Acta Genet Med Gemellol 31:64–116 (1982)
The Mendel Institute/Alan R. Liss, Inc.

Received 25 June 1981
Final 10 December 1981

Genetic and Environmental Influences on Reading Achievement:

A Study of First- and Second-Grade Twin Children

Emily Louann Harris

Department of Psychiatry, University of Iowa, Iowa City

CONTENTS

- Abstract
- 1. Introduction
  - 1.1. Reading
  - 1.2. Skills Underlying Reading
  - 1.3. Twin Studies of Reading
  - 1.4. Twin Studies of Related Abilities
  - 1.5. Relationship of Family Attitudes to Achievement in Young Children
  - 1.6. Summary
- 2. Materials and Methods
  - 2.1. Sample
  - 2.2. Testing Procedures
    - 2.2.1. Psychological Test Battery—Children

The original version of this work was submitted to the Graduate School faculty of Indiana University for completion of the requirements for the degree Doctor of Philosophy, which was awarded in May 1980. The advisory committee consisted of Drs. Richard J. Rose, Chairman, Joe C. Christian, P. Michael Conneally, Leo C. Fay, and James A. Norton. The dissertation work was supported by a Pre-doctoral Traineeship from the United States Public Health Service 5 T01 DE 00119, a Grant-in-Aid of Research from Sigma Xi, and a Doctoral Student Grant-in-Aid from Indiana University.

    - 2.2.2. Questionnaires—Parent(s)
    - 2.2.3. Questionnaire—School
  - 2.3. Methods of Analyses
    - 2.3.1. Reliability Estimates
    - 2.3.2. Adjustment
    - 2.3.3. Intercorrelations
    - 2.3.4. Factor Analyses
    - 2.3.5. Twin Analyses
- 3. Results
  - 3.1. Representativeness of the Sample
  - 3.2. Reliability Estimates
  - 3.3. Adjustments
  - 3.4. Interrelationships of the Children's Test Battery
    - 3.4.1. Intercorrelations
    - 3.4.2. Partial Correlations
    - 3.4.3. Factor Analysis
  - 3.5. Twin Analyses
    - 3.5.1. Woodcock Reading Mastery Tests
    - 3.5.2. Wechsler Tests
    - 3.5.3. ITPA Auditory Sequential Memory
    - 3.5.4. Auditory-Visual Integration Test Battery
    - 3.5.5. Factors
  - 3.6. Relationships of Environmental Measures to Woodcock Reading Mastery Tests
    - 3.6.1. Attitudes Toward Education
    - 3.6.2. Moos' Family Environment Scale
    - 3.6.3. Reading Instruction Approaches
- 4. Discussion
  - 4.1. Reading Models
  - 4.2. Constituent Factors in Reading Achievement
    - 4.2.1. Intercorrelations of Children's Test Battery
    - 4.2.2. Partial Correlations of AVI with Reading Achievement
    - 4.2.3. Factor Analyses of Children's Test Battery
    - 4.2.4. Environmental Variables
  - 4.3. Genetic Contributions to Variation in Children's Test Battery
    - 4.3.1. Reading Achievement
    - 4.3.2. WISC-R Tests
    - 4.3.3. Auditory Memory
    - 4.3.4. Auditory-Visual Integration Test Battery
    - 4.3.5. Factory Derived from Children's Test Battery
    - 4.3.6. Summary
- 5. Conclusions
- Acknowledgments
- References
- Appendices
- List of Tables, Appendices
- Abbreviations Used in Texts and Tables for Tests in Protocol
- Subject Index

Previous twin studies provide evidence for genetic contributions to individual differences in reading achievement, but the nature of those genetic effects is uncertain. Reading is a complex behavior composed of many lower-level skills including attention, memory, learning ability, and the integration of auditory and visual information; genetic influences may exert their effects through these lower-level skills. This study evaluated the relationship between auditory-visual integration (AVI) and reading achievement and assessed genetic variance in AVI and reading achievement using the conventional twin model (comparison of identical twins with fraternal twins).

The final sample consisted of 109 first- and second-grade volunteer twin pairs. Of the 89 same-sexed pairs, 57 pairs were classified as monozygotic, 31 pairs as dizygotic, and one remained unclassified; zygosity determination included examination of the intrapair similarity of genetic markers, dermatoglyphics, and/or physical appearance. The childrens' test battery included reading achievement tests (Woodcock Reading Mastery Tests) and the AVI task as well as estimates of general intellectual ability (Vocabulary and Block Design tests from the Wechsler Intelligence Scale for Children-Revised), auditory memory (Auditory Sequential Memory subtest of the Illinois Test of Psycholinguistic Abilities), comparison of auditory patterns, and comparison of visual-spatial patterns. Although the presence of heritable variation in reading achievement has been reported previously [22, 23, 30, 34, 50, 57], it does not account for all of the reliable variance; accordingly, two family questionnaires (Attitudes Toward Education and Moos' Family Environment Scale) and one dealing with first-grade method(s) of reading instruction were included to assess possible environmental factors.

