An Analysis of Academic Progress of Children Participating in the Core Knowledge Preschool Program in Baltimore County Head Start Centers

Susan Sonnenschein, Linda Baker, and Adia Garrett

University of Maryland, Baltimore County

August 2005

Correspondence can be directed to Susan Sonnenschein or Linda Baker, Department of Psychology, UMBC, 1000 Hilltop Circle, Baltimore MD 21250, 410-455-2361 or 410-455-2370, **sonnensc@umbc.edu** or **baker@umbc.edu**. We thank the following students for their assistance in administering and scoring the assessments: Andrea Anushko, Lisa Beall, Amy Benson, Sarah Davarya, Maria Finger, Elyse Grossman, Julie Grossman, Bret Hite, Faith Morse, Alicia Ritgert, Patricia Tenowich, and Stephanie Willett.

Executive Summary

This report presents results of an evaluation of the efficacy of the Core Knowledge Preschool Sequence implemented in the 10 Baltimore County Head Start Centers during the 2004-2005 academic year. The evaluation focused on the 4-year-olds participating in the program. The evaluation consisted of two components: (1) a comparison of pre-test and post-test performance on 9 subtests of the Woodcock Johnson Tests of Achievement- III and (2) a comparison of pre-test and post-test performance on 17 tasks of the Core Knowledge Preschool Assessment Tool (CK-PAT). The Woodcock Johnson is a nationally recognized, standardized measure of children's academic knowledge. The CK-PAT was developed by Core Knowledge staff to document children's attainment of skills taught in the program. Supplementary analyses examined relations among the two instruments and relations between performance on the academic assessments and ratings of social skills and problem behaviors provided by teachers using the Social Skills Rating System (SSSR).

Children at the10 Head Start centers in Baltimore County, MD participated in this evaluation. Teachers in each center were asked by Core Knowledge administrators to randomly pick a subset of children from the four-year-old classes and complete a social skills rating form for each child. Participants in our evaluation were children from that roster. Children were individually tested by trained graduate students or graduating seniors in developmental psychology. Testing took place at the Head Start centers during the school day. Pre-testing took place from the middle of December through the end of February. Post-testing took place from May through the middle of June. Eighty-seven children were available for pre-testing and post-testing on the Woodcock Johnson and 97 were available for the CK-PAT.

Results of analyses with scores on the Woodcock Johnson show that children receiving the Core Knowledge Preschool Sequence program made significant and moderate to moderately strong increases in performance from pre-test to post-test. Children's growth in knowledge was similar to the Woodcock Johnson standardization group, which consisted of children of different income levels and ethnicities comparable in age to the Core Knowledge children. Children receiving Core Knowledge instruction actually showed more growth compared to the standardization sample on a few of the subtests (Oral Comprehension, Quantitative Concepts, Oral Language Cluster). These results are particularly impressive given that the amount of time between the pre-tests and post-tests was really quite short, ranging from 3 to 5 months.

Using an instrument that is aligned directly to the goals and objectives of the Core Knowledge Preschool Sequence, the CK-PAT, we found that children gained significantly on 16 of the 17 individual tasks from pre-test to post-test and that they showed significant growth on the three composite areas of oral language, emergent literacy, and mathematics. Thus, these results indicate that children made progress in the specific objectives of the program.

A supplementary set of analyses related scores on the Woodcock Johnson to those on the CK-PAT. The moderate correlations between the two instruments indicate that CK-PAT has validity as a measure of children's achievement in the program.

Consistent with the literature showing the importance of social skills for children's success in school, we found that children's social skills, as rated by their teachers during the fall of the school year, were positively related to their pre-test and post-test scores on the Woodcock Johnson and CK-PAT. Children's social skills were also significantly related to their growth in oral language skills as assessed with the Woodcock Johnson.

The results of this evaluation show that the Core Knowledge Preschool Sequence as implemented in the Baltimore County Head Start centers is successful in providing low income children with the skills and knowledge that children of their age across the country are expected to master.

An Analysis of Academic Progress of Children Participating in the Core Knowledge Preschool Program in Baltimore County Head Start Centers

This report presents results of an evaluation of the efficacy of the Core Knowledge Preschool Sequence implemented in the 10 Baltimore County Head Start Centers during the 2004-2005 academic year. The evaluation focused on the 4-year-olds participating in the program. The evaluation consisted of two components: (1) a comparison of pre-test and post-test performance on 9 subtests of the Woodcock Johnson Tests of Achievement-III and (2) a comparison of pre-test and post-test performance on 17 tasks of the Core Knowledge Preschool Assessment Tool (Ck-PAT). Supplementary analyses examined relations among the two instruments and relations between performance on the academic assessments and ratings of social skills and problem behaviors provided by teachers using the Social Skills Rating System.

Information about Participants and Procedures

Participants

Children at the10 Head Start centers in Baltimore County, MD participated in this evaluation. The participating centers were: Back River, Campfield, Chase, Emily Harris, Fleming, Merrit Park, Reisterstown, Riverview, Towson, and White Marsh. Teachers in each center were asked by Core Knowledge administrators to randomly pick a subset of children from the four-year-old classes and complete a social skills rating form for each child. Participants in our evaluation were children from that roster. Table 1 provides demographic information about the participating children.

Children's ages as of September 1, 2004 ranged from 45 – 61 months, with a mean of 53.13 months. About 55% of the sample were girls, 45% were boys (N = 61 girls, 49 boys). Most of the children attended Head Start's half-day program, either in the morning or afternoon. Four of the centers offered a full-day class. About 16% of the children tested attended the full-day program.

Pre-testing took place from the middle of December through the end of February. One hundred and ten children were available for pre-testing on the Woodcock Johnson and the Core Knowledge Preschool Assessment Tool. One hundred and ten children were administered subtests from the Woodcock Johnson; 106 were administered tasks from the Core Knowledge Preschool Assessments Tool (CK-PAT). A fair number of children no longer attended Baltimore County Head Start centers (or attended so infrequently that they were never available for testing) when we administered the post-tests, from May through the middle of June. We made many attempts to test all the children in the sample. However, due to absences and school closings, several children were only available to complete some but not all of the various subtests on the two test batteries.

Table 1. Demographic Information about Participants at Each Center

Center	Number of Participants			Mean Age (months) as of 9/04
	Boys	Girls	Total	
Back River	5	11	16	53.94
Campfield	12	9	21	52.29
Chase	1	3	4	53
Emily Harris	6	9	15	52.07
Fleming	9	7	16	53.44
Merrit Park	2	4	6	51.33
Reistertown	2	3	5	55
Riverview	3	6	9	54.56
Towson	3	4	7	53.44
Whitemarsh	6	5	11	52.18
TOTAL	49	61	110	53.13

Eighty-seven children were available for testing on the pre-test and post-test Woodcock Johnson and 97 were available for the CK-PAT. Preliminary analyses showed no significant differences in mean pre-test scores on the Woodcock Johnson tests or CK-PAT between children who were available for post-testing and those who were not.

Testing Procedures

Children were individually tested at their Head Start centers during the school day. Given the length of each test battery, testing was conducted on two different days, typically about a week apart. The order of administration of the two test batteries was counterbalanced across children. Children were tested on one battery per session. Children were given stickers as a thank you for participating in the testing.

Testers were graduate students or graduating seniors in developmental psychology who received extensive training in test administration prior to beginning the testing. Once testing began, weekly team meetings were held to discuss any testing and scoring issues that arose.

Data Entry and Analytic Procedures

*Scoring of tests/entering of data***.** The validity of any set of findings depends upon the accuracy of the data. We used several precautions to minimize errors in scoring and data entry. All tests were initially scored on site by the tester and then checked after the completion of testing. Scores were then rechecked by another research assistant. Any

concerns about how items were to be scored were reviewed with the evaluators and other team members.

Raw scores from the Woodcock Johnson were entered into a computerized scoring program, reviewed for clerical errors, and then entered into the data base on the computer. Raw scores from the CK-PAT were entered directly into the computer. To check for errors we used a double entry system. That is, we entered data twice into two separate data bases and compared the scores from each set of data. Scores were entered into each of the two data sets by different research assistants.

Conventions for reporting data. Statistical analyses were conducted to compare pre-test and post-test scores on all measures. As is common in the field, statistically significant differences are defined here as those occurring less than 5 out of 100 times by chance ($p < .05$). For readability purposes, we do not present the statistics themselves in the body of the report but rather provide them in Appendix C. Also to increase readability, the primary data tables present only means and significance levels; the appendices include standard deviations as well.

The number of children tested in each of the 10 centers on the various tests ranged from 2 to 16. The variability in the number of children per center and, more importantly, the very limited number of children in several of the centers precludes the use of inferential statistics. Accordingly, we did not statistically compare children's performance across centers. However, Appendix C includes descriptive information about children's scores at each center on the two sets of assessments.

Woodcock Johnson Tests of Achievement

Description of the Assessment

The Woodcock Johnson Tests of Achievement - III (WJ-III) assess children's and adults' academic skills. The test is normed for people ages 2 through 90 years, although not all tests are appropriate for all ages. The Woodcock Johnson has 2 comparable versions, A and B, to allow for assessing growth in a person's skills without the risk of familiarity effects due to the repetition of items. Children in this evaluation were pre-tested with one version of the Woodcock Johnson and post-tested with the other.

