledge/Writing; (3) Science/Math; (4) Personal/Social Development; (5) Large/Small Motor Skills; (6) Art/Music; (7) Reach the Home. The program is organized into "lessons" which provide a daily menu of activities, which begin with a short full group activity (i.e., Meeting Circle), followed by a selection of individual and small group "Interest Center" activities. When one looks at a particular thematic unit (e.g., Explore with Mr. N), the teacher is directed to focus on activities associated with children's names (e.g., books that focus on letter recognition within the context of first names), phonological awareness (rhyming and alliteration via nursery rhymes and songs), integrated vocabulary (e.g., name, circle, capital letter, lowercase letter), and interest centers that reflect developmental areas such as oral language, listening, alphabet and story knowledge, math, science, art, drama, music, and social development. In terms of specific activities that relate to a given skill area (e.g., phonological awareness/phonemic awareness), The *Let's Begin* curriculum is structured to expose children to the increasing levels of phonological sophistication. During the first portion of the pre-K year, phonological activities focus on listening, rhyming, and word play. However, by the mid-point of the year, children are being exposed to alliteration activities, and by the end of the year teachers begin to concentrate on syllabication activities.

The *Doors to Discovery* (The Wright Group, McGraw Hill, 2001) curriculum also includes multiple aspects of early literacy development, but it has a major focus on the development of vocabulary and receptive and expressive language development. The topics and complexity of the thematic units are rich with information and opportunities for children's discussion. The teacher guide for *Doors to Discovery* provides specific open-ended statements and questions to promote discussion. This approach is consistent with research documenting greater vocabulary gains for children whose teachers followed book reading with extended discussion versus teacher-child interactions that focused on recall of factual information (Dickinson & Smith, 1994). *Doors to Discovery* was designed in ways to encourage children's literacy development across five areas: oral language, phonological awareness, concepts of print, alphabet knowledge, writing, and comprehension. For example, in order to improve children's oral language skills, *Doors to Discovery*, encourages teachers to use techniques to further children's language skills (e.g., cloze techniques, student retelling, think alouds, and scaffolding). *Doors to Discovery* utilizes eight distinct themes (e.g., Build it Big, Discovery

Street, Our Water Wonderland) that occur across multiple weeks of the school year. While there were many similarities between curricula, some differences are worthy of note. For instance, the *Doors to Discovery* curriculum includes a greater emphasis on development of a monthly theme. In contrast, *Let's Begin with the Letter People* is less reliant on establishment of well-defined themes. Other differences include the fact that *Doors to Discovery* has a stronger focus on small group activities in which teachers, along with help from the students in the classroom, build a scrapbook that details the theme. The *Let's Begin with the Letter People* curriculum was somewhat easier for teachers to implement as it included a structured teacher's guide that clearly specified learning objectives and classroom activities. In contrast, *Doors to Discovery* required teachers to use several sets of materials.

Professional development

The target teachers were trained in their randomly assigned curriculum across a 4-day workshop during the summer of 2002 by the publishing companies of each curriculum. Training occurred within small groups and included instruction/experience in all content areas. Research staff also attended all trainings to familiarize teachers with the multiple study requirements entailed in the local and national evaluation components of the grant. This report is concerned with only the local or "complementary" research questions as the national evaluation data is reported elsewhere. The Institute of Education Sciences (i.e., the research arm of the Department of Education) had the responsibility for completing the cross site evaluation of curriculum. Specifically, the national evaluator (i.e., RTI, Inc.) had responsibility for conducting classroom observations and child testing on a much smaller sample of the classrooms enrolled in this study (i.e., 42 vs. 82), as well as classrooms across six other universities. Additionally, the current report is interested in the effect of the use of mentoring and curriculum effects across three types of programs, two questions not addressed in the national study. The national evaluation component had no responsibility for training, mentoring, or ensuring that the curriculum was being implemented effectively and it is unlikely that the involvement of the national evaluator had any effect on the results reported below.

The training approach was learner-centered and knowledge-based, providing information in ways that build on previously learned experiences (Bransford, Brown, & Cocking, 2000). The training model highlighted aspects of the two target curricula that are important for supporting language and literacy development as well as responsive teaching practices that encouraged strong social/emotional skills.

Educational environment in control classroom

The main characteristic of control classrooms at both sites (Head Start and the School District) was the lack of a specified curriculum that included a scope and sequence of activities. Administrators at the school district reported that the control classrooms used a variety of classroom materials (e.g., children's literature from multiple publishers and materials that the district had developed over the course of several years). Personnel from the school district reported that materials provided to teachers in the control classrooms followed the state guidelines for pre-kindergarten classrooms and included attention to language and literacy.

The control classrooms in the Head Start site were much less systematic in terms of materials used. Some classrooms used pieces of outdated curriculum. Additionally, many of teachers of the control classrooms utilized worksheets (derived from different sources) and other materials that were developed within the particular centers. Similar to the control classrooms within the school district, Head Start control classrooms did not utilize a curriculum with a specified scope and sequence.

Mentoring

Curriculum mentors supported target teachers within the mentoring condition in terms of implementation of *Let's Begin* or *Doors to Discovery* Curriculum. Three mentors were able to provide ongoing technical assistance to teachers. Mentors were all senior level trainers, intimately familiar with the curricula, who encouraged the systematic implementation of the curriculum within a context that recognized children's social-emotional needs. Mentors worked with teachers for approximately 1.5 h (two times per month). During their visits, mentors assisted teachers with lesson planning, demonstration of specific curriculum components, room arrangement, curriculum fidelity issues, classroom schedules, provided assistance to teachers surrounding behavioral issues that might arise during classroom instruction, and provided side-by-side coaching while the classroom teacher was implementing the curriculum. The mentor documented improvements ("glows") and areas still needing improvement ("grows") and discussed these points with individual teachers.

Ensuring curriculum fidelity

Mentors for each curriculum completed Curriculum Fidelity Checklists (separate instrumentation was designed for each curriculum) three times

over the course of the academic year. Due to the fact that each curriculum has very specific components, mentors were not blind to experimental condition. However, mentors did not complete curriculum fidelity measures on teachers with whom they were currently mentoring. The curriculum fidelity checklists also were administered within the non-mentoring condition, as this was the only way to determine if classrooms in each curricula condition were effectively using the curriculum as designed. Mentors provided all teachers with information following completion of the fidelity checklist surrounding how well they appeared to implement particular components of the curriculum. To document the extent to which the curriculum was being implemented, all teachers received feedback surrounding their implementation throughout the course of the year (i.e., three feedback sessions for teachers in the non-mentoring condition). Therefore, another way to describe the mentoring versus non-mentoring conditions would be classifying them as one group receiving ongoing mentoring/coaching versus periodic feedback sessions for the non-mentoring group. While it is recognized that mentoring can take many forms and ranges from periodic consultation to intensive ongoing side-by-side coaching in classrooms, the mentoring approach used within the current study is more intense than teachers in the area receive prior to implementing a new curriculum. Researchers determined that the provision of feedback to the non-mentored group would assist in terms of overall fidelity to curriculum guidelines and allow for a fair comparison of the curricula.