The most unique aspects of this study were the young age of the children studied and the twin analyses of the auditory-visual integration test battery. Evidence for heritable variation in reading achievement was found in this young sample consistent with the results of studies using older children. The main thrust of this project was to investigate possible "lower-level" sources of this heritable variation, particularly auditory-visual integration. As a first step, it was necessary to establish the presence of a significant relationship between reading achievement and auditory-visual integration, independent of general intellectual ability. The partial correlations between the reading achievement measures and AVI (WISC-R test scores partialled out) ranged from 0.10 to 0.49 and were generally significant. Factor analysis of the children's test battery supported the idea that the integration aspect of the AVI task (auditory-visual and/or spatial-temporal) was the aspect of this skill related to reading achievement rather than auditory memory or visual-spatial ability, which also contribute to success on the AVI task. Twin analyses of the AVI task revealed evidence of the presence of heritable variation. However, the task relates to other skills exhibiting genetic variance, and it could not be determined whether aspects of the AVI task independently related to reading achievement were the sources of the heritable variation observed.

Although this study emphasized genetic relationships, environmental contributions to variation in reading achievement were also considered in a preliminary fashion. Family environmental data suggested that exposure to and emphasis on intellectual and cultural activities promote reading achievement.

In conclusion, a positive and significant relationship between reading achievement and auditory-visual integration, independent of general intellectual ability, was established for a first- and second-grade sample. One environmental variable, intellectual-cultural orientation of the family, was also related to reading achievement. Evidence indicative of the presence of genetic variance was presented for reading achievement, as previously reported in older twin samples, and for auditory-visual integration.

Key words: Reading achievement, Auditory-visual integration, Intellectual ability, Auditory memory, Woodcock Reading Mastery Tests, Wechsler Intelligence Scale for Children-Revised, Illinois Test of Psycholinguistic Abilities, Family Environment Scale, Parental Attitudes Toward Education Questionnaire, Genetic influences, Environmental influences, Twins

1. INTRODUCTION

1.1 Reading

Reading is a complex of skills that involves decoding of the printed symbols into spoken words and then comprehension of the decoded words. Efficient decoding depends upon the adequate functioning of memory, perception, and integration skills, along with other variables. It is unknown whether these skills are, in general, developed to the necessary level of functioning in beginning readers or whether individual differences in the rate of development and/or the final efficiency of these skills contribute to the variation observed in the achievement of beginning readers. In conjunction with these ideas, there has been some debate over whether reading retardation in children results from a developmental lag, with possible "catch up" in reading, or from specific deficits, where one would expect catch up only through compenstion [42]. One would expect that, in individual cases, either one of those mechanisms could cause reading deficit, although on a population level, one of the mechanisms could be primarily responsible for the cases of reading retardation seen in children. Individual variation in reading within normal limits, as well as specific forms of reading deficiency, may be due to both genetic and environmental influences.

Although previous research supports the idea of genetic contributions to variation in reading achievement in school-age children, little is known about the mechanisms of that influence on reading achievement. The primary objective of the present research was to further explore the nature of the genetic influences in reading achievement among beginning readers. To gain more insight into the nature of that influence, variation in several behaviors possibly related to reading achievement was examined for evidence of genetic contributions. The method employed was the classic twin study.

The specific objectives in this study were: 1) to establish the relationship of auditory-visual integration to reading achievement independent of general intellectual ability, 2) to examine the contributions of genetic factors to individual variation in reading achievement and auditory-visual integration through twin analyses, and 3) to explore, in a preliminary fashion, familial environmental factors possibly related to reading achievement.

1.2. Skills Underlying Reading

The literature from two types of approaches may be applied to the study of skills related to reading ability: 1) comparison of normal or adequate readers with poor or retarded readers and 2) studies of the development of normal reading ability. These approaches have implicated a number of factors in the development of reading ability and/or a reading deficit: auditory and visual memory, auditory-visual and spatial-temporal integration, auditory and visual perception, visual perceptual speed, vigilance, learning ability, and impulsivity [11, 13, 19, 29, 37, 43, 44, 49, 53]. Evidence implicating auditory and visual memory, auditory and visual perception, and auditory-visual and spatial-temporal integration is presented in Tables 1 and 2. Table 1 summarizes evidence accumulated through the comparison of good or normal readers with poor or retarded readers; Table 2 summarizes evidence gathered through the study of children reading within or above the normal range.

TABLE 1. Differences between Good or Normal Readers and Subnormal Readers in Measures of Auditory Memory, Auditory and Visual Discrimination, Auditory-Visual Integration, and Spatial-Temporal Processing

Study	Ages	Reading tests[a]	IQ measures	n_{Poor}	n_{Good}	Auditory memory[a,b]
Birch and Belmont, 1964	9-4 to 10-4 years (all males)	British Sentence Read. Test, Metropolitan Achieve. Test (word discrimination, word knowledge, reading)	NR (>80 to be included in sample)	150	50	NA
Blank and Bridger, 1966	9 years	NR	NR (matched according to IQ and Stanford-Binet Vocabulary Scores)	13	13	NA
Bryden, 1972	grade 6	Gates-MacGinitie Reading Tests	IPAT test of "g"	20	20	NA
Vande Voort et al, 1972	8-0 to 12-11 years	Wide Range Achievement Reading Age	Peabody Picture Vocabulary Test	48	48	NA
Vande Voort and Senf, 1973	8 to 10 years	Stanford Achievement Test Reading Age (retarded readers)	WISC (retarded readers)	16	16	NA
Badian, 1977	grades 3-5	Gray Oral Reading Test, Stanford Diagnostic Reading Test, Reading Comprehension Subtest	Lorge-Thorndike Intelligence Test	30	30	Reproduction of tapped patterns $p < 0.001$ ITPA $p < 0.001$