The Woodcock Johnson is a highly regarded norm-referenced test battery with excellent psychometric properties (internal consistency, test-retest, Sattler, 2001). Validity figures, although good, were not based on children of the age of those in this evaluation. The Woodcock Johnson was standardized on individuals selected to conform to the demographics of the population from the 2000 U.S. Census. The norming group contained over 1000 preschool children and was stratified according to factors such as parental education (often a proxy for income) and ethnicity. Although low income children were represented in the standardization sample consistent with their proportional representation in the population, middle income children also were included in the standardization sample. The inclusion of middle income children in the

norming sample becomes relevant when we discuss differences in standard scores from pre-test to post-test.

Twelve of the Woodcock Johnson subtests are considered appropriate for use with preschool children. Eight of the 12 subtests plus one additional subtest that is not part of the preschool set were selected by Core Knowledge personnel for this evaluation. The subtests are: Letter-Word Identification, Story Recall, Understanding Directions, Spelling, Applied Problems, Picture Vocabulary, Oral Comprehension, Academic Knowledge, and Quantitative Concepts (the latter is not part of the preschool subtests). A brief description of each subtest appears in Table 2.

Two sets of scores are available with the Woodcock Johnson, raw scores and standard scores. Raw scores reflect the number of points a child earns on a subtest. Such scores tell us nothing, however, about how well the child has performed compared to peers. Standard scores are derived by comparing the raw score to a normative national sample. The Woodcock Johnson provides two sets of possible norms for computing standard scores, one set based on age and the other based on grade. We used the age norms to document how the children receiving Core Knowledge instruction performed relative to a nationally representative group of children matched on age.

The standard score has a mean of 100 and a standard deviation of 15. Average scores fall between 90 and 110; low average scores fall between 80 and 89; high average scores fall between 111 and 120.

In addition to standard scores derived from individual subtests, several composite standard scores are available. The *Oral Language cluster* is a composite measure of expressive vocabulary, reasoning, listening comprehension, and memory. Subtest scores from story recall, understanding directions, picture vocabulary and oral comprehension form the composite. The *Math Reasoning cluster* is a composite measure of problem solving, analysis, reasoning and vocabulary. Subtest scores from applied problems and quantitative concepts form the composite. The *Academic Knowledge cluster* samples knowledge of science, social studies and culture. Scores from the three parts of the academic knowledge subtest form the composite.

Analyses of Change over Time in Children's Performance on the Woodcock Johnson

Table 3 shows the mean raw scores on the individual subtests, Table 4 shows the mean standard scores on the subtests, and Table 5 shows the mean cluster scores. These data are for the entire sample, aggregated across centers. Appendix C1 contains the same information grouped by center. We conducted paired sample *t* –tests to determine whether children's scores on the individual subtests and the cluster tests changed from pre-test to post-test. For the individual subtests, analyses were conducted twice, once with raw scores and once with standard scores as dependent variables. For the Oral Language and Math Reasoning clusters, only standard scores were available for analysis. Appendix C 2 - x contains tables showing results of the statistical tests.

Table 2. Brief Description of WJ-III subtests

Subtests:	Task Description: Constructs Assessed
Letter-Word Identification	Identifying printed letters and words: Decoding
Story Recall	Listening to and recalling details of stories: Language development, listening ability, meaningful memory
Understanding Directions	Listening to a sequence of instructions and then following directions: listening comprehension
Spelling	Spelling orally presented words: Spelling
Applied Problems	Performing math calculations in response to orally presented problems: Math
Picture Vocabulary	Identifying objects: Oral expression
Oral Comprehension	Identifying a missing key word that makes sense in a passage: Listening comprehension
Academic Knowledge	Responding to questions about science, social studies, and humanities: General information
Quantitative Concepts	Identifying math terms and formulae; identifying number patterns: Math knowledge and quantitative reasoning

Note. The description of the tests and the constructs each assesses is taken from the Technical Manual published with the Woodcock Johnson Tests of Achievement-III.

Children's language, math, and content knowledge increased from pre-test to post-test, as indicated by significant results on all subtests when analyses used raw scores as the dependent variable (see Table 3 for means). Inferential statistics allow us to determine whether differences between two means differ from random fluctuation. However, they tell us little about the magnitude of differences. Effect sizes calibrate the size of an effect. We use Cohen's *d* to report effect sizes. Effect sizes of .20 or less are considered small effects, .50 moderate, and .80 strong. Effect sizes generally fell in the moderate to moderately strong range, with the exception of effects for the Understanding Directions and the Academic Knowledge subtests. Effect sizes on those tests were small.

Table 3. Woodcock Johnson Mean Pre- and Post-Test Raw Scores

Woodcock Subtests	Pre-Test Mean Raw Scores	Post-Test Mean Raw Scores
Letter Word Identification	10.56	13.51**
Story Recall	10.57	14.92**
Directions	15.48	17.33*
Spelling	9.23	10.80**
Applied Problems	11.98	14.02**
Picture Vocabulary	14.37	15.58*
Oral Comprehension	6.49	9.14**
Quantitative Concepts	8.75	11.77**
Academic Knowledge	26.21	27.70*

Table 4. Woodcock Johnson Mean Pre- and Post-Test Standard Scores

Woodcock Subtests	Pre-Test Mean Standard Scores	Post-Test Mean Standard Scores
Letter Word Identification	110.74	112.56
Story Recall	104.73	108.28
Directions	109.30	110.34
Spelling	106.76	107.91
Applied Problems	99.33	99.74
Picture Vocabulary	100.74	101.77
Oral Comprehension	103.01	109.28**
Quantitative Concepts	96.98	99.82*
Academic Knowledge	97.43	97.10

*Note. N = either 86 or 87 on all subtests, with the exception of the story recall subtest where N = 74. The lower N on the story recall subtest was due to several children not having standard scores computed because they had earned raw scores of 0. A more complete table with standard deviations and t-values in included in Appendix C. *Difference between pre-test and post-test mean standard scores is significant, p< .05. ** Difference between pre-test and post-test mean standard score is significant, p< .001.*

Table 5. Woodcock Johnson Mean Pre- and Post-Test Standard Cluster Scores

Woodcock Clusters	Pre-Test Mean Standard Scores	Post-Test Mean Standard Scores
Oral Language	104.09	107.44*
Math Reasoning	97.07	98.33

*Note. N = 85. The Oral Language Cluster consists of the following subtests: Story Recall, Understanding Directions, Picture Vocabulary, and Oral Comprehension. The Math Reasoning cluster consists of the Applied Problems and Quantitative Concepts subtests. *Post-Test Mean Standard Score is Significantly Different from Pre-Test Mean Standard Score, p <.001.*

Learning that children's knowledge increased between pre-testing in December/January and post-testing in June is good but does not show whether the amount or rate of growth is normative compared to other children the same age. To learn whether the amount of growth displayed by Core Knowledge children is comparable to other same age-peers, we turn to standard scores.

There are three possible patterns of interest from results using standard scores. A decrease in standard scores from pre-test to post-test shows that children are not gaining knowledge at a rate commensurate with their peers. No change from pre-test to post-test indicates that children are continuing to gain knowledge at a rate commensurate with their peers. An increase in standard scores from pre-test to post-test shows that children are gaining knowledge at a faster rate than same-age peers. Either the second or third pattern is good, with the third being preferable.

When standard scores were used as the dependent variables in the analyses, there were no significant decreases in children's scores from pre-test to post-test; in other words, they did not demonstrate the undesirable Pattern 1. Most of the mean standard scores increased slightly but not significantly from pre-test to post-test, consistent with Pattern 2. However, children's scores increased significantly on the quantitative concepts and the oral comprehension subtests and on the Oral Language cluster, consistent with Pattern 3. Effects sizes were small for the Oral Language cluster and the quantitative concepts subtest (*d* = .28, .20, respectively) and moderate for the oral comprehension subtest (*d* = .54).These results show that children receiving the Core Knowledge preschool sequence not only increased their skills from pre-test to post-test but did so at a faster rate than the standardization sample. This is particularly remarkable because the standardization sample was not limited to low income children but included middle income children as well.

*Supplementary analyses***.** Supplemental analyses considered whether children's scores on the Woodcock Johnson were related to gender or age. Many researchers have found differences related to gender in children's academic performance. Using the standard scores as the dependent variable, we conducted analyses of variance (ANOVAs) with gender as a between-subjects variable and each of the three cluster scores (academic knowledge, oral language, and math reasoning) as a repeated-

measures factor (pre-test and post-test). There was a significant main effect for gender on Academic Knowledge and Math Reasoning. Consistent with what has been shown in the literature, girls received higher scores than boys on Academic Knowledge (mean: girls 100.02, boys 93.62) and Math Reasoning (mean: girls 100.27, boys 94.21). Girls' scores were descriptively higher on Oral Language but not significantly so (mean: girls 107.42, boys 103.51). Interactions between the gender of the child and scores on the pre-tests and post-tests were not statistically significant, indicating that changes in performance from pre-test to post-test were not differentially related to the child's gender.

We also considered whether the child's age was related to his or her performance on the Woodcock Johnson. We used the raw scores for these analyses because standard scores are corrected for age. Age was significantly and positively correlated with all of the Woodcock Johnson pre-test and post-test scores except for the Letter-Word Identification and Understanding Directions subtests (post-test). Correlations ranged from .20 to .44. Not surprisingly, older children earned higher pre-test and post-test scores than younger children.

