Appendix B:
Examining Children's School Readiness Outcomes: Effects of Enhancements to Early Childhood Programs

Appendix B.1:
What It Means and What It Takes to Prepare Children for School: A Synthesis of Evidence for the Impacts of Federally-Funded Research Initiatives in Early Childhood Education

Barbara Dillon Goodson
Abt Associates Inc.

Working paper prepared for *A Working Meeting on Recent School Readiness Research: Guiding the Synthesis of Early Childhood Research*
Washington, DC
October 21-22, 2008

This paper is part of a series of working papers prepared for a meeting sponsored by the U.S. Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation (ASPE) and the Administration for Children and Families, Office of Planning, Research, and Evaluation (OPRE). Abt Associates Inc and the National Center for Children in Poverty (NCCP) were funded to convene the meeting. The views represented in this paper are those of the author(s) and do not necessarily reflect the opinions of the U.S. Department of Health and Human Services.

In the last ten years, there has been a rapid increase in the number of research studies examining the impacts of preschool interventions on children's school readiness outcomes, with many of these studies using rigorous experimental methods that allow attribution of causal relationships. A substantial proportion of these studies have been supported through federal funding, as stand-alone evaluations of federal programs such as the Head Start Impact Study, as part of research initiatives such as the Preschool Curriculum Evaluation Research Initiative, or as individual studies funded through grants. State and local governments, as well as private funders, have also supported recent research studies of preschool interventions. One of the hallmarks of the current crop of research studies is the focus on the right hand side of the equation; that is, the studies are not simply concerned with demonstrating the size of the impacts on child outcomes but also with trying to understand the processes responsible for the impacts that are obtained.

The most recent meta-analyses of early childhood education programs (Jacob, Creps & Boulay, 2004; Nelson & Westhues, 2003; Gorey, 2001; Gilliam & Zigler, 2000) focus on the average size of the impacts of a range of early childhood education interventions. The meta-analyses either bypass the question of variation in instructional inputs as they relate to effect size or focus on programmatic features such as length of day, comprehensiveness of services or auspice rather than instructional methods. However, the research agenda in the past five to ten years has moved beyond proving that early childhood education can make a difference to children, especially at-risk children, to trying to build a body of knowledge about how to successfully intervene with at-risk children to improve their school readiness. The three summary papers discussed here are directly concerned with the most current evidence for instructional practices, interventions, curricula, and programs that have been shown to impact children's development in three domains: language and literacy, math, or socioemotional development.

The question being posed concerns the contributions of this emerging body of research as a source of new evidence or as an extension of what we know about effective interventions for school readiness. On the one hand, the three summary papers suggest that there are an increasing variety of types of early childhood education interventions and curricula that are effective at improving children's school readiness-related outcomes across domains. On the other hand, there are important limitations of the research. First, almost all of the interventions being tested encompass multiple components and the designs do not allow us to "unbundle" these components analytically to determine which programmatic factors make the biggest difference for children's outcomes. When the research is examined for lessons about variation in instruction, the interventions being compared differ on so many factors that it is impossible unable to link outcomes to specific characteristics of instruction or environmental changes. Just as in the past, this current research primarily consists of stand-alone studies, essentially unconnected to one another in any logical way nor connected to a systematic, integrated research plan. At the present time, the research does not go much further in helping us isolate the "potent" or "active" ingredients in instruction that are critical to different child outcomes.

The second limitation is that studies are not connected by a consistent definition of what in fact constitutes school readiness. Studies tend to use measures that align with the intervention and do not attempt to assess a more comprehensive set of outcomes across other domains. This limits our ability to compare the effects of different intervention strategies and to answer questions about whether focusing on one aspect of school readiness (e.g., self-regulation) has generalized impacts across other outcome domains.

Building a systematic knowledge base on effective practices, or a 'science of practice' for promoting school readiness will require an "infrastructure" to guide and link the research. More specifically, the infrastructure will be built on answers to two over-arching questions:

1. What constitutes school readiness?

    a. Do we understand the foundational skills/content knowledge/understandings that children need to develop by the time they enter school for academic success in both early elementary grades and longer?

    b. What is the developmental trajectory of these foundational skills?

A common definition of what is meant by "school readiness" will contribute to the ability to standardize the research on school readiness. A justifiable definition of school readiness will depend on evidence showing that skills developed during the preschool period have impacts on later school performance. While each of the three papers offers some rationale for linking the preschool outcomes in their domain to academic outcomes, the rationales are based on a mix of theory, opinions and correlational research. Even in the field of language and literacy, where the soon-to-be-released report from the National Early Literacy Panel will present a comprehensive summary of the research literature about the early or foundational skills/knowledge that are the strongest predictors of later reading achievement, the research base is correlational. Although some of the research reported in these papers will be able to test causal relationships between preschool and school outcomes, assuming long-term follow-up of children, for most of the interventions, it is too early to show long-term effects for children's academic performance and even longer-term social outcomes such as higher education and/or economic productivity. As such evidence is reported, it will be a basis for beginning to build a stronger research-based definition of school readiness. It is worth noting that there are other forces pushing us toward a measurable definition of school readiness. The large scale investments in early childhood through universal pre-kindergarten initiatives and quality improvement systems are being justified in terms of improvements in school readiness. In theory, a definition of school readiness should rest on research linking preschool skills/content knowledge/understandings to later school achievement.

Further, we need longitudinal evidence of the developmental trajectory of skills purported to be foundational.

2. What do we know about the contribution or influence of environmental factors in the development of the foundational skills, and can we build effective interventions based on this knowledge?

    a. Is there evidence that the skills are learnable or modifiable and therefore susceptible to intervention?

    b. Based on theory or basic research, can we develop effective interventions to enhance the development of these skills?

    c. Can we show a causal link between specific instructional practices and student school readiness-related outcomes?

Interventions aimed at enhancing children's school readiness are based on two premises: (a) that the skills being taught or supported by the intervention are learnable, and (b) that there is research or theory to justify the intervention strategies for changing the early childhood education experience so as to alter the developmental trajectory. Even when there is clear agreement on objectives for children at the end of preschool, there are typically alternative theories about effective intervention approaches, as reflected in the variability across intervention designs. This brings us back to the question of which intervention strategies are most powerful in creating changes in children.

The current research is insufficient for understanding the process by which these interventions lead to child impacts. We don't know which of the changes being created in early childhood environments through multi-faceted interventions are the causal factors in changing student outcomes. Even the experimental studies being conducted can't, in fact, establish that the teachers or classrooms that changed the most as a result of the intervention are the same sites where the child outcomes changed the most.

One approach that has been used to begin to build this information base is planned variation research, where the research is designed to systematically test different intervention strategies with similar children and a common set of outcomes, to attempt to isolate which models have the largest impacts. However, unless this type of planned variation research varies and compares the impacts of intervention components rather than multi-dimensional models, it is not possible for the research to provide us with the information we want about mechanisms of change.[1] Further, this kind of planned variation research focuses on the relative contribution of components of the instructional intervention. There are other aspects of the environment that are additional potential factors in the impact of instruction, such as how the classroom is managed (e.g., discipline, grouping), class size and heterogeneity of the child in terms of characteristics such as home language, special needs and the like. Research that allows us to disentangle the combined and individual effects of all of these factors will require complex designs and sample sizes that permit us to test multiple variations.

What follows is an overview of the conclusions from the three synthesis papers about where the current research stands in terms of the content and focus of the interventions being tested, what the findings tell us about effective instructional strategies, and what types of future research will be most informative. The two issues raised above—defining school readiness and the developmental

[1] The traditional method used to link interventions to outcomes, even in the context of randomized designs, is based on OLS regression, in which variation in implementation is correlated with variation in outcomes. Another experimental approach uses instrumental variable analysis. This approach has been applied in many contexts by economists and is becoming increasingly popular for use with randomized experiments (for example, Angrist, Imbens and Rubin, 1996). It does not compare existing levels of student achievement and instructional practice. Instead it leverages the fact that a high-quality randomized experiment (or a well-executed regression discontinuity analysis) can produce unbiased estimates of program impacts on classroom instructional practices and on student test scores. The approach thus examines the association between outcomes that is implied by a pattern of program impacts. For example, in a randomized study, the analysis compares program-induced changes in student outcomes with program-induced changes in classroom instruction, where both changes are estimated using the randomized design. Under certain conditions, this analysis can provide internally valid (statistically consistent) estimates of the causal effect of classroom instruction on student performance. This methodology overcomes some of the problems in relational analysis (omitted variables and attenuation bias), although instrumental variable analysis depends on being able to show that there are no mediators additional to the classroom instruction that could account for the relationships with child outcomes.

trajectory of foundational skills and identifying the active ingredients in interventions—will resonate through the overview.

The outcomes of these interventions are discussed in the context of the following domains: socioemotional, language and literacy, and then math. The paper starts with the socioemotional domain because the constituent skills are hypothesized as constituting the platform underlying the child's ability to negotiate successfully all other learning tasks, including early literacy and early math understandings. The second domain discussed is language and literacy. Although language and literacy are often paired, in many respects language should be considered in conjunction with socioemotional development, because of the broad central role language plays in children's learning. For the developing child, the ability to understand and use language is the primary mode by which he builds knowledge of the world and communicates his own ideas and feelings. In this sense, most aspects of socioemotional development are completely intertwined with language development: Children's internalized regulatory mechanisms are language-based, their social understanding is language-based, and their ability to interact and engage with others is primarily negotiated through language. Language development can be labeled as an "engine" of development.

Early literacy and early math are the final outcome areas discussed. As opposed to socioemotional and oral language outcomes, early literacy and math both represent specific skills and understandings. Literacy, for example, includes print knowledge, the alphabetic code and phonological processing (phonological memory, access, and awareness); math includes number and operations with numbers, geometric shapes, spatial relations, and measurement.

Socioemotional Domain

As Raver reports, the current conceptualization of the socioemotional domain distinguishes three major mechanisms or processes that support children's development: self- regulation (emotional and cognitive), social cognitions, and prosocial skills. Raver also describes a fourth area of socioemotional development, behavior problems (externalizing and internalizing), which factors into children's ability to learn and relate to other people. The behavioral manifestations of these processes, taken together, form a picture of a child socially and emotionally ready for school. This child is able to:

- follow adult directions;
- control his/her own emotions, attention, and impulses independent of adult regulation;
- establish positive social relationships with peers and adults;
- successfully solve social problems without being disruptive or aggressive;
- attend in a sustained way to learning tasks in the environment;
- evaluate his/her own behavior and make corrections; and
- demonstrate "cognitive flexibility."

Raver's description of the intervention research in the socioemotional domain clusters studies based on which of the three underlying processes the interventions are designed to effect. The research on interventions in the socioemotional domain is most consistent in the area of self-regulation and social skills. Evidence of the ability to reduce aggressive behavior in the classroom is more mixed. Further, all of the data reported represents short-term findings, with no evidence to date of longer-term benefits for school performance. Further gaps include: evidence of whether and how the various components of socioemotional functioning are inter-connected; and evidence of the

relationship of children's development of self-regulatory and social relationship skills in preschool to their oral language development or to the acquisition of early literacy or math skills. As discussed above, the fact that disparate intervention strategies all appear to have impacts raises the question of the mechanisms leading to child impacts.

Impacts on Children's Self-Regulation Skills

In the area of self-regulation, Raver cites evidence that children's attentional processes can be enhanced through a variety of intervention mechanisms. For example, Raver cites three interventions as having impacts on children's self-regulatory skills:

- **Project REDI:** uses small-group lessons focused on understanding emotions to help children regulate behavior and successfully negotiate social relationship; trains teachers on classroom management strategies that create a positive learning climate; uses instructional strategies in early literacy to build oral language skills and phonemic awareness that promote teacher/child interaction (scripted dialogic reading exercises to promote conversation and build vocabulary and small-group phonemic awareness activities to teach sounds and words).
- **Chicago School Readiness Project (CSRP):** focuses on improving the emotional climate of the class by providing teachers with training in behavior management and in-class coaching by mental health consultants on implementing positive behavior management strategies.
- **Tools of the Mind:** uses role play as a central mechanism to help children develop "self-regulatory scripts" to guide their own behavior; thematic dramatic play is the central type of role play, but roles also are used in children's work with peers in reading and other content areas.

All three of these interventions have reported impacts on children's levels of attention and focused effort and persistence, as measured through direct observations,[2] despite the fact that the three interventions use very different approaches. REDI and CSRP use the teacher as the primary change agent for helping children develop self-regulation, while Tools uses children's own role play to help children develop their own self-regulatory scripts. The fact that all three interventions report impacts on children's development of self-regulation skills and all three use multiple avenues to affect these changes underlines the importance of systematic research to isolate the most important "levers." Further, data on the long-term effects of these curricula will be crucial for understanding whether the differences in the approaches of Tools versus REDI and CSRP have ramifications for the persistence of impacts over time, once children leave supportive early childhood environments. If the children in Tools of the Mind build internal self-regulatory structures while children in REDI or CSRP are more dependent on the actions of the teacher, then it is possible that Tools will have more robust long-term impacts.

[2] The fact that parallel effects were not demonstrated on teacher ratings of children's attention and impulsivity may be related to power rather than to inconsistency in outcomes. Teacher ratings have been found to have higher correlations among children in the same classroom and center (ICCs) than do cognitive measures such as the PPVT. This means that only relatively large impacts can be detected for the teacher-reported outcomes,

Long-term follow-up data on differences in school performance for children with stronger or weaker self-regulation at the end of preschool will also provide important information to prove or disprove the contention that self-regulation encompasses a skill set that influences learning across content areas and across ages. For the same reason, it is important that the research on these interventions includes measures of children's acquisition of skills in other curriculum areas, such as early literacy or early math at the end of preschool. (For example, in the research on *Tools*, children not only develop stronger attentional processes, they also score higher on standardized tests of math at the end of preschool.)

In general, the maintenance of gains in preschool may depend not only on the types of behavioral and/or attentional changes that children experience in preschool but also on the characteristics of their subsequent classroom environments in elementary school. Gains in preschool may be maintained or even enhanced if children experience classroom environments in elementary school that continue to support positive, regulated behavior.

Impacts on Children's Social Cognitions and Prosocial Skills

A second area of intervention research described by Raver focuses on the social cognitive mechanisms underlying children's ability to form and sustain positive interpersonal relationships with peers and adults in the classroom and to solve problems in social relationships. The social cognitive mechanisms include: children's knowledge of emotions—their own and other people's; knowledge of prosocial behaviors (e.g. helping, sharing, and taking turns); and the ability to generate and use more effective social problem-solving skills. In this area, the child who is ready for school:

- Can develop a positive, engaged social relationship with the teacher;
- Can form positive friendships with peers;
- Can successfully solve problems that arise in social interactions with peers;
- Demonstrates prosocial behavior in the classroom, such as helping other children, sharing, and taking turns;
- Does not act aggressively with other children or adults.
- Does not act disruptively in the classroom.

In the same way that self-regulatory skills are correlated with children's learning across domains, children's social skills and the quality of their relationship with teachers have been found to be correlated to their later social and academic competence in early elementary school.

Raver focuses on the results from three interventions:

- **Project REDI** trains teachers to provide more emotional coaching and support in the classroom and includes a socioemotional curriculum that helps children develop emotional knowledge and accurate social attributions, and prosocial behavior strategies for interactions with peers. REDI reports significant differences for children's emotion understanding and interpersonal problem-solving, and significant gains in children's social competence (teacher rated aggression and observer-rated social competence). The project also reports significant changes in teachers' use of emotion coaching, positive classroom management and behavioral support.

- **My Teaching Partner,** a web-based teacher training curriculum developed by Pianta, focused on improving teacher/student relationships to be more responsive and supportive. Results showed that intervention teachers demonstrated significantly more sensitivity, language modeling, and quality of instructional support to students.
- CSRP also worked with teachers to establish more positive classroom environments, and there was a significant impact of the on positive classroom climate (d = .52 to d = .89). Although there was no child-focused curriculum on emotional language or self-awareness, the gains in children's behavioral self-regulation were attributed to the enhanced classroom environment.

Impacts on Children's Behavior Problems

Fewer studies have measured impacts of interventions on children's behavior problems. *Project REDI*, a socioemotional learning curriculum, reported significant reductions of children's aggression, as reported by teachers. Similarly, CSRP reports reductions in children's externalizing and internalizing problems as reported by teachers. Across the PCER studies, there were no effects on children's behavior problems as reported by teachers.

Language and Literacy

In many respects, the conceptualization of the critical foundational skills to be acquired during preschool has moved furthest along in the area of language and literacy. There has been a wealth of theoretical writings, professional opinions, and best practice documents proposing which skills are the precursors or foundational skills for reading achievement, and, it is only in the field of language and early literacy that we [soon] will have a systematic empirical summation of research demonstrating which early literacy skills predict later conventional literacy (via the National Early Literacy Panel). There is beginning to be a structure for understanding the developmental precursors to later reading and writing abilities. Further, in the challenge of defining school readiness, this domain has the advantage of the widely-shared criterion of the critical long-term academic outcome—becoming a skilled reader (with strong decoding and comprehension skills, a strong vocabulary, automaticity in reading).

Before the NELP, the field was driven in its thinking about "readiness" skills by two documents that provided consensus or narrative summaries of a portion of the research literature concerning the relation between early precursor skills and later conventional literacy skills: Whitehurst and Lonigan (1998) identified skills in the domains of oral language, print and letter knowledge, and phonological processing as encompassing two aspects (outside-in and inside-out skills) of emergent literacy that are related to later conventional forms of reading and writing; and Snow, Burns, and Griffin (1998), in their report of the National Research Council's panel on preventing reading difficulties in young children, identified weaknesses in oral language, phonological awareness, and alphabet knowledge as prime targets of intervention to prevent the occurrence of significant reading problems. Neither of these documents, however, was based on a comprehensive summary of the published literature.

The NELP provides an evidence base about early or foundational skills/knowledge that are the strongest predictors of reading achievement, as well as a summary of the average effects of the number of interventions to improve early literacy/language skills. In the ensuing discussion, we start with oral language and then move to early literacy, for the reasons spelled out above.

Oral Language

Oral language skills can be conceptualized as including productive language skills (forming sounds correctly, using the right forms of words, forming correct sentence syntax), language use (using words to express thoughts or ideas or to transmit information); and language content (understanding of vocabulary and narrative). In describing a child who is ready for school in terms of his/her oral language skills, the following skills are included:

- Ability to express thoughts, ideas into spoken words;
- Ability to understand other people when they talk;
- Ability to carry on a back-and-forth conversation with another person;
- Ability to use correct versions of plural, past and future tenses.
- Ability to understand narrative sequence (logical order of events);
- Expressive vocabulary that includes knowledge of words likely to be encountered in early readers; understanding of superordinate words for categories of objects (silverware, clothes, tools, etc).

As described in the synthesis paper by Caswell and He, numerous research studies have demonstrated a relationship between early, well-developed oral language skills and later reading abilities. Despite the primary of oral language skills in a child's cognitive readiness for school and, ultimately, for learning to read, the evidence for intervention effects is somewhat disappointing. Across the large number of interventions concerned with children's oral language outcomes, most show small to medium effects.

The synthesis paper describes some of the variety in the oral language activities used to promote children's understanding of vocabulary, comprehension of concepts, and language use. The problem with the research is that in most instances, the intervention being examined includes more than one type of oral language activity, as well as other literacy-related activities, so it is impossible to isolate the impact of the any one type of oral language activity. For example, a number of programs use dialogic reading to promote children's oral language skills. This includes dialogic reading as the sole intervention activity and dialogic reading that is integrated into a broader curriculum with additional activities and goals. There were inconsistent results of these interventions on children's outcomes, although most did find at least a small effect on children's vocabulary. Again, where dialogic reading was just one activity in the curriculum, we cannot know whether it was the dialogic reading was responsible for the impacts on vocabulary that were found.

Most of the research on oral language effects comes from studies of comprehensive or multi-dimensional curricula that included some oral language activities but were not focused on language specifically. The findings for impacts on oral language skills were inconsistent across studies.

Phonological Awareness

Phonological awareness is a component of the broader skill area of phonological processing, which includes not only the child's awareness of sounds, but also the ability to hold sounds in memory and to be able to access sounds from memory. Phonological awareness refers to the child's understanding that words are made up of smaller sounds that can be manipulated, combined and separated. This

knowledge helps children understand the relationship between written language (letters) and spoken language (sounds). Research has established that phonological awareness develops in the preschool period starting with sensitivity to words and moving toward sensitivity to smaller and smaller units of sound (syllables, onset-rime, and phonemes).

Phonological awareness has been shown to be a strong predictor of reading success. At the same time, there is inconsistent evidence of our ability to impact children's phonological awareness skills. The strongest evidence comes from research on individual literacy curricula with some explicit attention to sounds in language, including two curricula studies in Project Upgrade and a couple of individual PCER studies of literacy curricula. In these evaluations, the children receiving the literacy curricula scored higher on a test of sound blending and elision. No evidence of an effect on phonological awareness was found in the Head Start Impact evaluation or the Early Reading First evaluation, possibly because local programs vary widely in the extent to which instruction incorporates an intentional and consistent focus on sounds.

Print Knowledge

Research indicates that print and letter knowledge are strongly related to later reading performance. Children's knowledge of the alphabet when they enter school is one of the single best predictors of later reading achievement, most likely because the ability to recognize and distinguish individual letters is a necessary precursor to learning the sounds that the letters represent. Overall, this component of early literacy is the one most often targeted by interventions.

The majority of the interventions reviewed targeted children's print knowledge as an essential skill and there was consistent evidence that the interventions were effective in improving children's print and letter knowledge. This included the large national early childhood studies, where there was substantial variation across sites in the programmatic activities and individual curricula.

Early Math

As argued by Ginsburg et. al, the fact that the intervention research on early math lags behind the research on early literacy can be explained at least partially by the long-held belief that young children are not able to understand mathematics in complex ways, and that even "everyday" mathematical skills cannot be cultivated in children as young as preschool. As research has built the case showing just the opposite, early math concepts are now central in early childhood education standards, and comprehensive early childhood curricula include deliberate, organized activities to promote understanding of mathematical concepts.

What are the early mathematical concepts that children should acquire in preparation for school? There does not appear to have been extensive conversation among math educators and researchers about what mathematical concepts constitute school readiness. Across the curricula that have been developed, there are similarities in the content areas, however, including:

- basic aspects of number and operations,
- geometric shapes,
- spatial relations,
- measurement, and
- patterns and logic.