	Equivalence tasks[a,b,c]						
Study	At-Vs	Vt-Vs	Vs-Vs	Vt-Vt	At-Vt	At-At	Correlations
Birch and Belmont, 1964	$p < 0.001$	NA	NA	NA	NA	NA	$r_{\text{At-Vs, WISC IQ}} = 0.38$
Blank and Bridger, 1966	NA	$p < 0.05$	ns	$p < 0.005$	NA	NA	
Bryden, 1972	ns	$p < 0.05$			ns		$r_{\text{Read., IQ}} = 0.36$ $r_{\text{Read., matching (poor)}} = 0.60$
	Vs-At ns	Vs-Vt $p < 0.05$	ns	ns	Vt-At $p < 0.05$	$p < 0.05$	$r_{\text{Read., matching (good)}} = 0.14$[d]
Vande Voort et al 1972	$p < 0.01$ At-Vs (simul.) $p < 0.01$	NA	$p < 0.01$	NA	NA	$p < 0.01$	
Vande Voort and Senf, 1973	$p = 0.16$	NA	$p < 0.05$	$p = 0.41$	NA	$p < 0.01$	[e]$r_{\text{Vs-Vs, Vt-Vt (poor)}} = -0.48$ $r_{\text{Vt-Vt, At-At (good)}} = 0.57$ $r_{\text{Read., Vs-Vs (poor)}} = -0.47$ $r_{\text{IQ, Vt-Vt (good)}} = 0.53$ $r_{\text{IQ, At-At (good)}} = 0.53$ $r_{\text{Read., Vt-Vt (good)}} = 0.49$
Badian, 1977	$p < 0.005$ At-Vs (simul.) $p < 0.025$ At-Vs (10-second delay) $p < 0.001$ At-Vs (distract.) $p < 0.001$	NA	ns	NA	NA	ns	

[a] NA = not administered, NR = not reported.

[b] In all instances of significant differences, normal readers performed better than subnormal readers.

[c] At = auditory-temporal; Vt = visual temporal; Vs = visual-spatial.

[d] Possible ceiling effect.

[e] Only significant correlations listed.

*TABLE 2. Relationships of Auditory and Visual Memory, Auditory and Visual Discrimination, Auditory-Visual Integration, and Spatial-Temporal Processing to Normal Variation in Reading Achievement**

Study	Ages	N	Reading tests	IQ measures	Equivalence patterns
Birch and Belmont, 1965	5–12 years (grades K–6)	approx 30 per grade	Metropolitan Readiness Test (grade 1), Stanford Achievement Test (grades 2–6)	Otis Quick-Scoring Tests of Mental Ability (grades 1–6)	At-Vs
Muehl and Kremenak, 1966	grade 1	108	Harrison-Stroud Reading Readiness Profiles, Metropolitan Achievement Tests	Lorge-Thorndike Intelligence Test	Vs-Vs, At-At, At-Vs, Vs-At
Kahn and Birch, 1968	grade 2–6 (all males)	70 per grade	Metropolitan Reading Achievement Battery	Lorge-Thorndike Intelligence Test	At-Vs
Reilly, 1971	grades 1–4	225 total 60 (1st) 56 (2nd)	Gates-MacGinitie Reading Test	NR	At-Vs
Blackman and Burger, 1972	grade 1	78	Metropolitan Achievement Test (MAT), Wide Range Achievement Test	Peabody Picture Vocabulary Test (PPVT)	At-Vs
Warren et al, 1975	grades 1–3	100 (1st) 84 (2nd) 110 (3rd)	Cooperative Primary Test	Lorge-Thorndike Intelligence Test (grades 2 and 3)	Vt-Vt, At-At
Whiton et al, 1975	grade 1 (all males)	64	Primary Reading Profiles	NR	Vs-Vs, At-At, Vs-At, At-Vs
Barker, 1976	grades 2–5	72	California Achievement Tests (vocabulary, comprehension)	California Achievement Test	NA
Hare, 1977	grade 2	81	Wide Range Achievement Test (WRAT) Stanford Achievement Test (SAT)	Lorge-Thorndike Intelligence Test	NA

Study	Results (Correlations unless otherwise specified)[a]					
Birch and Belmont, 1965		IQ, At-Vs	Reading, At-Vs	Reading, IQ		
	grade 1	0.56	0.70	0.56		
	grade 2	0.42	0.42	0.53		
Muehl and Kremenak, 1966		Vs-Vs	At-At	Vs-At	At-Vs	
	Reading	0.16	0.28	0.39	0.52	
Kahn and Birch, 1968	grade 2		At-Vs, Word Knowledge		At-Vs, Reading Comprehension	
	no control for IQ		0.37		0.44	
	verbal IQ controlled		0.38		0.20	
	nonverbal IQ controlled		0.38		0.21	
Reilly, 1971		At-Vs, Vocabulary	At-Vs, Comprehension	At-Vs, Total Reading		
	grade 1	0.23	0.14	0.21		
	males (31)	0.12	0.19	0.15		
	females (29)	0.34	0.04	0.25		
	grade 2	0.65	0.71	0.70		
	males (20)	0.75	0.81	0.80		
	females (36)	0.61	0.68	0.67		
Blackman and Burger, 1972		PPVT	MAT, Word Knowledge			
	At-Vs	0.06	0.11			
Warren et al, 1975		Vt-Vt, Reading	At-At, Reading	Vt-Vt, IQ	At-At, IQ	IQ, Reading
	grade 1	0.18	0.32			
	grade 2	0.57	0.50	0.49	0.44	0.52
	(with IQ partialled out, At-At and Vt-Vt accounted for significant portions of reading variance)					
Whiton et al, 1975		Vs-Vs	At-At	Vs-At	At-Vs	
	Word Knowledge	0.24	0.16	0.44	0.36	
	Word Attack	0.22	0.12	0.51	0.28	
	Comprehension	0.25	0.13	0.34	0.39	
	Total	0.29	0.16	0.48	0.44	
Barker, 1976	Grades 2–5 Combined	Reading Vocabulary	Reading Comprehension	Auditory Memory		
	Visual Memory	0.31	0.43	0.30		
	Auditory Memory	0.48	0.46			
Hare, 1977		WRAT-Reading	SAT-Paragraph Meaning	IQ		
	Auditory Memory	0.33	0.24	0.07		
	IQ	0.00	0.11			