Performance on the Core Knowledge Preschool Assessment Tool

Description of the Assessment

The Core Knowledge Preschool Assessment Tool (CK-PAT) is an assessment instrument that is aligned directly to the goals and objectives of the Core Knowledge Preschool sequence. It was designed to measure and document the progress of children participating in the Core Knowledge program. The CK-PAT includes assessments in all of the domains that are part of the program, including Autonomy and Social Skills, Oral Language, Early Literacy Skills in Reading and Writing, Math, Science, Art, Music, and Movement. The CK-PAT, designed to be used by classroom teachers, provides instructions for assessing each program goal using direct observation, portfolio collection, or activity probes, and it suggests criteria to evaluate each child's performance. Performance is assessed in terms of three behavioral levels: (a) *Not yet*: child does not yet demonstrate this skill, knowledge, or behavior; (b) *Progressing*: child sometimes demonstrates this skill, knowledge, or behavior; and (c) *Competent*: child consistently and independently demonstrates this skill, knowledge or behavior.

Core Knowledge staff selected 17 tasks from the CK-PAT for us to administer as pre-tests and post-tests. In making the selections, the staff took into consideration the specific objectives and instructions of the tasks. Many of the CK-PAT tasks require that the assessor observe the children over a period of time and/or while interacting in groups. Because we would be doing one-on-one assessments during a single test session, only those tasks that could be appropriately administered under those conditions were selected. All tasks were designed to be used with 4-year-olds.

The tasks were either in the language arts or mathematics content areas. Within the language arts area, seven tasks (those with the LA designation) assessed oral

language competencies. Six of these tasks were administered in pairs, such that the same materials were used for both. In these tasks, children were asked to point to size, shape, or quantity words and were also asked to use size, shape or quantity words in response to the same materials. The seventh language arts task required children to identify emotions based on the facial expressions of children and adults shown in color photographs.

An additional six tasks within language arts (those with the WR designation) assessed emergent literacy competences. Children were asked to identify upper case and lower case letters and to write their name. Phonological awareness skills were also assessed, requiring children to identify the sounds of letters, beginning sounds of words, and rhyming words. One additional language arts task (with the SR designation) assessed storybook reading. Children were asked to arrange a set of pictures in sequence for a story that they were exposed to in the Core Knowledge program.

Within the mathematics content area, four tasks (with the KDMR designation) assessed mathematical competencies. Children were asked to identify numerals, name numerals, count groups of objects, and continue a sequence of color patterns (e.g., using tokens, place two red squares and one white square to extend the sequence of two red squares, one white square, two red squares, one white square, etc.).

Table 6 lists all of the tasks within each domain and includes the CK-PAT identification code for ease of reference to the actual materials. The numeral 4 preceding each task indicates that the task was intended for use with 4-year-olds.

Table 6. Tasks from the Core Knowledge Preschool Assessment Tool (CK-PAT) Administered in the Evaluation

1. Oral Language Tasks

4LDLAD1E	Use size words
4LDLAD1E1	Point to size words
4LDLAD1J	Use spatial words
4LDLAD1J1	Point to spatial words
4LDLAD1F	Use quantity words
4LDLAD1F1	Point to quantity words
4LDLAB03:	Identify Emotions

2. Emergent Literacy Tasks

4LDWRE4	Write name
4LDWR00	Rhyming word
4LDWRC7	Identify capital letters
4LDWRC7	Identify lower case letter
4LDWR01	Sound of letters
4LDWRC5	Beginning sounds

3. Storybook Reading Task

4LDSRB2	Sequence story pictures

4. Mathematics Tasks

4KDMRE05	Count groups
4KDMRE07	Name numerals
4KDMRE071	Write numerals
4KDMRB1	Continue color pattern

Training and standardization of testing procedures. The evaluation team standardized the instructions and scoring procedures in cases where it wasn't altogether clear how to administer and/or score the tasks. This ensured that all examiners were consistent in their approach. Members of the team practiced administering and scoring the task and brought questions to the group for discussion and resolution. All students who administered the tests participated in discussions and practice sessions before going out into the field. Included in Appendix A and B are the revised standardized instructions for administration and scoring (A), along with copies of the scoring sheets (B).

*Creation of composite domain scores***.** Although we provide data on children's performance on all of the tasks individually, it is unwise to read too much into how children perform on one single task. Therefore, consistent with established research procedures, we analyzed the data to see if the tasks in the different domains held together as scales. To do this, we first created the data base, entering ratings of "not yet" as 0, "progressing" as 1, and "competent" as 2. We then conducted reliability analyses and determined that the items did hold together well enough for us to create scale scores and to interpret them meaningfully. The index of internal consistency reliability is Cronbach's alpha. The higher the value, the better the items measure the same construct. Values are affected by the number of items in the scale, so reliability is lower with fewer items. To illustrate, the reliability of the entire set of tasks, all 17, was .87 on the pre-test and .81 on the post-test. Reliabilities of the subscales ranged from .63 to .76.

Composite mean scores were created by summing all of the scores for the individual items and then dividing the total by the number of items in the scale. This approach provides continuous integer values ranging from 0 to 2. Greater precision in statistical analysis is possible when there are finer distinctions among participants' performances. For example, with this approach, one child could have a mean score of 1.66 and another a mean of 1.93; when discrete values corresponding to the original scores of 0, 1, and 2 are used, both of these children would receive a score of 2). The mean scores were used in the statistical analyses.

In addition, composite readiness scores were calculated that allow us to describe children using the same performance indicators on which they were originally scored,

not yet, progressing, and competent. For the oral language domain, total scores of 0-3 were coded 0 (not yet), scores of 4-10 were coded 1 (progressing), and scores of 11-14 were coded 2 (competent). For the emergent literacy domain, total scores of 0-3 were coded 0, scores of 4-9 were coded 1, and scores of 10-12 were coded 2. For the mathematics domain, total scores of 0-2 were coded 0, scores of 3-6 were coded 1, and scores of 7-8 were coded 2. These scores are reported descriptively only.

Analyses of Change over Time on Children's Performance on the CK-PAT

Change over time, from pre-test to post-test, was examined for each of the separate tasks as well as for the composite scores using paired-samples t-tests. On every task, with one exception, children improved significantly over time. The one exception was the task requiring children to point to spatial words, where growth was not observed. Statistically significant improvements also occurred over time on each composite measure, oral language, emergent literacy, and mathematics. Most of the effect sizes were in the moderate range. Table 7 shows the mean composite scores on each domain at pre-test and post-test. Table 8 shows the mean scores on each of the individual tasks. See Appendix C for more detailed reporting of relevant statistics.

Table 7. Mean Composite Scores on the CK-PAT Domains

CKPAT Domains	Pre-Test Mean Composite Scores	Post-Test Mean Composite Scores
Oral Language	1.36	1.56 ***
Emergent Literacy	1.38	1.68 ***
Mathematics	1.14	1.54 ***

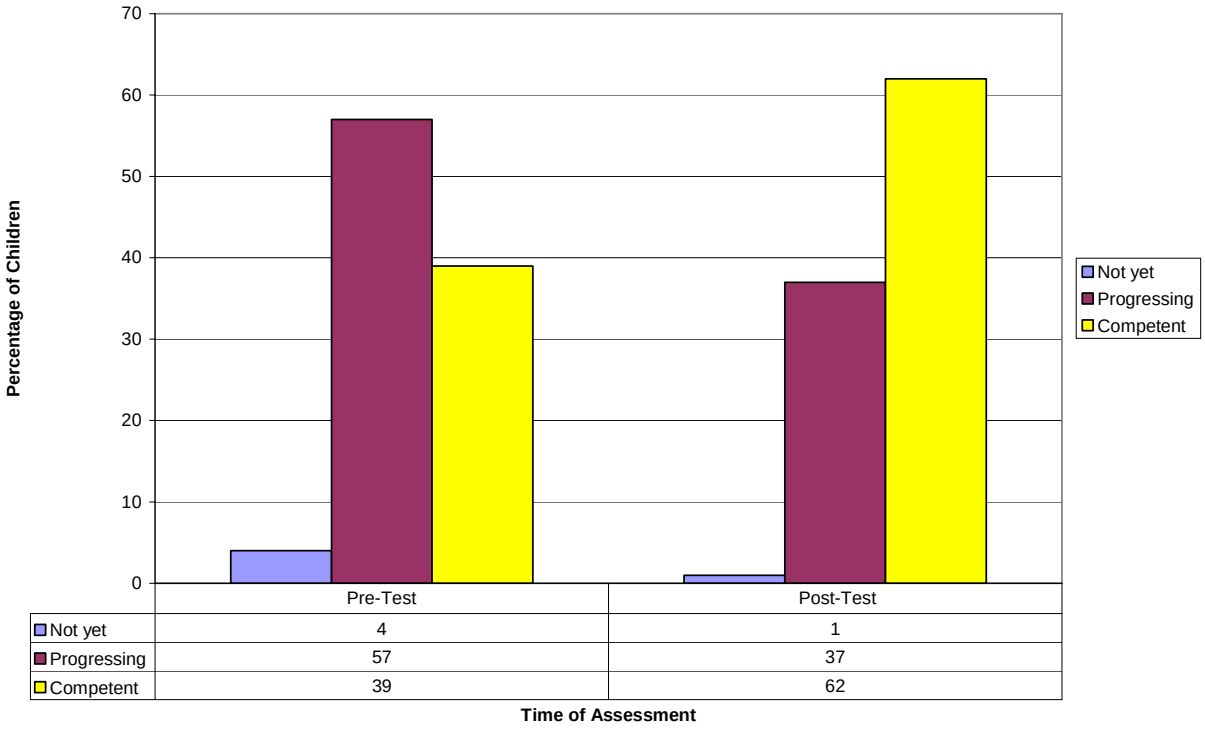
Note. N = 90 for oral language, and N = 97 for emergent literacy and mathematics.
*** $p < .001$*

Table 8. Mean Scores on Each CK-PAT Task

	Pre-Test	Post-Test
Oral Language		
Use size words	0.38	0.58
Point to size words	1.59	1.81
Use spatial words	0.98	1.27
Point to spatial words	1.88	1.91
Use quantity words	1.5	1.8
Point to quantity words	1.91	1.98
Identify emotions	1.28	1.46
Emergent Literacy		
Write name	1.6	1.81
Rhyming word	0.81	1.21
Identify capital letters	1.74	1.9
Identify lower case letters	1.44	1.85
Sound of letters	1.6	1.89
Beginning sounds	1.08	1.44
Storybook Reading		
Sequence story pictures	0.55	1.14
Mathematics		
Count groups	1.46	1.74
Name numerals	1.66	1.9
Write numerals	0.97	1.43
Continue color pattern	0.45	1.1

Figures1, 2, and 3 present the composite data in terms of readiness levels, rather than mean scores. What is shown is the percentage of children classified as *not yet*, *progressing*, and *competent* in each of the domains at pre-test and post-test. These visual representations reveal the patterns of transition, with more children initially classified at lower levels of readiness moving to higher levels of readiness by the end of the year. Statistical analyses on the readiness frequency data reveal the same pattern of growth over time as did the analyses on the mean scores.