To estimate reliability for the Fidelity Checklist, approximately 10% of observations were coded by a second rater. Using repeated measures analysis of variance, generalizability coefficients were calculated to determine inter-rater reliability for the Fidelity Checklist (Fleiss, 1986). This method is recommended for studies using continuous behavioral observational data and has the advantage of evaluating both the consistency across a variable for each rater and the variance across participants for those variables used in the analyses (Frick & Semmel, 1978). Coefficients above .50 indicate adequate reliability (Mitchell, 1979). Generalizability coefficients revealed that the reliability of the entire scale was adequate (i.e., *Let's Begin with the Letter People* = .80; *Doors to Discovery* = .76). In general, teachers were implementing key components of each curriculum at high levels. Additionally, results demonstrated that there was a great deal of growth in terms of curriculum fidelity scores. At the first evaluation, 29.8% of teachers using the *Let's Begin* curriculum scored at high levels (i.e., scores of 4 or 5 on a 5 point scale with higher scores indicating better implementation). By mid-year, 71.5% of teachers were receiving high scores on the *Let's Begin*

Curriculum Fidelity measure. In general, ratings of curriculum fidelity for *Doors to Discovery* were good, but not as strong as *Let's Begin*. For instance, 28.6% of teachers using *Doors to Discovery* scored in the top range on the Doors Curriculum Fidelity Checklist at the beginning of the year, and 59.6% of teachers scored at the highest levels by the mid-point of the year. Reports from teachers in the study indicated that *Let's Begin with the Letter People* was somewhat easier to implement because it included a more user friendly teacher guide. In contrast, it was necessary for teachers using *Doors to Discovery* to rely on several teacher guides in order to cover all skill areas.

Comparability of classroom environments and teaching practices across curriculum conditions

In order to better understand potential variance across classroom environments and conditions for control and target classrooms, we examined a range of teaching behaviors and classroom practices for a randomly selected group. Thirty-three classrooms evenly distributed across the three conditions and program sites were observed using the CIRCLE-Teacher Behavior Rating Scale (Landry, Crawford, Gunnewig, & Swank, 2002). This measure contains 11 subscales (total of 50 items) that capture responsive teaching practices, key language and literacy instructional areas, the use of lesson plans and progress monitoring, as well as classroom structure and organization. In order to complete this measure, observations of classrooms for approximately 1½ to 2 h were conducted by research staff blind to study conditions in the fall, winter, and spring. Inter-rater reliability using generalizability coefficients were high ranging from .80 to .98. Internal consistency was found to be high (i.e., .96). Although there were significant interrelations between subscales, they were not so highly correlated that the information was redundant. Results of these observations revealed that in spite of random selection, control classrooms showed higher overall scores on this measure compared to targets. Additionally, observations made in Head Start classrooms showed overall lower scores than those in Title 1 classrooms, or universal pre-K.

Child outcome measures

Child outcome measures included five standardized measures of pre-academic skills. As each curriculum could be considered to be focused on early literacy, attempts were made to build an assessment battery that that provided information surrounding key language/literacy outcomes.

During the start up phase of the project, there were limited options surrounding a comprehensive early language and literacy assessment. Therefore, a decision was made to use multiple assessments procedures that evaluated critical literacy components. As approximately half of the children enrolled in the project would also be enrolled in the national evaluation component of the PCER grant, efforts were made to avoid duplication of assessment procedures. The national evaluator was not able to begin assessing children until mid-November. In contrast, the results reported in this paper are based on pre-testing that was completed at the beginning of the school year. The selection of measures was based on what research has determined are the key skills that predict later reading competency (i.e., language, phonological awareness, and print knowledge) (e.g., Snow et al., 1998).

Pre-school language scale-IV edition—auditory comprehension subscale (PLS)

The PLS-IV's Auditory Comprehension subscale was used as measure of complex language understanding (Zimmerman, Steiner, & Pond, 2002). The Auditory Comprehension subscale specifically is used to assess a child's attention to people, sounds, and objects in the environment, play behavior, and comprehension of: basic vocabulary, gestures, quantitative, qualitative, and time/sequence concepts, morphological and syntactic structures, inferencing, categorizing objects, and phonological awareness skills. These are all specific skill areas that are highlighted to some degree within the target curriculum. The PLS-IV Auditory Comprehension subscale evaluates a child's understanding of complex language forms (e.g., structure, grammar, and syntax) as well as receptive vocabulary. Test-retest reliability of the PLS-IV's Auditory Comprehension subscale for children between the ages of 4 years and 5 years 11 months ranged from .83 to .91 with an average interval between testing of 5.9 days. Internal consistency reliability coefficients for the Auditory Comprehension Subscale for the same ages ranged from .83 to .90 (Zimmerman et al., 2002).

Expressive vocabulary test (EVT)

The EVT is an individually administered, norm-referenced assessment of expressive vocabulary and word retrieval skills appropriate for ages 2½ through adulthood. The EVT evaluates expressive vocabulary skills through two types of items: labeling (What do you see?) and synonym (Tell me another word for ____). The authors report test-retest reliability of .77 for the age range of children in this study with an interval between tests that ranged from 8 to 203 days (Williams, 1997).

Developing skills checklist (DSC)

The DSC is an individually administered assessment of the range of skills that children develop from pre-K through kindergarten. The DSC measures a full range of child skills and behavior that children typically develop (CTB/McGraw-Hill, 1990). The Auditory subscale of the DSC was used in this study as an evaluation of children's skills in the area of phonological awareness. The Auditory subtest evaluates the following skills: recognition of words that sound different, rhyming, sentence segmentation, and syllabication. The DSC technical manual reports a Kuder–Richardson 20 reliability coefficient of .84 for the Auditory subtest. Additionally the standard error of measurement for the Auditory subtest is 1.87 (CTB/McGraw-Hill, 1990).