*NA = not administered; NR = not reported. At = auditory-temporal; Vt = visual-temporal; Vs = visual-spatial.

[a] For correlation coefficients, underlined values not significantly different from zero (p ⩾ 0.05).

It has been generally accepted that children with reading disability represent a heterogeneous group in regard to the etiology of their disability, which may include biological as well as social, educational, and emotional factors. Birch [4] proposed that one cause of subnormal reading achievement could be a primary inadequacy in the ability to integrate auditory and visual stimuli. Shortly thereafter, Birch and Belmont [5, 6] published evidence consistent with this notion (see Tables 1 and 2); their results indicated that performance on an auditory-temporal-visual-spatial equivalence task was related to the level of reading achievement. The question of whether these findings were actually the result of auditory-visual integration differences is still unanswered and hotly debated. Possible confounding factors that were untested in Birch and Belmont's studies include short-term memory, perception, and spatial-temporal integration; research attempting to clarify matters has provided conflicting results. A few of these studies are described in Tables 1 and 2. Thus, several related tasks have been found to correlate with reading achievement, but the causal relationships remain unresolved.

1.3. Twin Studies of Reading

Evidence supporting the notion of a genetic contribution to variation in reading achievement is presented in Table 3. The studies presented in the table cover a period of almost 40 years and were conducted in two different cultures; in all cases, reading of the native language was studied. Perusal of the table reveals substantial familial resemblance for reading achievement with identical co-twins being more similar than fraternal co-twins; the data suggest that genetic factors make a substantial contribution to variation in reading achievement in school-age children.

1.4. Twin Studies of Related Abilities

Several twin studies have also included a measure of a variable that may be related to reading achievement—auditory short-term memory. Strandskov [48] and Vandenberg [50] both administered to adolescent children the Primary Mental Abilities Test, which includes memory as a primary ability. The F-ratio comparing the within-pair mean squares of the monozygotic (MZ) and dizygotic (DZ) twins for that ability was not significant in either study. Block [9] and Vandenberg [50] using the Digit Span test of the Wechsler Intelligence Scale for Children (WISC) with adolescent-aged twins, each reported a significant F-ratio of the within-pair mean squares. Pezzullo et al [38] administered a modified version of the WISC Digit Span to MZ and like-sexed DZ twins aged 10–15 years; a significant F-ratio of the within-pair mean squares was obtained. Thus, evidence exists for a genetic contribution to variation in auditory short-term memory, although the results of previous studies are conflicting; these conflicts may have arisen from the differing natures of the tasks used.

1.5. Relationship of Family Attitudes to Achievement in Young Children

In studies of cognitive abilities, few investigators have included information on environmental influences more specific than general environmental factors (eg, socioeconomic status, number of siblings, order of birth). However, Garfinkle [20], in a study of genetic and environmental influences on the development of mathematical concepts in young twin children (four to eight years old), included two parental questionnaires designed to

*TABLE 3. Twin Studies of Reading Achievement**

			Sample size			Intraclass correlation coefficients (r_I)[a]			Ratio of within-pair mean squares[b]
Study	Ages tested	Tests	MZ	DZ	FM	MZ	DZ	FM	
Newman et al, 1937	8–18 years	Stanford Achievement Test	50	50					
		Word Meaning				0.86	0.56		
		Spelling				0.87	0.73		
		Educational Age				0.89	0.70		
Wictorin, 1952	9–14 years grades 2,4&6	Grades in Reading (Swedish)	117	130		0.92	0.61		
Husen, 1959	7th grade	Grades in Reading	352	668		0.72	0.57		
Husen, 1960	12–13 years	Reading Achievement Test (Swedish)	134	180	180	0.89	0.62	0.45	
Vandenberg, 1962	high school age	Gray Oral Reading Paragraphs	45	36					1.97
Matheny and Dolan, 1974	9–12 years	California Achievement Test	44	26					
		Reading Vocabulary				0.84	0.65		1.84
		Reading Comprehension				0.85	0.52		2.53
		Total Reading				0.89	0.61		3.05

*MZ = monozygotic twin pairs, DZ = like-sexed dizygotic twin pairs, and FM = unlike-sexed dizygotic twin pairs.

[a] In cases when r_I's reported, MZ and DZ r_I's differed significantly from one another ($p < 0.05$).

