Abstract Title Page
Not included in page count.

Title:

Effects of a Literacy Focused Curriculum and a Developmental Curriculum on School Readiness and Subsequent State Achievement Test Outcomes in Rural Prekindergarten Classrooms

Author(s):

Mark W. Lipsey, PhD, Vanderbilt University
Dale C. Farran, PhD, Vanderbilt University
Sean M. Hurley, PhD, University of Vermont
Kerry G. Hofer, PhD, Vanderbilt University
Carol Bilbrey, PhD, Vanderbilt University

Contact information:

Mark W. Lipsey
VIPPS-Vanderbilt
1207 18th Ave. South
Nashville, TN 37212
Email: mark.lipsey@vanderbilt.edu
Phone: 615-343-2696
Fax: 615-322-8081

Abstract Body

Limit 5 pages single spaced.

Background/context:

Description of prior research and/or its intellectual context and/or its policy context.

Implementing pre-k programs is one of the major initiatives states have undertaken in recent years to improve educational outcomes for economically disadvantaged students (Barnett et al., 2007). Participation in formal pre-kindergarten does appear to improve some aspects of school readiness at kindergarten entry (Barnett et al., 2007b; Gormley et al., 2005), but evidence for longer term effects is mixed and a matter of some debate (Barnett, 1998; Magnuson, Ruhm, & Waldfogel, 2007). Other research suggests that, while preschool programs may improve basic pre-reading skills, their influence on complex language skills, mathematics, self-regulation, and social skills is less clear (Gormley et al., 2005; Jackson, et al., 2007; U.S. Department of Health and Human Services, 2005).

This situation makes the question of curriculum effectiveness important. An effective curriculum is one that tells teachers how to configure prekindergarten instruction to reliably promote school readiness and long-term school success. The most rigorous and comprehensive evaluation of preschool curricula to date is the IES funded PCER project, which launched 14 randomized trials around the country (Preschool Curriculum Evaluation Research Consortium, 2008). Most of the curricula tested had a literacy or general developmental focus (with one focused on math). Overall, 10 of these curricula showed no statistically significant differences from business as usual instruction in the control classrooms on any of the student-level outcomes, and only two showed significant differences on even one outcome measured in kindergarten. In light of these disappointing results, IES terminated the national evaluation of these curricula after the collection of the kindergarten follow-up data.

The question of greatest interest to school administrators, however, is whether any pre-k curricula improve students' scores on the state-mandated achievement tests that are typically not given until children are in the 2nd or 3rd grade. The PCER national evaluation did not follow children long enough to answer that question. This was understandable given the weak results on the early measures, but relied on the assumption that those early results were indicative of the potential for longer term effects. This paper reports the results of a particularly well-implemented PCER study that did follow the participating children through the third grade state administered achievement tests. It tests directly the effects of two contrasting curricula on the outcome of most relevance to schools and, in so doing, tests the assumption that weak effects on immediate pre-k outcomes are indicative of weak long term effects.

Purpose/objective/research question/focus of study:

Description of what the research focused on and why.

This research investigated the effects of two contrasting pre-k curricula, relative to practice as usual, on subsequent academic achievement. One curriculum had a strong literacy focus, the other was a less didactic "developmentally appropriate" curriculum that allowed children to have more influence on classroom activities. The purpose of the research was to determine if either of these curricula provided advantages for improving the academic performance of economically

disadvantaged children in rural Tennessee.

Setting:
Specific description of where the research took place.

The study was initiated at the beginning of the 2002-03 school year in seven school districts in six rural middle Tennessee counties. The preschool programs in these schools were funded by the state and enrolled students meeting criteria for economic disadvantage or other risks for school failure. Each classroom was required to have one certified teacher plus an assistant and a maximum of 20 children.

Population/Participants/Subjects:
Description of participants in the study: who (or what) how many, key features (or characteristics).

