IMPROVEMENTS IN MATH AND READING SCORES OF STUDENTS WHO DID AND DID NOT PARTICIPATE IN THE FOUNDATIONS AFTER SCHOOL ENRICHMENT PROGRAM DURING THE 2001-2002 SCHOOL YEAR

Stephen P. Klein, Ph.D. and Roger Bolus, Ph.D.

GANSK & ASSOCIATES
Santa Monica, CA
310-396-3030

December 2, 2002

Contents

List of Tables

Preface

The Foundations After-School Enrichment Program operates extended day programs during the school year and in the summer. During the 2001-2002 school year, this program administered tests in mathematics and reading in the fall and again in the spring to students in 19 schools. These schools were located in three counties in Pennsylvania (Philadelphia, Delaware, and Franklin); in Cumberland County, New Jersey; and in Palm Beach County, Florida. Almost all of the children tested were in grades 1- 5. This report uses the results from the fall and spring testing to describe how well students in the Foundations program performed relative to a national norm group and to students with comparable skills and background characteristics who were not in a Foundations program. A separate report describes the results at three schools where testing was done on a different schedule.[1]

[1] Emilio Matticoli and Rhe McLaughlin of the Foundations staff and Russ Dusewicz of CTB/McGraw-Hill were very helpful in providing the data and documentation that were needed for the analyses presented in these reports.

Summary

The students in the Foundations program plus a large sample of non-Foundations students took the CTB/McGraw-Hill CAT-5 Mathematics and Reading Comprehension tests in the fall of 2001 and again in the spring of 2002. Analyses of these data found that the Foundations students made substantial improvements in average scores between the fall pretest and spring posttest. In fact, their average score gains in mathematics were somewhat greater than what would be expected given the results obtained in CTB/McGraw-Hill's national norm sample. The Foundations students' gains in reading kept pace with those made in this national norm sample.

More importantly, the average score gains between the fall pretest and spring posttest that were made by students in the Foundations program were substantially (and statistically significantly) greater than those made by comparable students from the same schools and communities. For example, although the Foundations and non-Foundations students had nearly identical average pretest scores in each subject, the Foundations students' mean increase in effect size during the 2001-2002 school year was roughly twice as large as it was among non-Foundations students. The respective effect sizes for Foundations and non-Foundations students were .91 and .52 in math and .76 and .35 in reading.

In addition, the students who participated in the Foundations program during the 2001-2002 school year made greater gains in both math and reading between the pretest and posttest than did the students in this program during the 2000-2001 school year.

It may be concluded from these findings that the Foundations program continues to be very effective in developing the students' math and reading skills.

Background

There is growing desire for programs that offer before- and after-school child care, especially given the increase in the number of students living in households headed by single working parents or by dual-career parents. A large majority of mothers with children ages 6 to 13 are in the labor force (Vandell & Shumow, 1999), many of them working full time (Capella & Larner, 1999). Because the school day is shorter than the typical workday, there may be a two- to four-hour period when many children are left unsupervised. Approximately 44 percent of the children with working parents have no adult care after school, with low-income children more likely than their affluent peers to be left alone (Halpern, 1999).

The interest in establishing after-school programs stems in part from juvenile crime, violence, drug use, and sexual activity generally occurring during unsupervised hours. After-school programs are reported to be the fastest growing segment of child care (Seligson, Gannet, & Cotlin, 1992), with support from both federal and local initiatives. In 1999, Congress appropriated $200 million to create after-school programs in the schools, and the DeWitt Wallace-Reader's Digest Fund granted $1.2 million for intervention programs in several low-income areas of Boston, Chicago, and Seattle. The need for high-quality programs, however, continues to outstrip the supply, particularly for families whose economic resources make it difficult for them to pay even the modest tuition that is often charged. Halpern (1999) estimates that the available supply of programs meets only about one-quarter of the child-care needs.