[b] In cases where ratio reported, ratio differed significantly from one ($p < 0.05$).

measure more specific environment factors: Moos' Family Environment Scale (FES) [31] and Attitudes Toward Education (ATE, developed by Garfinkle, Claussner, and Vandenberg). She reported correlations significantly different from zero for the following: Moos' Intellectual Cultural Orientation subscale (Moos' ICO) and the Piagetian Mathematical Concepts Battery (PMCB), Moos' ICO and the Peabody Picture Vocabulary Test (PPVT), Moos' Achievement Orientation subscale (Moos' AO) and the PMCB (negative relationship), Moos' AO and Raven's Progressive Matrices (negative relationship), and the Moos' Cohesion subscale and the PPVT. Although a substantial amount of the individual variation remained unexplained by genetic factors and the environmental variables assessed, her results suggested that certain environmental factors were related to the development of cognitive abilities in young children and that environmental variables of a more specific nature than those usually assessed should be included in future research.

1.6. Summary

Reading is a very complex process, involving many skills or factors including memory, perception, integration of stimuli and learning ability. A severe deficit in one or more of these factors may drastically affect reading achievement, although some deficits may be compensated for by the other factors. Individual differences in one or more of these factors may produce the variation seen in reading achievement in school age children; these differences may be more readily detected in beginning readers, before compensation for relative deficiencies begins.

The evidence presented in this introduction suggests that there is a genetic contribution to individual variation observed in reading achievement. It is unlikely that the variation within the normal range of reading ability is directly affected by genes; rather, the effects of the genes are probably exerted through more basic skills involved in reading. This introduction suggests that proficiency in the integration of auditory and visual stimuli is related to reading achievement at a beginning reading level. For these reasons, the variation in auditory-visual integration and in skills related to it was examined to determine the extent of the influence of genetic factors.

2. MATERIALS AND METHODS

2.1. Sample

A volunteer twin sample was ascertained through the Indiana University Twin Panel, several Indiana Mothers-of-Twin Clubs (Anderson, Kokomo, New Castle, South Bend, Seymour, Charter Chapter-Indianapolis, Southside Chapter-Indianapolis), the Greater Flint Mothers of Multiples Club (Michigan), Dr. Ronald Wilson of the Louisville Twin Study, a number of Indianapolis area grade schools, and announcements released to Central Indiana newspapers. The sample was primarily middle class and Caucasian.[1] By history, none of the children had significant uncorrected vision or hearing deficits. The sample initially included 110 twin pairs. One pair was not included in the analyses due to lack of cooperation. Further, zygosity could not be determined for one twin pair, so only 108 pairs were included in the genetic analyses.

Zygosity determination of like-sexed twins was based on similarity of genetic markers (in blood, urine, and saliva), dermatoglyphics, and/or physical appearance; placental reports were also informative for several sets of twins. Fifty-seven pairs were classified as monozygotic (MZ) and 31 were classified as dizygotic (DZ); one set remained unclassified and was removed from twin analyses. The zygosity of the unlike-sexed twin pairs was based on the sex difference. Further information regarding zygosity determination in this sample may be found in Appendix A.

---

[1] The exceptions include one twin pair each of the following origins: Canadian Indian, Negro, Oriental-Caucasian, and Negro-Caucasian.

The major criteria for inclusion in the study were completion of at least the first semester of first grade and no more than the second semester of second grade. The actual range was, in tenths of a school year, 1.7–2.9; average: 2.26.[2] The age range of this sample was 6.71–9.26 years; average, 7.73 years. The sex ratio was significantly different from one with the sample being composed of 60% females. The sex, zygosity, and grade distribution may be found in Appendix A.

2.2. Testing Procedures

2.2.1. Psychological Test Battery—Children. The psychological testing included the Woodcock Reading Mastery Tests (form A), the Vocabulary and, when time permitted, the Block Design tests of the Wechsler Intelligence Scale for Children-Revised, the Auditory Sequential Memory subtest of the Illinois Test of Psycholinguistic Abilities, and the auditory-visual integration test battery. All tasks were individually administered. The co-twins were tested simultaneously by the researcher and a trained female assistant; the usual procedure was for one twin to be administered the reading achievement tests while the co-twin was given the auditory-visual integration test battery and the auditory memory task. The entire testing was completed in approximately 1½–2 hours. When possible, the zygosity testing was done at a time separate from the psychological testing.

The Woodcock Reading Mastery Tests (WRMT) [58] are a battery of individually administered tasks that includes tests of letter identification, word identification, word attack, word comprehension, and passage comprehension. A total reading score is derived from a combination of the five test scores. Two equivalent forms are published (A and B). The Letter Identification Test is designed to measure a child's ability to name letters of the English alphabet. Test items include a variety of upper- and lower-case cursive and printed letters. The Word Identification test measures a child's ability to name words; it is not assumed that the child knows the meaning of the word or that he/she has ever seen the word before. The Word Attack Test requires a child to pronounce nonsense words and measures the child's ability to apply phonic and structural analysis skills. The Word Comprehension Test is designed to measure a child's knowledge of word meanings through an analogy format. The analogy consists of two pairs of words with the last word missing for the child to complete the analogy. The Passage Comprehension Test consists of phrases and sentences that each have a missing word. The child is required to read the passage and tell the examiner an appropriate word to go in the blank space; the easier items are accompanied by pictures. Because of the variety of skills on which the child must rely to supply the missing word in a passage, Woodcock states that this test may be considered "an omnibus test of reading skills" [58: p. 4]. The Index of Total Reading results from the combination of the scores on the five tests and provides an index of overall reading skill. Several types of scores are available from the WRMT for the five individual tests and the composite index: Mastery Score, Grade Level, and Percentile at grade placement. The only unfamiliar type of score, the Mastery Score, is actually a transformation of the Raw Score developed to facilitate interpretation and prediction of success. Its derivation is described in the WRMT Manual [58]. The advantage of the Mastery Score over the Raw Score for this study is the availability of a composite score; this is simply the average of the Mastery Scores for the five individual tests. In the majority of the analyses, the variable used was the percentile score, which adjusts for grade level in tenths of a school year, rather than the Mastery Scores. Any deviations from this are noted.