The paper discusses six mathematics curricula for preschool on which impact research has been conducted in the United States and two with research from New Zealand. The paper also considers results from research on mathematics activities included as part of comprehensive curricula. As described in the paper, the curricula have different learning objectives and use a variety of materials and approaches, including games, story books, activities, manipulatives, and computer software; stand-alone activities and other activities. Across the various curricula and approaches, most had statistically significant impacts of at least moderate size. Since no two curricula studies used the same measure, it is difficult to compare effectiveness. Further, the research is not useful for determining which aspects of the instruction were most powerful in improving children's math knowledge. Long-term follow-up data also appear to be missing.

Final Thoughts/Future Research

The current set of research summarized in the three syntheses has moved the field forward in some respects. Until recently, there has been almost single-minded focus on language and literacy, which has conferred benefits in terms of the relative breadth and depth of knowledge we have for that domain. The current research reflects a new priority on socioemotional skills, especially self-regulation, and this has opened up new funding opportunities and new intervention designs, which are crucial for our ability to develop our knowledge base in this domain. Early math is also now receiving more scrutiny, although the research base is much more limited.

The current set of research studies does not address directly the critical over-arching issues of what constitutes school readiness, the developmental trajectory of the component skills in readiness, and the long-term benefits of early skill development in both the academic and social domains. The lack of a definition of readiness makes it difficult to summarize the findings from a large set of research studies, since different studies not only use different measures of the same construct but also assess a different set of constructs. Not only does this hinder comparisons, it also limits our ability to understand whether an intervention has broad or narrow effects on children's school readiness.

Nor is the research designed to yield supportable conclusions about the relationship between specific environmental inputs (intentional teaching, materials, technology) and child outcomes that go beyond simple correlations, for example, through systematic planned variation studies or through complex analyses such as instrumental variable analysis.

There also is a clear need for more longitudinal research on the development of children's early skills in all three domains, at least through preschool and into the early elementary grades.

All three synthesis papers note that future research will need to more clearly delineate the sources of variation in impact, as well as the overall impact. Potential factors include characteristics of students as well as of teachers and of the intervention itself.

The field is attempting to simultaneously develop effective, research-linked interventions, deliver them with high fidelity in a variety of education settings, use valid, reliable measures of what are often complex psychological constructs, and contribute to building a knowledge base on instructional practices. Despite the sometimes disappointing findings, we need to understand the difficulty of designing effective interventions to be implemented in real-life educational settings, with groups of at-risk children.

References

Angrist, J., Imbens, G, and Rubin, D. (1996). Identification of causal effects using instrumental variables." *Journal of the American Statistical Association*. 91(434) 444-55.

Gilliam, W.S. & Zigler, E.F. (2000). A critical meta-analysis of all evaluations of state-funded preschool from 1977 to 1998: implications for policy, service delivery and program evaluation. *Early Childhood Research Quarterly*, 15(4), winter 2000, 441-473.

Gorey, K.M. (2001) Early childhood education: A meta-analytic affirmation of the short- and long-term benefits of educational opportunity. *School Psychology Quarterly*, 16 (1), 9-30.

Jacob, R., C. Creps & B. Boulay (2004). *Meta-analysis of research and evaluation studies in early childhood education.* Abt Associates Inc., Cambridge, MA.

National Early Literacy Panel. (January 2008). Report of the National Early Literacy Panel. National Institute for Literacy & National Center for Family Literacy. Washington, DC.

Nelson, G. & A. Westhues (2003). A meta-analysis of longitudinal research on preschool prevention programs for children. *Prevention & Treatment, 6*, 31.

Snow, C. E., Burns, M. S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children*. Washington, D.C.: National Academy Press.

Whitehurst, G. J. & Lonigan, C. J. (1998). Child development and emergent literacy. *Child Development, 69, 848-872.*

Appendix B.2:
Approaches to Promoting Children's School Readiness: A Review of Federally-Funded Research Initiatives Aimed at Improving Young Children's Language and Literacy Skills in Early Education and Care Settings

Linda Caswell
Yeqin He
Abt Associates Inc.

Working Paper prepared for *A Working Meeting on Recent School Readiness Research: Guiding the Synthesis of Early Childhood Research*
Washington, DC
October 21-22, 2008

This paper is part of a series of working papers prepared for a meeting sponsored by the U.S. Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation (ASPE) and the Administration for Children and Families, Office of Planning, Research, and Evaluation (OPRE). Abt Associates Inc and the National Center for Children in Poverty (NCCP) were funded to convene the meeting. The views represented in this paper are those of the author(s) and do not necessarily reflect the opinions of the U.S. Department of Health and Human Services.

Introduction

An examination of research in the field of early childhood language and literacy development reveals substantial changes over the past two decades. Initially, a shift in the conceptualization of what constitutes literacy and when literacy begins resulted in a burgeoning corpus of research that examined children's literate experiences before the beginning of formal schooling. This perspective, termed emergent literacy, brought a new and vigorous focus to the developmental precursors of formal reading that originate in children's early years, thus broadening the scope of research to the years prior to formal schooling, that is, into the early childhood years.

Although research in the field of emergent literacy has been diverse both in topic and methodology, there is currently consensus about the key elements that are foundational to learning to read: oral language, phonological processing, and print awareness (Whitehurst & Lonigan, 2001). Research has provided empirical evidence of the relationships between these early skills and later reading abilities. For example, numerous research studies have demonstrated that early, well-developed oral language skills are a strong predictor of later reading abilities (e.g., ECCRN, 2005; Hart & Risley, 1995; Walker, Greenwood, Hart, & Carta, 1994; Storch & Whitehurst, 2002; Dickinson & Porsche, 2008; Spira, Bracken, & Fischel, 2005; Tabors, Roach, & Snow, 2001). Similarly, children who are sensitive to the sounds in words and are able to manipulate and use them are more likely to be successful in learning to read (Snow, Burns & Griffin, 1998; Pullen & Justice, 2003; Whitehurst & Lonigan, 2001) because these abilities are strongly related to decoding abilities. Finally, in terms of print awareness, studies have shown that a child's knowledge of the alphabet when they enter school is one of the single best predictors of later reading achievement (Snow, Burns & Griffin, 1998; Whitehurst & Lonigan, 2001). The ability to recognize and distinguish individual letters, as well as knowing the sounds of the language, together form the foundation for learning the sound-symbol association.

The importance of successfully mastering these skills for young children cannot be underestimated since limited early literacy skills tend to translate into persistent deficits. For example, Tabors, Snow, & Dickinson (2001) found stability between relative levels of reading performance in kindergarten and seventh grade, while Cunnigham & Stanovich (1997) found the same stability between first grade and the end of high school. Therefore, the effect of poor language and literacy abilities in early childhood can be cumulative, such that children who are behind early on continue to fall further and further behind more skilled readers in reading as well as in other academic areas (Chall, Jacobs, & Baldwin, 1990). Furthermore, evidence indicates that it is very difficult to remedy children's language and literacy difficulties with compensatory programs (McGill-Franzen & Allington, 1991), particularly after third grade (Good, Simmons, & Smith, 1998). Of particular policy relevance is the fact that children of lower socio-economic status are at high risk for reading difficulties. These children tend to begin school with less-developed abilities in the three foundational skills of early literacy than their more economically advantaged peers. Thus, interest in effective interventions to improve children's early language and literacy skills is motivated in large part by the possibility of narrowing the school readiness gap.

One argument for focusing on providing comprehensive support for children's development of early language and literacy skills comes from economists such as Lynch (2004), who have conducted cost-benefit analyses that support the idea that the benefits of substantial investment in early interventions, in terms of increased educational attainment and income earnings outweigh the costs of these

investments. Similarly, Reynolds (2005) found that early interventions are the most cost-effective method to make positive contributions to at-risk children's development.

Another argument comes from evidence that most children are able to achieve grade-level reading levels if they receive effective early reading instruction (Clay, 1985; Iverson & Tunmer, 1993; Pinnell, 1989; Snow, Burns, & Griffin, 1998; Wasik & Slavin, 1993). If this is indeed the case, then perhaps it is deficiencies in teachers' instruction, rather than in children's cognitive abilities that explains the large number of reading difficulties in U.S. schools (Dickinson, McCabe, & Clark-Chiarelli, 2004). Although parents are children's first and foremost teachers, more and more children are spending a large portion of their waking hours with adults in early childhood settings. Recent research has lent support for the idea that teachers' instructional practices can make a difference in children's outcomes. For example, Huttenlocher, Vasilyeva, Cymerman & Levine (2002) found a positive association between teachers' use of complex syntax and preschoolers' comprehension of complex syntax. More importantly, they found that classroom input made up for the lack of home input for children from disadvantaged backgrounds.

Thus, based on the benefits of attending to children's deficits in language and literacy before formal schooling, the lack of success remedying these difficulties after school entry, and the high cost of not doing so for later academic achievement, educators and policy makers have turned their attention to the possibilities of improving children's skills early on. Because over half of 3- to 5-year-old children in the United States – 57% in 2005 – spend time in early childhood care and education programs, including day care centers, Head Start programs, preschools, nursery schools, or prekindergartens (U.S. Department of Education, 2006), there has been a focus on reaching the many children who are in these settings.

However, despite substantial investments by federal and state governments in early childhood center-based programs such as Head Start, Even Start and public pre-kindergarten, until recently, little rigorous research had been conducted on the effectiveness of various curricula used to improve children's early language and literacy skills in these programs. It was against this backdrop that the federal government, through various agencies, funded rigorous evaluations of multiple curricula that focused on language and literacy, as well as other important school readiness skills. The Preschool Curriculum Evaluation Research (PCER) and the Interagency School Readiness Consortium (ISRC) consortia are two such federal sources that have provided funding for rigorous evaluations of curricula used in early childhood programs.

This review provides a synthesis of the emerging findings from this set of major federal research initiatives. We examine the evaluations of program enhancements funded through PCER, ISRC, and the Evaluation of Child Care Subsidy Strategies, as well as evaluations of federal early childhood programs – National Evaluation of Early Reading First and Head Start Impact Study – in terms of key issues in the field of young children's language and literacy development prior to formal schooling. For the PCER interventions, both the cross-site evaluation and individual papers (when available) were reviewed. For the ISRC interventions, the evaluations of which were funded later, there is no cross-site evaluation and most study teams had only reported initial findings in the form of conference presentations rather than journal articles. Therefore, the review of the ISRC interventions should be considered preliminary as findings are still emerging from this work. After synthesizing the set of studies, some possible directions for future research are suggested based on this body of research.

Key Issues in Early Childhood Language and Literacy

Below, we synthesize the findings from the studies reviewed, which examine the effects of different early care and education interventions on teacher and child outcomes. This paper focuses solely on child outcomes. We begin by discussing the evidence that federally-funded research on early childhood language and literacy-specific curricula has provided in terms of identifying effective interventions for improving young children's oral language, phonological sensitivity, and print knowledge skills – the three foundational skills upon which later literacy is based. We then discuss what this body of research has added to our understanding of some of the key factors that moderate the effectiveness of intervention programs. It should be noted that a challenge in reviewing this body of research was that most interventions were broad-based, encompassing many different components. This meant that, in this set of studies, when positive effects on child outcomes were found, it was often not possible to determine which of the many components was contributing to these effects. Fortunately, a more extensive review to be released soon – the National Evaluation of Early Literacy (NELP) – will be able to provide some insight into this question.

What Evidence Is Provided About Improving Children's Skills in Oral Language, Phonological Sensitivity, and Print Knowledge?

Oral Language

As more and more young children spend large portions of their time with teachers in early education settings, the quality of teacher language use plays a critical role in driving children's early language development. For example, studies have demonstrated that cognitively challenging conversations that address decontextualized or relatively abstract topics are particularly beneficial to children's language development (Dickinson, 2001a, 2001b; Dickinson & Smith, 1994). Unfortunately, not all teachers can provide high quality conversations, comments or questions. This is especially true with those underpaid or poorly prepared teachers serving low-income children in publicly funded programs. A descriptive study within the PCER initiative (Massey, Pence, and Justice, 2008) confirmed these prior findings about teacher talk by examining the quality and quantity of teacher questions in 14 preschool classrooms (both treatment—Language-Focused Curriculum—and control) serving economically disadvantaged 4-year-olds. They found that questions characterized one third of all teacher utterances, with management questions (e.g., "Are we ready?") occurring most frequently (44.8%), followed by more cognitively challenging questions (32.5%; e.g., "What do you think will happen next?") and less cognitively challenging questions (22.7%; e.g., "What was this called?") That is to say, more cognitively challenging questions represented only one tenth of all teacher utterances in the at-risk preschool classroom. They further examined the frequency of use for different types of questions across various classroom contexts and found that more cognitively challenging questions occurred most frequently in storybook reading. Unfortunately, according to Dickinson (2001a), only 1% to 4% of the total day is typically spent on storybook reading in early care and education settings.

The aforementioned findings naturally raise the following question: Are curricula that extend storybook reading time more effective in promoting children's language development? Several studies within the ISRC and PCER initiatives examined the effects of curricula that include interactive reading activities in the daily plan. The Head Start REDI (Research-Based Developmentally Informed) program developed a curriculum featuring interactive reading activities

based on shared reading and dialogic reading, providing teachers with scripted books and targeted vocabulary and instructing teachers to elicit children's language more effectively and to be more responsive. A randomized control trial was employed to compare 4-year-olds in the intervention condition and a similar group in non-intervention Head Start classrooms. The post-intervention tests showed that, after being exposed to the intervention for seven to eight months (September to March/April), children in the treatment group outperformed the comparison group on both vocabulary and language use at home (with effect sizes of .15 and .25, respectively), but no effect emerged on measures of children's grammatical understanding. Similar results were obtained from other curricula that integrated interactive reading activities into the curriculum, such as *Children's School Success (CSS)* and *Literacy Express (LE)*. HLM analyses demonstrated that *CSS* improved children's vocabulary and language use at home through changing teacher practice (e.g., more sensitive-responsive talk, richer talk, better instructional support). Teacher practice accounted for 53% and 67% respectively of the intervention effect on vocabulary and language use at home. *LE* was found to have a significant impact on expressive communication skills (ES = .30) and a potentially positive effect[1] on vocabulary (ES = .45). Over all three measures used in the oral language domain, the average effect size was .36.

In addition, dialogic reading or reading-aloud was an important component of three other early childhood curricula—*Breakthrough to Literacy (BTL*), Ready, Set, Leap! (RSL), and *Building Early Language & Literacy (BELL)*. Despite sharing common purposes, these three curricula differ in activity designs and implementation. Both *BTL* and *RSL* are comprehensive language and literacy programs that include activities throughout the day. *BTL* is built around a series of weekly books with a focus on interactive reading; while *RSL* utilizes interactive electronic technology and thematically-grouped children's trade books. In contrast, *BELL*, as an add-on pre-kindergarten literacy program, entails only two daily 15- to 20-minute lessons. The *Project Upgrade* study compared the curricula to each other and a business-as-usual control group. The results revealed that *RSL* and *BTL* had significant impacts on children's definitional vocabulary (ES = .30), even though the impacts were not large enough to reduce the gap (see below). On the other hand, *BELL*, the less intensive curriculum, yielded no significant impacts on any measures of early language and literacy. Taken together, findings from these studies may suggest that curricula with a focus on interactive reading activities do exert positive impacts on children's oral language development, given enough dosage of implementation.

Even though interactive reading seems to be an effective ingredient to improve oral language, not all curricula put an emphasis on interactive book reading. Instead, some PCER/ISCR curricula provide specific and explicit instructions to teachers to foster frequent and long high quality conversations that use complex syntax and address abstract concepts. For instance, the *Language-Focused Curriculum (LFC)*, designed for preschoolers with language limitations, identified specific linguistic targets (e.g., verb phrase structures, adjective, pronouns, etc.) in daily lesson plans and instructed teachers to use a set of *Language Stimulation Techniques (LSTs)* to foster the delivery of linguistically-responsive conversations with children. In a study with a random-control trial design, Justice, Mashburn, Pence, Wiggins (under review) analyzed children's 10–minute language samples gathered in the fall and spring, with the amount of professional development that teachers received matched in the intervention group and comparison group. However, they found no impacts of *LCF*

[1] Potentially positive is a rating given by the What Works Clearinghouse (WWC) indicating that although the difference between the treatment and control groups was not statistically significant, the effect size was large enough to be considered substantively important according to WWC criteria (i.e., at least .25).

and *LST* exposure on children's expressive language skills. Instead, the results demonstrated that children who attended preschool more frequently benefited more from the *LCF* curriculum and *LST* exposure compared to those with low attendance (no effect size reported). This finding is not unexpected: if a child does not go to class frequently, how can s/he benefit from the curriculum? From another angle, this finding aligns with the result of the *Project Upgrade* study that showed no impact of the lower dosage curriculum, *BELL*.

Two other curricula, *Let's Begin with the Letter People* and *Doors to Discovery*, also provide teachers with a detailed plan of the scope and activities that are developmentally appropriate to enhance literacy development. This plan provides specific instructions to help teachers determine group size, sequence instructional goals, and match appropriate materials with learning objectives. Both curricula are thematically based and involve the use of learning centers in the classroom. Despite the similarity, *Let's Begin with the Letter People* has a particularly strong emphasis on letter knowledge and phonological awareness while *Doors to Discovery (DD)* puts a strong emphasis on language. In an experimental study funded under PCER, Assel and his team (Assel, Landry, Swank, & Cunnewig, 2006) examined the effectiveness of the two curricula across three different settings (Head Start, Title 1, and Universal pre-K classrooms) and included a control group in each setting, in which teachers used teacher-developed, nonspecific curricula. The results revealed that both of the intervention curricula demonstrated similar effectiveness. The auditory comprehension and vocabulary skills of children in classrooms using either of these two curricula grew more than children in control classrooms, but this effect was moderated by program site (Head Start versus Title I versus Universal pre-K). [2] For auditory comprehension, children in Head Start showed the greatest gains compared to children in control classrooms, while for vocabulary, children in Head Start and Title I classrooms showed the greatest gains. Because their primary interest was to identify differences in the rates of growth of child skills over time, the authors acknowledge that their design did not control for differences in children's baseline scores. It was the case that universal pre-K children consistently showed higher initial scores than children in the other two programs, and Title I children outscored Head Start children. Therefore, differences in gains could be due to the fact that the Head Start children, who started with lower baseline scores, had more room to grow.

Two large national evaluations also demonstrated mixed results on children's oral language outcomes. *The National Evaluation of Early Reading First*, using a regression discontinuity (RD) design, evaluated the effect of additional funding for teacher professional development on teacher, classroom, and child outcomes. A variety of curricula were used in funded and non-funded early childhood sites, however, teachers in the funded sites received more professional development in all areas (language & literacy, assessments, and child development and behavior) than teachers in the non-funded sites. The program demonstrated positive impacts on teachers' language use and book reading practices in the funded classrooms. However, no significant impacts were found on children's oral language skills, as measured by the Expressive One-Word Picture Vocabulary Test or the Auditory Comprehension subscale of the Preschool Language Scale-IV. These findings mirror those from the recent Reading First Interim Study (Gamse, Bloom, Tepper, & Jacob, 2008), which also used an RD design to examine the effects of a federal funding stream at the K-3 level. Although the study found positive effects on teacher instructional practice, those effects did not translate into positive effects in student achievement. On the other hand, the National Head Start Impact Study

[2] The authors note that program site was confounded with child ethnicity (i.e., more Hispanic children in Head Start and Title I versus Universal pre-K) so that controlling for site in their design essentially controls for child ethnicity.

found small positive impacts on 3-year old children's vocabulary scores (effects sizes in the .10 to .20 range).

Results of the PCER cross-site evaluation, however, were disappointing with respect to oral language. It should be noted that the lack of effects in the PCER cross-site evaluation could be due in part to small sample sizes, to the timing of the baseline testing, which sometimes occurred later than the baseline testing done by the individual evaluations (Assel et al., 2006), or to differences in measures (Justice et al., under review). Only two of twelve curricula were found to have positive impacts on children's oral language skills in either pre-K or kindergarten: *DLM Early Childhood Express with Open Court Reading Pre-K (DLM)* and the *Early Language and Literacy Model (ELLM)*. For *ELLM*, effects were found only in kindergarten (not in pre-K), a surprising finding given that 11 of the 14 *ELLM* teachers were in their second year of implementation of the curriculum at the time of the evaluation. Effect sizes for both curricula were medium and similar in kindergarten for both curricula on the PPVT and the TOLD (Effect sizes range from .34 to .48), while in pre-K, *DLM's* effects were in the .40 range. One similarity across these curricula is the fact that both *ELLM* and *DLM* are implemented in combination with already comprehensive early childhood curricula and provide teachers with ongoing professional development and support, possibly indicating that the amount of curricular support to teachers needs to be fairly substantial in order to obtain effects on children's outcomes.

In sum, these recent federally-funded research initiatives, although far from conclusive, have provided some confirmatory evidence that children's oral language outcomes can be improved when teachers engage in and provide children with more complex language activities and opportunities. The fact that effect sizes for oral language were medium (according to Cohen, 1988), and not small, is a hopeful finding. The positive results, however, may be moderated by numerous factors, including the instructional support for the teacher, the dosage received by the child, and the program site, which in many cases serves as a proxy for other characteristics of children and teachers in those sites, including baseline test scores, poverty status, or teacher experience. Future research should focus on identifying more concretely the factors that need to be put in place to obtain consistent oral language gains, as well as the size of the effect that is needed to ensure success in reading comprehension.

Phonological Sensitivity

As stated above, the ability to distinguish and manipulate sounds is a strong predictor of reading success. Phonological awareness has been well documented for its critical role in learning to read (e.g., Gunning, 2000; Juel, 1994; Shu, Anderson, & Wu, 2000; Snow, Burns, & Griffin, 1998). Children who are more aware of the different sounds in words and are able to separate or combine sounds are more ready to learn to read and write. Studies have found that explicit instruction in phonological awareness can reduce the incidence of reading failure and thus improve the possibility of reading success (Adams, 1995; Stanovich, 1993; Snow et al., 1998).

In general, less evidence was found that the interventions studied through recent federally-funded research initiatives exerted positive impacts on children's phonological awareness skills than was found in terms of oral language. Neither of the two national evaluations included in this review, of Early Reading First and of Head Start, found effects on children's phonological awareness skills. Similarly, in the PCER cross-site study, 11 of 12 interventions showed no statistically significant effects in this domain (but note that possible limitations for the PCER cross-site evaluation listed above for the oral language domain apply for the phonological awareness domain as well). Only one

intervention – *DLM* – was found to have positive effects in pre-K and kindergarten as measured by the Pre-CTOPPP (pre-K) or the CTOPP (kindergarten) with effect sizes ranging from .32 to .38.