Woodcock–Johnson test of academic III tests of achievement (WJ-3)

The Letter Word Identification and Sound Awareness/Rhyming subtests were used from the WJ-3 (Woodcock, McGrew, & Mather, 2001). The WJ-3 series is a widely used measure with excellent psychometric properties. The standardization sample is large ($n = 8818$) and is weighted toward the younger ages. The test provides the user with a myriad of scores including standard scores, age equivalent, and grade equivalents. The Letter Word Identification subtest requires that children identify specific letters and words via both multiple choice and free response formats. Test–retest reliability for the Letter Word Identification subtest was reported to be .98 (McGrew & Woodcock, 2001). The WJ-3's Sound Awareness subtest was also used in this evaluation. This subtest was piloted in other ongoing studies of 4-year-old children and it was determined that the majority of 4-year-old children had significant struggles with the measure once it moved from the Rhyming section to the subsequent sections (i.e., Deletion, Substitution, and Reversal). Therefore, only the Rhyming section was administered during the pre- and post-test evaluations. WJ-3 Technical Manual reports test–retest reliability of the Sound Awareness subtest as being .71 for 4-year-old children.

Design and analysis plan

The $3 \times 2 \times 2$ factorial design for this study included site (HS, Title 1, universal pre-K), by curriculum type (*Let's Begin with the Letter People* and *Doors to Discovery*), mentoring versus no mentoring as another independent variable, and site as a control variable. A pre- and post-test design was utilized in the analyses, with children tested before the intervention and at the end of the year. Schools were randomly assigned to

treatment conditions in order to avoid contamination within sites. A repeated measures design was not used as children were different ages within even the same classroom and not all children were tested at equal intervals. Thus, in order to handle these conditions, we used a multilevel growth curve model where the child outcomes (dependent measures) are modeled as a function of the child's level of performance and rate of growth between pre- and post-testing. Multilevel analysis is a methodology for the analysis of data with complex patterns of variability, with a focus on nested sources of variability (e.g., pupils in classes, employees in firms). In the analysis of such data, it is important to take account of the variability associated with each level of nesting. As there is variability between pupils and also between classes, inappropriate conclusions might be made if either of these sources of variability is ignored (Snijders & Bosker, 1999). The level of the child's performance is represented by estimating an intercept (age centered at the mean age of 4.6 years) and rate of growth is estimated by a slope (rate of change at 4.6 years). Because there were only two time points assessed, the slope had to be estimated as a fixed effect whereas the intercept term was estimated as a random effect. When there are level (or intercept) differences in the absence of slope findings, these explain differences between the group at the 4.6 year level. However, when significant slope findings are present intercept findings are less meaningful because slope findings indicate that level differences are changing over time.

To address the effects of interest, four contrasts were specified for the grouping variable and two for the site variable. For the groups, all of the target groups were compared to controls, the *Doors to Discovery* curriculum to *Let's Begin*, the mentored condition to non-mentored condition, and the interaction of the curriculum by mentoring conditions. As child ethnicity was confounded with site (i.e., larger numbers of Hispanic children in HS and Title 1 versus universal pre-K) controlling for site in this design in essence controls for ethnicity. Although we were not interested in overall level differences in child skill areas across the three program types (HS, Title 1, universal pre-K), results consistently revealed that at the 4.6 year level (intercept), universal pre-K children showed higher scores than the other two programs, and Title 1 scores were higher than those found for Head Start children. Because the primary interest was in identifying differences in the rate of growth of child skills over time, when both intercept and slope effects were found, only slope effects are reported. For some analyses, sites were collapsed. That is, Title 1 and Head Start classrooms were compared to the universal pre-K classrooms. In some instances, Head Start classrooms were contrasted with Title 1 classrooms. SAS Proc Mixed (SAS, 2001) was used to complete all data analyses.

Results

Language comprehension findings

For the PLS-IV Auditory Comprehension subscale, there were significantly greater gains for children in classrooms utilizing a language-literacy curriculum than for those in the control group. However, this was moderated by program site. Overall, as hypothesized the groups using either *Let's Begin* or *Doors to Discovery* were increasing in language skills at a faster rate than controls, $t(1061) = 2.61$; $P = .0092$; $d = .18$, but as expected this was more true in the Head Start classes than in Title 1 classrooms or universal pre-K, $t(1061) = -7.20$; $P < .0001$; $d = .86$. Gains on the PLS-IV Auditory Comprehension subscale were not significantly different from controls in Title 1 and universal pre-K classrooms. Only the Head Start classes using one of the target curricula outperformed controls on the PLS-IV. The d statistic is Cohen's d (Cohen, 1988) and is determined based on the value of the contrast (with changes over time estimated for a 6 month difference) divided by the estimated standard deviation, which is based on the total of the random variance terms in the model.

There was also a significant curriculum by mentoring interaction between Head Start and Title 1 classes, $t(1061) = -2.98$; $P = .0030$; $d = 1.16$. On the average there were greater increases in language skills in the *Doors to Discovery* classrooms compared to *Let's Begin*, but Title 1 classes showed slower growth if they were mentored and using *Doors to Discovery* or non-mentored and using *Let's Begin* when compared to controls.

Thus, while the target Head Start classrooms had significantly lower language scores than the Title 1 and universal pre-K classrooms at 4.6 years of age, language skills were increasing more rapidly for Head Start target classrooms relative to controls than observed for Title 1 or universal pre-K classes. Secondly, while target classes were improving relative to controls overall, the difference was greater in the Head Start classrooms. Examination of the graphs indicates that there was less growth overall in the universal pre-K classrooms. Slower rates of growth in universal pre-K classrooms are likely attributable (in part) to the higher level of these classrooms at the pre-test. Figure 2 illustrates the rates of growth on child outcomes for Auditory Comprehension as well as the EVT and the Letter Word Identification Subtest, in the Head Start and Title 1 classrooms. Additionally, as there were no slope findings of interest for the universal pre-K classrooms on these variables, graphs for the universal pre-K classrooms are not presented.