Figure 1: Oral Language Readiness Levels



	Pre-Test	Post-Test
Not yet	4	1
Progressing	57	37
Competent	39	62

Figure 2: Composite Emergent Literacy Readiness Levels

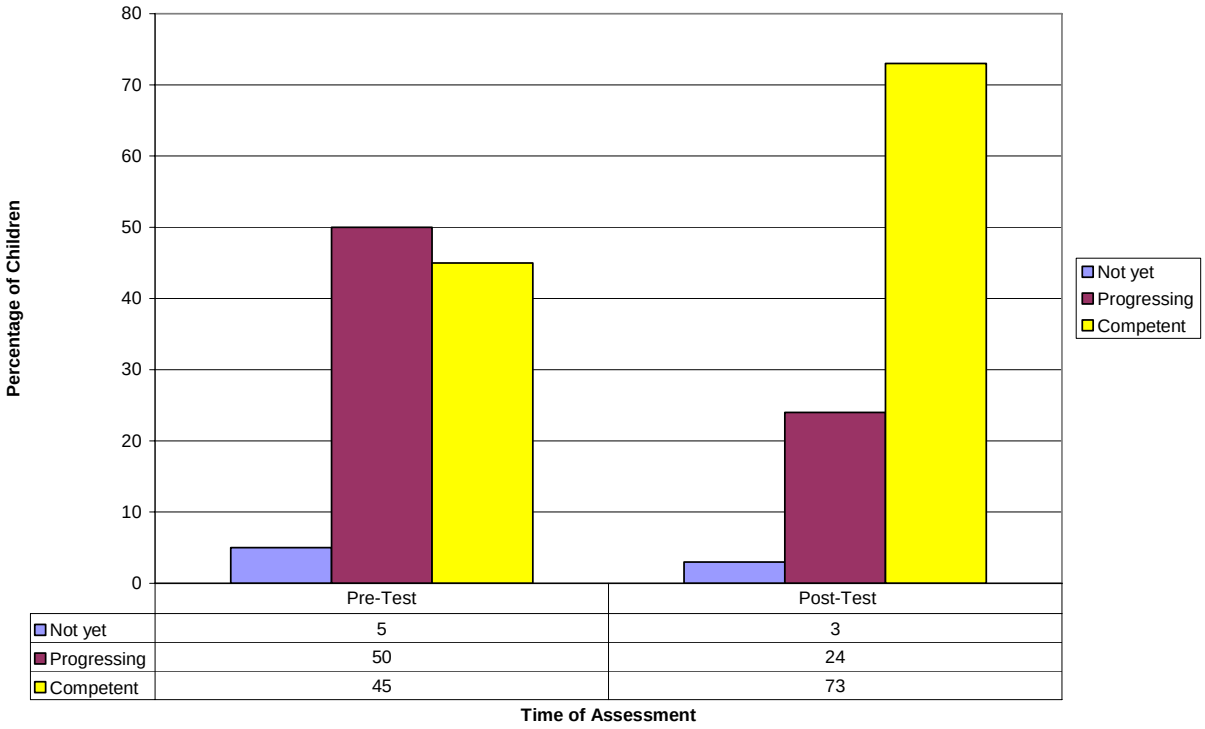
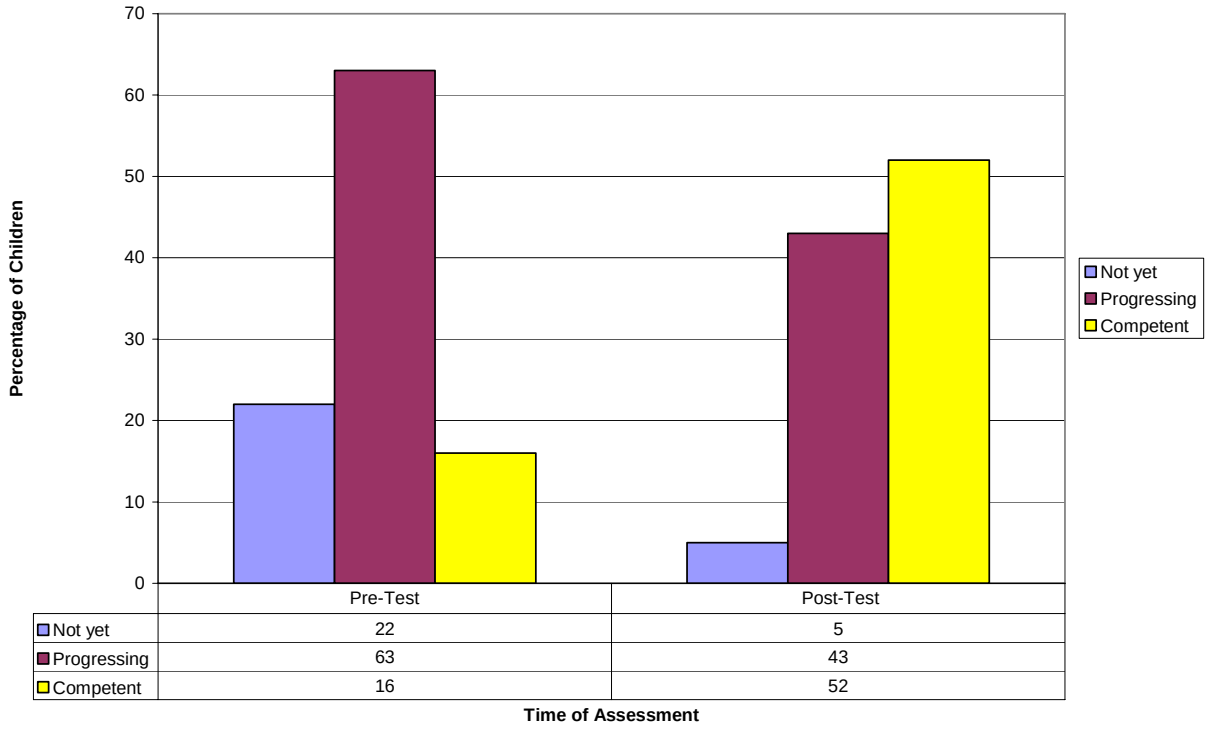


Figure 3: Composite Math Readiness Levels



Tables 9 through 12 show the readiness data on each of the CK-PAT tasks within each of the four domains (oral language, emergent literacy, mathematics, and storybook reading). Appendix E provides the readiness scores on the composite domains by center.

Table 9. Percentage of Children at Each Readiness Level on the Emergent Literacy Tasks.

Task		Not Yet	Progressing	Competent
Write name				
	Pre-Test	2.1	36.1	61.9
	Post-Test	0	18.6	81.4
Identify capital letters				
	Pre-Test	5.2	15.5	79.4
	Post-Test	1	8.2	90.7
Identify lower case letters				
	Pre-Test	12.4	30.9	56.7
	Post-Test	4.1	7.2	88.7
Beginning sounds				
	Pre-Test	35.1	21.6	43.3
	Post-test	21.6	12.4	66
Sound of letters				
	Pre-Test	15.5	9.3	75.3
	Post-Test	3.1	5.2	91.8
Rhyming word				
	Pre-Test	42.3	34	23.7
	Post-Test	23.7	32	44.3

Table 10. Percentage of Children at Each Readiness Level on the Oral Language Tasks.

Task		Not Yet	Progressing	Competent
Use size words				
	Pre-Test	64	34	2
	Post-Test	51	41	8
Point to size words				
	Pre-Test	3	65	62
	Post-Test	0	20	80
Use spatial words				
	Pre-Test	26	51	24
	Post-Test	14	43	42
Point to spatial words				
	Pre-Test	2	9	89
	Post-test	2	4	94
Use quantity words				
	Pre-Test	8	34	58
	Post-Test	2	16	83
Point to quantity words				
	Pre-Test	1	7	93
	Post-Test	0	2	98
Identify emotions				
	Pre-Test	7	58	35
	Post-Test	2	50	49

Table 11. Percentage of Children at Each Readiness Level on the Mathematics Tasks.