In contrast to the PCER cross-site evaluation findings, however, the individual evaluations of several curricula indicated some positive effects on children's phonological awareness skills. As mentioned above, differences in findings between the cross-site evaluation and the individual evaluations could be due in part to small sample sizes, differences in the timing of baseline testing, or to differences in measures. For example, *Literacy Express* was found to have an average positive effect size of .63 in the phonological processing domain, as measured by the P-CTOPPP Blending and Elision subtests at the end of pre-K (Lonigan, 2006). Similarly, the *Project Upgrade* study demonstrated that *Ready, Set, Leap! (RSL)* had a significant impact on children's phonological awareness skills at the end of pre-K as measured by the TOPEL (ES = .39, when compared to the control group jointly with another intervention, *Breakthrough to Literacy*). In a study of *Let's Begin with the Letter People* and *Doors to Discovery* (Assel et al., 2006), children in classrooms receiving either curriculum showed greater gains in rhyming skill than those in control classrooms, as measured by the Woodcock-Johnson-3 Sound Awareness subtest (d = .26). Additionally, there were differences in rates of growth by curricula that were moderated by program site, such that universal pre-K classrooms using *Let's Begin* had higher rates of growth than those using *DD* by an effect size margin of .85. No differences, however, were found in children's rates of growth between the two curricula in Head Start and Title I classrooms (Assel et al., 2006). The same caveats mentioned above apply to these findings, that is, since Head Start and Title I children began with lower baseline scores than those in Universal pre-K, they may have been more likely to gain at a faster rate .

Some preliminary findings from the ISRC consortium are in line with the aforementioned findings. For example, in the Head Start REDI study, significant impacts on phonological awareness were found (ES =.39 for Blending subtest of the TOPEL, .35 for Elision subtest). This curriculum provided professional development for teachers that focused on implementing sound games (three times per week). The evaluation of *Children's School Success* found an interaction effect between pretest scores and quality of implementation on children's early literacy outcomes, including phonological awareness. The study found that children who scored lower on pretest measures benefited more from high implementation and less from low implementation of the curriculum.

In sum, recent federally-funded research initiatives have provided mixed evidence of the studied curricula's effectiveness to improve children's phonological sensitivity skills. This lack of consensus could be due to methodological issues such as statistical power or differences in measurement of these skills. Or, it also could be the case that gains in this area are difficult to effect. Future research needs to address these methodological issues so as to produce more conclusive results. In addition, as with oral language, moderating factors – such as dosage, children's pre-test scores, and program site – are cited in these studies. Planned variation studies would be an important addition to further clarify the role of these moderators of intervention effectiveness.

Print Knowledge

In line with the core research about the essential role of print and letter knowledge for later literacy success (e.g., Clarke, 1988; Clay, 1991; Torgeson & Davis, 1996; Whitehurst & Lonigan, 2001), the majority of the interventions reviewed targeted children's print knowledge as an essential skill. The goal of these interventions was to improve children's print and letter knowledge skills through training teachers how to a) explicitly teach these skills, and/or b) provide children with opportunities

to practice these skills. Was there evidence that the interventions were effective in improving children's print and letter knowledge? Although not entirely consistent, the majority of interventions that targeted this area showed some evidence of positive effects. The national evaluations of ERF (U.S. Department of Education, 2007) and Head Start (U.S. Department of Health & Human Services, 2005) both had positive impacts on children's print knowledge. Head Start reduced, by almost half (47%) the gap in children's ability to recognize letters between Head Start children and the national average for all 3- and 4-year olds. Similarly, the impact of ERF on children's print and letter knowledge was 5.78 standard score points on the Pre-CTOPPP print awareness subtest (ES = .34).

The PCER cross-site evaluation conducted by Preschool Curriculum Evaluation Research Consortium (2008) indicated positive impacts for only two curricula of eleven that focused on children's language and literacy development – *Curiosity Corner (CC)* and *DLM*. The former curricula had an impact in kindergarten, while the latter had impacts in both pre-K and kindergarten. Of the three measures used, *CC* demonstrated positive impacts on the TERA and the WJ Letter Word Identification subtest (ES = .43 for both), while *DLM* had positive impacts on all three measures in pre-K (the TERA, the WJ Letter Word Identification subtest and the Spelling subtest) equaling effect sizes of .68, .51, and .46 respectively. In kindergarten, *DLM* had impacts only on the TERA and the WJ Letter Word Identification subtest (effect sizes equaled .76 and .50 respectively).

Of the nine remaining curricula that did not demonstrate statistically significant impacts in this domain in the cross-site evaluation, five were studied in individual evaluations and were found to have positive effects (*ELLM, Let's Begin, DD, Literacy Express,* and *Ready, Set, Leap!*). For example, the individual evaluation of *ELLM* suggests that the curriculum, which focuses on instructional strategies and learning materials for teachers to explicitly teach literacy skills and provide structured literacy experiences, had small, positive effects on measures of letter knowledge, print conventions, and meaning of print at the end of prekindergarten in favor of the intervention (effect sizes equaled .25, .28, and .26 respectively). By the end of kindergarten, positive effects were found only on letter knowledge (ES = .34). Similarly, Let's *Begin with the Letter People* and *Doors to Discovery* were both found to have positive effects on Head Start children's print knowledge skills, compared to children in Title I or Universal pre-K classrooms (ES = .53 for HS, versus .06 for Title I and .25 for Universal pre-K). The measure used in the study was the WJ-3 Letter Word Identification subtest. In the case of Literacy Express, the curriculum demonstrated statistically significant positive effects on children's skills in this domain, as measured by several assessments – the TERA-3 Alphabet subtest, the TERA-3 Meaning subtest, and the WJ-3 Spelling subtest (effect sizes equaled .57, .83 and .50 respectively). On two other measure – the P-CTOPPP Print Knowledge subtest and the TERA-3 Print Conventions subtest, impacts were not statistically significant, but were large enough by WWC standards to be substantively meaningful (effect sizes equaled .41 and .34 respectively). Finally, in the *Project Upgrade* study (U.S. Department of Health and Human Services, 2007), RSL, along with *BTL* had significant impacts on children's print knowledge skills, as measured by the Print Knowledge subtest of the TOPEL (ES = .63).

In summary, the majority of curricula evaluated seem to have been able to exert positive effects in the area of print knowledge across varied assessments and conditions, however there is much more to be done. The more extensive NELP review should provide more insight into common features across interventions that show effects on children's print knowledge. Future research would also benefit from moving beyond establishing the link, as done in the ERF evaluation, that more time spent by teachers on print awareness opportunities is related to children's higher print awareness scores, to

identifying more effective ways to teach children alphabetic knowledge. For example, in one non-experimental descriptive study funded by PCER (Justice, Pence, Bowles & Wiggins, 2006), findings based on children in classrooms using either the *Language-Focused Curriculum* or *High/Scope* indicated that the order of letter learning was not random and that some letters hold an advantage over others to influence their order of learning. The authors suggest that perhaps early care and education teachers should teach more difficult, less known letters first, since children are more likely to know more common letters. Teachers should also account for individual differences since children know different letters, depending on both extrinsic and intrinsic influences.

What Evidence Is Provided About Factors that Moderate Intervention Effectiveness?

A review of some of the interventions evaluated for this review points to the range of activities/components that are often implemented with the goal of producing positive changes in children's early language and literacy outcomes. For example, *ELLM* includes five components: research- and standards-based literacy curriculum, family involvement, professional development, working partners, and practice-focused research and evaluation. The interrelationships among these components and their interdependence were prominent, and were discussed in almost every study that was reviewed for this paper. When these comprehensive curricular approaches are implemented in early childhood settings, which are dynamic and complex learning environments in themselves, it becomes difficult to tease out the critical features for success from the wide range of possible influences. Yet, is important to understand what factors might be moderating the effectiveness of interventions. Because variation in these factors was not a focus of this body of research – the aim of which was to provide evidence of effectiveness of the interventions studied, *on average* – researchers were not always able to address questions about moderating factors. In addition, most analyses of moderators were conducted outside an otherwise experimental design, and as such, cannot be considered causal. Despite these limitations, in the research reviewed in this paper, some, mostly non-experimental evidence was provided regarding three possible critical factors that the studies suggest may be important moderators of intervention effectiveness: professional development, dosage of implementation, and child background characteristics.

Professional Development/Coaching

Before implementing the specific curriculum, teachers (and sometimes other educational staff) usually received professional development or training on how to deliver the intervention. Some interventions also provided ongoing coaching to monitor or refresh ideas and to solve problems rising during ongoing implementation. Professional development may affect the impact of an intervention through changing teachers' practice and fidelity. Using non-experimental methods, the *LFC* study showed that treatment teachers exhibited strikingly high fidelity to the curriculum immediately following a professional development workshop (Pence, Justice, & Wiggins, in press). This aligns with the findings of an evaluation of *Building Language and Literacy* in Montgomery County public schools (Ramey, Ramey, Kleinman, Lee, Farneti, Timraz, Nielsen, et al., 2008 unpublished manuscript), which compared two coaching conditions: weekly versus monthly. It revealed that weekly work-embedded coaching significantly improved implementation levels of the curriculum and yielded significant positive impact on children's literacy skills (ES = .44). These contrasts were tested within the experimental design and indicate that sufficient professional development may be related to the success of an intervention.

Professional development can even compensate for the insufficiency in teachers' educational background. The *Project Upgrade* study (U.S. Department of Health and Human Services, 2007) analyzed (outside of the experimental design) the observational data from study classrooms, and, surprisingly, instead of finding an educational background effect, the results demonstrated that the interventions eliminated the differences between better-educated teachers and less-educated teachers. Teachers in the treatment group all looked remarkably similar, regardless of their educational levels, compared with the dramatic differences among control group teachers. In other words, the professional development that treatment group teachers received and the well-specified curricula diminished the differences in teaching instruction due to teacher educational background. Similarly, another group of researchers (Lieber, Goodman-Jansen, Horn, Palmer, Hanson, Czaja, Butera, et al., 2007) examined 30 Head Start teachers in implementing the *CSS* curriculum and found that coaching and teachers' motivation to change, rather than teaching experience or degree status, affected curriculum implementation. These analyses were correlational, and outside of the experimental framework.

Dosage of Implementation

Program dosage can be measured in days of children's attendance during the academic year. When measured in this way, greater program dosage has been found to be related to stronger program impact. For example, in a study of two state public pre-K programs, Ramey, Ramey, and Stokes (year not provided) found that children who received the full day and full school year *LA4* program (Louisiana) gained nearly twice as much from the program as their peers who received only the half-year *LA4* program (pilot year) or the half-day full year Montgomery County Public Schools program (Maryland). Similarly, in an experimental study of *LFC*, researchers found that children who attended early care and education regularly benefited more from the intervention than those with low attendance rates (Justice, Mashburn, Pence, Wiggins, under review). It can be inferred that children who attended school more regularly were exposed to a higher dosage of the intervention compared to those who attended school less regularly.

Program dosage can also be thought of in terms of the amount of time that has been allotted for the curriculum to be implemented. In the *Project Upgrade* study (U.S. Department of Health and Human Services, 2007), the three curricula being compared—*RSL, BELL,* and *BTL*—all focused on the development of early literacy skills and knowledge. However, they were distinguished from one another in terms of instructional approach, materials provided, intensity and cost. Both *RSL* and *BTL* are full-day comprehensive curricula; *BELL* is an add-on literacy program entailing only two 15-20 min sessions daily. The finding that both *RSL* and *BTL* had significant impacts on all literacy measures compared to the lack of impacts of *BELL* suggests that dosage of the intervention should account for part of the differences in impacts on children's outcomes. This is more persuasive considering that *BELL* had a much stronger focus on phonological awareness than the other two curricula, yet had no impact on children's phonological awareness while *RSL* and *BTL* did.

Child Background Characteristics

Findings of some studies also revealed that child background characteristics (such as family economic status, pretest performance, personality, and language ability) may moderate the impacts of interventions. The research demonstrated that the interventions were effective for all children, but were particularly effective for *some* children. For example, children who were more economically or academically disadvantaged were found to have gained more from interventions than their more advantaged peers (Assel et al, 2006; Ramey, Ramey, & Stokes, 2008 unpublished manuscript; Odom,

Diamond, Hanson, et al., 2007). In a study that examined the contributions of child characteristics to the quality of teacher-child relationship, Rudasill, Rimm-Kaufman, Justice and Pence (2006), in their study of *LFC*, demonstrated that individual differences in child temperament and language skills affected teacher-child interactions, which ultimately contributed to intervention effect. This was especially true for early language and literacy curricula, in which teacher-child conversations are often key cornerstones of the implementation.

English Language Learner (ELL) status is also a very important factor when considering children's early language and literacy skills. Although most studies included ELLs in their study samples, results were not often reported by subgroup. This was perhaps due to power issues, since most studies were not powered to detect subgroup differences. An exception is the *LA ExCELS* (Los Angeles: Exploring Children's Early Learning Settings) study which explored ELL children' experience in early care and education settings and their school readiness outcomes (Fuligni, 2008). Preliminary results showed that low income bilingual Spanish children were behind monolingual peers in several language and school readiness domains during pre-kindergarten period. There were no differences in experiences of Spanish speaking and English monolingual children in early care and education programs at age 4. However, participation in early learning settings was particularly beneficial for Spanish speaking children. This is consistent with the aforementioned pattern that academically disadvantaged children benefited more from interventions or programs compared to their advantaged peers.

Child background in terms of family factors also includes family literacy environment and parent behavior. These issues were addressed by the *Getting Ready Nebraska* program. In this program, several studies examined the effects of home literacy and parents' belief or behavior on children's development. In a study investigating adolescent parents' participation in learning and their perceptions of professional support, Knoche, Woods, & Sheridan (2008) found that for children whose parents demonstrated low levels of parent learning behaviors, high levels of professional support were associated with higher scores in young children's language skills.

In sum, the federally-funded evaluation studies reviewed here provide support and replicate previous findings about factors that may be important as moderators of intervention effectiveness. However, many questions remain. For example, how much professional development is optimal? What amount of dosage of intervention is needed for children to progress? What interventions work best for which children? One way to address this in the future would be to conduct planned variation studies, in which hypotheses about "how much" and "for whom" can be tested. From this data, threshold levels for professional development and dosage, for example, could be more clearly understood and ultimately be used to inform intervention developers and policy makers.

Directions for Future Research

In addition to the specific suggestions for future research at the end of each section above, there are several comments regarding future research that apply in general based on the review of this corpus of research. From a substantive point of view, more focused attention should be paid to the needs of some subgroups of children, especially ELLs. As aforementioned, although most studies did a commendable job of including a diverse population of children in their studies, impacts on subgroups were seldom examined. In part, this may be an issue of research design, since effectiveness studies are designed to provide an overall mean, and are often not powered to be able to detect subgroup

differences. However, future research should certainly focus on the specific needs of these children, who make up more and more of the population of children in early care and education settings.

Another substantive issue that should be addressed in future research is trying to determine the active ingredients in those interventions for which positive effects were found. Because most of the PCER interventions, for example, incorporated multiple components, when effects were found, it was not possible to identify which component had led to the positive effect. More fine-tuned research would be able to disentangle the effects of various components and move the field forward in terms of identifying the most critical ingredients of interventions. In addition, the NELP, a much more extensive review, should be able to provide further insights into this question.

From a methodological perspective, it was quite remarkable that there were more than a dozen randomized controlled trial (RCT) studies of early childhood interventions to review. On the one hand, the national push for more rigorous research in the field has certainly increased the number of RCTs that have been implemented and, in this way, has improved the rigor of the research available. On the other hand, effectiveness research studies utilizing RCT designs have their own set of limitations. For example, in terms of statistical power, it is clear that in order to detect the types of effects on children that we would expect across one school year, sample sizes must be fairly large. Although RCTs require fewer units of randomization than say, regression discontinuity designs, it is still the case that in order to detect small effects, sample sizes must typically be in the range of 60 units with nested designs (observations within children within teachers, for example). Since randomization often occurs at the center level to avoid contamination across teachers within the same center, this can be quite a challenge for most researchers. One way to decrease sample size requirements is to conduct random assignment at the child level. This alternative, however, is not always practically or pragmatically feasible.

In addition, there is a trade-off between internal and external validity. Although the strength of RCTs is their high internal validity, they can suffer from low external validity. Especially in early care and education settings, when researchers are often limited to creating their study samples based on those who agree to participate from their overall recruitment efforts, generalizability can still be quite limited and therefore less policy relevant.

Meaningful detectable effect is another methodological issue that arises after reading these studies. In general, effect sizes were reported in terms of Cohen's *d*, and Cohen's guidelines for what is considered small (.20), medium (.50), and large (.80) are used. However, unless the author reports what the range of the assessment is and what the expected growth across a school year is, it is difficult to make a judgment about the substantive meaning of a .20 versus a .50 versus a .80 effect size. What does this mean in real world terms? What is a meaningful effect size? How does that vary by assessment or domain? Without diminishing the advances made in the field in the reporting of effect sizes, it would be helpful to also report a translation of Cohen's *d* into assessment-relevant terms, such as months of growth.

Finally, the PCER and ISRC initiatives have certainly made huge strides in terms of providing examples of conducting evaluations of programs and practices in real world settings. Lessons learned from these initiatives will make an important contribution to the field, both substantive and methodological. Lessons learned could address the wealth of knowledge of the implementers after having done these studies; suggest possible hypotheses for effects, or lack of effects, on child outcomes; and provide direction for future rigorous studies, of which there are certain to be more.

From a policy perspective, the issue of cost was not addressed in any of the studies that were reviewed. In line with the suggestion above regarding cost-benefit analysis in terms of achieving positive child outcomes, research on the cost of implementing the interventions would be useful for policy makers and educators.

Summary

This preliminary review of the published and unpublished papers on these federally-funded intervention evaluations suggests that there is evidence for positive effects of some of the selected interventions on some of the important early childhood language and literacy outcomes. However, evidence from these studies is not sufficient to inform policy makers about ways in which to assemble the critical ingredients necessary for more widespread and consistent success in raising young children's literacy outcomes. Many inter-related factors influence the effectiveness of interventions, some of which are just beginning to be understood. In addition, the experimental studies supported by these federal initiatives have proven that it is possible to conduct rigorous studies in early childhood settings and have moved the field forward in terms of methodology. Ongoing improvements and attention to new issues arising from these more rigorous methodologies will, however, be necessary. Both in terms of substance and methodology, therefore, the studies examined here constitute an important contribution to the knowledge base that informs early language and literacy education. Research on the characteristics of high quality programs that are both developmentally appropriate and successful in bridging the achievement gap will be in demand from legislators and policy makers as they are called upon to make informed decisions about early learning systems.

References

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.

Assel, M. A., Landry, S. H., Swank, P. R., & Gunnewig, S. (2007). An evaluation of curriculum, setting, and mentoring on the performance of children enrolled in pre-kindergarten. *Reading and Writing, 20,* 463-494.

Bierman, K. L., Domitrovich, C. E., Nix, R. L., Gest, S. D., Welsh, J. A., Greenberg, M. T., Blair, C., Nelson, K. E. & Gill, D. (in press). Promoting Academic and Social-Emotional School Readiness: The Head Start REDI Program (Final Revision). *Child Development.*

Bierman, K. L., Nix, R. L., Greenberg, M. T., Blair, C. & Domitrovich, C. E. (2008). Executive functions and school readiness intervention: Impact, moderation, and mediation in the Head Start REDI Program. D*evelopment and Psychopathology, 20,* 821-843.

Bryant, P., MacLean, M., & Bradley, L. (1990). Rhyme, language, and children's reading. *Applied Psycholinguistics, 11,* 237-252.

CEC, Liousville, K. Y. (2007, April 20). *Children's School Success (CSS): Child Outcomes from Three Years of Research*. Power point slides.

Chall, J. S., Jacobs, V., & Baldwin, L. (1990). *The reading crisis: Why poor children fall behind.* Cambridge, MA: Harvard University Press.

Clarke, L. K. (1988). Invented versus traditional spelling in first graders' writings: Effects on learning to spell and read. *Research in the Teaching of English, 22,* 281-309.

Clay, M. M. (1985). *The early detection of reading difficulties* (3rd ed.). Portsmouth, NH: Heinemann.

Clay, M.M. (1991). *Becoming literate: The construction of inner control*. Portsmouth, NH: Heinemann.

Cohen, J. (1988). *Statistical power analysis for behavioral sciences.* Hillsdale, NJ: Lawrence Erlbaum.

Cosgrove, M., Fountain, C., Wehry, S., Wood, J., and Kasten, K. (2006, April). *Randomized Field Trial of an Early Literacy Curriculum and Instructional Support System*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, California.

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology, 33,* 934-945.

Dickinson, D. K. (2001a). Large-group and free-play times: Conversational settings supporting language and literacy development. In D. K. Dickinson & P. O. Tabors (Eds.), *Beginning literacy with language: Young children learning at home and school* (pp. 223-255). Baltimore, MD: Brookes Publishing.

Dickinson, D. K. (2001b). Putting the pieces together: The impact of preschool on children's language and literacy development in kindergarten. In D. K. Dickinson & P. O. Tabors (Eds.), *Beginning literacy with language: Young children learning at home and school* (pp. 257-287). Baltimore, MD: Brookes Publishing.

Dickinson, D. K., & Porche, M. (2008). *The relationship between teacher-child conversations with low-income four-year olds and grade four language and literacy development*. Manuscript submitted for publication.

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*, 104-122.

Dickinson, D. K., McCabe, A. & Clark-Chiarelli, N. (2004). Preschool-based prevention of reading disability: Realities vs. possibilities. In C. A. Stone, E. R. Silliman, B. J. Ehren, & K. Apel (Eds.), *Handbook of language and literacy: Development and disorders* (pp. 209-227). New York: Guilford Press.

Dickinson, D., Darrow, C., Ngo, S., & D'Souza, L. A. (in press). Narrowing the gap between potential and possibility: The challenge of changing classroom conversations. Submitted for publication in: O.A. Barbarin, P. Frome, & D. Marie-Winn (Eds.), *The Handbook of Developmental Science and Early Schooling: Translating basic research into Practice.*

Fuligni, A. S. (2008, May). *School readiness of English-speaking and English-learning children: Links with experiences in early learning settings*. Power Point slides.