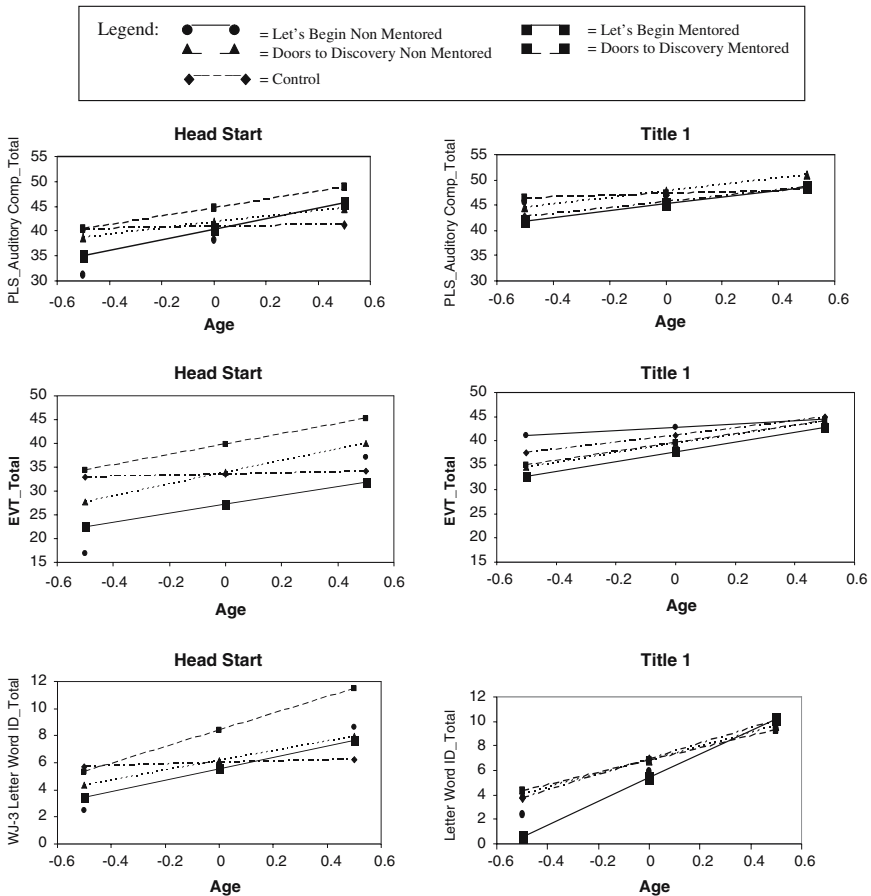


Figure 2. Rates of growth on child outcomes for PLS-IV Auditory Comprehension, EVT, and WJ-3 Letter Word Identification Subtests in Head Start and Title 1 classrooms.

Vocabulary findings

Results from the EVT indicate that the significant differences between classrooms using a language and literacy curriculum versus the controls on their overall growth rates in expressive vocabulary was moderated by the program site. That is, children in classrooms receiving either *Let's Begin* or *Doors to Discovery* grew at faster rates in vocabulary than those in control classrooms but this was particularly true in the Head Start and Title 1 classrooms versus universal pre-K classrooms, $t(1066) = -5.82$; $P < .0001$, and Head Start versus Title 1, $t(1066) = -4.33$; $P < .0001$. Title 1 classes also did better than the universal pre-K classes, $t(1066) = -3.00$; $P = .0028$. The effect size obtained for Head Start

classrooms was .68 compared to .04 for Title 1 classrooms and $-.52$ for Universal pre-K classrooms.

There was also a significant effect of mentoring versus non-mentoring but this depended on the type of curriculum and was moderated by program site, Head Start versus Title 1, $t(1066) = -2.30$; $P = .0215$. Specifically, while greater growth rates were seen in Head Start classrooms with the targeted curriculum ($d = .74$) compared to the other two program types (Title 1 ($d = .42$) and universal pre-K ($d = .08$)), the Title 1 classrooms receiving the *Doors to Discovery* condition showed greater growth irrespective of mentoring versus non-mentoring ($d = -.01$) while for *Let's Begin* the growth was greatest in the mentored condition ($d = .36$) (See Figure 2).

Print knowledge findings

On the Letter Word Identification subtest (WJ-3) there was a significant treatment versus control effect, $t(1062) = 4.10$; $P < .0001$; $d = .28$) but this was greater in Head Start ($d = .53$) than either the Title 1 ($d = .06$) or universal pre-K ($d = .25$) classrooms, $t(1062) = -3.20$; $P = .0014$. There was a curriculum effect of *Let's Begin* over *Doors to Discovery*, $t(1062) = -3.63$; $P = .0003$; $d = .31$, but this varied upon whether classrooms were mentored or non-mentored and this was moderated by program site, $F(8, 1062) = 3.82$; $P = .0002$. Specifically, the Title 1 and universal pre-K classrooms receiving the *Let's Begin* curriculum with mentoring did better on the print knowledge measure than those without mentoring. The classrooms in Title 1 and universal pre-K using *Doors to Discovery* did somewhat better without mentoring, while the Head Start classrooms with *Doors to Discovery* and mentoring did better, $t(1062) = -3.41$; $P = .0007$. The effect sizes obtained on the WJ-3 Letter Word Identification subtest were .41 for the Head Start classrooms but $-.38$ for the Title 1 and -1.13 for the universal pre-K classrooms. However, in Head Start classes the type of language-literacy curriculum and the presence or absence of mentoring was not as important as receiving such a curriculum. Thus in the Head Start Program, children with either language-literacy curriculum outperformed controls in print knowledge skills (See Figure 2).

Phonological awareness findings

For Rhyming (WJ-3 Sound Awareness) there was a significant difference between slopes for classrooms receiving a language and literacy

curriculum versus controls, $t(1062) = 3.54$; $P = .0004$; $d = .26$. There were also differences between the two language and literacy conditions, $t(1062) = -3.14$; $P = .0018$; $d = .25$, that were moderated by program site, $t(1062) = -4.12$; $P < .0001$. While children in classrooms receiving a language and literacy curriculum showed greater gains in rhyming skills than those in control classrooms, this was particularly true for *Let's Begin* compared to *Doors to Discovery* but only in the universal pre-K classrooms. Specifically, universal pre-K *Let's Begin* classrooms (i.e., mentored or non-mentored) outperformed *Doors to Discovery* classrooms in rhyming by an effect size margin of .85. In contrast, for both Head Start and Title 1 classrooms, children receiving either of these research-based curricula showed greater growth when compared to controls. Due to the fact that there were no significant differences in the rates of growth for Title 1 and Head Start classrooms on the WJ-3 Rhyming Subtest, these groups were combined to better illustrate the relations graphically. In Head Start and Title 1 programs, classrooms using either curriculum outperformed control classrooms in rhyming by an effect size margin of .20. Figure 3 shows the WJ-3 Rhyming results for the curriculum conditions by universal pre-K classrooms versus Title 1 and Head Start classrooms combined.