Task		Not Yet	Progressing	Competent
Count groups				
	Pre-Test	11	31	58
	Post-Test	3	20	77
Name numerals				
	Pre-Test	10	13	76
	Post-Test	3	4	93
Write numerals				
	Pre-Test	34	35	31
	Post-Test	17	24	60
Continue color pattern				
	Pre-Test	72	10	18
	Post-test	39	11	50

Table 12. Percentage of Children at Each Readiness Level on the Storybook Task

Task		Not Yet	Progressing	Competent
Sequence story pictures				
	Pre-Test	58	30	12
	Post-Test	31	24	45

*Supplementary analyses***.** Gender differences were examined on each of the CK-PAT tasks, as was the possibility that boys or girls showed greater growth over time. Analyses of variance revealed that on three tasks, girls outperformed boys overall (writing their name, using spatial words, and sequencing story pictures). On two tasks, boys showed greater growth over time than girls, having started out at lower levels (identifying capital letters and counting groups). On the three composite domains, girls' scores were generally higher than boys' scores, but the differences were not statistically significant.

Also examined were the effects of child's age on performance. As would be expected, older children performed significantly better in all domains than younger children, at both points in time.

Relations between Children's Performance on the Woodcock Johnson and the CK-PAT

Do the performance assessment tasks designed to measure specific skills and objectives of the Core Knowledge program relate to tasks that are part of a nationally-validated achievement test? This question is important because it provides evidence of validity of the CK-PAT, in this case concurrent validity. To address this question, we examined correlations between the Woodcock Johnson Oral Language Cluster and the CK-PAT Oral Language composite. The pre-test correlation was .51 and the post-test correlation was .41, both statistically significant at $p<.001$. We also examined the correlations between the Woodcock Johnson Mathematical Reasoning Cluster and the CK-PAT Mathematics composite. The pre-test correlation was .41 and the post-test correlation was .51, again both statistically significant at $p<.001$.

The Woodcock Johnson does not have a cluster score directly comparable to the Emergent Literacy composite that we created on the CK-PAT, so instead we examined correlations between two of the subtests, spelling and letter-word identification and the composite. Correlations with spelling and emergent literacy were .35 at pre-test and .56 at post-test; correlations with letter-word identification and emergent literacy were .50 at pre-test and .44 at post-test. All of these correlations were also statistically significant at the $p<.001$ level.

Correlations in the magnitude of .40 - .60 are generally considered of moderate size. Seven of the eight correlations were in this range, with one falling below that. Test developers interpret moderate correlations as evidence of validity. One would not expect two different tests to be strongly correlated unless they were designed to test identical skills under the same testing conditions. Clearly the CK-PAT was not designed to duplicate the Woodcock Johnson; rather, it was designed to assess whether children are attaining the skills that are the focus of the Core Knowledge curriculum. What the correlations do indicate is that the CK-PAT, at least when administered under the standardized conditions we adopted for the evaluation, provides valid information about children's early achievement.

Social Skills Rating System

Description of the Assessment

The Social Skills Rating System (SSRS; Gresham & Elliott, 1990) is a widely used rating system that provides information about student social behaviors that can affect teacher-student relations, peer acceptance, and academic performance. Different versions are available for preschool, elementary school, and secondary school. It has questionnaire forms for teachers, parents, and students (elementary and secondary only). The preschool version of the SSRS has two major scales, Social Skills and

Problem Behaviors. The Social Skills scale on the teacher questionnaire includes subscales for cooperation, assertiveness, and self control, and the Problem Behaviors scale has subscales for internalizing behaviors and externalizing behaviors.

The internal consistency reliability of the SSRS is strong, as reported by the test developers. On the teacher form of the preschool version, alpha coefficients were .94 for the Social Skills Scale and .82 for the Problem Behaviors scale. Norms are available based on a nationally representative sample of children. Because the social skills of boys and girls often differ quite dramatically, the norms were established separately for boys and for girls. The SSRS manual provides standard scores for each raw score. Standard scores relate one child's score to the scores of a comparison group. On the SSRS, the mean is 100, and the standard deviation is 15.

As part of the Core Knowledge evaluation, classroom teachers administered the SSRS to randomly selected children in their classrooms in the fall. We were provided with the rating sheets and entered the data into a data base for analysis and interpretation. Sufficiently complete data were available for 107 of the children, 59 girls and 48 boys. Using the sum of ratings on each scale, standard scores were obtained for each child.

Table 13 shows the mean scale scores for boys and girls separately on the Social Skills scale and the Problem Behaviors scale. As the figures show, children's social skills were within the normative "average" range for boys and girls at the preschool level on both scales. Statistical comparison of the standard scores showed that girls had significantly better social skills than boys, and boys had significantly more problem behaviors than girls.

Table 13. Mean Scale Scores on the SSSR

Social Skills		**Mean**
	Girls	100.58
	Boys	91.50
Problem Behaviors		
	Girls	94.66
	Boys	99.44

Relations between Children's Social Skills and Scores on the Woodcock Johnson

Research has shown that the social skills children exhibit in school are positively related to their learning availability and, in turn, to their academic progress (Raver, 2002). Consistent with such findings, we found that children's social skills, based on their scores on the social skills scale of the SSSR, were significantly and positively related to all raw and standard scores on the Woodcock Johnson pre-tests and post-tests, with the exception of the Understanding Directions subtest. Because the pattern with raw scores duplicated that with standard scores, we report only the results for

standard scores (see Table 14). Correlations generally were of small to moderate strength, ranging from .19 to .44. Similar analyses conducted with the Problem Behaviors scale were not statistically significant, showing that children's academic achievement was not related to behaviors such as fighting with others (externalizing) and acting sad (internalizing).

Table 14. Correlations Between Scores on the Social Skills Scale on the SSSR and the WJ-III Standard Scores

Subtest	***Pre-Test***	***Post-Test***
Letter Word ID	.32***	.30**
Story Recall	.27**	.29**
Understanding Directions	.17	.17
Spelling	.30**	.40***
Applied Problems	.29**	.35***
Picture Vocabulary	.19*	.22*
Oral Comprehension	.21*	.35***
Quantitative Concepts	.43***	.44***
Academic Knowledge	.32***	.27**
Oral Language Cluster	.27**	.38***
Math Reasoning Cluster	.26**	.44***

** p<.05, **p<.01, ***p<.001*

We next considered whether the social skills exhibited by children at the beginning of the year were related to the growth children displayed on the Woodcock Johnson tests, using the three cluster scores as indicators. Correlational analyses showed that children's social skills were significantly related to their growth on the oral language cluster (r = .25); this effect size is considered small. None of the other correlations reached statistical significance, either for the social skills scale or for the problem behaviors scale.

Relations between Children's Social Skills and the CK-PAT

Correlational analyses were conducted to determine if children's social skills and problem behaviors were related to their performance on the CK-PAT at pre-test and post-test. Table 15 shows the correlations among the measures. The data show that children with stronger social skills start out with higher performance levels than children with weaker social skills and that those initial social skills levels continue to be associated with their performance at the end of the year. In contrast, whether or not children exhibit problem behaviors in the classroom is unrelated to how well they perform on the various domains. None of the problem behavior correlations are significantly different from zero.

We also considered whether children's scores on the social skills or problem behaviors scales contributed to the amount of growth they made in the different domains. Analyses revealed that how children scored on these scales did not relate to how much growth they showed on any of the composites. (The largest correlation, between social skills and emergent literacy was .16, with a probability level of .11.)

Table 15. Correlations between the SSRS Scales and the CK-PAT Composites

	Social Skills	Problem Behaviors
Mathematics		
Pre-test	.45***	-.05
Post-test	.43***	-.09
Oral Language		
Pre-test	.46***	-.10
Post-test	.38***	-.19
Emergent Literacy		
Pre-test	.57***	.00
Post-test	.49***	-.07

*** $p<.001$

Summary and Conclusions

Results of analyses with scores on the Woodcock Johnson show that children receiving the Core Knowledge Preschool Sequence showed significant and moderate to moderately strong increases in performance from pre-test to post-test (based on raw scores). Children's growth in knowledge was comparable to the Woodcock Johnson standardization group, which consisted of children of different income levels and ethnicities comparable in age to the Core Knowledge children. Children receiving Core

Knowledge instruction actually showed more growth compared to the standardization sample on a few of the subtests (Oral Comprehension, Quantitative Concepts, Oral Language Cluster). These results are particularly impressive given that the amount of time between the pre-tests and post-tests was really quite short, ranging from 3 to 5 months.

It is important to note that children's mean standard scores on the Woodcock Johnson fell in the average range at both pre-test and post-test compared to same-age peers. We cannot say whether this is a testing artifact due to the pre-test being administered in the middle of the second year of what is probably a 2-year program for many of the children or the test not being maximally sensitive for children at the youngest ages in the allowed age span. However, we think the focus of interpretation should be on the comparison between pre-tests and post-tests and not on the initial score levels of the pre-tests.

Using an instrument that is aligned directly to the goals and objectives of the Core Knowledge Preschool Sequence, the CK-PAT, we found that children gained significantly on 16 of the 17 individual tasks from pre-test to post-test and that they showed significant growth on the three composite areas of oral language, emergent literacy, and mathematics. Thus, these results indicate that children made progress in the specific objectives of the program.

It is usually much easier to demonstrate growth when the measure is directly aligned to the program, as in the case of the CK-PAT, than it is when a more generic measure of academic achievement is used, as in the case of the Woodcock Johnson. The fact that significant progress was demonstrated with both instruments indicates that Core Knowledge is providing children with the skills and knowledge that children of their age across the country are expected to master. The moderate correlations between the two instruments indicate that the assessment tool itself has validity as a measure of children's achievement in the program.

References

Gresham, F. M., & Elliott, S. N. (1990). *Social skills rating system*. Circle Pines: MN: American Guidance Service.

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), 3-18.

