

Sustainability of State Pre-k Program Effects on
Children's Outcomes in Pre-k and Kindergarten

Ellen S. Peisner-Feinberg and Jennifer M. Schaaf
FPG Child Development Research Institute
University of North Carolina at Chapel Hill

Poster presented at
AERA Annual Meeting
Denver, CO
May, 2010

Introduction

Based on research over the past several decades, there is now a good understanding of the importance of high quality care and education programs for promoting young children's school readiness and academic success¹. The documented effectiveness of small-scale programs has helped to prompt the creation of larger programs serving greater numbers of children. Though many states or school districts have implemented such programs, questions about the impact of such programs as they are scaled up to serve more children and the extent to which program effects can be sustained over time remain largely unexplored.

The current study evaluates the effects of one statewide pre-kindergarten education initiative for at-risk children, the North Carolina More at Four Pre-kindergarten Program (MAF). Data were gathered across three cohorts over a 6-year period representing expansion of the program from serving 10,891 to 17,251 to 29,978 children. This particular program incorporates several other important characteristics: serving primarily at-risk children (risk factors include low family income, limited English proficiency, identified disability, chronic health condition, and/or developmental/educational need); serving children in a variety of public and private settings; and program guidelines derived from research-based quality indicators such as high teacher qualifications, high teacher-child ratios, and requirements for the quality of instructional practices.

Major Research Questions

The current study was designed to examine whether there were any changes over time in program quality and children's outcomes in this large and growing state pre-k program for at-risk children. The primary research questions addressed included:

- Were there any changes in classroom quality over time as the size of the program increased?
- What were the longitudinal outcomes from pre-k through kindergarten for children who attended the pre-k program?
- Were program factors (classroom quality) or child characteristics (English language proficiency) associated with better outcomes for children?
- Were there any differences in program effects over time as the pre-k program continued to expand?

Participants

Three cohorts of children in randomly selected MAF classrooms were followed from pre-k through kindergarten.

- Cohort 1 (2003-2005) -- 58 MAF classrooms (out of 599 statewide); 514 children in pre-k and 348 in kindergarten
- Cohort 2 (2005-2007) -- 57 MAF classrooms (out of 952 statewide); 478 children in pre-k and 400 in kindergarten
- Cohort 3 (2007-2009) -- 58 MAF classrooms (out of 1, 687 statewide); 321 children in pre-k and 280 in kindergarten

Procedure

- *Child Assessments*². Individual assessments of children's skills were gathered twice each year, in the fall and spring: Language skills (PPVT-III & PPVT-4); Math skills (WJ-III Applied Problems subtest; Counting Bears Task); General knowledge (Social Awareness Task); Teacher ratings of social skills and problem behaviors (SSRS)
- *Child Characteristics*². English language proficiency at program entry (PreLAS 2000), with five proficiency levels: Non-English speaker (1); Limited English speaker (2-3); Fluent English speaker (4-5)
- *Program Characteristics*². Observational assessments of the quality of classroom practices in pre-k (ECERS-R) were conducted mid-year

Analysis Strategy

Two series of analyses were conducted. First, to examine cohort effects on pre-k classroom quality, regression analyses using general linear models were conducted for the ECERS-R total and subscale scores. The effects of cohort on classroom quality were examined to evaluate changes in classroom quality over time as the size of the program increased.

Second, to examine children's development over time, a series of 3-level growth models were conducted for each outcome measure using a mixed model to account for repeated measures across each child and multiple children clustered within classrooms. Children's scores at each of four time points (fall pre-k, spring pre-k, fall kindergarten, spring kindergarten) were the dependent variables. Covariates included: cohort (1, 2, 3); grade (pre-k, kindergarten); assessment variations (age at first assessment, time between assessments/enrollment); More at Four dosage (days of attendance); English proficiency level; child gender; and pre-k classroom quality (ECERS-R total child items score). Analyses included Type I error adjustments to the p-values using the Benjamini and Hochberg³ correction.

Change over time was analyzed to examine the longitudinal outcomes from pre-k through kindergarten for children who attended the MAF program. Interactions between time and program factors (classroom quality) or child characteristics (English language proficiency) examined whether there were any moderating effects on outcomes of participation in MAF. Interactions between time and cohort examined whether there were any changes in children's longitudinal outcomes as the program expanded over time. Three-way interactions adding cohort were used to examine variations in moderating effects as the MAF program expanded.

Results

Classroom Quality

Classroom quality was generally high over time, with scores in the good or close to good range (See Table 1). However, there were some cohort differences, with some decreases in scores from the earliest cohort to later cohorts. Cohort 1 had higher scores than the other two cohorts in the total scores and five of the seven subscale scores. In contrast, Cohort 2 had higher scores than Cohort 1 on the Parents and Staff subscale and Cohort 3 scored higher than Cohort 2 on Space and Furnishings, Language-Reasoning, and Program Structure. In sum, although some cohort differences were present, classroom quality remained relatively high over time. Moreover,

the pattern of cohort differences suggests that there was not a consistent decline in quality in concert with the increasing program size.

Children's Outcomes

Longitudinal Growth

Results indicated that children exhibited significant growth over the two years in almost all areas of language, math, general knowledge, and behavior (Table 2). The gains on standardized measures of receptive language and applied math skills indicate a level of growth beyond that expected from increasing age alone. Teacher ratings of problem behaviors were the only area that did not change significantly over time and showed slightly fewer behavior problems than the norm. There were few differences among cohorts in the amount of growth, with no significant findings for most measures, including applied problems, counting, social skills and problem behaviors. For two measures, receptive language and social awareness, Cohort 3 evidenced less growth than earlier cohorts.