The participating pre-k teachers (N = 36) were credentialed public school teachers with a mean of 10.7 years of teaching experience. Twenty-two had Bachelor's degrees and 14 had earned a Master's degree. Eleven had an ECE teaching credential, 12 an Elementary credential, and 12 had both.

The sample of children (N=549) consisted of the students enrolled in the 36 pre-k classrooms who were age-eligible to attend kindergarten the next year (mean age = 4.4 years) and received parental consent to participate in the study. These students were 70% white, 19% African American, and 4% Hispanic. Slightly more than half were male (53%) and 10% had identified disabilities (IEPs).

Intervention/Program/Practice:
Specific description of the intervention, including what it was, how it was administered, and its duration.

Two different pre-k curricula were studied in comparison to practice as usual. The Bright Beginnings (BB) curriculum has a strong literacy focus with a didactic approach to instruction. Creative Curriculum (CC) has a developmental orientation and takes a constructivist approach to learning. Practice as usual in these school districts did not require any specific curriculum but, rather, left that decision to the discretion of the teachers. Teachers assigned to each of the experimental curricula received training the summer before implementation with a follow-up session during the school year. The curricula were implemented during the 2002-03 nine-month school year in full-day pre-k classrooms.

Research Design:
Description of research design (e.g., qualitative case study, quasi-experimental design, secondary analysis, analytic essay, randomized field trial).

The study design was a randomized field experiment. The participating schools were blocked by district and, within each district, randomly assigned to the BB curriculum, CC curriculum, or practice as usual. All the pre-k classrooms in each school were assigned to the same condition though, in most instances, there was only one classroom per school.

Data Collection and Analysis:
Description of plan for collecting and analyzing data, including description of data.

To assess the fidelity of implementation, observations were made in all the classrooms by trained observers near the beginning, middle, and end of the school year. The BB and CC fidelity checklists supplied by the curriculum developers were used for this purpose. Ratings using both these checklists were made in all classrooms, including those using the other curriculum and the control classrooms.

During the pre-k year, the 549 children in the sample were pretested at the beginning of the school year and posttested near the end of the year on the following measures:

- Peabody Picture Vocabulary Test (PPVT)
- Woodcock Johnson III Subtests
  - Letter Word Identification
  - Spelling
  - Picture Vocabulary
  - Story Recall
  - Understanding directions
  - Oral Comprehension
  - Applied Problems
  - Quantitative Concepts
- Adaptive Language Inventory (teacher rating)
- Instrumental Competence Scale (teacher rating)
  - Self-regulation subscale.

The children in the sample were tracked through kindergarten and first grade and these same measures, with minor variations, were collected at the end of each school year. The PPVT and WJIII assessments were all done individually for each child by a trained assessor.

For children reaching the third grade in school year 2006-7, the Tennessee State Department of Education provided scores from the statewide mandated achievement tests, known as TCAP (Tennessee Comprehensive Assessment Program). Four scores were available from this test: Reading/language arts, mathematics, science, and social studies. Of the 549 children in the initial sample, at least one of the posttest or follow-up achievement outcome measures (WJIII, PPVT, or TCAP) was obtained for 531 of them (97%).

Findings/Results:
Description of main findings with specific details.

Implementation fidelity. Across the three observational periods, the proportion of items passed on the Bright Beginnings checklist was 63% for the BB classrooms compared with 39% in the control classrooms. For Creative Curriculum, 66% of the items were passed in the CC classrooms compared with 49% in the control classrooms. Both these differences were statistically significant. Both the BB and CC implementation checklists, however, contained many items that were not distinctive to the respective curriculum but, rather, might characterize any pre-k classroom. The items from the combined checklists were therefore resorted into thematic categories to create scales for different aspects of the classroom. The results are shown

in Table 1. As can be seen there, the BB classrooms showed much more emphasis on language and literacy than the CC or control classrooms.