The Woodcock Reading Mastery Tests were chosen for this study for the following reasons: 1) designed for individual administration, 2) relatively short testing time, 3) wide age-range design permitting follow-up or longitudinal studies, 4) reliabilities exceeding 0.88 for all but the Letter Identification Tests for this age group (presented in Appendix B), 5) ease of administration, 6) availability of scores for several different skills rather than just a total reading score.

Two tests from the Wechsler Intelligence Scale for Children-Revised (WISC-R) [55] were administered in part to provide information on the representativeness of the twin children studied. The tests, chosen on the basis of their correlations with full-scale IQ, were Vocabulary from the Verbal scale and Block Design from the Performance scale. The test-scale correlations as well as the reliabilities of the Vocabulary and Block Design tests as reported in the WISC-R Manual for the appropriate age groups [55] are presented in Appendix B. The reliabilities for those two tests in the age range studied all exceed 0.68.

The Vocabulary test requires the child to verbally give the meaning of a number of age-appropriate words. In the Block Design test, the child is asked to reproduce a two-dimensional design with a set of painted cubes or blocks. Scaled scores, which are adjustments of the raw scores according to four-month age bands, were the variables used in the analyses unless otherwise noted.

---

[2] Grade level designations are those suggested by Woodcock [58], with the school year divided into tenths; eg, grade 1.0 refers to beginning first grade and grade 1.9 refers to the last month or completion of first grade.

The Auditory Sequential Memory subtest of the Illinois Test of Psycholinguistic Abilities (ITPA) [26] measures a child's ability to repeat a series of spoken digits. In contrast to the Digit Span test from the Wechsler series, this task requires only forward repetition of the numbers. Reliabilities for appropriate age groups, as reported by Paraskevopoulos and Kirk [36], exceed 0.80 and are presented in Appendix B.

The auditory-visual integration test battery consists of three tasks: the visual discrimination and the auditory discrimination tasks from the dissertation work of Badian [1] and the auditory-visual integration task adapted from the work of Kahn and Birch [25]. In the visual discrimination task, a child is first shown a visual-spatial (dot) pattern for approximately five seconds and is then asked to find the matching pattern from a choice of three. The auditory discrimination task is a same-different task in which a child hears two auditory-temporal patterns (tapped-out patterns) sequentially and must decide whether they are the same or not. The auditory-visual integration task involves an auditory-temporal stimulus with a choice of three visual-spatial patterns for response. Test-retest reliabilities were reported by Kahn and Birch [25] as 0.76 for a group of third-grade boys and 0.90 for a group of fifth-grade boys on this task.

2.2.2. Questionnaires-Parent(s). The parents of the children were asked to complete two questionnaires dealing with attitudes in their home—Moos' Family Environment Scale (FES) [31] and Attitudes Toward Education (ATE), developed by Garfinkle, Claussner, and Vandenberg [20]. The questionnaires can readily be completed in approximately 30 minutes.

Moos' FES "focuses on the measurement and description of the interpersonal relationships among family members, on the directions of personal growth which are emphasized in the family, and on the basic organizational structure of the family" [31: p. 3]. The parents were asked to complete form R, which consists of 90 true-false items. The test is divided into ten subscales under the three major areas:

1. Relationship dimensions: cohesion, expressiveness, conflict.
2. Personal growth dimensions: independence, achievement orientation, intellectual-cultural orientation, active recreational orientation, moral-religious emphasis.
3. System maintenance dimensions: organization, control.

The internal consistencies (Kuder-Richardson Formula 20) and the test-retest reliabilities for all the subscales are reported to be in an acceptable range, 0.64–0.79 and 0.68–0.86, respectively.

The (ATE) questionnaire consists of 15 statements; parents are instructed to rate the degree to which they agree or disagree with each statement (possible ratings, 1–5). Garfinkle [20] reports an internal consistency (coefficient alpha) reliability of 0.61 (excluding items 5 and 13). Items 5 and 13 were eliminated from Garfinkle's analyses because of low communalities (less than 0.09). She also reported results of factor analysis that suggested three meaningful factors: basic academic education (I), parental participation (II), and general utility of education (III). Analyses for this study will utilize the suggested factors.

2.2.3. Questionnaire-School. When possible, each child's first-grade and, if appropriate, second-grade teachers were asked (via the school principals) to complete a short questionnaire regarding the methods through which the child was taught to read. The questionnaire simply lists nine approaches to teaching reading; the teachers were asked to rate the contribution of each approach to his/her reading program. The ratings range from 1 (none) to 5 (very strong). This questionnaire was developed for this study and was based on information in a textbook of reading instruction [12]. A copy of this questionnaire may be found in Appendix C along with a description of the different approaches that was attached to each questionnaire.