Children who are unsupervised during non-school hours often spend that time watching television or engaging in other activities that do not contribute to intellectual or emotional growth. And while the most frequently cited purpose of after-school programs is to ensure adult supervision of children (Seppanen, deVries, & Seligson, 1993), many of these programs (including Foundations) also provide academic and other kinds of enrichment activities. For example, students in after-school programs spent more time in academic activities, enrichment experiences such as music lessons, and stimulating social interactions with adults and peers than did similar children who were unsupervised or who were cared for by their mothers or other relatives (Posner &Vandell, 1994).

Participation in after school programs has been found to be associated with fewer disciplinary problems (Petit, Laird, Bates, & Dodge, 1997), higher grades (Cooper, Valentine, Nye, & Lindsay, 1999; Pierce, Hamm, & Vandell, 1999), improved standardized test scores (Huang, et al., 2000), and better social relationships with their peers (Howes, Olenick, & Der-Kiureghian, 1987; Posner & Vandell, 1994). Results such as these are helping educators and policymakers to recognize the value of creating engaging learning opportunities to fill some of a student's non-school hours. For example, in 2002, California voters approved a statewide initiative that guarantees support for such programs.

Overview

This report describes the procedures that were used and the results obtained in our analysis of the test score gains in mathematics and reading test scores of the grade 1- 5 students in the Foundations program during the 2001-2002 school year. There were not enough middle school students in the program to provide meaningful results for them.

The remainder of this report begins with a brief description of the Foundations program and the measures, types of scores, and samples of students used in our analyses. Next, we discuss the improvement in performance between the fall and spring of the Foundations students at each grade level. We then report average gain scores by school. This is followed by a comparison of the scores of similarly situated students who did and did not participate in the program. The last section discusses our conclusions.

The Foundations Program and Previous Evaluations

Foundations is a non-profit organization established in 1992 to develop and operate before- and after-school enrichment programs. Almost all the children served by this program are in grades 1 - 5. Most of these children are eligible for free or reduced price lunches. Foundations programs operate during the school year (which is what we studied) and in select locations during the summer months.

The Foundations' programs include a curriculum that involves daily activities emphasizing academic subjects as well as experiences designed to foster physical and emotional development. Students also engage in other activities, including field trips and homework assistance.

The student-to-adult ratio in the program is approximately ten to one and all staff members have experience in working with elementary school children. Teachers and program coordinators have bachelor's or master's degrees in education or related fields. A sliding scale is used to determine fees based on each family's ability to pay. The program places special emphasis on family involvement, with frequent communication between staff and parents, and opportunities for families to volunteer in the classroom and participate in other ways. Staff also make an effort to communicate regularly with the teachers who teach students during the regular school day, thereby providing some coordination among the experiences students have throughout the day (this is particularly true for the homework assistance portion of the program).

There are many potential benefits to participation in the Foundations program, including improved social relations and attitudes toward school. Even among the academic subjects, improvement might be detectable in areas other than those tested. The program has been widely approved by parents (who welcome the supervision and appreciate the enrichment their children receive; Miller, 1997). Temple University evaluators who have visited the classrooms have applauded the program's apparent effects on student self-esteem (McCurdy, 1995). In 2000, an evaluation of four Foundations schools by Drexel University found favorable student attitudes toward the program, and significant improvements in reading at one of the school sites (Reisman, Bach, Whitelaw, 2001). RAND conducted analyses of math and reading test score gains for the 1997-98 through 2000-01school years (see Hamilton & Klein, 1998; Hamilton, Le, & Klein, 1999; and Le & Hamilton, 2001a and 2001b).

Mathematics and Reading Tests

From fall 1997 to spring 1999, the Foundations program administered the Mathematics Computation and the Reading/Language Arts sections of the Terra Nova (CTB/McGraw-Hill, 1997b). The Mathematics Computation section asks students to complete problems involving arithmetic operations that are appropriate to their grade level. All of these problems contain numbers only (i.e., there are no "word" problems).

The Reading/Language Arts portion of the Terra Nova consists of several reading passages followed by questions. These passages cover a variety of themes and are selected to engage students. This test also includes language, vocabulary, writing, and editing skills items (in multiple-choice format). Test length varies by grade level, with longer tests given to students in higher grades. Students receive separate scores for reading and language arts.