Pre-K outcomes. Multi-level analysis was used for the outcome data to account for the nesting of students within classrooms and classrooms within schools. School district was also included as a blocking factor along with covariates that included pretest measures for the respective posttests and selected student characteristics (age, sex, minority status, and disability status). Because of the small number of classrooms involved in the curriculum comparisons, the alpha level for statistical significance was set at .10. The results of this analysis are shown in Table 2. As can be seen there, only a scattered few of the comparisons between conditions reached statistical significance. It is especially notable that, despite the much greater emphasis on literacy, the children in the BB classrooms did not score significantly higher than the children in the control classrooms on any of the literacy and language related outcomes.

Kindergarten and first grade outcomes. Analyses similar to those above for pre-k outcomes were conducted on the analogous outcomes at the end of kindergarten and end of first grade. Again, there were only a few significant differences among all these comparisons, and none showed better outcomes on language or literacy measures for the BB curriculum.

State achievement test outcomes. The third grade TCAP scores were available for 68% of the 549 children in the original sample. The remaining children were no longer in the Tennessee school system or had been retained and were not yet in the third grade. There were no significant differences across curriculum conditions in the proportion of TCAP scores available, however. Multilevel analysis of the TCAP scores was conducted with students nested within pre-k classes and schools and with school district as a blocking factor. A composite pretest covariate was constructed from the principal components factor score for all the WJIII and PPVT pretest measures and included in the analysis with age, gender, minority status, and disability status as additional covariates.

Table 3 shows the results. Across all the TCAP subject areas, the children in the BB and CC curriculum conditions performed at least slightly better than the children in the control condition, as shown by the positive effect sizes. On the TCAP Reading/Language Arts achievement test, the BB children scored significantly higher than the control children (effect size=.27). On the TCAP Mathematics test, both the BB and CC children scored significantly higher than the control children (effect sizes = .34 and .41 respectively).

Because these positive effects on TCAP reading and math scores seem anomalous in light of the essentially null effects found on the direct assessment measures taken at the end of pre-k, kindergarten, and first grade, further analysis was done to probe their plausibility as pre-k effects. Note first that the TCAP scores changed the most in the subject areas given the most emphasis in pre-k—early reading and math—and not in the areas given little attention—science and social studies. The strong math findings, however, are unexpected because math is given much less attention than literacy and language in most pre-k classrooms. Further analysis showed that the gains made during pre-k on the WJIII and PPVT literacy/language and math measures were predictive of the TCAP reading and math scores. That is, children who made the most gains during pre-k were those who showed comparatively better TCAP scores with baseline levels,

age, gender, minority status, and disability status statistically controlled. The modest curriculum differences observed on the available measures during the early years thus do appear to reflect at least some gains related to later performance on the achievement tests. Further analysis is underway to more fully explore the nature of the relationship between pre-k experiences and outcomes and third grade TCAP performance.

Conclusions:

Description of conclusions and recommendations of author(s) based on findings and over study. (To support the theme of 2009 conference, authors are asked to describe how their conclusions and recommendations might inform one or more of the above noted decisions—curriculum, teaching and teaching quality, school organization, and education policy.)

Two important conclusions follow from this study. First, pre-k curriculum can make a difference in the later performance of economically disadvantaged children on the statewide standardized achievement tests that are so important in the current policy environment for schools. What is not so clear is the way in which curriculum matters. The Bright Beginnings effect on TCAP reading scores is understandable, given the emphasis on literacy and language skills in that curriculum. Its effects on TCAP math are not so easy to understand, however. Though early numeracy skills receive some emphasis in that curriculum, they receive much less attention than early reading skills. Creative Curriculum, which had the largest effect on TCAP math, is somewhat more balanced in its emphasis but, nonetheless, also gives relatively little attention to math compared to reading. The implementation measures taken from classroom observations bear this out—little difference appeared in math emphasis across the curriculum conditions (Table 1). The influence these curricula have on later math performance, thereore, does not seem to arise from direct attention to early math in the pre-k classroom. This is a topic that clearly warrants further investigation.