Gamse, B.C., Bloom, H.S., Kemple, J.J., Jacob, R.T., (2008). Reading First Impact Study: Interim Report (NCEE 2008-4016). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Good, R., III, Simmons, D. C., & Smith, S. (1998). Effective academic interventions in the United States: Evaluating and enhancing the acquisition of early reading skills. *School Psychology Review, 27*, 740-753.

Gunning, T. G. (2000) *Phonological awareness and primary phonics*. Boston: Allyn & Bacon.

Hart, B., & Risley, T. (1995). *Meaningful differences in the everyday lives of American children.* Baltimore, MD: Brookes Publishing.

Huttenlocher, J., Vasilyeva, M., Cymerman, E. & Levine, S. (in press). Language input and child syntax. *Cognitive Psychology*.

Iversen, S., & Tunmer, W. (1993). Phonological processing skills and the Reading Recovery program. *Journal of Educational Psychology, 85*, 112-126.

Juel, C. (1994). *Learning to read and write in one elementary school*. New York: Springer-Verlag.

Justice, L. M., Cottone, E. A., Mashburn, A., & Rimm-Kaufman, S. E. (2008). *Relationships between Teachers and Preschoolers Who Are At Risk: Contribution of Children's Language Skills, Temperamentally-based Attributes, and Gender*. Unpublished manuscript. University of Virginia.

Justice, L. M., Mashburn, A., Pence, K. L. & Wiggins, A. (2008). Experimental Evaluation of a Preschool Language Curriculum: Influence on Children's Expressive Language Skills. *Journal of Speech, Language, and Hearing Research, Vol.51*, 983-1001.

Justice, L. M., Pence, K., Bowles, R. B., & Wiggins, A. (2006). An investigation of four hypotheses concerning the order by which 4-year-old children learn the alphabet letters. *Early Childhood Research Quarterly, 21*, 374-389.

Knoche, L. L., Sheridan, S. M., Cline, K., Givens, J.A. & Fleissner, S. (2006, June). *Moderating the effects of risk on children's school readiness: What are the roles of family literacy and parent*

sense of competence? Poster session presented at the annual National Research Conference of Head Start, Washington, DC.

Knoche, L. L., Woods, K. E., & Sheridan, S. M. (2008). *Adolescent parents' participation in learning: Factors contributing to their children's development*. Manuscript submitted for publication.

Layzer, J.L., Layzer, C.J., Goodson, B.D., & Price, C. (2007). *Evaluation of child care subsidy strategies: Findings from Project Upgrade in Miami-Dade County.* Final report prepared for U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research, and Evaluation, Washington, DC.

Leather, C. V., & Henry, L. A. (1994). Working memory span and phonological awareness tasks as predictors of early reading ability. *Journal of Experimental Child Psychology, 58(1),* 88-111.

Lieber, J., Goodman-Jansen, G., Horn, E., Palmer, S., Mansen, M., Czaja C., Butera, G., Daniels, J., & Odom, S. (2007, April). Factors that influence the implementation of a new curriculum: Results from two years of implementation. Poster presented at the Biennial Meeting of the Society for Research in Child Development. Boston, MA.

Lonigan, C. J. (2006, July). *Impact of preschool literacy curricula: Results of a randomized evaluation in a public prekindergarten program.* Paper presented at the 13th annual meeting of the Society for the Scientific Study of Reading, Vancouver, British Columbia, Canada.

Lonigan, C. J., Farver, J. M., Clancy-Menchetti, J. & Phillips, B. M. (2005, June). *Promoting the development of preschool children's emergent literacy skills: A randomized evaluation of a literacy-focused curriculum and two professional development models.* Paper presented at the 12th annual meeting of the Society for the Scientific Study of Reading, Toronto, Ontario, Canada.

Lynch, R.G. (2004). *Exceptional returns: Economic, fiscal, and social benefits of investment in early childhood development*. Washington, DC: Economic Policy Institute.

Massey, S. L., Pence, K. L., & Justice, L. M. (2008). Educators' Use of Cognitively Challenging Questions in Economically Disadvantaged Preschool Classroom Contexts. *Early Education and Development, 19(2),* 340-360.

McGill-Franzen, A., & Allington, R. L. (1991). The gridlock of low reading achievement: Perspectives on practice and policy. *Remedial & Special Education, 12*(3), 20-30.

NICHD Early Child Care Research Network (2005). Pathways to reading: The role of oral language in the transition to reading. *Developmental Psychology, 41*, 428 - 442.

Odom, S. L., Diamond, K., Hanson, M., Lieber, J., Butera, G., Horn, E., Palmer, S., Marquist, J., & Fleming, K. (April, 2007). *Children's School Success: Treatment Dosage and Child Outcomes*. Poster presented at the Biennial Meeting of the Society for Research in Child Development. Boston, MA.

Pence, K. L., Justice, L. M., & Wiggins, A. K. (in press). Preschool Teachers' Fidelity in Implementing a Comprehensive Language-Rich Curriculum. *Language, Speech, and Hearing Services in Schools.*

Pinnell, G. S. (1989). Reading recovery: Helping at risk children learn to read. *Elementary School Journal, 90,* 161-184.

Preschool Curriculum Evaluation Research Consortium (2008). *Effects of Preschool Curriculum Programs on School Readiness* (NCER 2008-2009). Washington, DC: National Center for

Education Research, Institute of Education Sciences, U.S. Department of Education. Washington, DC: U.S. Government Print Office.

Pullen, P., Justice, L. (2003). Enhancing phonological awareness, print awareness, and oral language skills in preschool children. *Intervention in School and Clinic.* Austin: Vol. 39, Iss. 2; 87.

Ramey, C. T., Ramey, S. L., and Stokes, B. R. (2008). *Effective pre-K programs: Research evidence about program dosage and student achievement*. Unpublished manuscript. Washington, DC: Georgetown University.

Ramey, S. L., Ramey, C. T., Kleinman, B.E., Lee, L.M., Farnetti, C.C., Timraz, N.M., et al. (2008). *The Effects of Curriculum and Coaching Supports on Classrooms and Literacy Skills of Prekindergarten/Head Start students in Montgomery County Public Schools*. Unpublished manuscript. Washington, DC: Georgetown University.

Reynolds, A. J. (2005). Confirmatory program evaluation: Examples from early childhood programs. *Teachers College Record, 107*, 2401-2425.

Rudasill, K. M., Rimm-Kaufman, S. E., Justice, L. M., & Pence, K. (2006). Temperament and Language Skills as Predictors of Teacher–Child Relationship Quality in Preschool. *Early Education and Development, 17(2)*, 271-291.

Shu, H., Anderson, R. C., & Wu, N. (2000). Phonetic Awareness: Knowledge of orthography-phonology relationships in the Character acquisition of Chinese children. *Journal of Educational Psychology 92*(1), 56-62.

Snow, C. E., Burns, S., & Griffin, P. (Eds.). (1998). Preventing reading difficulties in young children. Washington, D.C.: National Academy Press.

Spira, E. G., Bracken, S. S., & Fischel, J. E. (2005). Predicting improvement after first-grade reading difficulties: The effects of oral language, emergent literacy, and behavior skills. *Developmental Psychology, 41*(1), 225-234.

Stahl, S. A., & Murray, B. A. (1994). Defining phonological awareness and its relationship to early reading. *Journal of Educational* Psychology, 86(2), 221-234.

Stanovich, K. (1993). Does reading make you smarter? Literacy and the development of verbal intelligence. A*dvance in Child Development and Behavior, 25*, 133-180.

Stanovich, K. E. (1986) Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 360-407.

Stevenson, H. W., & Newman, R. S. (1986). Long-term prediction of achievement and attitudes in mathematics and reading. *Child Development, 57*, 646-659.

Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology, 38*, 934-947.

Tabors, P. O., Roach, K. A., & Snow, C. E. (2001). Home language and literacy environment: Final results. In D. K. Dickinson & P. O. Tabors (Eds.). *Beginning literacy with language: Young children learning at home and school* (pp. 111-138). Baltimore, MD: Brookes Publishing.

Tabors, P. O., Snow, C. E., & Dickinson, D. K. (2001). Homes and schools together: Supporting language and literacy development. In D. K. Dickinson & P. O. Tabors, *Beginning literacy with language: Young children learning at home and school* (pp. 313-334). Baltimore, MD: Brookes Publishing.

Torgesen, J. K., & Davis, C. (1996). Individual difference variables that predict response to training in phonological awareness. J*ournal of Experimental Child Psychology, 63*, 1-21.

Torgesen, J. K, Wagner, R. K., & Rashotte, C. A. (1997). Prevention and remediation of severe reading disabilities: Keeping the end in mind. *Scientific Studies of Reading, 1*, 217-234.

U.S. Department of Education, National Center for Education Statistics. (2006). The Condition of Education 2006 (NCES 2006-071). Washington, DC: U.S. Government Printing Office.

U.S. Department of Education. (2007). *National Evaluation of Early Reading First.* (NCEE 2007-4007). Washington, DC: National Center for Education Evaluation and Regional Assistance. U.S. Government Print Office.

U.S. Department of Health & Human Services. (2005). *Head Start Impact Study: First Year Findings*. Administration for Children and Families. Washington, DC.

Vellutino, F. R., & Scanlon, D. M. (2001). Emergent literacy skills, early instruction, and individual differences as determinants of difficulties in learning to read: The case for early intervention. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of early literacy research* (pp. 295-321). New York: Guilford Press.

Wagner, R .K., & Torgesen, J. K. (1987). The natural of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin, 101*, 192-212.

Walker, D., Greenwood, C., Hart, B., & Carta, J. (1994). Prediction of school outcomes based on early language production and socioeconomic factors. *Child Development, 65*, 606-621.

Wasik, B. A., & Slavin, R. E. (1993). Preventing early reading failure with one-to-one tutoring: A review of five programs. *Reading Research Quarterly, 28*, 178-200.

Wehry, S., Cosgrove, M., & Fountain, C. (2006). *Preschool-to-Kindergarten: A Longitudinal Study of the Effectiveness of the Early Literacy and Learning Model (ELLM)*. Poster.

Whitehurst, G. J., & Lonigan, C. J. (2001). Emergent literacy: Development from prereaders to readers. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of early literacy research* (pp. 11-29). New York: Guilford Press.

Appendix B.3:
School Readiness and Early Childhood Education: What Can We Learn from Federal Investments in Research on Mathematics Programs?

Herbert Ginsburg
Teachers College, Columbia University

Ashley Lewis and Margaret Clements
Education Development Center, Inc.

Working Paper prepared for *A Working Meeting on Recent School Readiness Research: Guiding the Synthesis of Early Childhood Research*
Washington, DC
October 21-22, 2008

This paper is part of a series of working papers prepared for a meeting sponsored by the U.S. Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation (ASPE) and the Administration for Children and Families, Office of Planning, Research, and Evaluation (OPRE). Abt Associates Inc and the National Center for Children in Poverty (NCCP) were funded to convene the meeting. The views represented in this paper are those of the author(s) and do not necessarily reflect the opinions of the U.S. Department of Health and Human Services.

Programs designed to promote young children's school readiness have focused primarily on language and social emotional development. While these remain important skills for young children to acquire, there is a growing awareness that readiness for mathematics is also critical. Promoting school readiness for mathematics is particularly important for low-income and/or ethnic minority children who are at greater risk for beginning kindergarten with markedly lower math skills (Lee & Burkham, 2002). In fact, recent research shows that children's mathematics ability at kindergarten-entry is a better predictor of future academic success than their reading achievement (Duncan et. al., 2007). Preschool and prekindergarten programs can buffer children against school failure (Bogard & Takanishi, 2005) and prepare young children for success in primary school mathematics (Arnold & Doctoroff, 2003; Bogard & Takanishi, 2005; Goldbeck, 2001). Considerations like these have led many states to develop early learning standards for mathematics.

In spite of the evidence that early childhood education is the most promising and cost-effective way to positively affect the development of children at-risk for later school failure (Reynolds, 2005), there has been widespread reluctance to teach mathematical concepts to young children. This is because many mathematics educators were not convinced that young children could learn these concepts and because it was unclear how best to teach them (Perry & Dockett, 2002). In fact, some early childhood educators continue to resist the use of any planned teaching or curricula given their long held beliefs that young children need to learn on their own in a child-centered holistic environment and that deliberate teaching is not "developmentally appropriate" (Golbeck, 2001). In addition, many teachers' own fear of math is an obstacle to their willingness to teach mathematics (Ginsburg, Lee, & Boyd, 2008). The result has been that mathematics education has traditionally not figured prominently in early childhood education programs. For example, two major early childhood programs that account for a large portion of the market, Creative Curriculum and High/Scope, have traditionally not emphasized a comprehensive mathematics curriculum. However, both of these programs are in the process of expanding their mathematics offerings.

The historical reluctance to teach mathematics to young children stands in stark contrast to research showing that young children can understand mathematics in complex ways. While it was once thought that young children were incapable of abstract or logical thought because they were in Piaget's preoperational stage, recent research shows that young children can understand basic aspects of number and operations, geometric shapes, spatial relations, measurement, and patterns (Ginsburg, Lee, & Boyd, 2008; Perry & Dockett, 2002). Children's "everyday" mathematical skills can be cultivated and extended at this age level in ways that support a more advanced understanding of mathematical concepts (Ginsburg, Lee, & Boyd, 2008).

In response to the recent research findings demonstrating that young children are eager learners of everyday mathematics, leading mathematics and early childhood education professional organizations now stress the importance of deliberate early childhood mathematics education (National Association for the Education of Young Children and National Council of Teachers of Mathematics, 2002). Their position is that curricula providing organized activities designed to promote students' understanding of mathematical concepts can be used in a deliberate manner by teachers, while still allowing children the opportunity to play and explore the world flexibly (Ginsburg, Lee, & Boyd, 2008; Perry & Dockett, 2002). This approach to early mathematics education fits into prevailing views of quality early childhood education: children should play and be taught, and both should occur in a warm, and nurturing environment.

The goal of this paper is to examine the effectiveness of new research based mathematics curricula that attempt to respond to the call for organized programs of mathematics learning for young children. Given that relatively little rigorous research on preschool mathematics programs has been conducted—whether federally-funded or not—this paper will review research that has been supported by a number of different funding streams: federally-funded studies that were part of the PCER and ISRC initiatives; federally-funded Head Start research; studies funded through other federal programs, including the Institute for Education Science (IES) and the National Science Foundation (NSF); as well as foundation-funded research based in the U.S. and international research. All of the studies reviewed include pre-kindergarten or preschool-aged children (e.g., children who are approximately four years old) in their samples. These children may be attending organized programs like Head Start, or may be in other "preschool" or child-care center settings. In addition, all of the studies focused on improving the math skills of children from low-income families as these children are most at risk for beginning formal schooling with a poorer understanding of mathematics than their non-poor peers.

The first section of the paper focuses on mathematics-specific curricula whose development and/or evaluations have been supported by the federal government, as well as two curricula that were developed or evaluated by other funding sources. The second section will review federally-funded research on comprehensive curricula that include a mathematics component. In each of these sections, we will identify the funding stream and, when applicable, the research initiative supporting the research. The paper concludes with a discussion of what the research does and does not tell us at this point, and recommends directions for future research that would better illuminate the processes of teaching and learning that support mathematics learning in early childhood settings, as well as research designed to determine which underlying components of curricula and implementation are beneficial under the varying preschool and childcare settings that serve children most at risk for starting school with academic skills that lag behind those of their peers.

What Can We Learn from Federally-Funded Research on Early Childhood Mathematics Curricula?

Although leading professional organizations call for research-based curricula, the meaning of "research based" is a bit problematic. A restrictive definition might be that the curriculum should derive directly studies that focus on how mathematics should be taught. By this criterion, almost no programs would qualify. The designs of early childhood mathematics curricula are based on research investigating the development of children's mathematical thinking in the absence of instruction, not from teaching experiments. Thus, a more accurate definition of "research-based" curricula is one that is inspired by research on young children and attempts to translate the research into an organized program of teaching. The danger with this definition is that it can be over-inclusive. Publishers in particular may claim that their programs meet whatever standards are in place at the moment and, not surprisingly, will advertise that virtually any curriculum for young children is research-based (or developmentally appropriate or whatever the slogan of the day may be). Their goal is sales, not scientific rigor.

Our approach is not to take too seriously the claim of a basis in research. After all, the major question is not whether the program derived from research but whether it is effective. Sometimes, practical

applications precede and indeed inspire scientific investigation (Stokes, 1997). A creative curriculum developer may have a hunch, possibly based on some informal exploration, that an activity might work, and indeed it might. The issue is not whether the program is research-based but whether it has been evaluated and is shown to be effective in improving learning outcomes. While to date there have been few rigorous studies examining the effectiveness of mathematics curricula for young children (National Research Council, 2004; D. Clements & Sarama, 2008), the studies that have been conducted indicate that young children from low-income families can indeed benefit from curricula designed specifically to address mathematics learning.

Federally-Funded Cluster Randomized Studies of Mathematics-Specific Curricula: PCER, IES, and NSF Research

Federal dollars have supported the rigorous evaluation of three mathematics-specific early childhood curricula, although the evaluation of each has been supported by a different funding stream. An intervention consisting of the *Pre-K Mathematics Curriculum (PreK Math;* Klein, Starkey, & Ramirez, 2002) supplemented with the *DLM Early Childhood Express Math software (DLM;* D. Clements & Sarama, 2003) was evaluated as part of the Institute for Education Sciences' (IES) Preschool Curriculum Evaluation Research program (PCER; PCER Consortium, 2008). Development and evaluation of the *Building Blocks* curriculum (Sarama, 2004; D. Clements & Sarama, 2003, 2008) has been supported by the National Science Foundation. *Building Blocks* is a designed for use with children as young as three-years-old. The evaluation of the *Big Math for Little Kids* curriculum (BMLK; Greenes, Ginsburg, & Balfanz, 2004) was supported by a research grant from IES (M. Clements, Lewis, and Ginsburg, 2008). *BMLK* was developed for use by pre-kindergarten and kindergarten students.[1]

The three curricula share a number of characteristics, including the types of professional development offered to teachers, the contexts in which the curriculum is designed to be taught, and the scope of the curricula. It is important to note that the similarities noted here do not represent "precise" similarities across the curricula, but rather broad characteristics that they share. The specific representation of each of these characteristics certainly varies across the three curricula, possible in meaningful ways that result in differences in their effectiveness.

Professional development activities were a component of the treatment condition in the rigorous cluster randomized studies used to evaluate each of the curricula. All three of the evaluations included at least one "intensive" workshop on the curriculum before the beginning of the school year. Each of the interventions was also supported throughout the course of the study with regularly scheduled, periodic professional development sessions for teachers. These ranged from bi-weekly, one-on-one sessions in a teacher's classroom to bi-monthly "refresher" courses in which groups of teachers met to review particular aspects of the curriculum.

Another shared characteristic of the curricula is that all are designed to utilize multiple contexts for teaching mathematics. In terms of at school activities, the three curricula include whole class learning activities and small group activities. The curricula also incorporate information and activities designed to be sent home for parents and children to work on together at home.

[1] It should be noted that the authors of this paper conducted the evaluation of BMLK (Clements, Lewis, and Ginsburg, 2008) and that Ginsburg is one of the curriculum's developers.

A third characteristic shared by the these three curricula is that each was designed to be a comprehensive mathematics curriculum covering multiple important mathematics domains , such as numbers, counting, and operations; shapes (geometry); measurement; and pattern. It's important to note that here we are referring to very broad mathematical domains and that the specific content and emphasis of each curriculum may well vary. The major point is that each curriculum sets out to cover multiple important mathematical domains rather than just number and operations or just shapes. See Table 1 for a comparison of the domains covered (broadly defined) by curricula.

The *PreK Math/DLM, Building Blocks,* and *BMLK* curricula also differ in several possibly important ways. While it is true that the three curricula cover many of the same (broadly defined) mathematics domains, it is certain that the specific topics covered within each domain, the scope of coverage for each topic (in terms of depth and/or breadth), the types of activities and lessons developed to teach each topic, and the ways in which various topics and/or domains are integrated with each other varies across the curricula. Investigating the extent of this variation is beyond the scope of this paper. However, a review of published reports and descriptions of the curricula provide some information about these differences. For example, both *PreK Math/DLM* and *Building Blocks* incorporate regular use of computer software, while *BMLK* does not include a software component.

The findings of the rigorous evaluations of the developmentally appropriate mathematics-specific curricula stand in stark contrast to the findings from the Head Start Impact Study. The Head Start Impact study compared children randomly assigned to attend Head Start to a control group of children who, for the most part, attended some other type of center-based care on a number of cognitive domains. Among four-year-olds, the study found a statistically significant positive impact of Head Start attendance on four of eight language-related cognitive domains, but no difference in early math skills (U.S. Department of Health and Human Services, 2005). Given that two of the mathematics curricula reviewed above were evaluated in Head Start classrooms, it appears that the Head Start Impact Study's lack of significant findings regarding math is due to the dearth of effective early childhood mathematics curricula, not Head Start. In fact, the study's final report points to the need for effective early childhood mathematics curricula and teacher professional development in math education (US DHHS, 2005).

While all three of these curricula have been rigorously evaluated using cluster randomized trials, including variation in the length of the studies and the age of children in the study samples, mathematics outcome measures used in the evaluations, and the types of classrooms settings in which the mathematics curricula were evaluated. As we'll discuss below, these differences make it difficult to compare the relative effectiveness of the curricula, other than to conclude that all three demonstrate effectiveness in improving children's understanding of mathematics. The *PreK Math/DLM* and *Building Blocks* evaluation studies examined the impact of each curriculum over the course of children's pre-kindergarten year, and the research took place in a combination of Head Start classrooms and state-funded prekindergarten classrooms. The *BMLK* evaluation, on the other hand, examined the curriculum's impact over the course of children's pre-kindergarten and kindergarten years among children attending child care centers that are subsidized by the New York City Administration for Child Services and, thus, didn't include either Head Start classrooms or state-funded pre-kindergarten programs.