Moderating Factors

There was no moderating effect of classroom quality on children's growth in language, math, general knowledge or problem behaviors, and a very modest effect for social skills. Given this pattern of results (with only one effect for a teacher-rated variable), no further analyses were conducted with regard to this moderator. There were moderating effects for children's English proficiency levels (see Figures 1-6). Children with lower English proficiency levels (especially levels 1 and 2) scored lower than children with higher proficiency levels, but made greater progress over time in most areas: language, applied problems, counting, and social awareness. There were few cohort differences in outcomes by proficiency level, with those that were found showing no consistent pattern by cohort in terms of scores or growth.

Discussion

Results of the current study have several implications for understanding the sustainability of effects for large-scale pre-k programs. Importantly, results of this study indicate that MAF maintained relatively high quality over time, even as the program expanded substantially each year (from nearly 11,000 to 30,000 children over this time period). Although there was some decrease in quality from the first cohort, there was not a consistent pattern of decline over time. Across all cohorts, children exhibited significant longitudinal growth from pre-k through kindergarten across all domains (language, math, general knowledge, and behavior) with few differences by cohort. These results, both in terms of classroom quality and children's outcomes, indicate that positive effects can be maintained even as programs expand.

In conjunction with the relatively good quality of classroom practices over time, there was little moderating effect of this program factor on children's outcomes. In contrast, individual child characteristics related to English language proficiency did have a moderating effect. Children who were less proficient in English, especially those at the lowest levels, exhibited greater rates of growth in most areas, with similar effects across cohorts. These results provide evidence that the program continued to be most effective for the most at-risk children it was designed to serve.

In sum, the results of the present study indicate that a statewide pre-k program can maintain positive effects over time, even while vastly expanding in scale. Such sustained outcomes for children may be a reflection of the high quality maintained by the program. These results suggest

the importance of strong program standards incorporating research-based quality indicators for the design of future statewide programs to successfully serve at-risk populations.

Table 1. Quality of Classroom Practices (ECERS-R)

Item Description ¹	Cohort 1 2003-2004 n=57	Cohort 2 2005-2006 n=57	Cohort 3 2007-2008 n=50	Significant Cohort Differences
	Mean (SD) Range	Mean (SD) Range	Mean (SD) Range	
Total Score ²	5.3 (0.6) 3.4-6.4	4.3 (0.6) 2.8-5.8	4.6 (0.9) 2.8-6.4	1>2,3
Total Child Items Score ³	5.3 (0.6) 3.9-6.6	4.2 (0.7) 2.7-5.8	4.4 (1.0) 2.5-6.4	1>2,3
Space and Furnishings Subscale	5.1 (0.8) 3.4-6.8	3.9 (0.7) 2.6-5.8	4.5 (1.1) 2.4-6.4	1>3>2
Personal Care Routines Subscale	5.0 (1.1) 2.3-7.0	2.8 (0.9) 1.3-5.7	3.1 (1.2) 1.5-6.2	1>2,3
Language-Reasoning Subscale	5.8 (0.8) 3.5-7.0	4.8 (0.8) 3.3-7.0	5.2 (1.3) 1.8-7.0	1>3>2
Activities Subscale	4.8 (0.8) 2.8-6.5	4.5 (0.9) 2.2-6.9	4.6 (1.1) 2.3-7.0	NS
Interaction Subscale	6.2 (.8) 3.6-7.0	4.8 (1.2) 2.0-7.0	4.7 (1.7) 1.6-7.0	1>2,3
Program Structure Subscale	6.3 (0.8) 3.8-7.0	4.4 (1.4) 1.7-7.0	4.9 (1.2) 2.3-7.0	1>3>2
Parents and Staff Subscale	5.2 (0.9) 2.5-6.8	5.5 (0.8) 3.0-7.0	5.3 (0.9) 3.5-6.7	2>1

¹ Total and subscale scores could range from 1.0-7.0; item scores could range from 1-7.

² The Total Score includes all items on the ECERS-R (items 1-43).

³ The Total Child Items Score includes items from all subscales on the ECERS-R except the Parents and Staff subscale (items 1-37).

Table 2. Longitudinal Child Outcome Scores

Domain Outcome	PK Fall Mean (SD) Range	PK Spring Mean (SD) Range	K Fall Mean (SD) Range	K Spring Mean (SD) Range	Significance of Growth¹	Cohort Differences in Growth
Language						
<i>Receptive Language (PPVT)²</i>						
Cohort 1	85.4 (19.3) 40-124	89.9 (17.2) 40-126	94.7 (15.9) 40-127	96.3 (13.8) 40-132	***	** C1, C2>C3
Cohort 2	81.1 (20.9) 21-125	87.0 (19.7) 32-129	90.7 (17.4) 32-132	93.4 (15.8) 25-135		
Cohort 3	88.1 (17.9) 33-131	91.0 (17.2) 23-129	93.2 (15.9) 42-130	97.1 (13.3) 65-134		
Math						
<i>Applied Problems (Woodcock Johnson III)</i>						
Cohort 1	96.2 (13.0) 58-128	98.3 (11.4) 60-126	100.3 (11.1) 51-132	102.1 (11.0) 69-132	***	NS
Cohort 2	92.4 (15.2) 58-135	97.7 (12.5) 58-128	98.9 (11.5) 39-124	101.1 (11.9) 47-141		
Cohort 3	93.6 (14.7) 58-129	98.2 (12.2) 53-140	97.7 (12.2) 34-131	101.8 (11.5) 65-132		
<i>Counting Task</i>						
Cohort 1	11.3 (8.3) 0-40	18.9 (11.5) 1-40	28.2 (11.9) 1-40	33.7 (9.5) 1-40	***	NS
Cohort 2	11.2 (8.0) 0-40	18.8 (10.6) 0-40	24.4 (11.8) 2-40	34.7 (9.2) 4-40		
Cohort 3	11.6 (8.1) 0-40	18.0 (11.0) 0-40	25.0 (12.5) 2-40	35.5 (8.6) 3-40		