RAND staff advised Foundations to switch tests for the fall 1999 assessment because of concerns that the repeated use of the same measure across years might yield spurious score gains that were not associated with a commensurate increase in ability (Linn, 2000). This led Foundations to administer the CAT-5, which is also published by CTB/McGraw-Hill (1996). The Math Concepts and Application section of this multiple-choice test assesses a broader range of skills and knowledge than does the Mathematics Computation portion of the Terra Nova. However, because of time constraints only the Reading Comprehension section of the CAT-5 was administered. Thus, language arts skills and other aspects of reading proficiency (such as word analysis) were not assessed.

On both the Terra Nova and the CAT-5, there is a separate "form" (set of questions or "items") for each grade level. Some of the items on one form also appear on other forms. For example, the form for fourth graders might contain some of the more difficult items from the third grade form and some of the easier items from the fifth grade form. As noted below, this overlap allows putting the scores from different forms onto a common scale of measurement across grade levels.

Test Scores Used in Analyses

CTB/McGraw-Hill uses an item response theory (IRT) methodology to convert raw scores (i.e., the number of questions answered correctly) to scale scores on each section and form of the test (CTB/McGraw-Hill, 1996). This procedure puts the scores from different forms of the CAT-5 onto the same 100 to 900 point scale of measurement. As a result of this scaling, we can combine and average scale scores across grade levels.

CTB/McGraw-Hill also computes normal curve equivalent scores (NCEs) for each student on each section. The NCEs indicate how far above or below the mean the student's score falls relative to a national norm group of students who are at the same grade level. NCEs are computed separately for fall, winter, and spring testing. CTB/McGraw-Hill reports NCEs on a scale that has a mean of 50 and a standard deviation of 21.06 (Crocker & Algina, 1986). We converted the mean NCE within a group to its corresponding National Percentile (NP) scores to assess whether the improvements in the students' scores during the year were keeping pace with the improvement in the national norm group during a corresponding number of months (an increase in NPs suggests quicker progress than the national norm sample whereas a decrease in NPs indicates the opposite).

Analysis Samples

Foundations staff supervised the administration of the CAT-5 to students in 19 schools in October 2001 and May 2002. Foundations students were tested at all of these schools and non-Foundation students were tested at seven of them. The analyses used the data on just the students who had both a pretest and a posttest score in a subject. The number of students used in the analysis of math scores differed somewhat from the number used for reading due to absences and other factors.[2]

[2] Separate analyses were not conducted for 6th, 7th, and 8th graders because there were very small numbers of them and they were not distributed proportionately across schools.

Grade-by-Grade Analysis

This section of our report discusses the scale score and NP changes between the October 2001 pretest and the May 2002 posttest at each grade level in mathematics and reading. The major finding from this analysis is that while there was a highly statistically significant improvement in scores between the pretest and posttest, the amount of gain was only slightly higher than what would be expected given the results in the national norm group, although the amount of gain in mathematics was slightly greater than it was in reading. One factor that may have contributed to these findings was that the students' NPs (which were close to the national median) were higher than what would be expected on the basis of their demographic characteristics or their pretest scores in previous years (see prior reports). The remainder of this section documents these findings.

Table 1 shows the number of students who had a pretest and posttest scale score in mathematics, their means on these two occasions, and the difference between these means. Table 2 shows the corresponding data for reading. All the differences were statistically significant at $p < .001$. The last row of each table shows the mean of the values above it (all mean scale scores and NPs are rounded to integers).

Table 1
Mean Pretest, Posttest, and Difference Scale Scores by Grade Level
for Grade 1-5 Foundations Students: **Mathematics**

Grade	Number of Students	Mean Scale Score		Difference
		Pretest	Posttest	
1	94	572	640	68
2	92	631	691	60
3	90	652	690	38
4	64	686	703	17
5	44	693	735	42
Overall	384	647	692	45

Table 2
Mean Pretest, Posttest, and Difference Scale Scores by Grade Level
for Grade 1-5 Foundations Students: **Reading**

Grade	Number of Students	Mean Scale Score		Difference
		Pretest	Posttest	
1	101	596	648	52
2	100	638	689	51
3	95	665	694	29
4	65	683	714	31
5	45	693	717	24
Overall	406	655	693	38

Table 3 expresses the differences reported in Tables 1 and 2 as "effect" sizes. In this context, an effect size is the difference between the mean pretest and posttest scale scores at a grade level divided by the standard deviation of the pretest scores at that grade level. Effect sizes are a more appropriate indicator of gain than simple differences between means. As in the past, effect sizes were somewhat larger in the lower grade levels than in the upper grade levels. Effect sizes were generally larger in math than in reading.