Sattler, J. M. (2001). *Assessment of children: Cognitive Applications,* 4th edition. LMesa:CA: Jerome Sattler.

Appendix A

INSTRUCTIONS FOR CORE KNOWLEDGE ASSESSMENTS

4LD-WR-E4
Need to include a blank piece of paper. The kit should contain a "child" size pencil.
Part 1: Writing name
Instructions: "Here's a piece of paper and a pencil. Please write your name." If the child says, "I don't know how", say, "Can you write any letters? Just do the best you can. If the child still does not attempt the task after you have urged him/her 1 time, then just go on to the next part of the task.

Part 2: Copying Name
This part is to be done ONLY if child does not attempt to write his/her name OR is not able to correctly write one letter of the name. In that case, print the child's name (first letter upper-case, remainder lower-case). **Say**, "I have written your name (point to it). I want you to copy your name here (point lower down on the page)." If child does not attempt, urge child to try, "Can you copy one of the letters?"

Scoring:
NY: Unable to copy or write any letter or name
P: Copies or writes at least one letter from name (ok if letter is reversed)
R: Independently writes first name (upper- or lower-case letters ok, 1 reversed letter ok).

4LD-WR-01
4LD-WR-C7
Note. We are combining the presentation of the two tasks; however, they need to be scored separately. That is, score the upper-case and lower-case letters separately.
Instructions:
Upper-case letters: Take the letters of the child's name and "randomly" arrange them so that first letter of the name is first on left, followed by last letter. Others are randomly arranged. Point to the letters in order from left to right.
Given the letters we have, use each letter in the child's name only once. That means in some cases, the child's name will not be completed spelled out.

Point to the first letter in the child's name (should be on the left from the child's view). "What letter is this?" Proceed to ask what each letter in the child's name is going from left to right. After completing the child's name, ask about what sounds each letter in the child's name makes. Point to first letter in name, "What sound does this letter make?" If child repeats, the name of the letter, "That is the name of the letter. Now, can you tell me its sound?" Continue through from left to right the remaining letters in the child's name.

After completing the sounds of the letters in the child's name, randomly arrange the remaining letters. Point to them. "Look at all these letters. Do you know what any of these letters are? (If so, ask the child to name as many as possible). After a child labels a letter ask, "What sound does this letter make?" When a child can no longer label any more letters, "Do you know what sound any of these other letters makes?"

After completing this section of the task, go back and ask the sounds for any letter not correctly identified. That is, reshow letter and ask what sound it makes.

Lower-case letters. Repeat the above with lower-case letters. This time totally randomly arrange letters in child's name. Make sure the first letter in the child's name is NOT the one on the left.

4KD-MR-E07
Note: Contrary to what the task requires, it makes more sense to first have child label each number then copy them. Need unlined paper and pencil. You need to provide the paper, pencil should be in the kit.

Instructions: Order of cards: 2, 4, 6, 1, 3, 5

Naming task. Point to the number "2". "What number is this?". Continue for each of the remaining numbers.

Copying task: Use same ordering of numbers. Point to the number "2". "Look at this number. I want you to write the number on your paper (point to the paper)." Repeat for each of the 6 numbers. Remove the writing paper.

Score the naming and copying tasks separately.

4LD-LA-D1F
4LD-LA-D1F1
NOTE. If child gets 9 quantity words wrong discontinue and proceed to 4LD-LA-D1F1. If child scores at ready range on 4LD-LA-D1F do not administer the 4LD-LA-D1F1 task but credit it as ready.
Instructions: Organize as a cluster of groups. Don't have child look while putting together.

4KD-MR-E05
Instructions:

Order of presentation of blocks:

a. 2 - - horizontal line
b. 5 - - dice pattern
c. 3 - - vertical line
d. 6 - - throw them out
e. 4 - - square
f. 5 - -random
g. 6 - - 2 horizontal lines
h. 2- - vertical line

After displaying blocks, ask child. “How many blocks are there?” Child is allowed to use fingers to count. However, if child counts and then does not indicate the total number, ask child, “So how many are there?”

4LD-LA-DIE
4LD-LA-DIE1
Note. If child misidentifies 9 size words, go to 4LD-LA-D1E1. If child receives ready rating on 4LD-LA-DIE, skip 4LD-LA-D1E1 but credit with rating of “ready” on that task.
Instructions: Follow instructions as written through the pictures of the sofa and feather.

String: take the 2 middle length strings. For horizontal task, left justify the strings with shorter one on top. Point to the one on top, “This string is ….” And this string is…..” If necessary, use probe. “Tell me about their length, how long they are. This string is (point to top) and this one is (point to bottom one)…..”

Rearrange vertically with shorter one on left. “When I put the yellow string this way, this one is short and this one is ….(point to the tall string)”

Blocks: Put the thin one of left and the thick one on right. Child is allowed to pick up blocks, if wants.

4-LD-LA-D1E1
Instructions: Arrange in following order as in a W:
Empty basket, bigger rock, feather, sofa, full basket, other rock, empty basket.

“Point to the picture that shows something large.” (Note. This is slightly different wording than on the instruction sheets. It seemed to me that their wording may confuse the child and therefore not get at the actual concept.)

Arrange string as in prior task. Take the 2 middle length strings. For horizontal task, left justify the strings with shorter one on top. Point to the one that is short. Now point to the one that is long.”

Rearrange strings vertically with shorter one on left. “Point to the one that is tall. Point to the one that is short.”

Blocks: Blocks: Put the thin one of left and the thick one on right, row one. Reverse order for row 2. “Point to the thin one, wide one, narrow one, thick one.”

4LD-LA-D1J
4LD-LA-D1J1
Note. A child who scores at the ready range for 4LD-LA-D1J should not be given the 4LD-LA-D1J1 task but credited with a ready on that.

Instructions: Follow written instructions. Cover all but the row being tested. Note that row 5 does not have a picture representing “inside”.

For 4LD-LA-D1J1, administer as recognition task . “Point to the one that….” Administer in same order. Follow instructions on sheet.

4LD-LA-B03
Instructions: Do sheet 1, sheet 2, sheet 3. Sheets have number and emotion label on back.

4LD-SR-B2
Instructions:
Goldilocks story: Sequence 3, 2, 1, 5, 4, 6
Gingerbread: Sequence: 4, 3, 1, 5, 2

4KD-MR-B1
Instructions: Arrange the blocks of two colors so that they are intermingled.
The pattern cards should be placed horizontal position with the writing to the child’s left. Point to the right side for the child to add onto the pattern. In order to be correct, the child needs 1 repeat of the pattern.

4LD-WR-C5
Instructions. Go through instructions, one picture at a time (top left, top right, etc.)
Sheet 1, 2 3, (number of sheet is on back)
Sheet 1. Fork, pumpkin, table, sink
Sheet 2. banana, leaf, cat baby
Sheet 3. mitten, car, scissors, pan

4LD-WR-OO
Instructions. Label each picture for the child.

Sheet 1. Mop (target), clap, milk, hop
Sheet 2. hen (target), pen, hat, can
Sheet 3. Truck (target), tree, duck, back
Sheet 4. bee (target), baby, boat, tree
Sheet 5. star (target), car, sign, cat
Sheet 6. cat (target), can, hat, boat

Appendix B
Core Knowledge BCHS Level II UMBC Assessments
Student Response Sheet

Child's Name_____ **ID#** _____
Center/Teacher_____ **AM PM FULL**
Examiner's Name _____ **Date** _____

Skill Code	Score	Description	Notes
4LD-WR-E4 Green	NY P R	Write one's first name	
4LD-WR-01 Green	NY P R	Give the sound of at least 3 letters (see next page too)	
4LD-WR-C7 Green	NY P R	Identify at least 10 capital letters (see next page too)	
4LD-WR-C7 Green	NY P R	Identify at least 10 lower case letters (see next page too)	
4KD-MR-E07 Orange	NY P R	Name numerals 1-6	
4KD-MR-E07 Orange	NY P R	Write numerals 1-6	
4LD-LA-D1F Grey	NY P R	Use quantity words	
4LD-LA-D1F1 Beige	NY P R	Point to quantity words	
4KD-MR-E05 Orange	NY P R	Count groups of objects -6	
4LD-LA-D1E Grey	NY P R	Use size words	
4LD-LA-D1E1 Beige	NY P R	Point to size words	
4LD-LA-D1J Grey	NY P R	Use spatial words	
4LD-LA-D1J1 Beige	NY P R	Point to spatial words	
4LD-LA-B03 Yellow	NY P R	Identify emotions	
4LD-SR-B2 Blue	NY P R	Sequence story illustrations	
4KD-MR-B1 Orange	NY P R	Continue a 2 color pattern	
4LD-WR-C5 Green	NY P R	Give the beginning sound of a word	
4LD-WR-00 Green	NY P R	Provide a rhyming word	

4LD-WR-E4

NY: Unable to copy or write any letter or name
P: Copies or writes at least one letter from name (ok if letter is reversed)
R: Independently writes first name (upper- or lower-case letters ok, 1 reversed letter ok).