¹ * $p < .05$, ** $p < .01$, *** $p < .001$. Significance levels indicate results of testing of the parameter estimates for the adjusted gains over time based on longitudinal growth model estimations.

² The PPVT-III was used for Cohorts 1 and 2 and the PPVT-4 was used for Cohort 3.

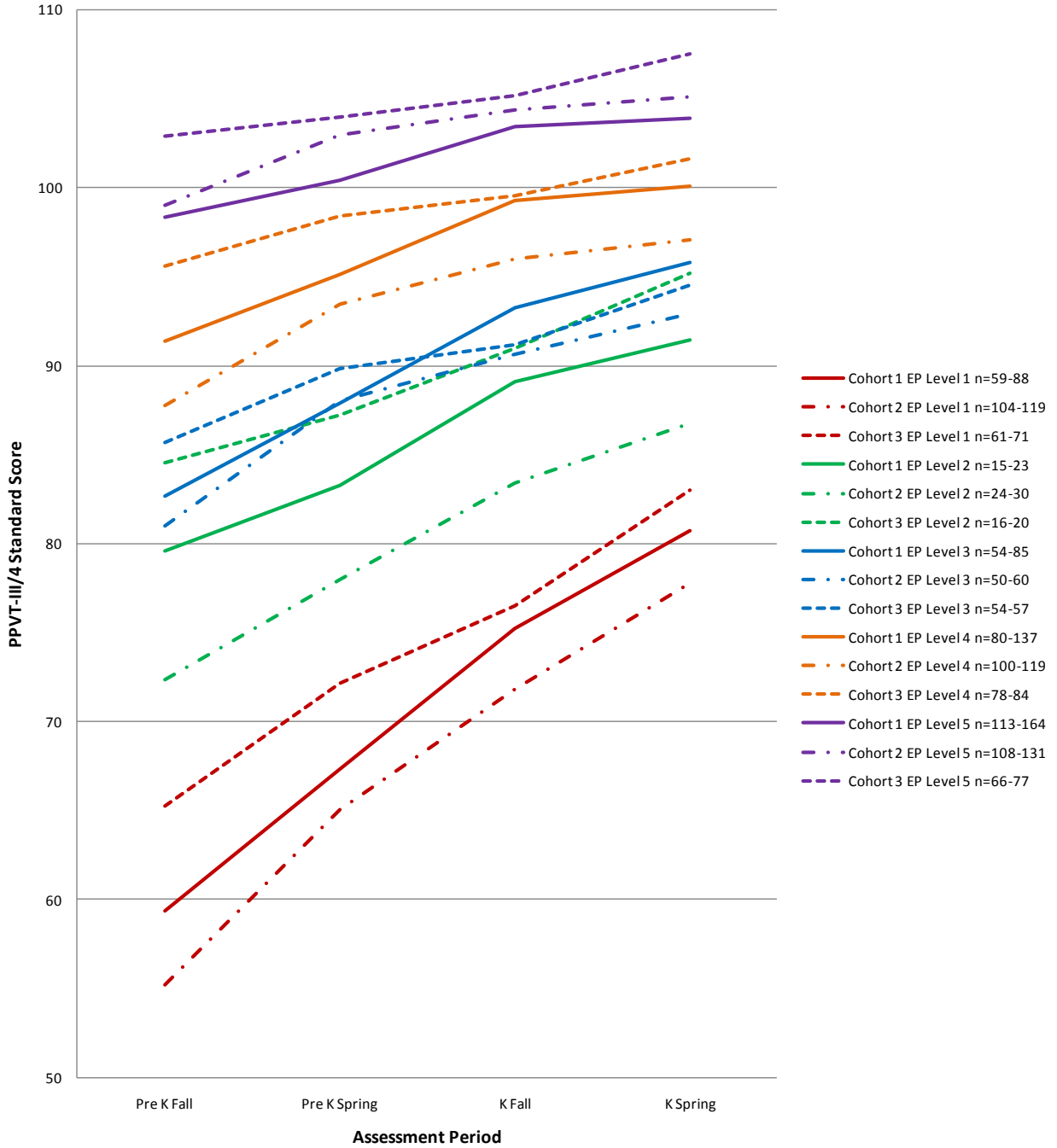
Table 2. Longitudinal Child Outcome Scores