Table 3
Scale Score Effect Sizes by Subject for Grade 1-5 Foundations Students

Grade	Mathematics	Reading
1	1.20	0.88
2	1.38	1.06
3	0.74	0.64
4	0.34	0.67
5	0.88	0.55
Mean	0.91	0.76

Tables 4 and 5 show the pretest and posttest means expressed as national percentiles (NPs). These data suggest that the Foundations students were progressing faster than CTB/McGraw-Hill's national norm group in math skill development and they were keeping pace with the norm group in reading.

Table 4
Mean Pretest, Posttest, and Difference in National Percentile Equivalents
for Grade 1-5 Foundations Students: **Mathematics**

Grade	Mean Percentile		Difference
	Pretest	Posttest	
1	50	67	17
2	64	70	6
3	46	56	10
4	55	54	- 1
5	46	61	15
Mean	51	61	10

Table 5
Mean Pretest, Posttest, and Difference in National Percentile Equivalents
for Grade 1-5 Foundations Students: **Reading**

Grade	Mean Percentile		Difference
	Pretest	Posttest	
1	74	65	- 9
2	57	67	10
3	51	58	7
4	46	58	8
5	38	41	3
Mean	56	58	2

School-by-School Analysis

Although the Foundations programs at each school generally follow a common curriculum and approach, it is likely that the nature of the instruction varies across schools as a function of instructor characteristics and other factors that are unique to each school. Consequently, we might see variation in the average amount improvement across schools. To investigate this matter, we combined all the grade levels within each school, computed the difference between the school's mean pretest and posttest scores, and then transformed this difference to an effect size by dividing the difference by the standard deviation of the school's pretest scores. The results of this analysis appear in Table 6.

Table 6
Mean Scale Score Effect Sizes by Subject and School for Foundations Students*

School	Number of Students		Mean Effect Size	
	Mathematics	Reading	Mathematics	Reading
1	17	17	0.89	0.73
2	19	20	0.66	0.57
3	14	14	0.81	1.16
4	18	18	0.23	0.66
5	52	52	0.85	0.83
6	27	28	0.21	- 0.04
8	16	25	0.59	0.49
9	21	21	1.49	1.15
10	23	24	1.32	1.06
11	24	25	0.52	0.28
14	48	48	0.45	0.33
15	23	22	1.02	1.04
17	26	30	1.30	1.18
18	21	23	1.30	0.93
19	11	10	0.80	0.61

* Includes schools with 10 or more students in each subject.

All but five of the effect sizes in Table 6 were statistically significantly greater than zero at the .05 level (the ones that were not significant were schools #4 and #6 in math and schools #6, #11, and #19 in reading). Some schools showed more gain than others. For example, school #9 had over a one standard deviation unit increase in both math and reading whereas schools #6 and #14 had much smaller increases. Other schools showed a large improvement in one area but only a modest gain in the other. We do not have the data that would be needed to investigate the source of this variation among schools, but it is not related to how many Foundation students they had.

Comparison of Score Gains for Foundations and Non-Foundations Students

As noted above, one way to assess the effects of the Foundations program on the development of student math and reading skills is to compare their progress to that of students in the national norm sample (see Tables 4 and 5 for this comparison). A second method, and the one discussed in this section, involves comparing the progress of students in the Foundations program to that of non-program participants from the same or similar schools. Although this method also has it limitations (e.g., program participants may differ in important ways from non-participants), it is still better than just relying on the national norm sample.[3]

Consequently, at seven schools, Foundations staff administered the same math and reading tests to non-program participants as they gave to participants. Table 7 shows the number of non-Foundation students taking the math pretest and posttest at each grade level and their mean scores on this test. Table 8 shows the same data for reading. These tables correspond to Tables 1 and 2 for the Foundations students.