Another important difference across the studies is that each evaluation utilized a different mathematics assessment as the outcome variable. Both *PreK Math/DLM* and *Building Blocks* used assessments developed by the curriculum's developer (and evaluators), neither of which is nationally

normed. Both sets of authors clearly articulate that their assessment is not overly aligned with the curriculum; they are designed to evaluate children's understanding of the concepts taught, but do not use the same activities and materials that are part of the curriculum. At the suggestion of *IES, BMLK* used the mathematics assessment developed for the Early Childhood Longitudinal Study, Birth Cohort (ECLS-B; National Center for Education Statistics) which is nationally normed. Each of the assessment procedures (using versus not using an assessment designed to evaluate a particular curriculum) has its strengths and weaknesses. On the one hand, using an assessment designed to evaluate a particular curriculum is likely to be able to provide a more nuanced understanding about what concepts the curriculum was more (or less) successful at teaching. On the other hand, an assessment that has not been nationally normed will not provide information about how children compare to their peers throughout the country both before and after being taught with the curriculum.

Results from the studies indicate that all three curricula were effective in promoting children's mathematics learning. The effect sizes for the studies were .43 for *BMLK* (compared to control classrooms; M. Clements et al., 2008), .55 for *PreK Math/DLM* (compared to control classrooms; Klein, Starkey, D. Clements, Sarama, & Iyer, in press), and for *Building Blocks* the effect sizes were 1.07 (for the comparison with control classrooms) and .47 (for the comparison with classroom using *PreK Math/DLM*[2] ; Clements & Sarama, 2008). See Table 2 for additional details on the evaluation studies. In all three studies, the control classrooms used a variety of curricula, including Creative Curriculum, High/Scope, Montessori, or other local curricula.

We should note that the difference between the *BMLK* and control students did not become statistically significant until the second year of the study (the year following pre-kindergarten[3]), while the *Building Blocks* and *PreK Math/DLM* studies found statistically significant differences at the end of the pre-kindergarten year. This could be due to several factors. One possibility is that our use of the ECLS-B math assessment (a standardized test that was designed as a general assessment of mathematics learning, and not developed to test a particular mathematics curriculum) resulted in a stricter test of the curriculum's effectiveness and, as a result, additional months of exposure to the curriculum were necessary before differences in children's learning could be detected by this assessment. A second possibility is that more than seven months of exposure to the curriculum were necessary before group differences emerged. In our opinion, the fact that the overall math achievement of children in the study was near the national median score throughout the course of the

---

[2] There are several reasons we do not draw any conclusions about the relative efficacy of Building Blocks and *PreK Math* even though results from the *Building Blocks* evaluation study found that children using it scored significantly higher than children using *PreK Math.* First, the evidence from the federally funded evaluations of the two curricula indicates that both *PreK Math* and *Building Blocks* are more effective than the control curricula used in the control conditions. Second, both evaluations were conducted using mathematics assessments designed by the curriculum developers/evaluators. As we said before, this is not to suggest that that either assessment tool is overly aligned with a specific curriculum or to question either assessment's validity. Rather, we suggest that it is premature to draw conclusions regarding the relative effectiveness of the two curriculum based on two studies, each of which used a different mathematics assessment. In fact, both curricula have been evaluated by the What Works Clearinghouse and received its highest rating: "strong evidence of a positive effect with no overriding contrary evidence".

[3] ACS child care centers offer a kindergarten year for their students, and many students choose to complete the kindergarten year at the child care center rather than transitioning into a public elementary school. Many kindergarten programs in New York City are half-day, while the ACS child care centers offer full-day care and are intended to meet the needs of working parents.

study rules out a third possibility: that the curriculum is too advanced for preschoolers and not appropriate until children reach kindergarten.

An advantage of using the ECLS-B math assessment in the *BMLK* evaluation is that we were able to determine the extent to which *BMLK* helped reduce the achievement gap between children from low-income families and the national average. Specifically, in the fall of pre-kindergarten the average student score on the ECLS-B was at the 48th percentile, but increased to the 56th percentile by the end of the prekindergarten year, and was at the 55th percentile at the end of the kindergarten year.

In summary, evaluations of all three curricula demonstrated that they are effective, with effect sizes ranging from moderate to large. Furthermore, the fact that their effectiveness was demonstrated across a variety of classroom contexts (Head Start, state-funded prekindergarten, and NYC ACS subsidized child care centers) suggests that these comprehensive mathematics curricula are likely to be effective in promoting mathematics-related school readiness among children from low-income families.

A Federally-Funded Cluster Randomized Study of a Comprehensive Curriculum: ISRC research

Among the studies that were funded as part of the Interagency School Readiness Consortium, only the *Children's School Success* curriculum (CSS) included a component designed specifically to advance children's mathematics knowledge. Odom and his colleagues (Leiber et al., 2007; Odom et al., 2007a and 2007b) refer to CSS as an early childhood education model designed to combine science, language, literacy, math, and social skills into a "meaningful learning experience". The mathematics component of the curriculum is described as being adapted from D. Clements and Sarama's *Building Blocks* curriculum, but details regarding the extent to which the curriculum was modified are not provided.

Based on research reports (consisting of slides and posters from conference presentations), it is difficult to discern whether or not CSS was effective in promoting more advanced mathematics knowledge among the children attending Head Start centers where it was implemented. Analyses for this study are still underway, and, to date, none of the presentations provide statistical results demonstrating the curriculum's effectiveness. However, there are multiple presentations that examine the impact of treatment fidelity and children's attendance rates on children's math achievement scores. As would be expected, fidelity of implementation is associated with higher student math scores at the end of the school year.

Other Early Childhood Mathematics Research

In light of the fact that only three early childhood mathematics curricula have been subjected to federally-funded rigorous evaluations, this research review will briefly review additional curricula. They include curricula that were evaluated by non-federal funds and/or by study designs that were not as methodologically rigorous or extensive as those for *Building Blocks, PreK Math/DLM,* and *BMLK*.

The National Science foundation has funded recent research on the '*Round the Rug* curriculum (Casey, 2004), which is a supplementary language arts-based curriculum designed to promote children's understanding of key mathematical concepts including pattern, geometry (shape), measurement, and graphs. The curriculum consists of six books that teachers use to lead lessons that

integrate oral story-telling with hands-on mathematics. The impact of one of the lessons (on geometry) has been evaluated in a smaller-scale random assignment study involving six kindergarten teachers (Casey, Erkut, Ceder & Young, 2008). This study found that a lesson taught using both the story-telling and hands-on components promoted greater mastery of the material than the hands-on lesson alone.

The Head Start-University Partnership, a program of the US DHHS Administration for Children and Families, has supported research on a preschool mathematics curriculum that Katherine Sophian developed for use with three- and four-year olds. The curriculum consists of weekly activities that parents and teachers are to complete with children. The emphasis of the program is on measurement with various units and exploring the relationships between shapes rather than identifying features of shapes (i.e., the number of sides or angles). The curriculum has been evaluated in a study of three Head Start centers with children ranging in age from 2.5 years old to 4.6 years. This study found that use of the curriculum had a small positive effect on the math scores of children at the end of the year.

Discussion

We consider several sets of questions concerning the effectiveness of the programs and what can be learned from the evaluations of them. We conclude with suggestions for a research agenda.

Questions Concerning the Current Programs

How successful are the programs? A basic finding is that math education, as exemplified by the programs described above, can "work" for young children. Studies of different curricula find relatively large effect sizes, as indicated above. They were at least fairly successful in accomplishing their various and sometimes diverse goals. There is little doubt that early education can promote early mathematics learning in different areas, including number, shape, space, and pattern. This is valuable information, and it gets the enterprise started: there should be no doubt that early childhood mathematics education can be effective, at least in the short term.

At the same time, there are many questions remaining to be addressed and much that still needs to be learned. One question refers to the differential effectiveness of the programs under consideration. Do some achieve better results than do others? The answer is probably yes, but it is hard to compare programs directly. As we showed, the research studies used a wide variety of outcome measures for evaluation. As a result, it is hard to examine the relative effectiveness of programs (even using effect size) when they are trying to accomplish different goals. One program may be effective in promoting spatial reasoning and another effective in teaching the reading of numerals. It is good that both are effective, but it is hard to compare programs when goals and subject matter differ.

Further, it is important to note that many evaluations use outcome measures developed in conjunction with the goals of the curriculum (e.g., *PreK Math/DLM*), whereas other programs (e.g., *BMLK*) use measures that do not align with the curriculum itself. In a sense, the aligned outcome measures can be considered near transfer tasks and standardized measures, far transfer tasks. The use of an outcome measure that aligns with the curriculum increases the likelihood that the evaluation will find positive effects, but does not indicate whether the treatment group would perform better on mathematical topics not emphasized in the curriculum. The use of far transfer tasks can provide insight into general aspects of learning but provide little useful detail about the specifics. Each approach has strengths and limitations that need to be recognized.

We also need to be clear about the inevitable limitations of the various outcome measures. Although most have reasonably sound psychometric properties, it is fair to say that of necessity the measures generally focus on relatively easy to measure aspects of performance. The results of such an approach are valuable in establishing that some learning has occurred, but the approach often fails to illuminate that learning in any detail. It is conceivable, of course, that a curriculum "works," in the sense of promoting high test scores on these kinds of evaluations, but that it does not promote thinking or enhance long-term motivation for learning mathematics. It is conceivable that teachers may teach to the evaluation and in the process fail to promote meaningful learning. High stakes assessment may have negative effects at the preK and kindergarten levels, just as it does at higher levels of education.

How successful are the programs at teaching various topics within the mathematics curriculum? Mathematics is a complex subject, even in preschool. It involves far more than teaching rote aspects of number. The discipline is both wide and deep (Ginsburg & Ertle, 2008), and includes topics ranging from the invariance of cardinal number across various transformations to the idea of mapping physical space. Following the advice of the NAEYC/NCTM, many of the curricula present mathematics as a broad array of topics, including number, measurement, space, shape and pattern. At the same time, the program evaluations generally present little information concerning children's learning in each of these specific areas. Consequently, we need to know much more about program effectiveness in teaching the very different topics of mathematics, ranging from number to shape and pattern.

In particular, we need to learn much more about a very special topic, namely mathematical thinking and reasoning. Children need to learn to understand why a figure is a triangle, not a rectangle, and to reason about why one operation (like 2 + 3) yields the same result as another (like 3 + 2). Some of the programs seem to promote such mathematical thinking and reasoning, but in general, the evaluations do not attempt provide in depth information concerning thinking and reasoning processes, strategies employed, and understanding of important ideas. One reason is that random assignment studies involving large numbers of children need to employ tests that are easy to administer on a large scale and relatively short. Such tests, although useful for their purpose, are not optimal for measuring cognitive phenomena as subtle and complex as reasoning and understanding. Another reason is that the field lacks appropriate and practical measures of mathematical thinking and reasoning.

In brief, we need to know much more than that a program "works." We need to know how it works in the different substantive areas of mathematics, and how it works in the key area of mathematical thinking and reasoning. This kind of information can be of great value for researchers, teachers, and curriculum developers alike.

What aspects of the programs' pedagogical methods or materials are most powerful in promoting children's mathematical learning? The programs employ various methods and materials. Sometimes they use small groups, and sometimes the use large ones. Sometimes the approach is relatively didactic and sometimes more open-ended. Sometimes they use games, and sometimes stories. Sometimes they use computers, and sometime they do not. Sometimes they do mathematics as a stand-alone activity, and sometimes it is integrated into other activities.

There are many questions to ask about these practices. How effective are the various methods—games, manipulative, stories, and the like, under various circumstances? How should the various methods be used in presenting the material? These of course are the primary issues of interest to teachers who work every day on teaching mathematics.

A crucial set of questions revolves around teaching. Many of the studies attempt to ensure the fidelity of instruction, in the sense of determining whether teachers teach the material more or less as intended. But the studies pay very little, if any, attention to the ways in which teachers implement the activities, incorporate them into their own teaching styles, find some topics easier to teach than others, interpret the materials, adjust teaching to meet student needs, and understand (or misunderstand) the competence of their students. Teachers are at the heart of any program and curriculum, yet the present studies tell us little about their roles in the enterprise.

In general, because of their broad focus on student outcomes, the evaluations typically provide no information about the strengths and weaknesses of various aspects of the programs, or about intentional teaching. As a consequence, the questions about methods, materials and teaching—the questions of most interest to teachers (and creators of professional development programs)—remain unanswered.

What have we learned about group, individual, and developmental differences in children's mathematics learning? There are substantial differences between SES groups in mathematics achievement. As is well known, low SES children generally perform more poorly than their middle SES peers. It appears that preschool instruction can be effective for both groups, although it may not eliminate the initial gap between them. But it is important to know whether, how, and to what extent the groups differ in their reactions to and learning from various programs. How do the different groups of children interact with the teachers and activities and does that contribute to the outcomes?

There are also wide individual differences in preschool children's psychological functioning, language and mathematical knowledge. Some children enter preschool knowing little English. Some have poor executive function. Some may be stronger than others in number (Dowker, 2005). It is conceivable that some children may benefit more than others from a particular pedagogical method or curriculum.

Similarly, there may be important developmental differences in learning mathematics. The old view that preschoolers in general are "concrete" thinkers, or "preoperational" and therefore cannot learn an abstract subject like mathematics has been discredited. Nevertheless, there may be important differences between typical 3-year-olds and 4-year-olds in their learning of mathematics. What is the nature of these differences?

In general, the evaluation studies, focused as they are on the measurement of broad outcomes, do not provide information useful for addressing issues of group, individual or developmental differences in learning mathematics.

What can we conclude about effectiveness? The evaluation research has shown that the various programs are effective in varying degrees in achieving their varied goals. That is important to know, but the research tells us little more than that, perhaps in part because of the very nature and demands of large-scale random assignment research. The research has little to say about relative effectiveness of different programs, about their success in teaching specific topics, about the relative power of different pedagogical techniques and materials, about how teachers teach, and about group, individual and developmental differences in learning.

A Research Agenda

The current evaluation paradigm has taught us a great deal, and has taken a useful first step in the direction of sound early childhood mathematics education. Yet, as we have shown, the paradigm is limited in its ability to answer key questions. The productive solution is not simply more and bigger RCT studies. Instead, we need a new and wide research agenda dealing with several issues fundamental to early mathematics education.

What and how should we evaluate? One set of issues concerns further evaluation of mathematics programs. Now that we know that many of them work, it is important to conduct research targeted to more specific issues, like the relative effectiveness of different kinds of programs for teaching specific content. What are some effective ways for teaching 4-year-olds the analysis of geometric forms or 3-year-olds some fundamental properties of number? How effective are particular materials or pedagogical methods?

In conducting work of this type, the field can benefit from improved outcome measures that tap into essential aspects of learning across the various topics that comprise the content of early mathematics. We need to get beyond using measures because they are convenient or have sound internal or test-retest reliability. The fundamental question is whether they measure what is important to measure. Fortunately, NIH is now funding the development of new research based measures of mathematics knowledge and other topics relevant to early childhood.

And as we go forward, here's a topic that should not receive much research attention: the long-term effects of early mathematics curriculum. Children's later mathematics outcomes must be influenced by the education children receive after preschool. We know that much of that education, particularly for poor children is lacking, with the likely result that children receiving good preschool math education may not do very well later in school. This outcome is entirely to be expected and does not reflect on the children's abilities or what is possible to achieve. Hence not much effort need be put into studying it. A more effective approach is to work at improving and evaluating education at all levels.

What are the processes involved in mathematical teaching and learning? A second set of issues revolves around the processes of teaching and learning. Mathematics has seldom been taught at the early childhood level. Consequently we know little about how to teach it or how children learn it. Most of the cognitive developmental research that has provided a revolution in the way we conceptualize young children's mathematical abilities does not focus at all on teaching or on how children learn from teaching and in an educational context. The various curricula are "research-based" mostly in the sense that they are inspired by research on children's mathematical competence, and not in the sense that they derive from the research any particular guidance on how to present or teach any topic. Therefore, we need research, some of which needs to be exploratory, that focuses on teaching and on children's learning from teaching in an organized setting. Because so little is known about these topics, this kind of research will ultimately be of great practical value to teachers. By contrast, current evaluation research does not speak to teachers about these issues, except to tell them that effective early math education is possible.

How can we effectively implement math curricula? Many early childhood teachers have no interest in early mathematics, fear it, and do not want to teach it, sometimes because of outmoded notions of developmental appropriateness. School districts, preschools, and childcare organizations typically

give the teachers little help in their efforts to implement mathematics programs. Several questions then arise: What are the obstacles that stand in the way of successful implementation? How can they be overcome? How can one help teachers to cope with their fears of mathematics and learn effective teaching methods (assuming we learn what those are)? What kind of supports—especially professional development—do teachers need over the long term to implement early mathematics education? In general, the problem is first to set up and then examine the effectiveness of an infrastructure for promoting early mathematics education. In the end, everything boils down to helping and supporting teachers to do good work over the long term.

References

Arnold, D. H. & Doctoroff, G. L. (2003). The early education of socioeconomically disadvantaged children. *Annual Review of Psychology, 54,* 517-545.

Bogard, K. & Takanishi, R. (2005). PK-3: An aligned and coordinated approach to education for children 3 to 8 years old. *Social Policy Report: Society for Research in Child Development, 19* (3), 3-23.

Casey, B. (2004). *Mathematics problem-solving adventures: A language-arts-based supplementary series for early childhood that focuses on spatial sense.* In Clements, Sarama, and DiBiase (Eds.), Engaging Young Children in Mathematics (377-389). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Casey, B., Erkut, S., Ceder, I., & Young, J. M. (2008). Use of a storytelling context to improve girls' and boys' geometry skills in kindergarten. *Journal of Applied Developmental Psychology, 29* (1), 29-48.

Casey, B., Kersh, J. E., & Young, J. M. (2004). Storytelling sagas: An effective medium for teaching early childhood mathematics. *Early Childhood Research Quarterly, 19,* 167-172.

Clarke, D. (2001). Understanding, assessing, and developing young children's mathematical thinking: Research as a powerful tool for professional growth. In J. Bobis, B. Perry, & M. Michelmore (Eds.) *Numeracy and beyond* (Proceedings of the 24th Annual Conference of the Mathematics Education Research Group of Australia, vol. 1, pp. 9-26). Sydney: MERGA.

Clark, B. & Shinn, M. R. (2004). A preliminary investigation into the identification and development of early mathematics curriculum-based measurement. *School Psychology Review, 33* (2), 234-248.

Clements, M., Lewis, A., and Ginsburg, H. (2008). *Preliminary Findings from the Big Math for Little Kids Evaluation Study*. Unpublished manuscript.

Clements, D. H. & Sarama, J. (2003). *DLM Express math software*. In *DLM Express Math Resource Package.* Columbus, OH: SRA/McGraw-Hill.

Clements, D. H. & Sarama, J. (2008). Experimental evaluation of the effects of a research-based preschool mathematics curriculum. *American Education Research Journal, 45* (2), 443-494.

Dodge, D. T., Colker, L. J., & Heroman, C. (2002). The Creative Curriculum for Preschool (4th Ed). Washington, DC: Teaching Strategies, Inc.

Dowker, A. (2005). *Individual differences in arithmetic: Implications for psychology, neuroscience and education*. Hove, UK: Psychology Press

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P. Pagani, L. S., Feinstein, L., Duckworth, K., Brooks-Gunn, J., Sexton, H., & Japel, C. School readiness and later achievement. *Developmental Psychology, 43* (6), 1428-1446.

Fantuzzo, J., Bulotsky-Shearer, R., McDermott, P. A., McWayne, C., Frye, D., & Perlman, S. (2007). Investigation of dimensions of socio-emotional classroom behavior and school readiness for low-income urban preschool children. *School Psychology, 36* (1), 44-62.

Ginsburg, H. P., & Ertle, B. B. (2008). Knowing the mathematics in early childhood mathematics. In O. N. Seracho & B. Spodek (Eds.), *Contemporary perspectives on mathematics in early childhood education* (pp. 45-66). Charlotte, NC: Information Age Publishing, Inc.

Ginsburg, H. P., Lee, J. S., & Boyd, J. S. (2008). Mathematics education for young children: What it is and how to promote it. *Social Policy Report: Society for Research in Child Development, 22* (1), 3-22.

Greenes, C., Herbert, H. P., & Balfanz, R. (2004). Big math for little kids. *Early Childhood Research Quarterly, 19,* 159-166.

Griffin, S. (2004a). *Number worlds: A research-based mathematics program for young children.* In Clements, Sarama, and DiBiase (Eds.), Engaging Young Children in Mathematics (325-342). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Griffin, S. (2004b). Building number sense with Number Worlds: A mathematics program for young children. *Early Childhood Research Quarterly, 19,* 173-180.

Goldeck, S. L. (2001). *Instructional models for early childhood: In search of a child-regulated/teacher-guided pedagogy*. In S. L. Goldbeck (Eds.), Psychological Perspectives on Early Childhood Education: Reframing Dilemmas in Research and Practice (3-34). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Klein, A. & Starkey, P. (2004). *Fostering preschool children's mathematical knowledge: Findings from the Berkeley math readiness project.* In Clements, Sarama, and DiBiase (Eds.), Engaging Young Children in Mathematics (343-360). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Klein, A., Starkey, P., & Ramirez, A. B. (2002). *Pre-K mathematics curriculum.* Glenview, IL: Scott Foresman.

Klein, A., Starkey, P., Clements, D., Sarama, J., & Iyer, R. (in press). Effects of a pre-kindergarten mathematics intervention: A randomized experiment. *Journal for Research on Educational Effectiveness.*

Lieber, J., Goodman-Jansen, G., Horn, E., Palmer, S., Hanson, M., Czaja, C., Butera, G., Daniels, J., & Odom, S. (March, 2007). *Factors that influence the implementation of new curriculum: Results from two years of implementation*. Poster presented at the Annual Meeting of the Society for Research on Child Development, Boston, MA.

National Research Council. (2004). *On evaluating curricular effectiveness: Judging the quality of K-12 mathematics education.* Washington, DC: Mathematical Sciences Education Board, Center for Education, Division of Behavioral and Social Sciences and Education, National Academic Press.

NICHD Early Child Care Research Network (2002). Early child care and children's development prior to school entry: Results from the NICHD study of early child care. *American Education Research Journal, 39* (1), 133-164.

Odom, S., Diamond, K., Hanson, M., Lieber, J., Butera, G., Horn, E., Palmer, S., Marquist, J., & Fleming, K. (March, 2007a). *Children's school success: Treatment dosage and child outcomes*. Poster presented at the Annual Meeting of the Society for Research on Child Development, Boston, MA.

Odom, S., Butera, G., Schneider, R., Lieber, J., Sarpatwari, S., Horn, E., Palmer, S., Goodman-Jenson, G., Diamond, K., Czaja, C., Hanson, M., & Ceja, M. (April 20, 2007b). *Children's school*

success (CSS): Child outcomes from three years of research. Paper presented at the Council for Exceptional Children, Louisville, KY.