Domain Outcome	PK Fall Mean (SD) Range	PK Spring Mean (SD) Range	K Fall Mean (SD) Range	K Spring Mean (SD) Range	Significance of Growth¹	Cohort Differences in Growth
General Knowledge						
<i>Social Awareness</i>						
Cohort 1	3.7 (1.8) 0-6	4.5 (1.5) 0-6	4.8 (1.2) 1-6	5.4 (1.0) 1-6	***	** C2>C3
Cohort 2	3.3 (1.9) 0-6	4.2 (1.5) 0-6	4.6 (1.4) 0-6	5.3 (1.0) 1-6		
Cohort 3	3.5 (1.8) 0-6	4.2 (1.6) 0-6	4.5 (1.4) 0-6	5.3 (1.1) 2-6		
Classroom Behavior						
<i>Social Skills (SSRS)</i>						
Cohort 1	100.8 (15.3) 56-130	107.8 (15.3) 62-130	101.6 (14.3) 64-130	106.4 (14.3) 61-130	***	NS
Cohort 2	100.4 (15.7) 53-130	109.7 (14.7) 60-130	101.8 (15.2) 49-130	107.7 (15.5) 54-130		
Cohort 3	101.0 (16.1) 54-130	109.4 (14.6) 57-130	99.2 (14.4) 64-130	105.4 (14.3) 64-130		
Classroom Behavior						
<i>Problem Behavior (SSRS)</i>						
Cohort 1	98.6 (11.9) 85-138	99.3 (12.8) 85-145	99.0 (12.8) 85-135	99.0 (13.0) 85-137	NS	NS
Cohort 2	98.2 (13.1) 85-142	97.2 (12.0) 85-135	96.9 (12.7) 85-137	98.0 (13.4) 85-141		
Cohort 3	99.8 (13.1) 85-140	99.5 (13.2) 85-145	99.1 (12.7) 85-134	98.3 (12.8) 85-138		

¹ * $p < .05$, ** $p < .01$, *** $p < .001$. Significance levels indicate results of testing of the parameter estimates for the adjusted gains over time based on longitudinal growth model estimations.

Figure 1:

Growth in Receptive Language Skills (PPVT-III/4) by Cohort and English Proficiency Level

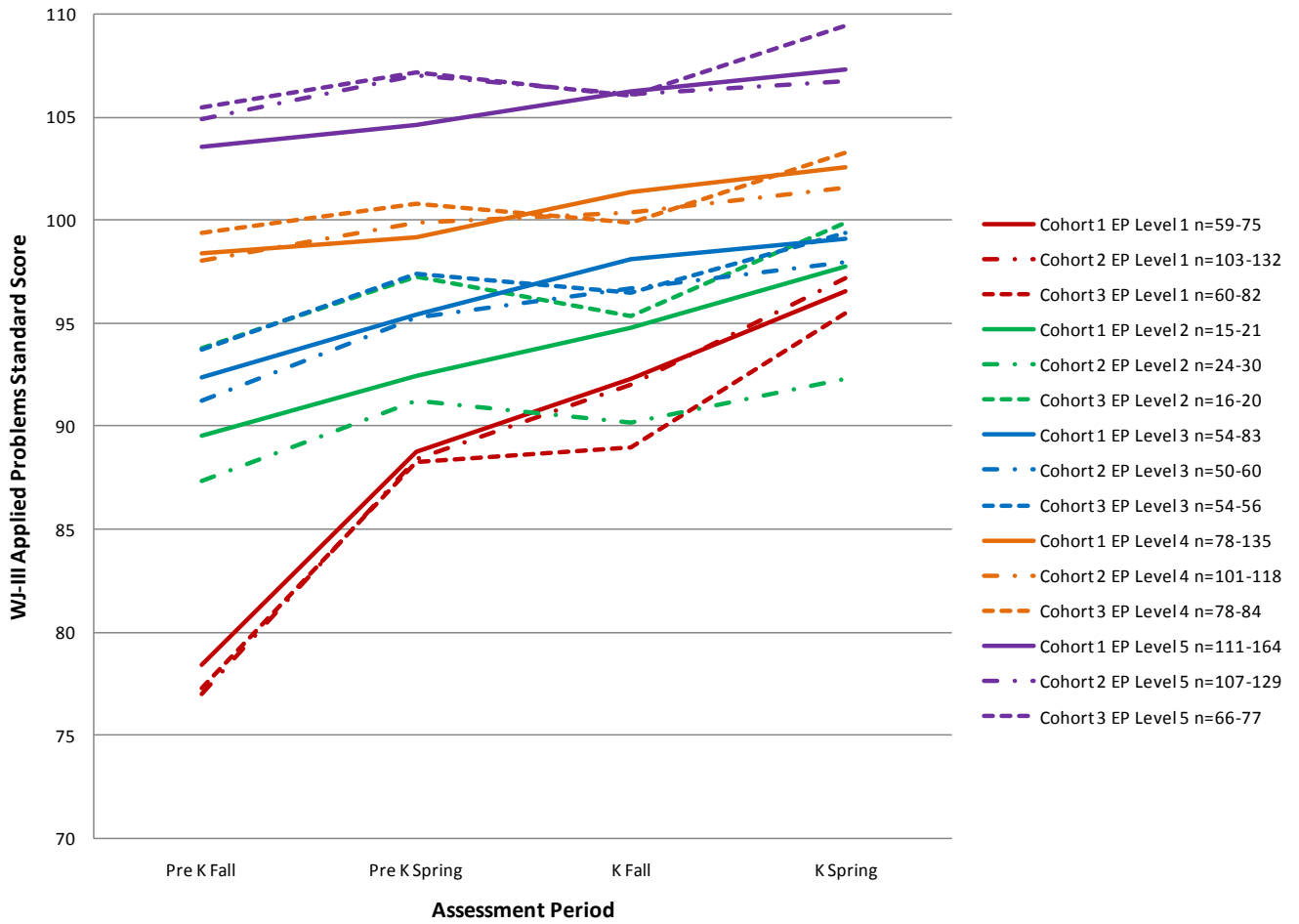


Significant differences by English proficiency level: Pre-K Fall: 1<2<3<4<5. Pre-K Spring: 1<2<3<4<5. K Fall: 1<2<3<4<5. K Spring: 1<2<3<4<5. Growth over Time: 1>2,3>4>5.