Second, this study raises questions about the measures conventionally used to assess the effects of pre-k interventions. The WJIII and PPVT measures selected for this study are widely used as outcomes in pre-k studies and represent validated measures of early decoding skills, vocabulary, comprehension, numeracy, and the like. Their failure to show convincing curriculum differences when such differences appear on later state achievement measures is surprising. It may be that these measures are not sufficiently sensitive to respond to pre-k differences and thus understate the effects. Or, it may be that these measures do not tap into the skills most important for later achievement test performance that are influenced by pre-k. This topic also clearly warrants further investigation.

Appendixes

Not included in page count.

Appendix A. References

References are to be in APA format. (See APA style examples at the end of the document.)

Barnett, W. S. (1998). Long-term cognitive and academic effects of early childhood education on children in poverty. *Preventive Medicine, 27,* 204-207.

Barnett, W. S., Hustedt, J. T., Friedman, A. H., Boyd, J. S., Ainsworth, P. (2007). *The state of preschool 2007: State preschool yearbook*. National Institute for Early Education Research: Rutgers University.

Barnett, W. S., Jung, K., Wong, V., Cook, T., & Lamy, C. (2007b). *Effects of five state prekindergarten programs on early learning.* Rutgers University, NJ: The National Institute for Early Education Research.

Gormley, W. T., Gayer, T., Phillips, D., & Dawson, B. (2005). The effects of universal Pre-K on cognitive development. *Developmental Psychology, 41,* 872-884.

Jackson, R. McCoy, A. Pistorino, C., Wilkinson, A., Burghardt, J., Clark, M., Ross C., Schochet, P., & Swank, P. (2007). *National Evaluation of Early Reading First: Final Report,* U.S. Department of Education, Institute of Education Sciences, Washington, DC: U.S. Government Printing Office.

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

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.

U.S. Department of Health and Human Services, Administration for Children and Families (May 2005). *Head Start impact study: First year findings*. Washington, DC.

Appendix B. Tables and Figures
Not included in page count.

Table 1. Mean Proportion (and Standard Deviation) of Items on the Implementation Checklists Scored as Passing for All Pre-K Classrooms over Three Waves of Observation

Implementation Scales	BB: Bright Beginnings (N=11)	CC: Creative Curriculum (N=11)	C: Control (N=14)	Statistically Significant Contrasts[a]
Developer's checklists				
BB checklist, all items	.63 (.10)	.49 (.15)	.39 (.11)	BB>CC>C
CC checklist, all items	.64 (.07)	.66 (.18)	.49 (.11)	BB, CC> C
Scales from combined checklists				
A. Literacy & language				
Expressive Language	.81 (.17)	.57 (.26)	.40 (.16)	BB>CC>C
Vocabulary Building	.89 (.13)	.64 (.22)	.48 (.24)	BB>CC[b]> C
Letter-Word-Writing	.80 (.10)	.38 (.32)	.40 (.17)	BB>CC, C
Reading	.76 (.11)	.40 (.22)	.37 (.14)	BB>CC, C
Spelling	.68 (.13)	.49 (.27)	.46 (.15)	BB>CC, C
Literacy Materials-Reading	.47 (.17)	.29 (.17)	.19 (.10)	BB>CC, C
Literacy Materials-Writing	.53 (.22)	.24 (.12)	.25 (.12)	BB>CC, C
Environmental Print	.43 (.11)	.53 (.26)	.23 (.13)	BB, CC>C
B. Math and other subjects				
Math Instruction	.60 (.19)	.69 (.24)	.66 (.22)	None
Math Materials	.59 (.15)	.66 (.14)	.52 (.21)	CC>C
Natural & Social Science	.70 (.14)	.68 (.22)	.57 (.23)	None
C. Teacher interactions				
Emotional Warmth	.87 (.20)	.74 (.22)	.62 (.18)	BB>C
Positive Behavior Management	.85 (.22)	.68 (.22)	.62 (.16)	BB>CC[b], C
Extends Learning	.87 (.11)	.64 (.23)	.56 (.15)	BB>CC, C
D. Classroom organization				
Small Group Instruction	.77 (.20)	.73 (.19)	.64 (.20)	None
Learning Centers	.82 (.08)	.79 (.23)	.55 (.09)	BB, CC>C
Center-based Teaching	.94 (.07)	.79 (.25)	.63 (.20)	BB[b]>CC[b]>C
Outdoor time	.65 (.23)	.69 (.14)	.57 (.17)	None