A comparison of these two sets of tables indicates that the Foundations and Non-Foundations students had nearly identical mean pretest scores. Their overall pretest means in math were 647 and 649, respectively. The corresponding values in reading were 655 and 653. Thus, the two groups had comparable math and reading skills in the fall. However, the Foundations students gained an average of 45 points in math and 38 points in reading compared to only 26 and 17 points for non-Foundations students.

[3] Because of the voluntary nature of the program, it was not possible to randomly assign students to Foundations versus non-Foundations groups.

Table 7
Mean Pretest and Posttest Scores for Non-Foundation Students: **Mathematics**

Grade	Number of Students	Mean Scale Score		Difference
		Pretest	Posttest	
1	108	574	607	33
2	84	630	664	34
3	115	661	677	16
4	151	674	701	27
5	125	707	726	19
Overall	583	649	675	26

Table 8
Mean Pretest and Posttest Scores for Non-Foundation Students: **Reading**

Grade	Number of Students	Mean Scale Score		Difference
		Pretest	Posttest	
1	119	593	617	24
2	87	623	641	18
3	117	655	669	14
4	184	682	699	17
5	139	710	722	12
Overall	646	653	670	17

A regression analysis that controlled for grade level and school found that the Foundation students' overall average improvement in scale scores between the pretest and posttest was statistically significantly greater than it was among the non-Foundations students ($p < .001$). This held for both math and reading. A t-test comparison of the differences in mean gain scores produced statistically significant effects (at $p < .05$) for almost all grade levels and subjects. The only exceptions were Grade 4 in math and Grade 5 in reading.

Tables 9 and 10 report the effect sizes for Foundations and non-Foundations participants as well as the differences between these effect sizes.[4] To put these values in perspective, an effect size of about one-quarter of a standard deviation unit is comparable to that found for the effect of reduced class size in Tennessee (Finn & Achilles, 1999).

Table 9
Effect Sizes Within and Between Groups by Grade Level: **Mathematics**

Grade	Foundation	Non-Foundation	Difference
1	1.20	0.58	0.62
2	1.38	0.79	0.59
3	0.74	0.29	0.45
4	0.34	0.55	-0.21
5	0.88	0.41	0.47
Mean	0.91	0.52	0.39

Table 10
Effect Sizes Within and Between Groups by Grade Level: **Reading**

Grade	Foundation	Non-Foundation	Difference
1	0.88	0.41	0.47
2	1.06	0.36	0.70
3	0.64	0.31	0.33
4	0.67	0.38	0.29
5	0.55	0.27	0.28
Mean	0.76	0.35	0.41

[4] The effect size was computed by dividing the difference between the pretest and posttest mean scale scores by the standard deviation of the pretest scores among all the students in the analysis (i.e., Foundation and non-Foundation students combined). This was done separately for each grade level.

Comparing the 2001-2002 Results to Those of Previous Years

Tables 11 and 12 present the national percentiles and effect sizes for gains from the pretest to the posttest for Foundations participants in grades 1-5 from fall 1997 through spring 2002 in math and reading.[5] A comparison of the data in Table 12 with those in Table 3 indicates that the mean effect sizes in the 2001-2002 school year were larger than those found in the analysis of the 2000-2001 school year data. Specifically, the mean math and reading effect sizes across all five grade levels during the 2001-2002 school year were .91 and .76, respectively. The corresponding values for the 2000-2001 school year were .51 and .49.

Consistent with prior findings (Vandell & Shumow, 1999), younger children tended to show larger improvements (as measured by effect sizes) than older students. For example, across all five years studied, the math effect sizes for first- and second-graders were generally larger than those of fourth-and fifth-graders.

Discussion

Mean pretest national percentiles tended to be higher in fall 2001 than in previous years. This difference may stem in part from any carry over benefit of participating in the Foundations program in prior years. However, this hypothesis does not explain the especially high pretest percentiles among first graders in the fall of 2001. We could not investigate this issue further because there was no way to link a student's scores across years.