Significant cohort differences by English proficiency level: Pre-K Fall: EP 1 C3>C1>C2; EP 2 C3>C2; EP 4 C3>C1,C2; EP 5 C3>C1. Pre-K Spring: EP 1 C3>C1,C2; EP 2 C3>C2; EP 4 C3>C2. K Fall: EP 1 C3>C2; EP 2 C3>C2. K Spring: EP 1 C3>C2; EP 2 C3>C2; EP 4 C3>C2. Growth differences: EP 1 C2>C3; EP 3 C1>C3; EP 4 C2>C3.

Figure 2:

Growth in Math Skills (WJ-III Applied Problems) by Cohort and English Proficiency Level

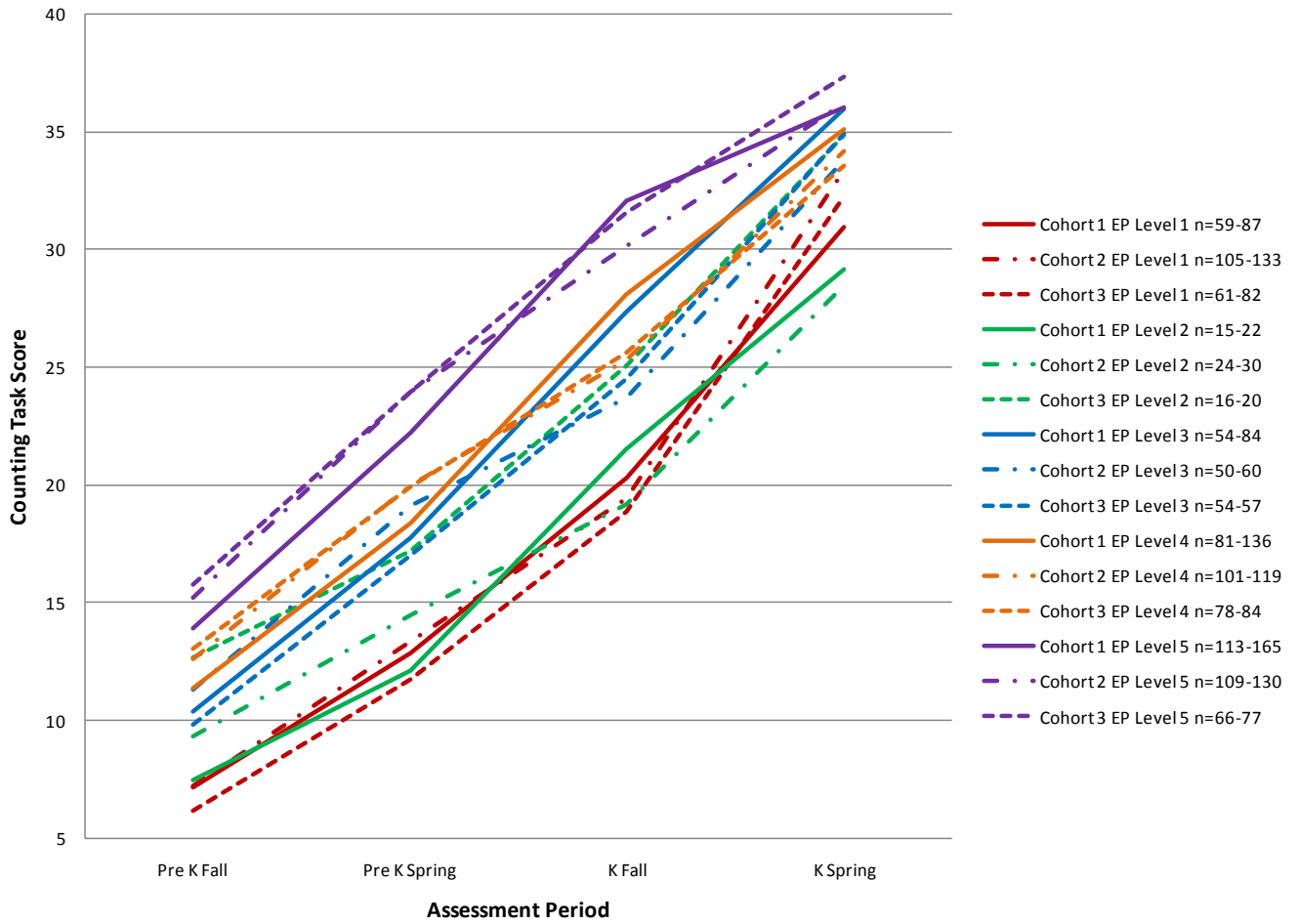


Significant differences by English proficiency level: Pre-K Fall: 1<2,3<4<5. Pre-K Spring: 1<2<3<4<5. K Fall: 1,2<3<4<5; 1<3. K Spring: 1,2,3<4<5. Growth over Time: 1>2,3,4,5; 3>5.

Significant cohort differences by English proficiency level: Pre-K Fall: EP 2 C3>C2. Pre-K Spring: NS. K Fall: NS. K Spring: EP 2 C3>C2. Growth differences: NS.