Preschool Curriculum Evaluation Research Consortium (2008). *Effects of Preschool Curriculum Programs on School Readiness* (NCER 2008-2009). Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education. Washington DC: U.S. Government Printing Office

Reynolds, A. J. (2005). Confirmatory program evaluation: Applications to early childhood interventions. *Teachers College Record, 107* (10), 2401-2425.

Rock, D. A., & Pollock, J. M. (2002). *Early childhood longitudinal study – Kindergarten class of 1998-99 (ECLS-K), psychometric report for kindergarten through first grade* (NCES Working Paper No. 2002-05). Washington, DC: U.S. Department of Education. Retrieved December 17, 2003, from http://nces.ed.gov/pubsearch/pubsinfor.aps?=200205.

Sarama, J. (2004). *Technology in early childhood mathematics: Building Blocks as an innovative technology-based curriculum*. In Clements, Sarama, and DiBiase (Eds.), Engaging Young Children in Mathematics (361-373). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Sarama, J. & Clements, D. H. (2004). Building blocks for early childhood mathematics. *Early Childhood Research Quarterly, 19,* 181-189.

Sophian, C. (2004a). *A prospective developmental perspective on early mathematics instruction*. In Clements, Sarama, and DiBiase (Eds.), Engaging Young Children in Mathematics (253-266). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Sophian, C. (2004b). Mathematics for the future: Developing a Head Start curriculum to support mathematics learning. *Early Childhood Research Quarterly, 19,* 59-81.

Starkey, P., Klein, A., & Wakeley, A. (2004). Enhancing young children's mathematical knowledge through a pre-kindergarten mathematics intervention. *Early Childhood Research Quarterly, 19,* 99-120.

Thomas, G. & Ward, J. (2002). *Exploring issues in mathematics education: An evaluation of the early numeracy project 2001*. New Zealand Ministry of Education, New Zealand.

Young- Loveridge, J. M. (2004). Effects on early numeracy of a program using number books and games. *Early Childhood Research Quarterly, 19,* 82-98.

Table 1. Mathematics Domains Covered by Each Curriculum

Program	Ages	Number	Shape Geometry	Measurement	Pattern	Sorting Sequencing	Logic	Spatial	Data
Big Math for Little Kids	PreK & K	X	X	X	X		X	X	
Building Blocks	PreK through Grade 2	X	X	X	X	X			X
Pre-K Mathematics	PreK	X	X	X	X	X	X		
Children's School Success	PreK				*not reported*				
Sophian's Curriculum	3 & 4 year olds	X	X	X		X			
'Round the Rug Math	PreK through Grade 2	X	X	X	X			X	X

Table 2. Comparison of Curriculum Evaluation Studies

Program	Research Funding	Study Type	Sample	Control Condition(s)	Measure	Effect Size[4]
Big Math for Little Kids	IES Research Grant	RCT	Treatment • 16 preK classrooms • 10 K classrooms Control 16 preK classrooms 10 K classrooms	Prevailing curriculum (e.g., Creative Curriculum, "home grown" curriculum)	ECLS-B Mathematics	.43
Building Blocks (BB)	NSF	RCT	BB • 8 classrooms PMC • 7 classrooms Control • 8 classrooms	PreK Math or Prevailing curriculum (e.g., Creative Curriculum, Montesorri, "home grown")	Early Mathematics Assessment (EMA)	BB vs. *PreK Math*: .47 BB vs. Control: 1.07
Pre-Kindergarten Mathematics Curriculum with DLM Express Software	PCER	RCT	Treatment • 20 classrooms Control • 20 classrooms	Prevailing curriculum (e.g., Creative Curriculum, Montesorri, High/Scope, "home grown")	Child Mathematics Assessment (CMA)	.55
Children's School Success	ICSR	RCT	*not reported*	*not reported*	Woodcock Johnson (WJ), subtest 10 and 18	*not reported*

[4] All mathematics curricula reviewed except for Children's School Success reported a positive statistically significant impact on children's mathematics knowledge.

Additional Information on Mathematics Curricula Reviewed by Ginsburg, Lewis, & Clements

Big Math for Little Kids

The Big Math for Little Kids (*BMLK*) is a mathematics curriculum designed to facilitate mathematics learning for pre-kindergarten and kindergarten students (Greenes, Ginsburg, & Balfanz, 2004). The program includes six units (number, shape, measurement, constructing and partitioning numbers, patterns and logic, and navigation and spatial concepts) containing a sequence of enjoyable activities designed to promote both mathematical understanding and language (Greenes, et al., 2004). The program is designed to be used in whole-class and small-group settings, as well as with individual students. Early field-testing suggested that children taught using the curriculum achieved a high level of mathematical understanding, learned to count to high numbers, were able to take the perspective of others, and anticipated further events and predicted outcomes (Greenes et al., 2004).

The effectiveness of the curriculum has been examined using a two-year randomized controlled trial (RCT) that was funded by the US Department of Education Institute of Education Sciences. The study, which focused on low-income children attending subsidized child care centers in New York City for pre-kindergarten and kindergarten, compared the mathematics achievement of children whose teachers either used the *BMLK* curriculum or continued to teach mathematics using the Creative Curriculum (Dodge, Colker, & Heroman, 2002) or a home grown early childhood curriculum. The treatment teachers attended monthly workshops to deepen their understanding of young children's mathematical learning, as well as to demonstrate important components of the curriculum.

Student achievement was assessed using the mathematics assessment developed for the Early Childhood Longitudinal Study, Birth Cohort (ECLS-B; National Center for Education Statistics) at the beginning and end of their pre-kindergarten year and then again in kindergarten, yielding scores at four time points. The advantage of using the ECLS-B is that it is (1) nationally normed and standardized and (2) that the assessment is not directly aligned with the content of the intervention, providing a stricter test of impact. The norming was conducted using a large stratified random sample including 14,000 children born in 2001 and the measure has high internal reliability (Rock & Pollock, 2002). The test itself is adaptive, meaning that the accuracy of responses determine whether the test taker receives easier or more difficult items, and allows for precise estimation of ability with fewer administered questions.

Preliminary results, using latent growth modeling, are currently available for this study and suggest that children in the *BMLK* group demonstrate a larger increase in mathematics achievement compared to children in the control group. There were no significant differences between the two groups at the beginning of the study, but by the end of kindergarten these differences emerge with a medium effect size (Cohen's d=.43). It should be noted that this study was conducted by the authors of this paper (Clements, Lewis, and Ginsburg, 2008).

Building Blocks

The *Building Blocks* (funded by the National Science Foundation) mathematics curriculum, designed for pre-kindergarten through 2nd grade children, is designed specifically to develop competencies

detailed in the National Council of Teachers of Mathematics "Principals and Standards for School Mathematics" (Sarama, 2004). To this end, the curriculum focuses on developing spatial and geometric competencies, as well as numeric and quantitative concepts (Sarama & Clements, 2004). Within these two areas, three mathematical themes are integrated including patterns, data, and sorting and sequencing (Sarama, 2004). In addition to classroom activities, the curriculum relies heavily on the use of computer software, designed as part of the curriculum, to meaningfully engage children as young as 3 years of age in mathematical concepts (Sarama, 2004; Clements & Sarama, 2003). As a result, teachers are required to provide guidance within and between formats.

A randomized controlled trial (RCT) was conducted, comparing three groups of teachers, namely a group using the *Building Blocks* Curriculum, a second group using the Pre-Kindergarten Mathematics Curriculum (PreK Math; Klein, Starkey, & Ramirez, 2002), and a third control group that experienced whatever teaching was involved in "business as usual". These groups included equal numbers of classrooms, some serving low-income students only and other classrooms serving both low-income and middle-income students.

Researchers assessed the impact of the three mathematics curricula using a measure of mathematics ability that was constructed by Clements and Sarama and includes many of the same mathematics activities that are part of the *Building Blocks* curriculum. The Early Mathematics Assessment (EMA) is administered individually to children during two 10-20 minute interviews, which include detailed protocol, coding, and scoring for the interviewer to follow (Clements & Sarama, 2008). In this study, the interviews were videotaped and recoded to ensure reliability. EMA is a comprehensive assessment of mathematical knowledge, is not aligned with any particular curriculum, and is has high internal reliability (Clements & Sarama, 2008).

Results showed that both the *Building Blocks* and *PreK Math* curriculum groups performed significantly better on the EMA measure than the control group and the *Building Blocks* group performed significantly better than the *PreK Math* intervention group (Clements & Sarama, 2008). *Building Blocks* outperformed the control group with a large effect size of 1.07 and outperformed the *PreK Math* curriculum with a medium effect size of .47. The Pre-Kindergarten Mathematics Curriculum outperformed the control group with a medium effect size of .64. Overall, the program effects for *Building Blocks* were the same regardless of program type (i.e., Head Start or a state-funded program), classroom socioeconomic (SES) composition, and child-level SES. In other words, there was no evidence that the impact of *Building Blocks* varied for different groups of students.

In addition to the impact of the curricula on the composite scores, subscore analyses demonstrated that some skills benefited more from *Building Blocks* than *PreK Math* (count higher without committing errors, describing counting errors, and explaining how to correct counting errors), while for other skills the *Building Blocks* and *PreK Math* students performed equally as well (object counting, verbal counting, comparing numbers, sequencing, shape identification and representation, and identifying counting errors) (Clements & Sarama, 2008).

Pre-Kindergarten Mathematics Curriculum

The Pre-Kindergarten Mathematics Curriculum (PreK Math), originally designed as part of the Berkeley Math Readiness Project has been evaluated as part of the Preschool Curriculum Evaluation Research (PCER) Program. *PreK Math*, was developed for children in grades [XX through XX] (Klein & Starkey, 2004). The curriculum is organized around seven units: enumeration and number

sense, arithmetic reasoning, spatial sense, geometric reasoning, pattern sense and unit construction, nonstandard measurement, and logical reasoning. The small-group activities included in the curriculum use concrete materials and are designed to improve mathematical knowledge, specifically numerical and spatial-geometric thinking (Klein & Starkey, 2004).

An effectiveness study compared children in equivalent numbers of low- and middle-income classrooms using *PreK Math* to a comparison group (Klein & Starkey, 2004). Both income levels were included in order to test the researchers' hypothesis that because the curriculum provides experiences to low-income children that middle-income children were likely to receive at home, the impact of *PreK Math* would be more pronounced among low-income children (Starkey, Klein, & Wakely, 2004). In addition to classroom activities, the authors of *PreK Math* developed a home component of the curriculum, which includes parent classes three times per year designed to teach parents how to use the activities with their children (Starkey, Klein, & Wakely, 2004).

Researchers administered the Child Math Assessment (CMA; Klein & Starkey, 2004; Starkey, Klein, & Wakely, 2004) to both groups of students in the fall and spring of their PreK year. The CMA assesses a wide variety of mathematical concepts using 16 separate tasks, which are administered in two 20-30 minute individual testing sessions. For this study, the assessments were videotaped and coded for reliability (Starkey, Klein, & Wakely, 2004). Half of the children received the first section during the first testing session and the other half received the second section of the test during the first testing session.

The results demonstrated that that mathematics ability for middle-income children in both study groups was significantly higher than that of their low-income peers, and that their mathematics ability grew at a faster rate over the course of the study (Klein & Starkey, 2004). The results also indicated that there was a significant main effect for *PreK Math*, with the intervention group having significantly higher CMA scores. The researchers conclude that while *PreK Math* was effective for both low- and middle-income children, it was particularly beneficial to the low-income students (Klein & Starkey, 2004; Starkey, Klein, & Wakely, 2004).

Researchers conducted a second study (also involving random assignment of classrooms) of *PreK Math* in two early childcare settings—Head Start and state-funded preschools—representing 40 pre-kindergarten classrooms (Klein, et al., in press). Teachers in the treatment group implemented *PreK Math* and the DLM Early Childhood Express Math software (Clements & Sarama, 2003), part of the *Building Blocks* curriculum, while the control group continued their regular curriculum, which included Creative Curriculum, High Scope, Montessori, and other local curricula (Klein, et. al., in press). As in the study described above, the children were assessed using the Child Math Assessment (CMA) and coded from videotapes. As, expected, the math scores of the *PreK Math/DLM* and control groups did not differ between groups in the fall, but by spring the intervention group scored significantly higher than the comparison groups with a medium effect size of .55 (Klein, et. al., in press). This study used a second mathematics outcome measure; this composite score consisted of the CMA, a Shape Compositions task, and the Woodcock Johnson Applied Problems score. Analyses using the composite score also demonstrated a significant difference between the treatment and control groups with an effect size of .62 (Klein, et. al., in press).

Children's School Success (ISRC)

The Children's School Success (CSS) Program is a comprehensive curriculum for preschool children implemented with at-risk children (low income families, students with disabilities, and/or ELL), which focuses on oral language and literacy, science, math, and social competence (Lieber, et. al., 2007). The program views young children as "active, self-motivated learners" and includes student choice, family involvement and individualization into the program's conceptual framework (Odom, et. al., 2007b). The curriculum utilizes "linked learning", or activities that build upon the previous lesson's content, integrates curricular domains across activities, includes a problem solving process, and capitalizes on children's interests and experiences (Odom, et. al., 2007b). The mathematics aspect of the program was adapted from Douglas Clements' *Building Blocks* curriculum and includes number and operations, geometry and spatial sense, measurement, pattern/algebraic thinking, and displaying and analyzing data (Odom, et. al., 2007b).

Three years of research was conducted with approximately 800 at-risk children in Head Start or state pre-k or private childcare centers, in which the majority of enrolled children were of Caucasian/Non-Hispanic dissent (Odom, et. al., 2007b). Student achievement was measured used the Woodcock Johnson Math Subtest. Although the authors did not provide information on the characteristics of the math subtest, a study conducted by the NICHD Early Child Care Research Network (2002) found that the Woodcock Johnson Applied Problems subtest has an internal consistency of .91. This assessment does not align directly with the curriculum itself and as such is less biased in favor of the curriculum.

Presentations on the research have not included analyses comparing the treatment and control groups. Instead, the focus of the presentations thus far has been on the impact of treatment fidelity on children's assessment scores, as well as their initial ability levels. These presentations have presented analyses that show that treatment fidelity has a positive significant association with children's post test scores (after controlling for their pretest scores) on many (but not all) of the outcome measures. The presentations have also shown that, not surprisingly, children with lower test scores at the beginning of the study learned more in high fidelity classrooms than initially low-achieving children in low-fidelity classrooms. The lack of research findings regarding the treatment and control group comparisons, combined with the focus on treatment fidelity in the majority of research conference presentations leads us to wonder whether the evaluation of the CSS curriculum model did not find a significant difference in the treatment and control children on study outcome measures.

'Round the Rug Math: Adventures in Problem Solving

'Round the Rug Math: Adventures in Problem Solving is a supplementary program for pre-K through 2nd grade classrooms that uses stories to teach problem-solving (Casey, 2004; Casey, Kersh, & Young, 2004). This approach teaches mathematics concepts within a language rich medium that extends over the course of many lessons (Casey, 2004). The program specifically focuses on spatial and analytical skills, which can help address learning gaps, so it is not meant to be a comprehensive curriculum (Clements & Sarama, 2008). However, the focus on developing spatial skills is also intended to achieve equity between girls and boys, who consistently show better spatial and geometry skills (Casey, 2004). The program does two things simultaneously: (1) integrates mathematical content into the theme-based approach generally used throughout early childhood curricula, and (2) teaches mathematics content systematically with sequenced lessons (Casey, Kersh, & Young, 2004). Specifically, the 'Round the Rug Math curriculum teachers mathematical concepts in a "systematic, hierarchical progression" through the use of long epic stories, which allow characters to have multiple

adventures the expose students to mathematical problems or concepts (Casey, Kersh, & Young, 2004). Students must solve the problem before going on to the next part of the story, which includes progressively more difficult concepts.

In the first evaluation of the effectiveness of one story on students' geometric understanding was conducted with Kindergarten students, comparing the 'Round the Rug Math curriculum to a control group. The initial results indicate that the students who learned the content with the storybook approach improved significantly more than students who learned the content without the storybook approach, although details what this control group received were not described (Casey, Kersh, & Young, 2004). However, no information on the outcome measure or any statistical information was provided on this study.

A second study comparing the effectiveness of the program by gender suggests that in Kindergarten girls benefit more than boys from learning the mathematical content in a storytelling format (Casey, Erkut, Cedar, & Young, 2008). In this experimental study, six kindergarten teachers were randomly assigned to either the treatment or control group, with 76 students in the treatment group and 79 students in the control group. Two measures were used for pre- and post-test, including Triangles subtest of the Kaufman Assessment Battery for Children (K-ABC) and the Tangram test (Casey, Erkut, Cedar, & Young, 2008). The overall reliability of the K-ABC using a split-half procedure is .86-.93, with the Triangles subtest's factorial loading at .70 for boys and .76 for girls (Casey, Erkut, Cedar, & Young, 2008).

There were higher pretest scores for the intervention group on the Triangle test, but no differences on the Tangram test. For the Triangle test, a repeated measures ANOVA showed a significant improvement from pretest to posttest ($p<.001$), as well as a significant difference between the treatment and control group ($p<.003$), particularly for the girls ($p<.001$; partial $eta^2 = .141$). A comparison of the boys by condition did not yield a significant difference (Casey, Erkut, Cedar, & Young, 2008). For the Tangram test, a repeated measures ANOVA showed a significant improvement from pretest to posttest ($p<.001$), but no other effects.

Mathematics Curriculum Developed by C. Sophian

Sophian (2004a) differentiates the "developmentally appropriate curriculum" as one that matches the cognitive abilities of the learner from what she coined the "prospective developmental perspective", meaning that some mathematical skills are important for learning at a later developmental point. This is often the unspoken goal of early education: teaching students enough so that they are ready and able to learn effectively in later grades and for low-income children the hope is that this preparation closes the achievement gap (Sophian, 2004b). While the development of social competence has long been a goal of Head Start, recent trends in accountability have broadened the focus of early childhood educators generally, and Head Start specifically, to include reading and math skills needed for school success (Fantuzzo, et. al., 2007). In addition, research suggests that low-income children, such as those in Head Start, have less mathematical understanding compared to wealthier counterparts (Sophian, 2004b).

Sophian developed a mathematics curriculum specifically for 3- and 4- year old children attending Head Start centers, which focuses heavily on measurement, object properties, and geometry (Sophian, 2004b). The curriculum is meant to be integrated within the rest of the Head Start program rather than

as a stand alone curriculum. The curriculum was organized into weekly project activities and parents and teachers were given specific activities to complete with the children.

There is great emphasis on combining shapes in new ways and measurement with various units (Sophian, 2004b). Specifically, Sophian (2004b) describes the program as exploring the relationships between shapes rather than identifying features of those shapes (i.e. number of sides or angles). Rather than including measurement as a separate unit within the curriculum, Sophian (2004b) used measurement throughout the curriculum with a specific focus on measuring the same objects with different units of measurement; something she claims is not present in other similar programs.

An evaluation of the program was conduced to determine whether this math program could improve the readiness of low-income young children. Specifically, three Head Start centers, two classrooms within each center, served in the treatment group. Then six Head Start centers were matched on center characteristics and served in the control group; three centers conducted a literacy intervention and three centers continued their regular curriculum (Sophian, 2004b). In this case, the treatment group was provided the mathematics curriculum while the comparison group received either a literacy curriculum or no intervention (i.e. "business as usual" group). Children were assessed in the fall and spring of their pre-K year using an assessment procedure intended to align closely with the curriculum: the Developing Skills Checklist (DSC) and a supplemental measure developed for the study. The mathematics portion of the DSC assesses:

> naming shapes, reproducing and extending patterns, counting, identifying numerals, matching sets and numerals, joining and separating sets, identifying original positions, and logically operations (classification, conservation of number, estimation, and seriation) (Sophian, 2004b, pp. 69).

Using the DSC score, the mathematics intervention group scored significantly higher than either the literacy intervention group or the no-intervention group and there was a significant difference between conditions, using pretest scores as covariates, with an effect size (partial η^2=0.092; Sophian, 2004b). The supplemental score also showed the mathematics group scored significantly higher and there was a significant effect for the mathematics intervention (partial η^2=0.083).

Appendix B.4:
Promoting Children's Socioemotional Development in Contexts of Early Educational Intervention and Care: A Review of the Impact of Federally-Funded Research Initiatives on Young Children's School Readiness

C. Cybele Raver
New York University

Working Paper prepared for *A Working Meeting on Recent School Readiness Research: Guiding the Synthesis of Early Childhood Research*
Washington, DC
October 21-22, 2008

This paper is part of a series of working papers prepared for a meeting sponsored by the U.S. Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation (ASPE) and the Administration for Children and Families, Office of Planning, Research, and Evaluation (OPRE). Abt Associates Inc and the National Center for Children in Poverty (NCCP) were funded to convene the meeting. The views represented in this paper are those of the author(s) and do not necessarily reflect the opinions of the U.S. Department of Health and Human Services.

Abstract

Recent findings in applied developmental science highlight ways that children's socioemotional development may play an important foundational role for later chances of school success. Children's social skills and emotional and behavioral adjustment have been identified as particularly important sources of support for low-income children facing higher risk of school failure. The following report reviews selected models and methods in applied developmental science that focus on young children's socioemotional development. It then reviews recent findings from a large number of randomized trials as well as nonexperimental studies and places those findings in the context of applied developmental science. Lessons learned regarding ways to strengthen children's school readiness will be discussed.

Three decades of research in the fields of developmental psychology and early childhood have suggested that children's socioemotional development is clearly associated with their school readiness (see Blair, 2002; Zaslow et al., 2003). Children have been argued to draw upon positive styles of self-regulation and social skill as key sources of support when navigating new contexts of school (Raver, 2002). Conversely, children who are persistently emotionally dysregulated and behaviorally disruptive have been found to receive less instruction from teachers and to have fewer opportunities for learning from peers (see Arnold et al., 2006; McClelland & Morrison, 2003). However, claims of the role of socioemotional competence for children's later academic achievement have recently received greater scrutiny (Duncan et al., 2007). In addition, recent analyses using the nationally representative Early Childhood Longitudinal Survey-Kindergarten (ECLS-K) data set suggest that preschool experience may pose both risks and benefits to children's long-term chances of success in school (Magnuson, Ruhm, & Waldfogel, 2007). It is against this backdrop that a new set of federally funded research initiatives funded by the U.S. Department of Health and Human Services and the U.S. Department of Education were to test innovative models of program improvement and support for children's school readiness. Findings from these sets of research initiatives are particularly timely from both the standpoints of science and social policy.