We recognize that students may benefit from participation in the Foundations program in ways that are not captured by their CAT-5 math and reading test scores, such as in terms of their social development and interest in learning. But again, we did not have the data that would be needed to investigate such outcomes.

[5] We include only those grade levels that were represented across all years from 1997-2002. We do not compare Terra Nova and CAT-5 scale scores across these tables because these scores are on different scales of measurement.

Table 11
Pre- and Post-Test National Percentiles and Effect Sizes for Foundations Participants in Math and Reading by Grade Level and Academic Year: Fall 1997 to Spring 2000

Year	Grade	Math					Reading				
		Number of Students	Effect Size from Pretest to Posttest	Pr > |t|	Mean Pretest NP	Mean Posttest NP	Number of Students	Effect Size from Pretest to Posttest	Pr > |t|	Mean Pretest NP	Mean Posttest NP
97-98	1	66	1.07	.0001	32	60	68	1.40	.0001	21	53
	2	55	1.17	.0001	36	54	54	.33	.0003	39	40
	3	37	1.15	.0001	35	65	37	.97	.0002	32	52
	4	36	.77	.0007	43	60	36	.63	.0021	31	42
	5	18	-.08	.7701	17	13	18	.06	.8001	11	10
98-99	1	53	1.37	.0010	26	61	53	1.31	.0548	22	55
	2	43	.91	.0055	24	45	46	.59	.0037	24	38
	3	26	.34	.3166	29	39	30	.49	.0297	24	36
	4	33	.88	.0324	33	53	37	.32	.0937	32	41
	5	16	.18	.1975	31	38	21	.18	.5321	31	35
99-00	1	64	.83	.0028	31	46	51	1.10	.0226	38	51
	2	26	.55	.2530	69	61	71	.35	.0018	61	53
	3	45	.41	.0558	31	36	45	.12	.3215	35	24
	4	37	.27	.0587	51	52	38	.10	.2674	51	43
	5	35	.38	.3980	45	55	34	.25	.0178	30	30

Table 12
Pre- and Post-Test National Percentiles and Effect Sizes for Foundations Participants in Math and Reading by Grade Level and Academic Year: Fall 2000 to Spring 2002

Year	Grade	Math					Reading				
		Number of Students	Effect Size from Pretest to Posttest	Pr > |t|	Mean Pretest NP	Mean Posttest NP	Number of Students	Effect Size from Pretest to Posttest	Pr > |t|	Mean Pretest NP	Mean Posttest NP
00-01	1	108	.78	.0000	29	50	113	.86	.0000	45	56
	2	90	.56	.0029	47	53	110	.52	.0000	37	51
	3	93	.46	.0000	45	51	97	.46	.0000	44	47
	4	78	.26	.0073	45	51	85	.29	.0000	38	47
	5	32	.48	.0009	45	63	32	.30	.0198	44	51
01-02	1	94	1.20	.0000	50	67	101	0.88	.0000	74	65
	2	92	1.38	.0000	64	70	100	1.06	.0000	57	67
	3	90	0.74	.0000	46	56	95	0.64	.0000	51	58
	4	64	0.34	.0000	55	54	65	0.67	.0000	46	58
	5	44	0.88	.0000	46	61	45	0.55	.0000	38	41

Conclusions

The students who participated in the Foundations program during the 2001-2002 school year made substantial and statistically significant gains in both math and reading scores between the fall pretest and spring posttest. These gains kept pace with those found in CTB/McGraw-Hill's national norm sample for reading and showed more rapid growth than the national norm sample in math. More importantly, even though Foundations and comparable non-Foundations students from the same or similar schools had nearly identical mean pretest scores, the Foundations students had significantly higher mean spring posttest than the non-Foundations students. This was true for both math and reading.

In addition, the Foundations students' test score gains during the 2001-2002 school year were over a quarter of a standard deviation greater than those obtained by Foundations students during the 2000-2001 school year. Thus, when taken together, ***the foregoing findings show that the Foundations program continues to significantly help students develop their math and reading skills.***

Foundations staff may find it productive to investigate why the test score gains at some schools were much greater than they were at other schools. This type of an analysis may help to determine which aspects of the program are especially important in fostering student learning. Foundations and other organizations that operate after-school enrichment programs might then be able to increase their effectiveness by insuring that these features are present.