Figure 3:

Growth in Counting Skills (Counting Task) by Cohort and English Proficiency Level

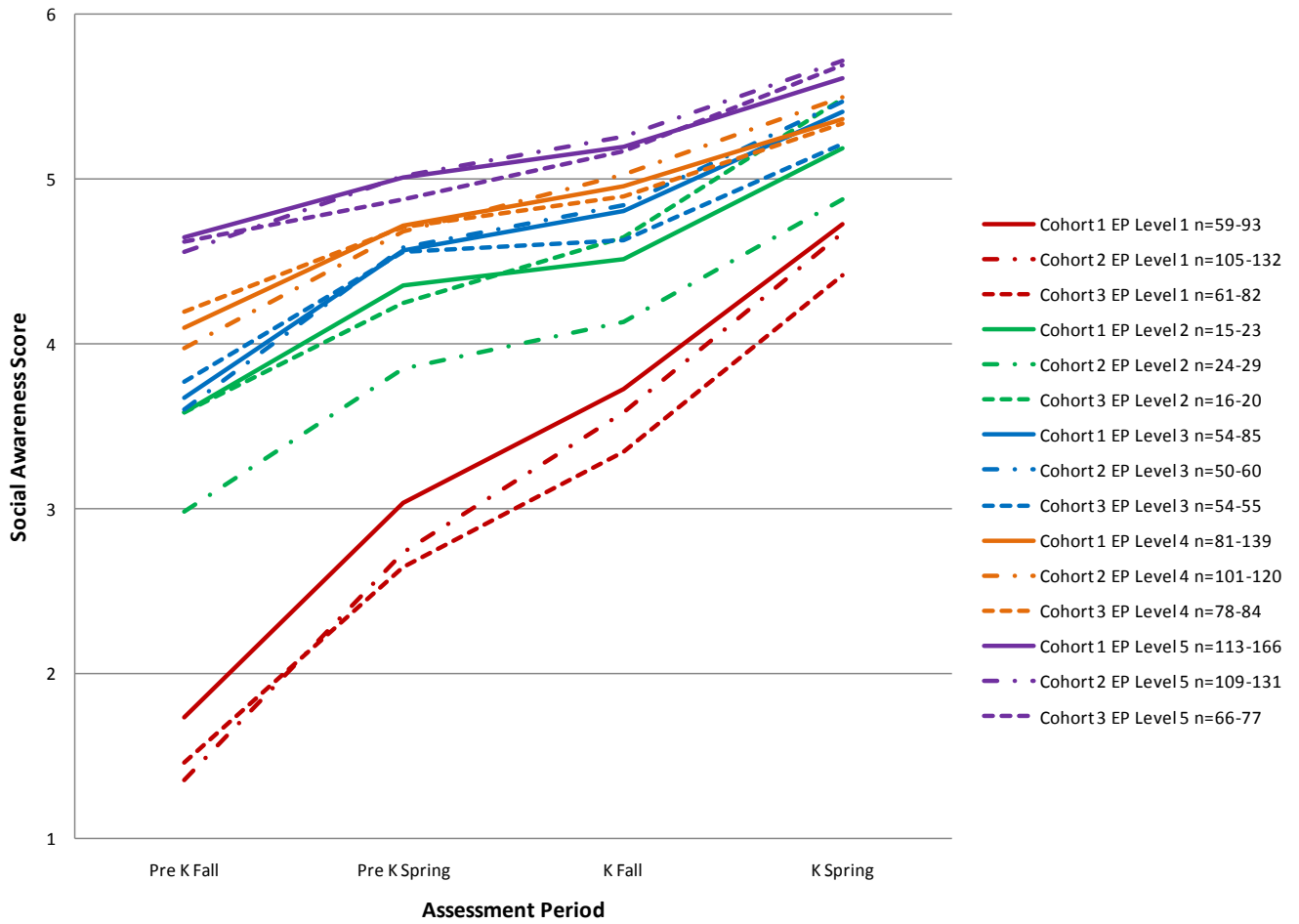


Significant differences by English proficiency level: Pre-K Fall: 1<2,3<4<5. Pre-K Spring: 1,2<3,4<5. K Fall: 1,2<3,4<5. K Spring: 1,2<3,4<5. Growth over Time: 1>2,4,5; 3>4.

Significant cohort differences by English proficiency level: Pre-K Fall: NS. Pre-K Spring: NS. K Fall: EP 2 C3>C2; EP 3 C1>C2. K Spring: EP 2 C3>C2. Growth differences: NS.

Figure 4:

Growth in Social Knowledge (Social Awareness Task) by Cohort and English Proficiency Level