2.3. Methods of Analyses

2.3.1. Reliability Estimates. Reliability estimates for the three tests in the auditory-visual integration test battery were calculated through the Statistical Package for the Social Sciences (SPSS) Subprogram Reliability [46]. Estimates of internal consistency (Cronbach's alpha coefficient and Spearman-Brown split-half coefficient) are the coefficients appropriate for these data. An odd-even split was used to calculate the split-half coefficient because the items within each test tend to increase in difficulty. The SPSS Subprogram Reliability usually divides the items so that the first N/2 items are compared with the second N/2 items. However, an odd-even split may readily be achieved through rearrangement of the items on the procedure card [47: p. 5].

2.3.2. Adjustment. Stepwise regression (BMD2R from the UCLA BMD series[3]) was used to determine which

---

[3] Programs, published by the Health Science Computing Facility, UCLA, were revised for use on the computing systems at Indiana University; these revised programs are described in the Research Computation Center's BMD Manual.

variables (grade, sex, age, order of administration) needed to be adjusted for in the nonstandardized tasks. The standardized tests were also examined through this method to determine if suggested standardizations were adequate for this sample.

Adjustment was accomplished through the use of a short FORTRAN program using the following general equation for each variable to be adjusted:

$$Y_{adj} = Y + b_{age} (8.0 - X_{age}) + b_{grade} (2.0 - X_{grade}) + b_{sex} (2 - X_{sex}),$$

where Y_{adj} is the score adjusted for the appropriate independent variables, Y is the original score, b_i's are the regression coefficients obtained from stepwise regression at the appropriate step, and X_i's are the individual's values for the variables for which adjustment was deemed necessary. Only the appropriate independent variables were included in the adjustment equation for a particular dependent variable.

2.3.3. Intercorrelations. BMD3D (from the UCLA BMD series) was used to investigate the interrelationships of appropriate variables. BMD2R (stepwise regression) was used to determine partial correlations when desired.

2.3.4. Factor Analyses. BMDX72 (from the UCLA BMD series) was utilized to examine the factor structure of the tasks administered to the children. Replication of the factor analysis results of the Attitudes Toward Education parental questionnaire reported by Garfinkle [20] was attempted. This program includes an option to consider blanks as missing data, which is useful for analyses of the children's test battery data. The correlation matrix, which is the basis for the factor analysis, accounts for missing data by deleting a case only for the correlation coefficients in which the variable with the missing value is involved (eg, if the Block Design test score is missing, the individual is deleted only for calculations involving Block Design). In computing factor scores, the program substitutes the mean for the particular variable when the value is missing.

The Varimax solution was the option chosen for rotation of the factor matrix. This is an orthogonal rotation and tends to produce factors that each have high loadings for a few variables and near zero loadings for the remaining variables. A more detailed explanation of this method for rotation may be found in Cooley and Lohnes [17].

2.3.5. Twin Analyses. The twin analyses (comparison of MZ twins with DZ twins) were performed through the use of TANOVA and TWNAN from the Department of Medical Genetics at Indiana University Medical Center. The model used is described in Christian et al [14]. The output from these programs provides statistics useful in evaluation of the assumption that the total variance of MZ twins equals that of DZ twins and the assumption that the covariance among environmental effects within pairs of MZ twins equals that within pairs of DZ twins. Also included in the output are means, standard deviations, and intraclass correlation coefficients for the MZ and DZ samples, as well as tests of the significance of estimates of genetic variance and heritability estimates.[4]

3. RESULTS

3.1. Representativeness of the Sample

Means and standard deviations for the children's test battery are presented in Table 4. Representativeness of the sample was examined through comparison of the means and variances of the sample with those of standardization samples for the WISC-R tests and the Woodcock Reading Mastery Tests. Results of these comparisons are presented in Tables 5 and 6.

The raw score means and standard deviations are reported in the WRMT Manual for the Pre-A form in the grade 1.9 sample and for form A in the grade 2.9 sample. Table 5 shows that the means in the present sample were consistently greater than the means for the appropriate comparison groups; only one difference was not significant ($p < 0.05$). The sample variances were larger in all but one case, but the differences were significant for only three tests in the first-grade sample.

The Manual for the WISC-R reports means and standard deviations for the scaled scores in age groups within six weeks of a half-year age; appropriate for this sample are

---

[4] A user's guide for the twin analysis programs TWNAN and TANOVA, compiled by Mary M. Evans, is available from the Department of Medical Genetics, Indiana University.

*TABLE 4. Means and Standard Deviations for Children's Test Battery**

	$\bar{x}$	SD
WRMT (percentile)		
LI	73.2	22.3
WI	63.7	29.3
WA	61.3	29.3
WC	51.9	27.1
PC	65.1	23.8
Total	68.4	30.4
WISC-R (scaled scores)		
Vocabulary	9.3	2.7
Block Design	10.9	2.9
ITPA		
Auditory Memory	31.0	9.0
AVI Battery		
VV	8.4	1.7
AA	6.0	1.7
AVI	10.7	4.2

*N = 218 except for Vocabulary (N = 216), Block Design (N = 172), and Auditory Memory (N = 217).