Appendix C

Supplementary Analyses and Detailed Statistical Reports

1. Woodcock Johnson Mean Pre- and Post-Test Scores by Center

Back River Head Start Center, N = 7

Woodcock Subtests	Pre-Test Mean Raw Scores	Post-Test Mean Raw Scores
Letter Word Identification	12.00	16.14
Story Recall	13.14	13.14
Directions	11.14	9.57
Spelling	10.43	15.00
Applied Problems	12.14	17.00
Picture Vocabulary	13.86	16.71
Oral Comprehension	6.57	13.14
Quantitative Concepts	9.43	14.43
Academic Knowledge	27.86	24.00

Woodcock Subtests	Pre-Test Mean Standard Scores	Post-Test Mean Standard Scores
Letter Word Identification	112.14	115.00
Story Recall	108.43	111.29
Directions	96.43	98.00
Spelling	109.57	121.29
Applied Problems	96.57	104.86
Picture Vocabulary	97.86	104.14
Oral Comprehension	101.29	121.86
Quantitative Concepts	97.57	104.71
Academic Knowledge	98.57	85.71
Woodcock Clusters		
Oral Language	98.86	109.43
Math Reasoning	95.71	104.29

Campfield Head Start Center, N=16

Woodcock Subtests	Pre-Test Mean Raw Scores	Post-Test Mean Raw Scores
Letter Word Identification	8.75	11.94
Story Recall	8.25	14.69
Directions	16.69	14.25
Spelling	7.80	9.53
Applied Problems	10.13	12.06
Picture Vocabulary	13.69	14.69
Oral Comprehension	4.94	8.44
Quantitative Concepts	6.75	10.75
Academic Knowledge	24.75	26.75

Woodcock Subtests	Pre-Test Mean Standard Scores	Post-Test Mean Standard Scores
Letter Word Identification	105.38	110.00
Story Recall	102.83	107.75
Directions	110.63	108.44
Spelling	100.47	104.00
Applied Problems	93.31	94.38
Picture Vocabulary	99.06	98.94
Oral Comprehension	97.93	108.13
Quantitative Concepts	92.31	95.00
Academic Knowledge	93.81	96.44
Woodcock Clusters		
Oral Language	101.13	104.81
Math Reasoning	92.19	92.56

Chase Head Start Center, N = 4

Woodcock Subtests	Pre-Test Mean Raw Scores	Post-Test Mean Raw Scores
Letter Word Identification	11.75	9.75
Story Recall	10.75	15.00
Directions	10.00	8.50
Spelling	9.50	11.25
Applied Problems	14.75	11.00
Picture Vocabulary	15.75	14.00
Oral Comprehension	6.00	6.00
Quantitative Concepts	10.50	10.25
Academic Knowledge	26.75	21.75

Woodcock Subtests	Pre-Test Mean Standard Scores	Post-Test Mean Standard Scores
Letter Word Identification	115.50	105.50
Story Recall	109.50	108.75
Directions	98.25	97.75
Spelling	109.25	112.00
Applied Problems	108.75	91.75
Picture Vocabulary	105.75	96.00
Oral Comprehension	101.75	99.25
Quantitative Concepts	104.25	97.50
Academic Knowledge	99.50	82.75
Woodcock Clusters		
Oral Language	103.25	98.50
Math Reasoning	107.00	92.50

Emily Harris Head Start Center, N =15

Woodcock Subtests	Pre-Test Mean Raw Scores	Post-Test Mean Raw Scores
Letter Word Identification	10.86	14.53
Story Recall	12.13	17.67
Directions	14.47	22.27
Spelling	10.47	10.73
Applied Problems	12.40	14.27
Picture Vocabulary	15.01	16.13
Oral Comprehension	8.40	9.80
Quantitative Concepts	11.07	10.73
Academic Knowledge	26.93	28.53

Woodcock Subtests	Pre-Test Mean Standard Scores	Post-Test Mean Standard Scores
Letter Word Identification	113.00	116.20
Story Recall	102.71	110.86
Directions	110.87	114.33
Spelling	113.60	108.87
Applied Problems	104.64	103.36
Picture Vocabulary	103.60	104.87
Oral Comprehension	110.67	111.40
Quantitative Concepts	105.40	99.13
Academic Knowledge	100.93	99.13
Woodcock Clusters		
Oral Language	107.93	111.27
Math Reasoning	101.20	99.33

Fleming Head Start Center, N = 13

Woodcock Subtests	Pre-Test Mean Raw Scores	Post-Test Mean Raw Scores
Letter Word Identification	11.54	13.23
Story Recall	7.62	11.31
Directions	15.54	17.85
Spelling	8.46	9.54
Applied Problems	11.46	14.31
Picture Vocabulary	13.85	15.62
Oral Comprehension	5.69	7.92
Quantitative Concepts	8.46	11.54
Academic Knowledge	26.08	28.31

Woodcock Subtests	Pre-Test Mean Standard Scores	Post-Test Mean Standard Scores
Letter Word Identification	114.92	112.08
Story Recall	104.56	103.56
Directions	111.92	112.00
Spelling	104.00	101.23
Applied Problems	97.77	100.38
Picture Vocabulary	99.00	101.85
Oral Comprehension	100.62	104.92
Quantitative Concepts	96.08	99.92
Academic Knowledge	96.62	98.15
Woodcock Clusters		
Oral Language	104.62	105.38
Math Reasoning	99.31	99.46

Merrit Park Head Start Center, N =6

Woodcock Subtests	Pre-Test Mean Raw Scores	Post-Test Mean Raw Scores
Letter Word Identification	9.17	12.50
Story Recall	16.67	14.33
Directions	15.17	17.83
Spelling	7.50	9.17
Applied Problems	11.00	13.33
Picture Vocabulary	17.20	16.40
Oral Comprehension	8.83	9.00
Quantitative Concepts	7.00	8.00
Academic Knowledge	31.00	29.83

Woodcock Subtests	Pre-Test Mean Standard Scores	Post-Test Mean Standard Scores
Letter Word Identification	108.00	113.83
Story Recall	108.67	101.50
Directions	111.50	112.50
Spelling	100.83	105.00
Applied Problems	98.83	100.67
Picture Vocabulary	112.00	105.80
Oral Comprehension	113.00	110.33
Quantitative Concepts	93.83	90.00
Academic Knowledge	113.33	104.83
Woodcock Clusters		
Oral Language	113.67	109.00
Math Reasoning	95.00	94.50

Reistertown Head Start Center, N = 4

Woodcock Subtests	Pre-Test Mean Raw Scores	Post-Test Mean Raw Scores
Letter Word Identification	12.75	18.50
Story Recall	7.50	22.25
Directions	17.50	25.25
Spelling	10.25	12.50
Applied Problems	12.00	17.25
Picture Vocabulary	15.50	16.75
Oral Comprehension	8.75	11.00
Quantitative Concepts	11.75	15.25
Academic Knowledge	29.00	32.00

Woodcock Subtests	Pre-Test Mean Standard Scores	Post-Test Mean Standard Scores
Letter Word Identification	120.25	122.75
Story Recall	110.50	124.50
Directions	118.00	115.75
Spelling	114.50	113.50
Applied Problems	100.25	108.00
Picture Vocabulary	104.50	105.50
Oral Comprehension	113.00	114.00
Quantitative Concepts	106.00	110.75
Academic Knowledge	107.50	106.75
Woodcock Clusters		
Oral Language	114.33	116.00
Math Reasoning	108.00	109.00

Riverview Head Start Center, N = 7

Woodcock Subtests	Pre-Test Mean Raw Scores	Post-Test Mean Raw Scores
Letter Word Identification	11.71	14.43
Story Recall	15.43	16.29
Directions	20.71	17.14
Spelling	10.57	12.29
Applied Problems	14.71	15.71
Picture Vocabulary	14.43	16.00
Oral Comprehension	7.43	10.29
Quantitative Concepts	9.71	15.43
Academic Knowledge	27.00	29.00

Woodcock Subtests	Pre-Test Mean Standard Scores	Post-Test Mean Standard Scores
Letter Word Identification	113.00	112.43
Story Recall	110.43	106.43
Directions	114.29	111.00
Spelling	111.29	112.00
Applied Problems	106.00	102.29
Picture Vocabulary	99.86	101.71
Oral Comprehension	104.86	112.43
Quantitative Concepts	98.57	110.00
Academic Knowledge	97.50	99.17
Woodcock Clusters		
Oral Language	108.00	109.50
Math Reasoning	103.17	106.67

Towson Head Start Center, N = 5

Woodcock Subtests	Pre-Test Mean Raw Scores	Post-Test Mean Raw Scores
Letter Word Identification	8.80	11.00
Story Recall	16.00	17.40
Directions	20.20	19.80
Spelling	9.60	10.60
Applied Problems	14.20	15.00
Picture Vocabulary	14.40	14.00
Oral Comprehension	5.80	9.40
Quantitative Concepts	8.60	12.60
Academic Knowledge	27.00	28.40

Woodcock Subtests	Pre-Test Mean Standard Scores	Post-Test Mean Standard Scores
Letter Word Identification	101.20	102.60
Story Recall	108.80	113.20
Directions	112.00	114.80
Spelling	103.40	104.20
Applied Problems	101.20	98.40
Picture Vocabulary	98.80	94.60
Oral Comprehension	97.40	108.00
Quantitative Concepts	92.40	96.60
Academic Knowledge	95.20	96.40
Woodcock Clusters		
Oral Language	103.60	107.60
Math Reasoning	96.40	96.00

WhiteMarsh Head Start Center, N = 10

Woodcock Subtests	Pre-Test Mean Raw Scores	Post-Test Mean Raw Scores
Letter Word Identification	10.30	13.70
Story Recall	5.40	12.30
Directions	13.60	18.60
Spelling	9.10	10.70
Applied Problems	11.30	13.00
Picture Vocabulary	13.00	15.60
Oral Comprehension	4.70	7.70
Quantitative Concepts	6.90	11.90
Academic Knowledge	21.40	28.00