For the Auditory Subscale of the DSC, also evaluating aspects of phonological awareness skills, there was a significant slope effect for classrooms receiving a language and literacy curriculum versus those who did not, $t(1070) = 3.44$; $P = .0006$; $d = .26$, but this was moderated by program site, $t(1070) = -2.61$; $P = .0091$. Children in classrooms receiving either language–literacy curriculum in Head Start ($d = .48$) and Title 1 programs ($d = .34$) showed greater gains in phonological

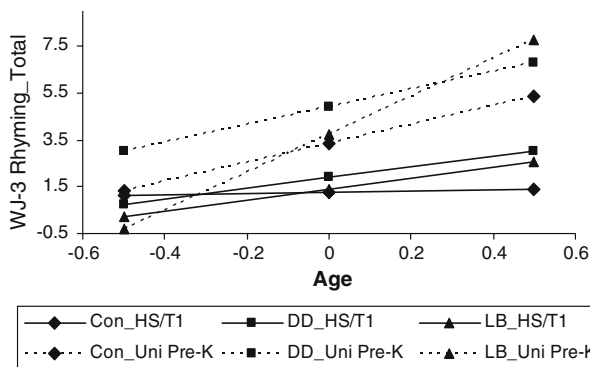


Figure 3. WJ-3 Rhyming results for the curriculum conditions by universal pre-K classrooms versus Title 1 and Head Start classrooms combined.

awareness than controls in these programs when compared to target versus control classrooms in universal pre-K classrooms ($d = -.04$). There was also a significant effect of curriculum, $t(1070) = -2.61$; $P = .0092$; $d = .24$, with *Let's Begin* outperforming *Doors to Discovery*. In addition, there was a mentoring effect that was moderated by program site, $t(1070) = 3.35$; $P = .0008$. Title 1 classrooms that received mentoring showed significantly greater gains, irrespective of type of curriculum, than classrooms with a curriculum without mentoring ($d = .45$). The mentoring effect was smaller in the universal pre-K classrooms ($d = .26$) and was in the opposite direction in the Head Start classrooms using a language-literacy curriculum ($d = -.27$). In relation to the acquisition of phonological awareness skills, children in Head Start performed less well if they were enrolled in either curricula condition with mentoring. Figure 4 presents these results graphically. As shown in the first figure, Head Start and Title 1 classrooms are combined. In the second figure, rates of growth in phonological awareness skills for mentored versus non-mentored Head Start and Title 1 classrooms are presented. Finally, as all graphs are based upon raw scores, Appendix A contains standard scores for several of the outcome measures.

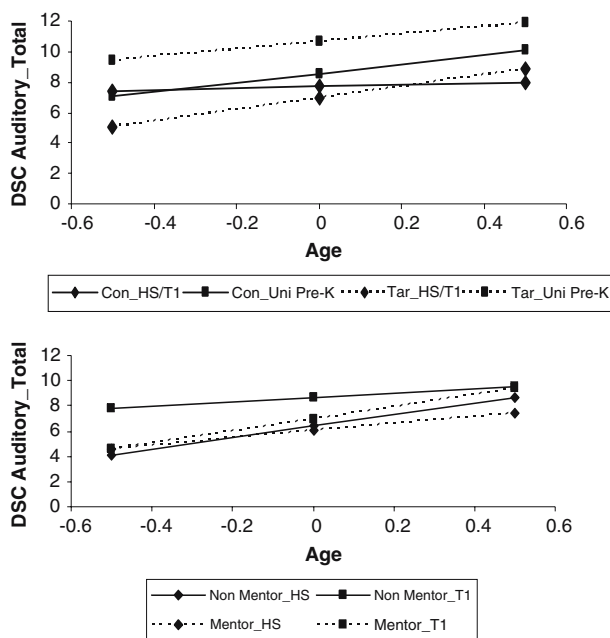


Figure 4. Rates of growth on the DSC Auditory comprehension subscale.

Discussion

A major objective of this study was to evaluate the impact of multiple factors that influence curriculum effectiveness. Of particular interest was whether two research based language and literacy curricula would be effective across multiple early childhood settings when compared to classrooms using general stimulation activities without a focus in the language and literacy area. Although we expected both research based curricula to show stronger gains in children's skills, we hypothesized that our target curricula would be more effective in sites where poorly detailed teaching objectives and training were more likely to be the norm. Additionally, we were interested in whether the intensity of training/professional development associated with curriculum implementation would help explain curriculum efficacy. As predicted, for most language and literacy skills assessed, the findings demonstrated that children in classrooms receiving the two research-based curricula were more likely to show stronger gains on language and literacy outcome measures than children in control classrooms. For example, there were significant main effects for curriculum as compared to control classes when examining the rate of growth on child skills on four out of five outcome measures. However, in all cases the effectiveness of the curriculum was dependent on a variety of factors.

As predicted, program site appeared to be one of the primary factors that moderated curriculum efficacy. For instance, greater gains in skills were seen in Head Start classrooms independent of whether or not the curriculum was *Doors to Discovery* or *Let's Begin with the Letter People*, or whether or not the curriculum was mentored. While curriculum effects were found in the other two sites (Title 1 and universal pre-K), effects were less widespread and more dependent on factors such as the type of curriculum and mentoring. For example, in terms of children's outcomes on the Expressive Vocabulary Test, while greater growth was seen in Title 1 and Head Start programs, the rate of growth was significantly greater in the Head Start Programs. This was similar for language comprehension skills and letter knowledge. However, the phonological awareness skills were impacted by curriculum to a similar degree for both Title 1 and Head Start. In the universal pre-K setting, curriculum effects appeared to be less strong and when they were seen (e.g., Rhyming on the WJ-3), they were only seen under certain conditions (e.g., *Let's Begin with the Letter People* classrooms outperforming *Doors to Discovery* classrooms).

One reason that Head Start classrooms may show greater benefit is the fit between the teacher's need for instructional supports given their lower levels of training in how to enhance cognitive skills associated with early

literacy. Both curricula provide teachers with a detailed plan of the scope and sequence of developmentally appropriate activities to enhance literacy outcomes. Their structured nature (e.g., specification of small versus large group instruction, and sequencing of phonological awareness activities) provides teachers with the specificity they need to make determinations about group size, sequence instructional goals, and/or match appropriate materials with learning objectives. This may be particularly important for teachers without a strong background in instructional approaches in these early childhood skill areas.