Tests of the role of children's socioemotional development for their later chances of success in school become even more pressing in the context of income poverty. Specifically, young children in poverty are more likely to be exposed to multiple ecological stressors such as higher levels of neighborhood and family violence, greater psychological distress among adult caregivers, and a range of other "co-factors" that appear to place children's ability to regulate their emotions and behavior in jeopardy (Brooks-Gunn, Duncan, & Aber, 1997; Li-Grining, 2007; Raver, 2004). Policy contexts (such as early childhood education) that provide direct services to children have been argued to be the most effective means of supporting low-income children's optimal outcomes (Magnuson & Duncan, 2003). This context underscores the significance of major federal investments in evaluations of the impact of interventions targeting low-income children's school readiness (such as the interventions within the ISRC and PCER consortia).

This review provides the opportunity to briefly review emerging findings from this set of major federal research initiatives. After providing a brief definition for each relevant socioemotional construct, this review summarizes the rationale for targeting that domain. Models of program impact mediated through improvements in "instructional support" (such as changes in teachers' use of emotionally and behaviorally supportive classroom practices) are also reviewed, with the recognition that children within this set of interventions were hypothesized to be affected primarily through improvement in the quality and quantity of teachers' instruction. (It is important to note that interventions such as Head Start and Early Head Start have invested in more comprehensive approaches that include provision of family supports and services, but those more comprehensive approaches will not be discussed, here). This review also discusses some of the potential tradeoffs in implementing new curricula in early childhood settings. Specifically, this review examines whether there is any evidence for any unexpected benefits or of any unanticipated negative consequences for children's socioemotional development or for emotionally supportive classroom practices from the implementation of a large number of interventions in preschool settings. Finally, new directions for applied developmental science in early childhood educational settings are briefly outlined.

Contrasting models of the role of socioemotional development for children's school readiness

The empirical "case" for the importance of children's socioemotional development in classroom contexts has emerged from several different traditions in developmental, clinical, and educational psychology. From developmental perspectives, converging lines of inquiry from social developmental and neurobehavioral literatures suggest that children enter schools with distinct profiles of emotional reactivity, regulation and executive functioning that appear to facilitate or hinder their engagement with other learners, teachers, and the process of learning (Blair, 2002; Fantuzzo et al., 2007; Howse, Calkins, Anastopoulos, Keane, & Shelton, 2003; Raver, 2002). Similarly, drawing from a tradition of attachment theory, developmental researchers have highlighted ways that some children establish and maintain relationships with teachers that are characterized by a high degree of mutual positive engagement while other children engage in relationships with teachers that are characterized by a high level of conflict (for review, see Pianta, Justice, Cottone, Mashburn, & Rimm-Kaufman, symposium presentation). Third, clinical and educational psychological studies have highlighted the extent to which children's disruptive, aggressive, and withdrawn behaviors have serious implications for short-term opportunities as well as long-term opportunities for learning, both for children manifesting behavioral difficulty and for their peers (Campbell, Shaw, & Gilliom, 2000). A fourth tradition of observational research in classrooms has highlighted ways that teachers also bring their own regulatory and interpersonal profiles of strength and difficulty to classroom interactions and instruction with their students (LoCasale-Crouch et al., 2007). These four mechanisms are likely to be transactionally, bidirectionally related as children with varying self-regulatory profiles elicit differing patterns of responsiveness versus conflict with teachers. These variables are also likely to be highly confounded by "omitted variables" or unmeasured characteristics across children, teachers, and settings (Duncan, 2003). For these reasons, investigators across developmental, clinical, and educational fields have come to consensus that experimental and quasi-experimental approaches are integral to our ability to draw causal inferences on the roles and modifiability of these processes as predictors of children's school readiness.

In each of the sections below, a brief literature review is provided for each of these four possible mechanisms supporting low-income children's school readiness. Findings from federally funded research initiatives are then considered, with close attention to whether those interventions yielded clear evidence of significant impacts on children's socioemotional development (see Table 1 for summary of interventions' designs, samples, and findings).

Self-regulation: How children handle their emotions, attention, and behavior in classroom contexts

Preschool has long been viewed as an important social context where children learn to follow adults' directions, to handle their own emotions, attention, and impulses with increasing independence from adult regulatory support. Imagine any one of a number of routine classroom scenarios, where children are expected to sit attentively through circle time, line up for trips to the playground or bathroom without pushing or shoving peers, and to follow teachers' directions to gather materials for a writing activity, clean up, or share a favorite book even when children feel tired, bored, or frustrated. Investigators have identified individual differences and growth trajectories in children's ability to handle these regulatory challenges, based on a research tradition focusing on reactivity and regulation (see McClelland et al., 2007; Graziano, Reavis, Keane, & Calkins, 2007 for recent reviews). More recently, children's ability to handle classroom challenges has been examined through a second

Table 1. Brief overview of selected RCT efficacy trials targeting school readiness

Title of Intervention	Principal Investigator	Targeted Sample	Synopsis of intervention/ treatment	Synopsis of control	Analytic approach	Evidence of school readiness benefit?
Project REDI	Bierman	356 urban and suburban/rural southeastern PA HS children (25% African American, 17% Hispanic)	Teacher-delivered, curriculum-based lessons; SEL and literacy enrichment ; teacher training; parent materials	"usual practice" Head Start curricula	HLM, Level 1: child sex and race Level 2: center site, cohort, intervention status	Yes
Chicago School Readiness Project (CSRP)	Raver	90 teachers (71% African American, 20% Hispanic); 602 low-income, ethnic minority children (% African American, % Hispanic) in Chicago HS	30 hours of teacher training, coaching, and mental health consultancy for teacher and children	Teacher aide rather than mental health consultant	HLM, Level 1: child characteristics Level 2: classroom characteristics Level 3: site-level characteristics + randomized status in treatment vs. control	Yes
Tools of the Mind	Diamond	147 low-income, urban students (78% annual income <$25,000)	Teacher training on Vygotskyan emphasis on activities that promote executive functioning	District's version of *Balanced Literacy* curriculum	Multiple regression analyses with age, gender, curriculum, years in curriculum as IV	Yes
Project Approach	Powell	13 teachers with at least a BA in urban Midwest serving 204 ethnic minority children (40% African American, 17% Hispanic)	48 hours of teacher training and support (18 introductory, 12 follow-up, 12 individual consultation)	Teacher-developed, nonspecific curricula	ANCOVA and repeated measures analyses	No, iatrogenic impact reported.
My Teaching Partner (MTP)	Pianta	113 early childhood teachers with at least BA (24% African American, 4% "multiracial") in Virginia serving "at-risk" children in state-funded pre-K	Traditional materials; access to planning materials through website; interactive, web-based consultancy	Materials and website only resource	HLM growth trajectories accounting for observer influence, teacher education and experience, number of students, % of students in poverty	Yes

Title of Intervention	Principal Investigator	Targeted Sample	Synopsis of intervention/ treatment	Synopsis of control	Analytic approach	Evidence of school readiness benefit?
Building Language for Literacy	Ramey	24 classrooms of at-risk, mostly ethnic minority children in Louisiana and Maryland	Job-embedded coaching with literacy skills emphasis and quality of classroom environment	Existing MCPS supports		Yes
N Florida ELLM	Fountain	28 teachers (64% African American) serving 297 children (71% African American, 8% Hispanic) in Florida	5-day training session for literacy coaches, 2-day follow-up months later; teacher training with focus on materials and curriculum; weekly literacy coach visits	Assorted curricula: *Creative Curriculum, Beyond Centers and Circletime, High Reach Learning Pre-K, High/Scope*	ANCOVA; repeated measures analyses	Yes
Pre-K Mathematics		316 children (45% African American, 23% Hispanic) in California and New York	4-day teacher training workshops, ongoing on-site training twice per month, feedback after bimonthly observations	Assorted curricula: *Creative Curriculum, Montessori, High Scope, BPS Benchmarks*	ANCOVA	Yes
Language-Focused Curriculum	Justice	14 teachers and 205 children (21% African American, 5% Hispanic) in rural and suburban Virginia	3-day teacher training workshop and two follow-up sessions over school year, with focus on language stimulation	*High/Scope* curriculum materials	ANCOVA and repeated measures analyses	No statistically significant findings
Doors to Discovery/Let's Begin with the Letter People	Assel	603 pre-kindergarten children (21% African American, 42% Hispanic) in greater Houston area	Teacher training and materials, focus on small group activities and scaffolding/Teacher training and materials, focus on "responsive teaching practices" to encourage strong socioemotional skills; both curricula utilized mentors	Comparison school	Multilevel growth curve modeling	yes

neurobehavioral "lens" with research on children's executive functioning emphasizing the roles of children's working memory, attention deployment, and ability to inhibit prepotent impulses in order to meet external demands (Diamond & Taylor, 1996; Greenberg, Riggs, & Blair, 2007). In applied developmental contexts, investigators have considered children's modulation of positive affect, attention, and behavior in classroom contexts as important "approaches to learning" that are correlated with teacher reports and direct assessments of children's academic skill (Fantuzzo et al., 2007; McDermott, Leigh, & Perry, 2002; Rimm-Kaufman, Fan, Chiu, & You, 2007)

Evidence from a small, extant literature on self-regulation and executive functioning among low-income children suggests that exposure to more poverty-related risks is associated with children's greater difficulty in their executive functioning and self-regulation skills (Li-Grining, 2007; Lengua, 2002). Evidence from recent neurobehavioral research suggests that executive functioning skills are late-developing through early childhood, suggesting an important "window of opportunity" or sensitive period for the development of competent regulation of attention, impulses, and use of working memory in early childhood (Diamond & Taylor, 1996). On the basis of this model of self-regulation and school readiness (see Greenberg, 2006), several federally funded interventions in the ISRC consortium posited that children in treatment group would show significant gains in this domain of school readiness as compared to their control group counterparts (Bierman, Nix, Greenberg, Blair, & Domitrovich, in press-b; Fantuzzo, in preparation, Raver et al., revised and resubmitted).

Was there evidence from these federally funded research initiatives of significant impact of interventions on children's self-regulatory skills? Several studies within the ISRC have found that children would specifically gain in self-regulatory skills when in classrooms that provided greater regulatory support. These have included Project REDI (Bierman et al., in press-a, reporting effect size of d = .29 on direct assessments of task engagement) and the CSRP (unpublished findings). Across these two studies, children in the treatment group were found to demonstrate stronger levels of attention, engagement, or focused effort on a direct assessment of attention and impulsivity at post-test, compared to children in the treatment group. In contrast, no statistically significant differences were found on teacher reports of children's attentiveness, persistence, and other learning-related skills, on the Preschool Learning Behavior Scale (McDermott, Green, Francis, & Stott, 1996; PCER final report, 2008). These null findings are interpreted with caution in this review. This caution is based on concerns for power and correspondingly, the relatively high values that effect sizes would have to achieve, in order to be minimally detectable (see cell sizes and MDEs listed in PCERS final report, pp. 31)

Findings from REDI and CSRP are in line with prior work by Greenberg and colleagues (e.g., Riggs, Greenberg, Kusché, & Pentz, 2006) with older children, suggesting significant program impact on children's executive function, and by recent findings by Diamond, Barnett, Thomas, & Munro (2007) where children assigned to the treatment group receiving the Tools of the Mind curriculum demonstrated significant benefits on a directly-assessed executive function task (the flanker task) relative to their control group assigned counterparts. In short, these findings suggest substantial evidence for the modifiability of children's self-regulatory skills across the preschool year.
What are the implications of these hypothesized and demonstrated short-term gains in children's executive function or self-regulation skills? An optimistic hypothesis might be that children with improved self-regulatory skills may be placed on a more positive developmental trajectory, better able to capitalize on future opportunities for learning in kindergarten and early elementary years. A less optimistic hypothesis is that these behavioral gains will be sustained only as long as children continue to have access to the conditions and classroom practices that supported the development of

executive function and adaptive self-regulation within the intervention year. Future research is needed to learn whether these early gains in children's ability to regulate their engagement, attention, and behavior are sustained into early elementary school years.

Children's social cognitions and prosocial skills in classroom contexts

A parallel area of research has focused on what children know about their emotions and the negotiation of interpersonal problems, emphasizing the social cognitive mechanisms revealed in children's successes versus failures to get along with peers and adults (see classic work by Dodge, Pettit, & Bates, 1994; Conduct Problems Prevention Research Group, 2002). Additional research on children's attachment relationships with teachers, with the development of relationships characterized by closeness versus conflict also informs several interventions funded by the ISRC and PCER initiatives (Hamre & Pianta, 2001). Children's social skills and quality of relationship with teachers have been found to be correlated to their later social and academic competence in early elementary school (see Raver, Garner, & Smith-Donald, 2007 for review). Both of those research areas suggest that children develop relatively stable social cognitions or attributions regarding strategies of getting along with peers and adults in classroom contexts. These attributions appear to be built on a foundation of children's knowledge of emotions, knowledge of prosocial behaviors (e.g., helping, sharing, and taking turns), and the ability to generate and use more effective social problem-solving skills (see Domitrovich, Cortes, & Greenberg, 2007).

Past correlational research has faced the persistent problems of omitted variables bias and reverse causality (or bidirectional influence). For example, children who are temperamentally prone to be more sociable have been found to elicit more positive responses from peers and teachers than do children who express more anger and distress in the classroom (see for example, Justice, Cottone, Mashburn, & Rimm-Kaufman, under review). In the context of those relationships, more well-liked children may have greater opportunities to talk about, process, and remember information about their own and others' feelings, and about strategies for successfully navigating social relationships than might children who are less well-liked. Similarly, children's placement in classrooms with more emotionally supportive teachers and their negotiation of academic as well as social challenges are likely to be at least partially influenced by time-invariant individual and contextual variables that are often "omitted" from models (see O'Connor & McCartney, 2007 for exception and methodological solutions using longitudinal data).

It is within this framework that the federally funded research initiatives targeting children's SEL skills are likely to be of major impact to the field. In this area, randomized trials represent a key opportunity to test causal claims of the role of Social Emotional Learning (SEL) curricula for children's knowledge, attributions, and behaviors regarding prosocial versus aggressive behavior with peers. Interventions targeting teachers' practices also offer the opportunity to test the modifiability of children's relationships with adults in classroom contexts. Outcomes that are commonly tapped in interventions that target children's social problem-solving with peers and positive relationships with teachers include direct assessments of children's emotion understanding, of children's selection of adaptive versus maladaptive strategies in hypothetical vignettes of conflict with peers. Outcome variables also include more general teacher reports of children's social skills as well as teachers' reports of the quality of their relationships with individual children.

With that brief review as an empirical "backdrop," was there evidence from the federally funded research initiatives of significant impact of interventions on children's social problem-solving skills and their ability to get along with peers? Evidence from Project REDI suggests that the intervention, comprised of cognitive and socioemotional curricula as well as teachers' provision of emotion coaching and support was associated with moderate to medium-sized program impacts for children's emotion understanding and interpersonal problem-solving (*d*s ranging from .15 to .39; Bierman et al., symposium presentation). These gains in children's socioemotional skill acquisition were paralleled by substantial gains in treatment enrolled children's generalized social competence, with effect sizes of d = -.28 for teacher rated aggression, d = .26 for observer-rated social competence (p<.08) (Bierman et al., in press-a). These findings are in keeping with prior randomized trial research by Bierman and colleagues (see Greenberg et al., 2007 and Domitrovich et al., 2007 for comparison) and by other senior leaders in the area of low-income children's socioemotional development (see for example, Izard, Trentacosta, King, Morgan, & Diaz, 2007).

Was there evidence from the federally funded research initiatives of significant impact of interventions on children's relationships with teachers? Building on their hallmark program of observational research across large samples in preschool and elementary school contexts, Pianta et al specifically targeted teacher-student relationships as a key socioemotional outcome for their web-based intervention, with evidence of improved teacher-student relationship using observational measures (see below). Similar findings of program impact on the teacher reports of the quality of teacher-student relationship have been informally discussed, but not yet submitted for publication from Project REDI and CSRP. These findings (should they be robust to sensitivity checks using alternative model specifications) would suggest that teacher-child relationships are modifiable. Additional analyses are also currently underway in both the REDI and CSRP labs to detect whether improvements in teachers' relationships with children are bidirectionally related to children's improvements in self-regulation (the teams are constrained from making causal claims regarding those linkages, however; see Raver et al., submitted, for further discussion).

Children's behavior problems

While most of the studies in the ISRC consortium have highlighted children's reductions in their risk for manifesting behavior problems, only two of the seven teams have submitted evidence of significant impact of intervention in this domain. These two studies include Project REDI, reporting reductions of children's aggression by teachers (d = -.28) and by parents (d = -.13, at trend level of significance) (Bierman et al., in press-a). These findings are similar to those yielded by the CSRP team, suggesting significant reductions in children's externalizing and internalizing problems as reported by teachers, and trend-level reductions in children's observed aggressive disruptive behavior in the classroom (Raver et al., revised and resubmitted). Review of the PCER final report suggests that there were null impacts on children's behavior problems in the pre-Kindergarten year, with point estimates of program impact using the SSRS Problem Behaviors Scale) reported to be small in magnitude and signed in inconsistent directions. Of concern is the finding that one intervention (Project Approach) appears to have yielded evidence of negative impact on children's behavior problems in the Kindergarten year, with children in the treatment group showing significantly higher numbers of behavior problems than the control group. It is important to highlight however that that finding has not been replicated in any of the other 20 studies in the two consortia.

Mechanisms of improvement in children's socioemotional development through improvement in the quality and quantity of instruction

How were these child-focused program impacts achieved? Consistent across all interventions reviewed was a clear emphasis on multi-day trainings for teachers, followed by extensive "coaching" support and attention to fidelity of implementation. Some studies (but not others) have also published findings of proximal improvement in classroom practices as a result of the implementation of the interventions planned. That smaller set of studies is reviewed below.

Findings from My Teaching Partner suggest that teachers who received web-based consultancy as well as web-based access to information on ways to improve instructional strategies made significant improvements in their classroom practices, as compared to teachers with access to web-based information, only (Pianta, Mashburn, Downer, Hamre, & Justice, submitted). Teachers in the treatment group were found to show significant gains in sensitivity, language modeling, and quality of instructional support to students, as compared to teachers in the control group. Effect size estimates are reported and therefore must be understood in terms of change over time: The investigators report unstandardized regression coefficients of B = .07 to .09 per unit of time (30 days). Briefly, this means that treatment group programs averaged .42 to .54 of a point gain (on the CLASS 7-point scale) relative to programs in the control group, in a six month period. Importantly, gains were substantially larger for programs with very high proportions of poor children enrolled in their classrooms (see figures).

Similarly, Project REDI targeted teachers' generalized classroom practices and induction strategies as well as their use of SEL curricular lessons to increase the level of emotional support and contingency to children's emotional and social experiences (Bierman et al., in press-b; Bierman, personal communication, May 2008). Teachers' use of emotion coaching and improvements in overall classroom management and behavioral support were significantly improved by the REDI intervention (Domitrovich et al., revised and resubmitted). Importantly, results from the REDI team suggest that these changes in classroom processes were powerful predictors (and likely mediators) of children's language and socioemotional gains (Bierman et al., presentation). From a congruent theoretical framework, CSRP aimed to improve children's self-regulation and opportunities for learning by increasing teachers' use of emotionally supportive classroom practices where teachers maintained clear, firm yet warm patterns of limit-setting (see Raver et al., 2008). In contrast to project REDI, no specific child-focused curricula on emotional language or self-awareness were specifically targeted in CSRP. Findings from the CSRP intervention suggested that classroom climate was significantly benefited (d = .52 to d = .89). CSRP findings of intervention impact on positive classroom climate support the hypothesized mechanism of influence for intervention-enrolled children's observed gains in self-regulation, relative to their control group enrolled counterparts.

Findings from some of the PCER studies provide sparse but congruent evidence of improved emotionally supportive classroom processes as a result of intervention. The University of Virginia team, for example, targeted both teachers' increased use of language-rich classroom activities and the complexity of the language that teachers use when conversing with children (Pence, Justice, & Wiggins, in press). Analyses of the impact of this intervention suggest that teachers made changes in their activities most quickly, but were able to improve the quality of their conversations (described as a "relational process") with the children in their classrooms, also (Pence et al., in press). Ramey et al. (submitted) also primarily targeted teachers' language and literacy instruction using two different levels of coaching (weekly and monthly) in the Building Language for Literacy intervention trial, but also collected independent observations of teachers' time spent engaged in emotionally less supportive practices such as "placing restrictions on

children" and "negative/harsh treatment" of children. In the report included for this review, the investigators chose not to analyze whether difference between intervention conditions on these measures were statistically significant (see pp. 21), but inspection of the means on both measures suggests that point estimates of differences between the groups appear to favor treatment assigned classrooms.

Building relationships between teachers and intervention staff

All the intervention models reviewed above (e.g., MTP, REDI, CSRP) as well as most other models in the ISRC that are currently analyzing their data for evidence of treatment impact (led by Fantuzzo, Kupersmidt, Odom, Sheridan) have relied on significant investments in "coaching" of teachers in supporting gains in classroom climate. Similar levels of investment in training and coaching were found in all studies reviewed from the PCER consortium (e.g., Ramey et al., submitted; Assel, Landry, Swank, et al., in press; Cosgrove, Fountain, Wehry, Wood, & Kasten, submitted; Klein, Starkey, Clements, Sarama, & Iyer, in press).

Across all interventions using "coaching" or consultation approaches in the ISRC consortium, levels of coaching were commensurate with levels used in the language- and literacy interventions in the PCER group (e.g. N Florida ELLM used two days of intensive training followed by 1 hour weekly coaching sessions across the school year while training for Pre-K Mathematics included 2 4-day trainings and 15 on-site coaching sessions). Comparison of models across all ISRC and PCER studies that employed coaching suggests several commonalities, including emphasis on "job embedded," collaborative models (including cycles of modeling, observation and feedback) between teachers and coaching staff (see Cosgrove et al., submitted; Raver et al., 2008). In short, intervention staff focused substantial levels of effort in building trusting, collaborative relationships with teachers (see Brown, Knoche, Edwards, & Sheridan, submitted for case study).