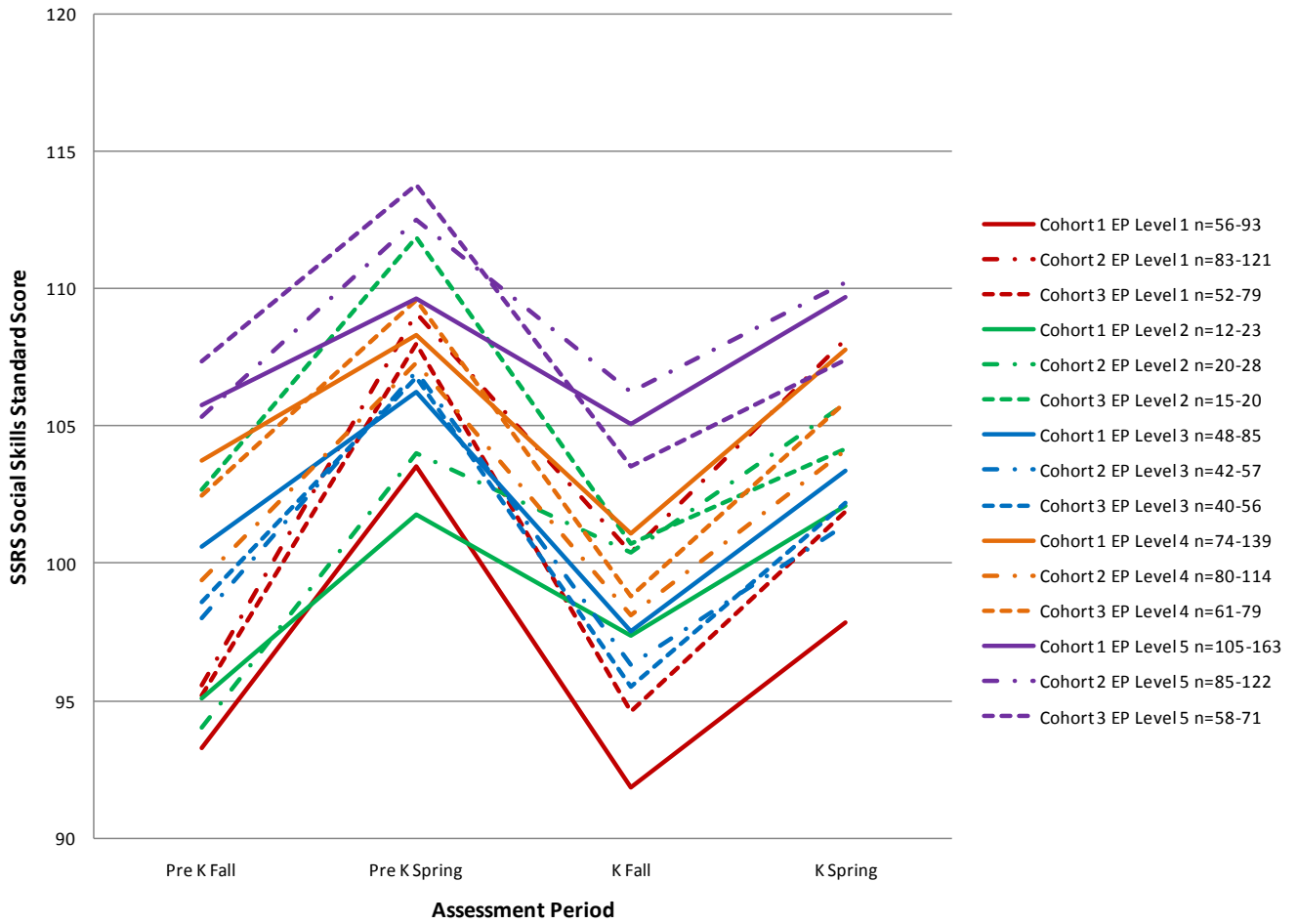


Significant differences by English proficiency level: Pre-K Fall: 1<2,3,4<5; 2,3<4. Pre-K Spring: 1<2<3<4<5. K Fall: 1<2<3,4<5. K Spring: 1<2,3,4<5. Growth over Time: 1>2,3,4,5; 3<4.

Significant cohort differences by English proficiency level: Significant cohort differences: Pre-K Fall: NS. Pre-K Spring: NS. K Fall: NS. K Spring: NS. Growth differences: NS.

Figure 5:

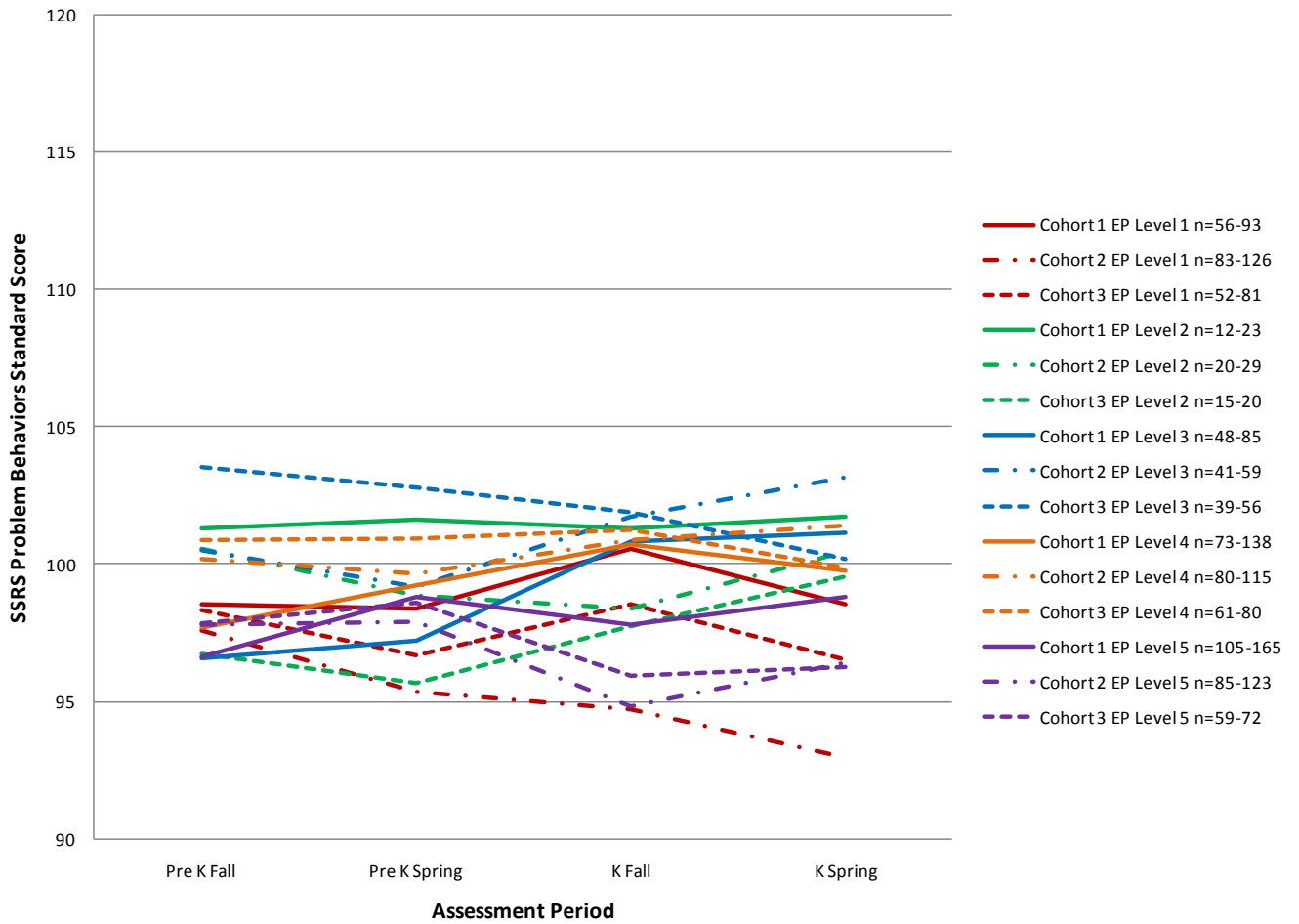
Growth in Social Skills (SSRS) by Cohort and English Proficiency Level