Teachers in the Title 1 program also had children at risk for delays in language and literacy skills, who would be expected to benefit from good curricula. However, they may have not depended on specific curricula as much as Head Start teachers because of having more education and backgrounds that included training in supporting early literacy development than Head Start teachers. Thus, Title 1 teachers may have been more comfortable with identifying appropriate instructional approaches for language and literacy, and therefore not as dependent on this being provided by the curriculum. Additionally, the independent school district where the Title 1 program existed had a longer history of understanding the importance of school readiness in contrast to the Head Start program that held a broader set of goals (e.g., physical and mental health issues). Therefore, the aspects of the curriculum that were important for supporting gains for the Head Start program may not have been as critical in Title 1 classrooms. This, however, was not completely true as phonological awareness skills were responsive to the curriculum to similar degrees across both program types.

As building phonological awareness skills for young children has not been a focus for young children in education programs, the specificity for instruction within this area seemed to be important for teachers regardless of educational or training background. Generally, effects of curriculum were not apparent across target versus control classrooms in the universal pre-K setting, with the exception of rhyming and letter identification but this only occurred with one curriculum. Effects of curriculum might not have been as great in universal pre-K programs on both language measures and the Auditory subscale of the DSC because the children in this program come from more highly educated families, and therefore, start with higher language skills. In general, because these children were scoring in the high average range at the pre-test, strong gains were not as likely to occur.

Curricula are frequently trained in multi-day workshops prior to the start of school and often ongoing support for implementation is not present. A frequent consideration when intervention programs do not

show effects is the potential for poor implementation because teachers cannot sustain a program with support across the year. Because of this concern, there was an interest in determining whether ongoing support for the implementation of these two curricula would be an important factor for understanding their efficacy. Contrary to expectations, the added advantage of mentoring in combination with a curriculum was not apparent across all skills and program sites. It was predicted that mentoring might be particularly helpful for supporting Head Start teacher's ability to use the curriculum to support learning but this was not the case. Contrary to prediction, when mentoring showed a positive impact, it was only in the Title 1 or universal pre-K classrooms. It is interesting that when mentoring was helpful, its effect was noticed for literacy rather than language skills. For example, children in mentored Title 1 classrooms showed greater gains for both letter knowledge and phonological awareness (Developing Skills Checklist, Auditory) compared to children in non-mentored classrooms. The advantage of mentoring in letter knowledge also was seen in the universal pre-K classes. However, the advantage of mentoring for letter knowledge across Title 1 and universal pre-K settings was specific for the *Let's Begin* curricula.

There is not a clear explanation as to why mentoring was beneficial in some instances in the Title 1 and universal pre-K programs and not as necessary in the Head Start program classrooms. There are a number of possibilities for this finding including the fact that we did not incorporate an extensive mentoring program in classrooms (i.e., 2×s per month). Without varying the amount of mentoring within the study, one limitation is that we cannot determine whether a program with more regular mentoring visits would have shown stronger results. Another possibility includes the high salience of a well-specified curriculum for Head Start but not the public school programs. This importance may have overridden any added impact of mentoring for curriculum efficacy for this particular program. Additionally, teachers in the Title 1 and universal pre-K programs may have been more responsive to the mentoring process as a mechanism for supporting their efforts to appropriately implement the curriculum. Also, while there were educational supports (e.g., educational coordinators) inherent in all three settings, the Head Start teachers across mentored and non-mentored conditions may have received some benefit from these supports that differed from the other programs. It is also important to note that all target teachers received specific feedback surrounding those aspects of the curriculum they were effectively implementing and those areas that needed additional attention. This may have served as a form of mentoring that might have lessened the overall impact of the mentoring condition.

A final objective was to determine if there was any evidence that the curricula under investigation would impact child outcomes in different ways. We hypothesized that because each curricula had differential strengths and weaknesses, there would be differential effects on language versus literacy skills related to their use. For instance, *Let's Begin* was expected to lead to greater gains in skills in the letter knowledge and phonological awareness areas, while *Doors to Discovery* was expected to show the strongest impact on language skills. There was some evidence for *Let's Begin* leading to more robust growth for letter knowledge and rhyming when compared to *Doors to Discovery*. Given the strong focus in *Let's Begin* on playing with sounds during rhyming activities and book reading as well as the use of the letter people "huggables" and their integration into letter knowledge activities, the advantage of this curriculum for these skills is not surprising. However, for language skills, both curricula showed similar effectiveness. While building language skills is a particularly strong focus of *Doors to Discovery*, lower levels of implementation fidelity might have explained why stronger language outcomes were not seen when this curriculum was used. It may be that when teachers have a second year of implementation, they might be able to more effectively use *Doors to Discovery* to support language development.

There are a number of additional factors that may be important for understanding the results of this study. For example, across three program types there were differences in the children's skill level at the pre-test that were expected given the differences in the family educational backgrounds of the children associated with programs. However, this was a question of interest as it allowed us to evaluate whether these curricula would be effective for children across all socio-economic levels. Our results suggest that the greatest benefit is seen, as predicted, for those with the greatest need. However, children from higher SES backgrounds benefited from the research based curricula in letter knowledge and phonological awareness areas.

A limitation of this study is the lack of systematic examination of the effect of these curricula on language and literacy outcomes for ELLs versus Monolingual English speaking children. This could not be incorporated into the statistical design because of extreme discrepancy between the numbers of ELL children across different pre-school sites. However, because children who were monolingual Spanish speakers were not in this study, all conditions included children learning English as a second language and thus, this was controlled for with random assignment. It is important to note that in both the public school and Head Start classrooms, supports were put in place to assist children learning English

(e.g., assistant teachers who were bilingual and reinforced concepts presented by the teachers in both languages).

While this design utilized random assignment, there were some differences at pre-test between target and control classrooms. However, our statistical approach of evaluating gains over time via growth modeling techniques should help correct for these differences. There were of course a large number of unmeasured factors that may contribute to understanding these results. For instance, instructional practices in the control classrooms were not under the control of the study, thus there may have been good language and literacy activities going on in these classrooms that may explain the lack of curriculum efficacy. This may have been particularly true in the public school classrooms due to the school district's value of cognitive readiness. However, as this was mostly true for language development, the fact that we see more curriculum effectiveness for the literacy areas suggests that there was an added benefit of the curriculum for these particular skills.