Woodcock Subtests	Pre-Test Mean Standard Scores	Post-Test Mean Standard Scores
Letter Word Identification	108.60	113.20
Story Recall	93.75	106.63
Directions	105.20	112.80
Spelling	105.50	108.50
Applied Problems	96.00	97.10
Picture Vocabulary	95.90	102.30
Oral Comprehension	95.50	104.60
Quantitative Concepts	89.10	101.90
Academic Knowledge	84.90	98.10
Woodcock Clusters		
Oral Language	95.50	105.90
Math Reasoning	91.30	98.00

2. T-tests Comparing WJ-III Pre-Tests and Post-Tests

Raw Scores

Subtests	**Pre-Test Mean (SD)**	**Post-Test Mean (SD)**	**df**	***t*-value**	***p*-value**
Letter Word ID	10.56 (4.02)	13.51 (4.51)	86	8.39	.001
Story Recall	10.57 (9.85)	14.92 (11.32)	86	4.41	.001
Understanding Directions	15.48 (6.90)	17.33 (7.40)	86	2.26	.026
Spelling	9.23 (2.97)	10.80 (3.40)	85	5.01	.001
Applied Problems	11.98 (3.70)	14.02 (3.85)	86	5.28	.001
Picture Vocabulary	14.37 (2.60)	15.58 (2.32)	85	4.74	.001
Oral Comprehension	6.49 (3.35)	9.14 (3.32)	86	8.05	.001
Quantitative Concepts	8.75 (4.30)	11.77 (3.99)	86	7.36	.001
Academic Knowledge	26.21 (4.90)	27.70 (4.57)	85	2.61	.011

Standard Scores

Subtests	**Pre-Test Mean (SD)**	**Post-Test Mean (SD)**	**Df**	***t*-value**	***p*-value**
Letter Word ID	110.74 (12.19)	112.56 (10.32)	86	1.54	ns
Story Recall	104.73 (14.15)	108.28 (15.78)	73	1.85	ns
Understanding Directions	109.30 (13.54)	110.35 (10.73)	86	.83	ns
Spelling	106.76 (13.44)	107.91 (14.15)	85	.71	ns
Applied Problems	99.33 (10.17)	99.74 (10.50)	85	.35	ns
Picture Vocabulary	100.74 (8.61)	101.77 (7.79)	85	1.15	ns
Oral Comprehension	103.01 (12.47)	109.28 (10.86)	85	4.77	.001
Quantitative Concepts	96.98 (14.13)	99.82 (12.61)	86	2.07	.041
Academic Knowledge	97.43 (13.02)	97.43 (11.31)	85	.21	ns
Math Reasoning Cluster	97.07 (13.32)	98.33 (12.11)	84	1.11	ns
Oral Language Cluster	104.09 (11.72)	107.44 (9.41)	84	3.53	.001

3.. Gender and WJ-III Standard Scores: Analyses of Variance

Clusters	**df**	***F*-value**	***p*-value**
Academic Knowledge	1,84	9.62	.003
Math Reasoning	1,83	5.99	.017

4. Correlations (*r*) of Age with WJ-III Raw Scores:

	Pre-Test N=110	**Post-Test** N=87
Subtest:		
Letter Word Id	.18	.26
Story Recall	.29	.37
Understanding Directions	.24	.08
Spelling	.33	.35
Applied Problems	.30	.35
Picture Vocabulary	.26	.25
Oral Comprehension	.21	.44
Quantitative Concepts	.20	.26
Academic Knowledge	.26	.25

Note. All correlations were statistically significant ($p < .05$ or less) with the exceptions of Letter Word ID and Understanding Directions.

5. Means, Standard Deviations, and Statistical Test Results for Each CK-PAT Task

		Pre-Test		Post-Test				
Oral Language	**N**	**Mean**	***SD***	**Mean**	***SD***	***t***	***p***	***d***
Use size words	97	0.38	*0.53*	0.58	*0.64*	3.02	*0.003*	*0.31*
Point to size words	95	1.59	*0.56*	1.81	*0.39*	3.37	*0.001*	*0.34*
Use spatial words	97	0.98	*0.71*	1.27	*0.71*	3.73	*<.001*	*0.38*
Point to spatial words	94	1.88	*0.39*	1.91	*0.35*	0.94	*not sig.*	*0.09*
Use quantity words	97	1.5	*0.65*	1.8	*0.45*	5.39	*<.001*	*0.55*
Point to quantity words	93	1.91	*0.32*	1.98	*0.15*	2.16	*0.033*	*0.21*
Identify emotions	97	1.28	*0.59*	1.46	*0.54*	3.13	*0.002*	*0.32*
Emergent Literacy								
Write name	97	1.6	*0.53*	1.81	*0.39*	4.62	*<.001*	*0.46*
Rhyming word	97	0.81	*0.79*	1.21	*0.8*	4.31	*<.001*	*0.44*
Identify capital letters	97	1.74	*0.55*	1.9	*0.34*	3.9	*<.001*	*0.4*
Identify lower case letters	97	1.44	*0.71*	1.85	*0.46*	6.53	*<.001*	*0.66*
Sound of letters	97	1.6	*0.75*	1.89	*0.41*	4.2	*<.001*	*0.43*
Beginning sounds	97	1.08	*0.89*	1.44	*0.83*	5.11	*<.001*	*0.52*
Storybook Reading								
Sequence story pictures	97	0.55	*0.71*	1.14	*0.87*	6.93	*<.001*	*0.7*
Mathematics								
Count groups	97	1.46	*0.69*	1.74	*0.51*	4.64	*<.001*	*0.47*
Name numerals	97	1.66	*0.66*	1.9	*0.39*	4.53	*<.001*	*0.46*
Write numerals	97	0.97	*0.81*	1.43	*0.76*	6.45	*<.001*	*0.66*
Continue color pattern	97	0.45	*0.78*	1.1	*0.94*	6.92	*<.001*	*0.7*

6. Percentages of Children at each Readiness Status on Pretest and Post-Test Domain Composites by Center

Center	N	Readiness	Mathematics		Oral Language		Emergent Literacy	
			Pre-Test	*Post-Test*	*Pre-Test*	*Post-Test*	*Pre-Test*	*Post-Test*
Back River	14	Not yet	14	0	7	0	0	0
		Progressing	71	57	43	14	57	14
		Competent	14	43	50	86	43	86
Campfield	15	Not yet	27	0	13	0	7	0
		Progressing	60	40	47	33	53	20
		Competent	13	60	40	67	40	80
Chase	4	Not yet	0	0	0	0	0	0
		Progressing	100	100	0	0	75	0
		Competent	0	0	100	100	25	100
Emily Harris	15	Not yet	13	0	7	7	7	0
		Progressing	40	40	47	33	53	33
		Competent	47	60	47	60	40	67
Fleming	14	Not yet	21	7	0	0	7	7
		Progressing	79	36	71	43	36	21
		Competent	0	57	29	57	57	71

Center	N	Readiness	Mathematics		Oral Language		Emergent Literacy	
			Pre-Test	*Post-Test*	*Pre-Test*	*Post-Test*	*Pre-Test*	*Post-Test*
Merrit Park	6	Not yet	67	33	0	0	17	17
		Progressing	17	50	83	83	50	50
		Competent	17	17	17	17	33	33
Reisterstown	5	Not yet	20	0	0	0	0	0
		Progressing	60	40	60	60	0	0
		Competent	20	60	40	40	100	100
Riverview	8	Not yet	88	0	0	0	0	0
		Progressing	13	38	50	25	50	0
		Competent	0	63	50	75	50	100
Towson	5	Not yet	20	0	0	0	0	0
		Progressing	80	40	100	60	80	60
		Competent	0	60	0	40	20	40
White Marsh	11	Not yet	36	18	0	0	9	9
		Progressing	55	27	82	64	46	36
		Competent	9	54	18	36	46	55

7. Mean Scale Scores on the SSSR, with Standard Deviations and Statistics

Social Skills		**Mean**	**Standard deviation**	**Statistical test**
	Girls	100.58	15.59	t(105)=3.03, p=.003
	Boys	91.50	15.19	
Problem Behaviors				
	Girls	94.66	10.08	t(105) = -2.15, p=.034
	Boys	99.44	12.92	

8. Correlations Between Scores on Social Skills Scale on SSSR and WJ-III Pre-Test Standard Scores

Subtest	**Df**	***r*-value**	***p*-value**
Letter Word ID	105	.32	.001
Story Recall	93	.27	.008
Understanding Directions	105	.17	Ns
Spelling	105	.30	.002
Applied Problems	104	.29	.003
Picture Vocabulary	105	.19	.052
Oral Comprehension	104	.21	.032
Quantitative Concepts	105	.43	.001
Academic Knowledge	105	.32	.001
Oral Language Cluster	105	.27	.005
Math Reasoning Cluster	105	.26	.006

9. Correlations Between Scores on Social Skills Scale on SSSR and WJ-III Post-Test Standard Scores

Subtest	**Df**	***r*-value**	***p*-value**
Letter Word ID	83	.30	.005
Story Recall	78	.29	.010
Understanding Directions	83	.17	Ns
Spelling	82	.40	.001
Applied Problems	83	.35	.001
Picture Vocabulary	83	.22	.044
Oral Comprehension	83	.35	.001
Quantitative Concepts	83	.44	.001
Academic Knowledge	82	.27	.012
Oral Language Cluster	81	.38	.001
Math Reasoning Cluster	81	.44	.001

10. Partial Correlations Between Scores on Social Skills scale on SSSR and WJ-III Clusters

Cluster	**Df**	***r*-value**	***p*-value**
Oral Language Reasoning Post-Test (Pre-Test as Covariate	80	.25	.022