With variations on this coaching and training model, multiple teams demonstrated significant improvements in teachers' classroom practices (see above). Building of positive, supportive coaching relationship may be particularly important given that interventions may be asking teachers to be reflective, self-critical, and willing to take the risk of trying new approaches in the ways that they run their classrooms. In one study, for example, teachers in the treatment group reported increasing levels of efficacy in implementing language stimulation techniques over the school year (Justice et al., under review). Importantly, teachers in the treatment group were also found to report lower, rather than higher levels of self-efficacy and comfort when compared to teachers' ratings of self-efficacy in an untreated control group. These findings, though drawn from a single intervention trial, are congruent with other studies that document the challenges that teachers face as well as the gains that they are capable of making in programs emphasizing professional development and quality improvement (see Li-Grining et al., submitted; Brown et al., submitted). Extensive focus group and evaluation surveys conducted by Pianta's team suggest that teachers generally reported feeling supported by consultancy services, even when they are web-based (Whitaker, Kinzie, Kraft-Sayre, Mashburn & Pianta, 2007).

An obvious next question is whether there is a threshold level to the amount of coaching needed to support improvements in the quality and quantity of instruction. Ramey, Ramey, and Stokes (in preparation) raise this by pointing to contrasting models of "coaching" in weekly versus monthly delivery schedules, with no clear evidence that more frequent coaching yields substantially greater benefit than less frequent coaching. This represents an important new direction for future research.

Checking to determine whether there were unanticipated benefits or drawbacks of early intervention for children's socioemotional development

One fair question might be whether there are unanticipated "spillover" benefits from focusing on child language, literacy and math outcomes on children's socioemotional outcomes. One hypothesis might be that children may gain increasingly strong regulatory skills through more cognitively demanding and engaging curricula, where the content of teachers' lessons helps to entrain and strengthen children's attentional and memory skills (see Doctoroff, Greer, & Arnold, 2006). A contrasting hypothesis might be that children might respond negatively to more cognitively demanding and firmly structured classroom practices and curricula, showing increased behavioral difficulty that might offset language, literacy, or math gains.

Several ISRC interventions used "hybrid" models combining foci on language/literacy as well as children's socioemotional development and analyses of treatment impact will elucidate whether there were consistent benefits or costs to children's behavioral development, across interventions (see interventions led by Pianta, Fantuzzo, Odom, Kupersmidt, and Bierman). Of the ISRC hybrid models tested, Project REDI provided data to support improvements, rather than decrements in children's socioemotional development as well as in their language development (see above). Across 13 of the 14 interventions in the PCER evaluation, teachers in the intervention groups and teachers in the control group did not differ on the level of their students' behavioral difficulty or social skills (using the SSRS; Gresham & Elliott, 1990). Again, these null findings should be interpreted with caution. The one exception was that children in the Learning Approaches treatment group were found to fare less well on socioemotional measures than were children in the control group, as rated by kindergarten teachers (see above). With that exception noted, there was no clear evidence of negative consequences for teacher-child interaction. Nor is there evidence for negative behavioral or emotional consequences for children's socioemotional development, in almost all studies where teachers were extensively trained and monitored to implement significantly more cognitively demanding interventions.

Another way to explore this question is to consider whether teachers' training, time, or curricular focus on academically focused outcomes might inadvertently lead classrooms to become too tightly structured, overly cognitively demanding, or somehow less emotionally or behaviorally supportive. Descriptive data from many of the non-experimental studies submitted for this review, however, suggest that the risk of preschool classrooms becoming overly cognitively demanding is relatively low. For example, descriptive work by the Howes & Fuligni team (Fuligni, revised and resubmitted) as well as work by Justice et al. (under review) on the preschool activity contexts and preschoolers' exposure to language suggests that relatively low percentages of class time are spent engaged in instructional effort. Similarly, Massey, Pence, Justice and Bowles (2008) report that teachers' use of more cognitively challenging questions is limited to approximately 11% of their utterances directed to the low-income children in their classrooms (pp. 12). While speculative, it does not appear that those classrooms included in this broad range of studies were already too tightly paced or cognitively demanding, prior to implementation of the intervention. Put another way, there may be significant regulatory benefits, and possibly fewer regulatory "costs" to raising the "bar" for teachers' structure and pacing of cognitively demanding material in classrooms serving low-income children.

The PCER 14-study evaluation offers limited but important opportunity to examine this question: Data on the quality of teacher-child interaction were collected three times during the school year

across all 14 studies (as rated by observers using Arnett scales) (Preschool Curriculum Evaluation Research Consortium, 2008). Overall, statistically significant evidence of beneficial "spillover" effects in improving the classroom climate were found for the Creative Curriculum intervention, where treatment-assigned teachers were observed to be less detached and more positive in spring than were teachers in control group classrooms. Though non-significant, evidence from seven of the exclusively literacy/language oriented curricula demonstrated point estimate differences between treatment and control groups that were in the right direction (e.g., with point estimates of effect sizes equal to .38 or higher) (see Preschool Curriculum Evaluation Research Consortium, 2008, pp. xliv). In sum, measured indicators of classroom quality across all studies but one suggest that placing higher demands on teachers' instructional practices using either language/literacy or "hybrid" intervention models did not lead to measurably negative impacts and in one case (mentioned earlier), the implementation of these interventions led to clear benefits regarding the socioemotional climate of the classroom.

Directions for future research in promote children's readiness for school

The role of child, family, classroom, and context characteristics as moderators

Increasingly, randomized trials have been analyzed with attention to moderating roles of "person" and "place," where interventions may fit the needs of some children, in some contexts more than the intervention might for other children, in other contexts (Gorman-Smith & Tolan, 1998). The role of moderators was explored in some studies reviewed here, but not in others, and they represent a very promising direction for future research.

A small number of studies considered the role of child characteristics, such as child gender, race/ethnicity, English-language-learner status, and risks for self-regulatory or expressive language difficulty. For example, children at higher levels of behavioral and cognitive risk (e.g. those children who are more temperamentally or neurocognitively prone to high levels of shyness, impulsivity, or distractability) might be expected to benefit more greatly or less greatly from interventions (see Bierman et al., in press-a for review). Yet this review suggests that few of the socioemotionally–oriented. "hybrid," or cognitively-oriented interventions (in ISRC and PCER) considered whether intervention impacts were greater or smaller for children with greater proneness to regulatory skill or difficulty. One exception was the nonexperimental finding that children with greater proneness to shyness had significantly more difficult time establishing positive relationships with teachers in nonexperimental analyses of one PCER- funded intervention (Justice et al., under review). Importantly, child temperament moderated relations between children's language skills and student-teacher relationship, where children who were temperamentally prone to anger and had low expressive language abilities were at particularly high risk of conflictual relationship with their preschool teachers (Justice et al., under review). Additional findings of moderation of intervention impact by child risk were found for Raver's team for observational measures of child behavioral problems (Raver et al., revised and resubmitted). In future, it will be important to carefully consider whether program impacts are larger or smaller for children with differing profiles of strength versus risk.

Family level risk may also be important and parsimonious way to consider fit of different intervention models for families with substantially differing economic and psychosocial resources. Findings by Pianta's team of clear, larger benefit of the MTP program for serving very high-poverty classrooms as compared to programs serving proportionally fewer poor children highlights the importance of

including family-level income poverty and related risks in models. A third important set of moderators are those of program type and program resources. For example, an intervention targeting the emotional climate of classrooms may be difficult to implement in settings that are chaotic or disorganized, or under-resourced (see Raver et al., 2008 for review). In contrast, programs that have mental health consultants on staff, on-site personnel to address teacher training, quality improvements, etc. may already be sufficiently resourced that they are likely to show little, if any benefit of additional services implemented through our intervention efforts. In short, it is important to include some observable indicators of level of program resources as covariates and as moderators, to detect whether programs with higher organizational capacity are able to benefit from intervention more so than others (see Assel, Landry, Swank, & Gunnewig, 2007 for examples of heterogeneity of child level program impacts across program type).

The importance of socioemotional measures in study analyses

Past reviews have highlighted the importance of including socioemotional measures as well as cognitively oriented measures when benchmarking intervention impact (e.g. Raver & Zigler, 1997). There are several key benefits (highlighted earlier) for including socioemotional measures at both child- and classroom levels, even when interventions are targeted toward children's language and literacy. The inclusion of child social skills and behavior problem measures in the PCER evaluation and some individual PCER studies (e.g., Klein et al., in press) helps to rule out concern, for example, that there may be iatrogenic sequelae from the introduction of interventions targeting language and literacy. Similarly, the inclusion of children's language and math skills in interventions that target only classroom socioemotional processes offers the opportunity to test whether there are costly tradeoffs (in terms of lower instructional time) or unanticipated benefits (in terms of children's language gains) when focusing program improvement efforts on socioemotional processes. This cross-domain integration of measures at child- and classroom levels represents an important area of future collaboration and future research.

The importance of modeling cluster-randomized status in study analyses

From a methodological standpoint, the impact of a number of these interventions on children's socioemotional development was difficult to interpret for this review because of variability in the ways that data were analyzed and reported. A substantial number of studies provided careful, sophisticated analyses of program impact, using Intent-to-treat analyses, multi-level modeling (e.g., HLM), and clear description of model specification so that the role of cluster-randomized status to treatment versus control groups could be clearly identified. In contrast, a smaller number of studies limited their reports to analyses of program fidelity as a predictor of child-level or classroom-level outcomes, effectively reintroducing selection bias into designs that were initially randomized. Future research in this area would be substantially strengthened by a tiered reporting process, whereby intention-to-treat (ITT) analyses and treatment-on-treated/ dosage analyses could both be encouraged.

Summary

At this early stage of review, most research teams have only recently wrapped up final stages of data collection and completion of preliminary data analyses. Few research teams have completed the full set of ITT analyses that are needed to be able to determine the individual and collective impacts of preschool intervention on children's socioemotional outcomes (A "full set of ITT analyses" would include tests of moderation and sensitivity checks regarding whether program impact estimates are sensitive to model specification). With that caveat in mind, preliminary review of the current set of

published and unpublished papers suggests clear evidence for the benefits of several intervention approaches in supporting low-income children's socioemotional development across their preschool year. Findings of improved classroom instructional processes and improved classroom emotional climate across both types of interventions suggest that interventions using teacher training and coaching models yielded substantial improvements in program quality. Children in treatment groups were found to show lower behavioral problems, increased self-regulatory skills, and greater prosocial skills with peers and with teachers, than their counterparts in control group classrooms, in a smaller number of interventions. As these trials are completed, they are likely to make a major contribution to our knowledge of the ways that scientists and policy makers can best support the school readiness of our nation's low-income children.

References

Arnold, D. H., Brown, S. A., Meagher, S., Baker, C. N., Dobbs, J., & Doctoroff, G. L. (2006). Preschool-based programs for externalizing problems. *Education & Treatment of Children, 29*, 311–339.

Assel, M. A., Landry, S. H., Swank, P. R., & Gunnewig, S. (2007). An evaluation of curriculum, setting, and mentoring on the performance of children enrolled in pre-kindergarten. *Reading and Writing, 20,* 463–494.

Bierman, K. L., Domitrovich, C. E., Nix, R., Gest, S., Welsh, J. A., Greenberg, M. T., et al. (in press-a). Promoting academic and social-emotional school readiness: The Head Start REDI program. *Child Development*.

Bierman, K., Nix, R. L., Greenberg, M. T., Blair, C., & Domitrovich, C. (in press-b). Executive functions and school readiness intervention: Impact, moderation, and mediation in Head Start REDI program. *Development and Psychopathology*.

Blair, C. (2002). School readiness: Integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *American Psychologist, 57,* 111–127.

Brooks-Gunn, J., Duncan, G., & Aber, L. (Eds.). (1997). *Neighborhood poverty: Context and consequences for children (Vol. 1).* New York: Russell Sage Foundation.

Brown, J. R., Knoche, L. L., Edwards, C. P., & Sheridan, S. M. (submitted). Professional development: A case study of early childhood programs in the Getting Ready Project.

Campbell, S. B., Shaw, D. S., & Gilliom, M. (2000). Early externalizing behavior problems: Toddlers and preschoolers at risk for later maladjustment. *Development and Psychopathology, 12,* 467–488.

Conduct Problems Prevention Research Group (2002). Using the Fast Track randomized prevention trial to test the early-starter model of the development of serious conduct problems. *Development and Psychopathology, 14,* 925–943.

Cosgrove, M., Fountain, C., Wehry, S., Wood, J., & Kasten, K. (submitted). Randomized field trial of an early literacy curriculum and instructional support system.

Diamond, A., Barnett, S., Thomas, J. & Munro, S. (2007). Preschool program improves cognitive control. *Science, 318,* 1387–1388.

Diamond, A., & Taylor, C. (1996). Development of an aspect of executive control: Development of the abilities to remember what I said and to "Do as I say, not as I do." *Developmental Psychobiology, 29,* 315–334.

Doctoroff, G. L., Greer, J. A., & Arnold, D. H. (2006). The relationship between social behavior and emergent literacy among preschool boys and girls. *Journal of Applied Developmental Psychology, 27,* 1–13.

Dodge, K. A., Pettit, G. S., & Bates, J. E. (1994). Socialization mediators of the relation between socioeconomic status and child conduct problems. *Child Development, 65,* 649–665.

Domitrovich, C. E., Cortes, R. C., & Greenberg, M. T. (2007). Improving young children's social and emotional competence: A randomized trial of the preschool "PATHS" curriculum. *Journal of Primary Prevention, 28,* 67–91.

Domitrovich, C. E., Gest, S. D., Gill, S., Bierman, K. L., Welsh, J. & Jones, D. (revised and resubmitted). Fostering high quality teaching with an enriched curriculum and professional development support: The Head Start REDI program.

Duncan, G. J. (2003). Modeling the impacts of child care quality on children's preschool cognitive development. *Child Development, 74,* 1454–1475.

Duncan, G., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., et al. (2007). School readiness and later achievement. *Developmental Psychology, 43,* 1428–1446.

Fantuzzo, J., Bulotsky-Shearer, R., McDermott,P. A., McWayne, C., Frye, D., & Perlman, S. (2007). Investigation of dimensions of social-emotional classroom behavior and school readiness for low-income urban preschool children. *School Psychology Review, 36,* 44–62.

Fuligni, A. (revised and resubmitted). Children's experiences in early childhood programs for low-income children: Influence of program type and curriculum use. *Early Education and Development.*

Gorman-Smith, D., & Tolan, P. (1998). The role of exposure to community violence and developmental problems among inner-city youth. *Development and Psychopathology, 10,* 101–116.

Graziano, P. A., Reavis, R. D., Keane, S. P., & Calkins, S. D. (2007). The role of emotion regulation in children's early academic success. *Journal of School Psychology, 45,* 3–19.

Greenberg, M. T., Riggs, N. R., & Blair, C. (2007). The role of preventive interventions in enhancing neurocognitive functioning and promoting competence in adolescence. In D. Romer & E. F. Walker (Eds.), *Adolescent psychopathology and the developing brain: Integrating brain and prevention science* (pp. 441–462). New York: Oxford University Press.

Gresham, F. M., & Elliot, S. N. (1990). *The Social Skills Rating System.* Circle Pines, MN: American Guidance Services.

Hamre, B. K., & Pianta, R. C. (2001). Early teacher-child relationships and the trajectory of children's outcomes through eighth grade. *Child Development, 72,* 625–638.

Howse, R. B., Calkins, S. D., Anastopoulos, A. D., Keane, S. P., & Shelton, T. L. (2003). Regulatory contributors to children's kindergarten achievement. *Early Education and Development, 14,* 101–119.

Izard, C., Trentacosta, C., King, K., Morgan, J., & Diaz, M. (2007). In G. Matthews, M. Zeidner, & R. Roberts (Eds.), *Emotions, emotionality, and intelligence in the development of adaptive behavior: Knowns and unknowns. Series in affective science* (pp. 127–150). New York: Oxford University Press.

Justice, L. M., Cottone, E. A., Mashburn, A., & Rimm-Kaufman, S. E. (under review). Relationships between teachers and preschoolers who are at risk: Contribution of children's language skills, temperamentally-based attributes, and gender. *Journal of Speech, Language, and Hearing Research.*

Klein, A., Starkey, P., Clements, D., Sarama, J. & Iyer, R. (in press). Effects of a pre-K mathematics intervention: A randomized experiment. *Journal of Research on Educational Effectiveness.*

Lengua, L. J. (2002). The contribution of emotionality and self-regulation to the understanding of children's response to multiple risk. *Child Development, 73,* 144–161.

Li-Grining, C. P. (2007). Effortful control among low-income preschoolers in three cities: Stability, change, and individual differences. *Developmental Psychology, 43,* 208–221.

Li-Grining, C. P., Raver, C. C., Smallwood, K. M., Sardin, L., Metzger, M. W., & Jones, S. M. (submitted). Classroom quality and teacher support in Head Start classrooms: The role of teachers' psychosocial stressors. *Early Education and Development.*

LoCasale-Crouch, J., Konold, T., Pianta, R., Howes, C., Burchinal, M., Bryant, D., et al. (2007). Profiles of observed classroom quality in state-funded pre-kindergarten programs and associations with teacher, program, and classroom characteristics. *Early Childhood Research Quarterly, 22,* 3–17.

Magnuson, K., Ruhm, C., & Waldfogel, J. (2007). The persistence of preschool effects: Do subsequent classroom experiences matter? *Early Childhood Research Quarterly, 22,* 18–38.

Magnuson, K., & Duncan, G. (2003). Parent- versus child-based intervention strategies for promoting children's well-being. In A. Kalil & T. DeLeire (2004), *Family investments in children's potential: Resources and parenting behaviors that promote success. Monographs in parenting* (pp. 209–235). Mahwah, NJ: Lawrence Erlbaum Associates.

Massey, S. L., Pence, K. L., Justice, L. M., & Bowles, R. P. (2008). Educators' use of cognitively challenging questions in economically disadvantaged preschool classroom contexts. *Early Education and Development, 19,* 340–360.

McClelland, M. M., Cameron, C. E., Connor, C. M., Farris, C. L., Jewkes, A. M., & Morrison, F. J. (2007). Links between behavioral regulation and preschoolers' literacy, vocabulary, and math skills. *Developmental Psychology, 43,* 947–959.

McClelland, M. M., & Morrison, F. J. (2003).The emergence of learning-related social skills in preschool children. *Early Childhood Research Review Quarterly, 18,* 206–224.

McDermott, P., Green, L. F., Francis, J. M., & Stott, D. H. (1996). *Preschool Learning Behaviors Scale*. Philadelphia: Edumetric and Clinical Science.

McDermott, P. A., Leigh, N. M., & Perry, M. A. (2002). Development and validation of the Preschool Learning Behaviors Scale. *Psychology in the Schools, 39,* 353–365.

O'Connor, E., & McCartney, K. (2007). Examining teacher-child relationships and achievement as part of an ecological model of development. *American Educational Research Journal, 44,* 340–369.

Preschool Curriculum Evaluation Research Consortium (2008). Effects of preschool curriculum programs on school readiness (NCER 2008–2009). Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education.

Pence, K. L., Justice, L. M., & Wiggins, A. K. (in press). Preschool teachers' fidelity in implementing a comprehensive language-rich curriculum.

Pianta, R. C., Mashburn, A. J., Downer, J. T., Hamre, B. K., & Justice, L. (submitted). Effects of web-mediated professional development resources on teacher-child interactions in pre-kindergarten classrooms.

Ramey, C. T., Ramey, S. L., & Stokes, B. R. (in preparation). Effective pre-K programs: Research evidence about program dosage and student achievement.

Ramey, S. L., Ramey, C. T., Kleinman, B. E., Lee, L. M., Farneti, C. L., Timraz, N. M., et al. (submitted). The effects of curriculum and coaching supports on classrooms and literacy skills of prekindergarten/Head Start students in Montgomery County Public Schools.

Raver, C. C. (2002). Emotions matter: Making the case for the role of young children's emotional development for early school readiness. *Social Policy Report, 16,* 3–6.

Raver, C. C. (2004). Placing emotional self-regulation in sociocultural and socioeconomic contexts. *Child Development, 75,* 346–353.

Raver, C. C., Garner, P., & Smith-Donald, R. (2007). The roles of emotion regulation and emotion knowledge for children's academic readiness: Are the links causal? In R. C. Pianta, M. J. Cox, & K. L. Snow (Eds.), *School readiness and the transition to kindergarten in the era of accountability* (pp. 121–147). Baltimore: Brookes Publishing.

Raver, C. C., Jones, A. S., Li-Grining, C. P., Metzger, M., Smallwood, K., & Sardin, L. (2008). Improving preschool classroom processes: Preliminary findings from a randomized trial implemented in Head Start settings. *Early Childhood Research Quarterly.*

Raver, C. C., Jones, S. M., Li-Grining, C. P., Zhai, F., Bub, K., & Pressler, E. (submitted). CSRP's impact on low-income preschoolers' pre-academic skills: Self-regulation and teacher-student relationships as two mediating mechanisms. *Child Development.*

Raver, C. C., Jones, S. T., Li-Grining, C., Zhai, F., Metzger, M., & Solomon, B. (revised and resubmitted). Multi-component intervention targeting children's behavior problems in Head Start-funded preschool classrooms: A cluster-randomized controlled trial. *Journal of Consulting and Clinical Psychology.*

Raver, C. C., & Zigler, E. F. (1997). Focus section: New perspectives on Head Start. Social competence: An untapped dimension in evaluating Head Start's success. *Early Childhood Research Quarterly, 12,* 363–385.

Riggs, N. R., Greenberg, M. T., Kusché, C. A., & Pentz, M. A. (2006). The mediational role of neurocognition in the behavioral outcomes of a social-emotional prevention program in elementary school students: Effects of the PATHS curriculum. *Prevention Science, 7,* 91–102.

Rimm-Kaufman, S. E., Fan, X., Chiu, Y.-J., & You, W. (2007). The contribution of the Responsive Classroom Approach on children's academic achievement: Results from a three year longitudinal study. *Journal of School Psychology, 45,* 401–421.

Whitaker, S., Kinzie, M., Kraft-Sayre, M. E., Mashburn, A., & Pianta, R. C. (2007). Use and evaluation of web-based professional development services across participant levels of support. *Early Childhood Education Journal, 34,* 379–386.

Zaslow, M., Reidy, M., Moorehouse, M., Halle, T., Calkins, J., & Margie, N. G. (2003). Progress and prospects on the development of indicators of school readiness. In *Child and youth indicators: Accomplishments and future directions* (in press). Bethesda, MD: National Institutes of Health.