When considering many of the characteristics of state funded pre-school programs and Head Start classrooms, the impact of curriculum is complex. Most importantly, the results of this study demonstrate that it is important to consider characteristics of a program including the training and educational background of the teachers, the type of children served, and the fit between the curriculum's focus and targeted skill areas. Language and Literacy curricula, regardless of the specificity of its focus, appeared most important in programs without strong emphasis on cognitive readiness, less detailed teacher objectives, where teachers had lower levels of education/professional development, and served the most high-risk children (i.e., Head Start). The next greatest benefit of curriculum was in Title 1 classrooms who serve higher risk children than Universal pre-K classrooms. In programs with the least need (i.e., universal Pre-K) curriculum appeared more effective in more specialized skill areas (e.g., phonological awareness). When mentoring was important it was more likely to be within the public school system in classrooms utilizing *Let's Begin* and specifically within the literacy domain. Thus, these results demonstrate how critically important it is to attend to the diversity of educational characteristics when examining the efficacy of a pre-school curriculum. As this study examined the implementation of new curriculum at the end of one year, it will be important to also consider the impact of additional time for teachers to find ways to seamlessly incorporate curriculum goals into all aspects of the day. In future studies attention should be paid to examining additional characteristics of participants and programs, as well as

components of a curriculum, that may be important in understanding why and under what conditions particular curricula can be effective.

Appendix A

Pre- and post-test means, standard deviations, and sample sizes for standardized child outcome variables.

		PLS-IV			
	Site	Group	(Auditory Comprehension)	Expressive vocabulary test	DSC auditory (percentiles)
Pre-test	HS	LBM	73.43 (16.59)/28	70.66 (23.61)/32	27.94 (16.56)/18
		LBN	66.02 (15.86)/47	64.02 (21.50)/48	21.26 (16.69)/34
		DDM	76.67 (11.31)/30	85.23 (20.26)/30	27.00 (16.28)/19
		DDN	78.96 (12.68)/26	79.23 (21.43)/31	33.71 (15.40)/24
		Control	76.90 (13.25)/73	81.95 (19.44)/84	34.11 (14.26)/55
	T1	LBM	86.08 (14.95)/37	88.69 (18.62)/39	29.26 (16.80)/39
		LBN	89.38 (16.98)/34	96.88 (14.52)/34	36.00 (12.42)/33
		DDM	91.63 (19.78)/40	90.45 (15.92)/40	35.70 (18.62)/40
		DDN	95.05 (18.90)/38	90.55 (17.31)/40	44.63 (18.64)/40
		Control	85.14 (15.69)/58	92.73 (15.28)/60	42.06 (15.92)/60
	UPK	LBM	108.53 (13.02)/15	110.20 (8.27)/15	47.40 (11.80)/15
		LBN	103.13 (9.38)/30	102.70 (9.48)/30	45.63 (13.58)/30
		DDM	98.33 (12.31)/33	107.42 (9.50)/33	45.48 (11.24)/33
		DDN	103.97 (13.92)/32	109.66 (9.38)/32	49.26 (13.99)/31
		Control	91.41 (12.73)/59	94.69 (14.47)/59	39.45 (14.04)/58
Post-test	HS	LBM	73.06 (11.08)/31	75.74 (22.98)/31	30.37 (14.15)/27
		LBN	74.81 (11.90)/48	76.15 (22.35)/48	36.12 (10.31)/49
		DDM	77.78 (14.97)/27	88.54 (13.71)/28	31.85 (12.31)/26
		DDN	77.57 (11.19)/28	82.61 (17.93)/28	36.50 (11.25)/24
		Control	79.00 (10.42)/77	85.39 (15.49)/77	33.98 (12.21)/65
	T1	LBM	85.15 (14.82)/34	91.41 (11.41)/34	36.38 (10.96)/34
		LBN	83.96 (16.12)/28	92.89 (13.52)/28	38.21 (16.63)/28
		DDM	84.62 (18.00)/34	92.47 (10.23)/34	37.74 (10.77)/34
		DDN	91.44 (18.15)/34	92.74 (16.10)/34	38.85 (12.82)/34
		Control	82.63 (14.12)/54	92.74 (10.98)/54	38.24 (11.22)/54
	UPK	LBM	109.50 (8.62)/14	102.29 (4.68)/14	47.00 (9.70)/15
		LBN	102.97 (13.51)/29	101.39 (7.60)/28	43.00 (13.37)/28

Appendix Continued

Site	Group	PLS-IV		
		(Auditory Comprehension)	Expressive vocabulary test	DSC auditory (percentiles)
	DDM	97.25 (13.93)/32	96.53 (14.59)/30	48.59 (8.92)/32
	DDN	105.55 (13.15)/29	102.14 (10.72)/29	42.24 (12.45)/29
	Control	92.86 (16.73)/51	99.34 (10.66)/50	39.10 (10.68)/51

Note: LBM, Let's Begin with the Letter People Mentored Classrooms; LBN, Let's Begin with the Letter People Non-Mentored Classrooms; DDM, Doors to Discovery Mentored Classrooms; DDN, Doors to Discovery Non-Mentored Classrooms.

References

- Abrams & Company (2000). *Let's begin with the letter people*. Waterbury, CT: Abrams & Company.
- Arnold, D. H., Lonigan, C. J., Whitehurst, G. J., & Epstein, J. N. (1994). Accelerating language development through picture book reading: Replication and extension to a videotape training format. *Journal of Educational Psychology, 86*, 235–243.
- Barnett, W. S. (1995). Long term effects of early childhood programs on cognitive and school outcomes. In R. E. Behrman (Ed.), *The future of children: Long-term outcomes of early childhood programs* (pp. 25–50). Los Altos, CA: The Center for the Future of Children, The David & Lucile Packard Foundation.
- Biemiller, A. (2003). Using stories to promote vocabulary. Paper presented at the International Reading Association Symposium, Orlando, FL, May, 2003.
- Bowman, B. T., Donovan, M. S., & Burns, M. S. (Eds.) (2001). *Eager to learn: Educating our preschoolers*. Washington, DC: National Academy Press.
- Box, J. A., & Aldridge, J. (1993). Shared reading experiences and Head Start children's concept about print and story structure. *Perceptual & Motor Skills, 77*, 929–930.
- Bradley, R., Caldwell, B., & Rock, S. (1988). Home environment and school performance: A ten year follow-up and examination of three models of environmental action. *Child Development, 59*, 852–867.
- Bransford, J., Brown, A., & Cocking, R. R. (Eds.) (2000). *How people learn: Brain, mind, experience, and school*. Committee on the Developments in the Science of Learning, National Research Council. Washington, DC: Academic Press.
- Burchinal, M. R., Peisner-Feingberg, E., Bryant, D. M., & Clifford, R. (2000). Children's social and cognitive development and child-care quality: Testing for differential associations related to poverty, gender, or ethnicity. *Applied Developmental Science, 4*, 149–165.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. (2nd Edn.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- CTB/McGraw-Hill (1990). *Developing skills checklist*. Monterey, CA: CTB/McGraw-Hill.
- Dickinson, D. K., & Smith, M. W. (1994). Long-term effects of preschool teachers' book readings on low-income children's vocabulary and story comprehension. *Reading Research Quarterly, 29*, 1–25.