References

Capella, E., & Lerner, M.B. (1999). America's schoolchildren: Past, present, and future. *The Future of Children*, *9(2)*, 21-29. Available at www.futureofchildren.org.

Cooper, H., Valentine, J. C., Nye, B., & Lindsay, J. L. (1999). Relationships between five after-school activities and academic achievement. *Journal of Educational Psychology, 91,* 369-378.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Fort Worth: Harcourt Brace Jovanovich.

CTB/McGraw-Hill (1996). CAT-5 Technical Manual. Monterey, CA: Author.

CTB/McGraw-Hill (1997a). Beyond the numbers: A guide to interpreting and using the results of standardized achievement tests. Monterey, CA: Author.

CTB/McGraw-Hill (1997b). Teacher's guide to Terra Nova. Monterey, CA: Author.

Finn, J.D., & Achilles, C.M. (1999). Tennessee's class size study: Findings, implications, and misconceptions. *Educational Evaluation and Policy Analysis*, 21, 97-109.

Halpern, R. (1999). After-school programs for low-income children: Promises and challenges. *The Future of Children*, *9(2)*, 81-95. Available at www.futureofchildren.org.

Hamilton, L.S., & Klein, S. (1998). Achievement test score gains among participants in the Foundations school-age enrichment program. PM-858-EDU. RAND: Santa Monica, CA.

Hamilton, L.S., Le, V., & Klein, S. (1999). Foundations school-age enrichment program: Evaluation of student achievement. PM-998-EDU. RAND: Santa Monica, CA.

Howes, C., Olenick, M., & Der-Kiureghian, T. (1987). After-school child care in an elementary school: Social development and continuity and complementarity of programs. *Elementary School Journal*, *88*, 93-103.

Huang, D., Gibbons, B., Kim, K. S., Lee, C., & Baker, E. L. (2000). *A Decade of Results: The Impact of the LA's BEST After School Enrichment Program on Subsequent Student Achievement and Performance*. Los Angeles; UCLA Center for the Study of Evaluation.

Le, V., & Hamilton, L.S. (2001a). Examining test score gains among participants of the Foundations after-school program. PM-1178-EDU. RAND: Santa Monica, CA.

Le, V., & Hamilton, L.S. (2001b). Achievement gains in math and reading by participants of the Foundations after-school enrichment program. PM-1265-EDU. RAND: Santa Monica, CA.

Linn, R.L. (2000). Assessments and accountability. *Educational Researcher*, 29(2), 4-14.

McCurdy, B. (1995, May). Foundations program gets A+. *News of Delaware County*.

Miller, K. C. (1997, March 3). Solid "Foundations". *Delaware County Daily Times*, p. 4.

Pettit, G.S., Laird, R.D., Bates, J.E., & Dodge, K.A (1997). Patterns of after-school care in middle childhood: Risk factors and developmental outcomes. *Merrill-Palmer Quarterly*, *43*, 515-538.

Pierce, K.M., Hamm, J.V., & Vandell, D.L. (1999). Experiences in after-school programs and children's adjustment in first-grade classrooms. *Child Development*, *70*, 756-767.

Posner, J. K., & Vandell, D. L. (1994). Low-income children's after-school care: Are there beneficial effects of after-school programs? *Child Development, 65,* 440-456.

Reisman, F.K., Bach, C.N., Whitelaw, L.A. (2001). *21st Century Schools Evaluation*. Philadelphia, PA.: Drexel University.

Seligson, M., Gannet, E., & Cotlin, L. (1992). Before- and after-school child care for elementary school children. In B. Spodek & O.N. Saracho (Eds.), *Issues in child care* (pp.125-142).

Seppanen, P. S., de Vries, D. K., & Seligson, M. (1993). *National study of before- and after-school programs: Executive summary.* Washington, D.C.: U.S. Department of Education, Office of Policy and Planning.

Vandell, D.L., & Shumow, L. (1999). After-school child care programs. *The Future of Children*, *9(2)*, 64-80. Available at www.futureofchildren.org.