[a] $p<.05$ unless otherwise indicated. [b] $p<.10$.
BB=Bright Beginnings; CC=Creative Curriculum; C=Control.

Table 2. Curriculum Effects at the End of Pre-K: Covariate-Adjusted W or Raw Score Means and Effect Sizes for Comparison with the Control Condition

	Bright Beginnings (N=147 to 153)		Creative Curriculum (N=156 to 160)		Control (N=192 to 198)
	Mean (sd)	Effect Size	Mean (sd)	Effect Size	Mean (sd)
Decoding					
WJIII Letter-Word	339.6[a] (23.4)	0.20	331.0[a] (24.9)	-0.14	334.6 (26.3)
WJIII Spelling	379.6[b] (27.3)	0.15	373.6[b] (23.9)	-0.08	375.6 (26.0)
Vocabulary					
WJIII Picture Vocabulary	468.5 (14.0)	-0.02	468.6 (9.7)	-0.02	468.8 (15.2)
PPVT (raw score)	64.1[a] (17.0)	0.05	60.8*[a] (16.2)	-0.14	63.2 (17.1)
Oral Language					
WJIII Story Recall	486.5 (10.2)	0.06	486.3 (10.3)	0.04	485.9 (10.4)
WJIII Understanding Directions	461.1 (12.5)	-0.06	460.8 (10.7)	-0.09	461.9 (12.5)
WJIII Oral Comprehension	457.0 (15.9)	-0.17	457.3 (13.5)	-0.16	459.6 (15.4)
Math					
WJIII Applied Problems	413.6 (19.2)	0.13	410.8 (19.3)	0.00	410.9 (22.6)
WJIII Quant Concepts	421.0 (14.4)	-0.02	420.5 (15.0)	-0.05	421.3 (15.2)
Teacher Ratings					
ALI Total	57.2[b] (15.9)	-0.02	61.6*[b] (17.9)	0.25	57.5 (14.7)
ICS Total	52.7 (7.5)	0.07	52.0 (7.8)	-0.01	52.1 (8.8)
ICS Self Regulation	2.82 (.59)	0.05	2.78 (.62)	-0.02	2.79 (.62)

** p <.05 compared to control. * p <.10 compared to control.
[a] p <.05 between BB and CC. [b] p <.10 between BB and CC.

Table 3. Curriculum Effects on Third Grade TCAP Scores: Covariate-Adjusted Means and Effect Sizes for Comparison with the Control Condition

	Bright Beginnings (N=116)		Creative Curriculum (N=109)		Control (N=148)
TCAP Subject	Mean (sd)	Effect Size	Mean (sd)	Effect Size	Mean (sd)
Reading	494.7* (24.2)	0.27	492.1 (23.1)	0.16	488.4 (22.3)
Mathematics	482.1* (26.2)	0.34	485.0** (32.3)	0.41	472.5 (29.3)
Science	206.7 (16.8)	0.19	204.2 (12.3)	0.04	203.5 (17.6)
Social Studies	203.8 (16.0)	0.15	202.4 (13.7)	0.08	200.8 (22.1)

** p <.05 compared to control. * p <.10 compared to control.
[a] p <.05 between BB and CC. [b] p <.10 between BB and CC.