- Eisenhower National Clearinghouse for Mathematics and Science Education (1998). *Ideas that work science professional development*. Columbus, OH: The Ohio State University.
- Fleiss, J. L. (1986). *The design and analysis of clinical experiments*. New York: Wiley.
- Frede, E. (1995). The role of program quality in producing early childhood program benefits. In R. E. Behrman (Ed.), *The future of children: Long-term outcomes of early childhood programs* (pp. 115–132). 5 Los Altos, CA: The Center for the Future of Children, The David & Lucile Packard Foundation.
- Frick, T., & Semmel, M. (1978). Observer agreement and reliabilities of classroom observational methods. *Review of Educational Research*, 48, 157–184.
- Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Paul H. Brookes.
- Helburn, S. W. (Ed.) (1995). *Cost, quality, and child outcomes in child care centers* (Tech. Rep.). Denver, CO: University of Colorado at Denver, Department of Economics, Center for Research in Economic and Social Policy.
- Howes, C., Phillips, D., & Whitebook, M. (1992). Thresholds of quality: Implications for the social development of children in center based child care. *Child Development*, 63, 449–460.
- International Reading Association, & National Association for the Education of Young Children (1998). Learning to read and write: Developmentally appropriate practices for young children. *Young Children*, July, 30–46.
- Juel, C., Griffith, P. L., & Gough, P. B. (1986). Acquisition of literacy: A longitudinal study of children in first and second grade. *Journal of Educational Psychology*, 78, 243–255.
- Kontos, S. C., Howes, C., & Galinsky, E. (1997). Does training make a difference to quality in family child care?. *Early Childhood Research Quarterly*, 12, 351–372.
- Landry, S. H., Crawford, A., Gunnewig, S., & Swank, P. R. (2002). *Teacher behavior rating scale*. Center for Improving the Readiness of Children for Learning and Education, University of Texas Health Science Center at Houston, unpublished research instrument.
- Landry, S. H., Smith, K. E., Swank, P. R., Assel, M. A., & Vellet, S. (2001). Does early responsive parenting have a special importance for children's development or is consistency across early childhood necessary?. *Developmental Psychology*, 37, 387–403.
- Lee, E., Brooks-Gunn, J., & Schnur, E. (1988). Does Head Start work? A 1-year follow-up comparison of disadvantaged children attending Head Start, no preschool, and other preschool programs. *Developmental Psychology*, 24, 210–222.
- Legislative Office of Education (1998). *Head Start's impact on school readiness in Ohio: A case study for kindergarten students*. Report for the Ohio State Legislative Office of Education Oversight, Columbus, ERIC: ED 421 2237.
- Lonigan, C. J., Burgess, S. R., Anthony, J. L., & Barker, T. A. (1998). Development of phonological sensitivity in two- to five-year-old children. *Journal of Educational Psychology*, 90, 294–311.
- MacLean, M., Bryant, P., & Bradley, L. (1987). Rhymes, nursery rhymes, and reading in early childhood. *Merrill-Palmer Quarterly*, 33, 255–282.
- McGrew, K. S., & Woodcock, R. W. (2001). *Technical manual. Woodcock-Johnson III*. Itasca, IL: Riverside Publishing.
- Mitchell, F. (1979). Interobserver agreement, reliability, and generalizability of data collected in observational studies. *Psychological Bulletin*, 86, 366–370.

- Molfese, D. L. (1989). Electrophysiological correlates of word meanings in 4-month-old human infants. *Developmental Neuropsychology*, 5, 79–103.
- Neville, H. J., Bavelier, D., Corina, D., Rauschecker, J., Karni, A., & Lalwani, A. *et al.* (1998). Cerebral organization for language in deaf and hearing subjects: Biological constraints and effects of experience. *Proceedings of the National Academy of Science*, 95, 922–929.
- Reynolds, A. J. (1995). One year of preschool intervention or two: Does it matter?. *Early Childhood Research*, 10, 1–31.
- SAS Institute Inc., SAS OnlineDoc[®], Version 8 (2001). Cary, NC: SAS Institute Inc.
- Senechal, M. (1997). The differential effect of storybook reading on preschoolers' acquisition of expressive and receptive vocabulary. *Child Language*, 24, 123–138.
- Smith, S., Blank, S., & Collins, R. (1992). Pathways to self-sufficiency for two generations: Designing welfare-to-work programs that benefit children and strengthen families. New York: Foundation for Child Development.
- Snijders, T., & Bosker, R. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. Thousand Oaks, CA: Sage.
- Snow, C. E., Burns, M. S., & Griffin, P. (Eds.) (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.
- Spodek, B. (1996). The professional development of early childhood teachers. *Early Child Development and Care*, 115, 115–124.
- Whitehurst, G. C., & Lonigan, C. J. (1998). Child development and emergent literacy. *Child Development*, 68, 848–872.
- Whitehurst, G. J., Epstein, J. N., Angell, A. L., Payne, A. C., Crone, D. A., & Fischel, J. E. (1994). Outcomes of an emergent literacy intervention in Head Start. *Journal of Educational Psychology*, 86, 542–555.
- Williams, K. T. (1997). *Expressive vocabulary test*. Circle Pines, MN: American Guidance Service.
- Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock–Johnson III tests of achievement*. Itasca, IL: Riverside Publishing.
- Wright Group, McGraw Hill (2001). *Doors to discovery: A new pre-kindergarten program*. Bothell, WA: Wright Group/McGraw-Hill.
- Zevenbergen, A. A., Whitehurst, G. J., Payne, A. C., Crone, D. A., Hiscott, M. D., & Nania, O. C. *et al.* (1997). Outcomes of an emergent literacy intervention in Head Start homes and classrooms. *National Head Start Association Research Quarterly*, 1, 137–147.
- Zill, N., & West, J. (2001). *Entering kindergarten: A portrait of American children when they begin school*. U.S. Department of Education, OERI, NCES 2001-035.
- Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2002). *Preschool language scale-fourth edition*. San Antonio, TX: The Psychological Corporation.

Address for correspondence: Michael Andrew Assel, Department of Pediatrics, University of Texas Health Science Center-Houston, 7000 Fannin, Ste. 2300, Houston, TX 77030, USA Phone: +1-713-500-3714; Fax: +1-713-500-3705; E-mail: michael.a.assel@uth.tmc.edu