Significant differences by English proficiency level: Pre-K Fall: 1<3,4<5; 2<4,5. Pre-K Spring 1<3<4<5; 2<4,5. K Fall: 1,2,3,4<5; 1,3<4. K Spring: 1,3,4<5. Growth over Time: NS
 Significant cohort differences by English proficiency level: Pre-K Fall: EP 2 C3>C2. Pre-K Spring: EP 1 C2>C1; EP 2 C3>C1. K Fall: EP 1 C2>C3,C1. K Spring: EP 1 C2>C1. Growth differences: EP 1 C2>C1.

Figure 6:

Growth in Problem Behaviors (SSRS) by Cohort and English Proficiency Level



Significant differences by English proficiency level: Pre-K Fall: NS. Pre-K Spring: NS. K Fall: NS. K Spring: NS.
 Growth over Time: NS.
 Significant cohort differences by English proficiency level: Pre-K Fall: NS. Pre-K Spring: NS. K Fall: NS. K Spring: NS. Growth differences: NS.

¹ Previous Research

- Burchinal, M., Howes, C., Pianta, R., Bryant, D., Early, D., Clifford, R., & Barbarin, O. (2008). Predicting child outcomes at the end of kindergarten from the quality of pre-kindergarten teacher-child interactions and instruction, *Applied Developmental Science, 12*, 140-153.
- Campbell, F. A. & Ramey, C. T. (1994). Effects of early intervention on intellectual and academic achievement: A follow-up study of children from low-income families. *ChildDevelopment, 65*, 684-698.
- Gormley, W. T., Gayer, T., Phillips, D., & Dawson, B. (2005). The effects of universal pre-k on cognitive development, *Developmental Psychology, 41*, 872-884.
- Reynolds, A. J., & Temple, J. A. (1998). Extended early childhood intervention and school achievement: Age thirteen findings from the Chicago Longitudinal Study. *Child Development, 69*, 231-246.
- Schweinhart, L. J., & Weikart, D. P. (1998). High/Scope Perry Preschool Program effects at age twenty-seven. In J. Crane (Ed.), *Social programs that work* (pp. 148-162). New York: Russell Sage Foundation.
- Wong, V. C., Cook, T. D., Barnett, W. S., & Jung, K. (2008). An effectiveness-based evaluation of five state pre-kindergarten programs. *Journal of Policy Analysis and Management, 27*, 122-154.

² Measures:

- Duncan, S. E., & De Avila, E. A. (1998). PreLAS 2000. Monterey, CA: CTB/McGraw-Hill.
- Dunn, L. M., & Dunn, L. M. (1997). Peabody Picture Vocabulary Test Third Edition. Circle Pines, MN: American Guidance Service.
- Dunn, L. M., & Dunn, D. M. (2007). Peabody Picture Vocabulary Test Fourth Edition. Minneapolis, MN: NCS Person, Inc.
- FACES Research Team, Social Awareness Task, modified from the Social and Communicative Competence tasks in: J. M. Mason and J. Stewart (1989). The CAP Early Childhood Diagnostic Instrument (prepublication edition). Iowa City, IA: American Testronics.
- Gresham, F., & Elliott, S. (1990). Social Skills Rating System. Circle Pines, MN: American Guidance Service.
- Harms, T., Clifford, R. M., & Cryer, D. (1998). Early Childhood Environment Rating Scale Revised Edition. New York: Teachers College Press.
- National Center for Early Development and Learning (2001). Counting Numbers. Unpublished instrument.
- Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). Woodcock-Johnson III Tests of Achievement. Itasca, IL: The Riverside Publishing Company.

³ Benjamini, Y. & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B (Methodological)*, 57, 289-300.