TABLE 5. WRMT: Raw Scores Means and Standard Deviations

	Grade 1.9[a] Pre-A		Present study grade 1[b]		Grade 2.9[a] A		Present study grade 2[c]	
	$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD
LI	31.7	2.8	34.1	3.1	38.0	3.7	40.5	4.0
WI	40.3	20.7	58.3	28.6[e]	73.8	20.3	87.9	22.1
WA	5.8	7.4	15.7	13.3[e]	18.8	13.3	28.7	14.7
WC	7.9	7.6	12.9	9.2[e]	18.0	8.4	21.9	8.2
PC	18.3	11.1	19.4[d]	11.9	25.0	10.5	33.1	12.3

[a] From WRMT Manual [58: p. 57].

[b] Actual range = 1.7–2.0; mean = 1.90.

[c] Actual range = 2.7–2.9; mean = 2.87.

[d] Sample mean and mean of standardization sample not significantly different ($p \geqslant 0.05$).

[e] Sample variance significantly different from variance of standardization sample ($p < 0.05$).

TABLE 6. WISC-R Scaled Scores Means and Standard Deviations

	Age 7½ years[a]		Present study grade 1[b]		Age 8½ years[a]		Present study grade 2[c]	
	$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD
Vocabulary	9.8	2.8	9.0[d]	2.7	10.3	3.2	9.7	2.6[e]
Block Design	10.1	3.0	10.8[d]	3.0	10.1	3.1	11.1[d]	2.5[e]

[a] From WISC-R Manual [55: pp. 37–38].

[b] Mean age = 7.4 years.

[c] Mean age = 8.2 years.

[d] Sample mean significantly different from mean of standardization sample ($p < 0.05$).

[e] Sample variance significantly different from variance of standardization sample ($p < 0.05$).

comparisons with the 7½ year (± 6 weeks) and 8½ year (± 6 weeks) age groups. The scaled scores were designed so that the mean and standard deviation of the individual tests are 10 and 3, respectively. These, of course, may differ slightly in selected age groups as shown in Table 6. The means for the Vocabulary test were lower than those of the normative group, although the difference is significant only in the grade one sample. Mean Block Design scores significantly exceeded the normative data for both grade groups. The sample variance was significantly less than the published norms in the grade-two sample.

3.2. Reliability Estimates

Estimates of reliability for the auditory-visual integration test battery revealed, first, that the auditory comparison task was not very reliable; Cronbach's alpha coefficient was 0.25. The other two tasks were somewhat more reliable: Cronbach's alpha was 0.65 for the visual comparison task and 0.76 for the auditory-visual integration task. Further information regarding these results and separate grade statistics are presented in Appendix B. Results of subsequent analyses involving the auditory comparison task should be interpreted with great caution due to its low reliability.

3.3. Adjustments

The variables in the children's test battery were examined for effects of age, sex, grade, and order of administration. Adjustments were deemed appropriate for all of the tasks except the reading tests. Rationale and details of these analyses are presented in Appendix D. The variables used in subsequently reported analyses were the adjusted scores; for reading tests, percentile scores have been used.

3.4. Interrelationships of the Children's Test Battery

3.4.1. Intercorrelations. To examine the relationships between the tests administered to the children, co-twins were treated as individuals. This is obviously not the ideal procedure since the sample is not made up of random independent observations, but this method was deemed adequate for superficial examination of the interrelationships of these variables. The intercorrelations are presented in Table 7.

Note first that the tests of the WISC-R, used as estimates of general intellectual ability, correlated significantly with all the other variables, except in the case of Block Design and Auditory Memory. Auditory Memory exhibited only low correlations with any of the variables. Perusing the results for the AVI battery, one quickly notices the low magnitude correlations for the AA task versus the moderate correlations for the VV and AVI tasks. The low correlations for the AA task were, at least in part, due to the low reliability of this task. In Appendix E, correlation matrices for the grade-one and grade-two samples are presented separately.

3.4.2. Partial Correlations. One could argue, on the basis of the intercorrelations, that scores from the reading tests correlated significantly with the auditory-visual integration test battery scores due to their common relationship with general intellectual ability, here estimated by the Vocabulary and Block Design tests of the WISC-R. One way to test that hypothesis is through an evaluation of partial correlation coefficients. Table 8 lists the partial correlations of the reading tests and the tasks from the AVI battery with verbal, nonverbal and both the verbal and nonverbal "IQ" tasks partialled out. The partial correlations of the AA task with the reading tests were generally nonsignificant, while those for the AVI and VV tasks were all significantly different from zero and ranged from 0.16 to 0.45.

TABLE 7. *Intercorrelations of Children's Test Battery**

	WRMT						WISC-R		Auditory memory	AVI Battery		
	LI	WI	WA	WC	PC	Total	Vocab.	Blocks		VV	AA	AVI
LI		0.57	0.51	0.48	0.58	0.66	0.26	0.18	0.11	0.33	0.18	0.26
WI			0.82	0.85	0.89	0.96	0.45	0.30	0.13	0.38	0.16	0.45
WA				0.73	0.76	0.88	0.48	0.34	0.14	0.43	0.16	0.41
WC					0.84	0.90	0.56	0.40	0.18	0.43	0.13	0.44
PC						0.92	0.55	0.33	0.13	0.41	0.23	0.48
Total							0.50	0.36	0.16	0.45	0.19	0.48
Vocab.								0.30	0.18	0.36	0.22	0.38
Blocks									0.12	0.40	0.18	0.39
Memory										0.12	0.11	0.25
VV											0.15	0.34
AA												0.33
AVI												

*N = 218 except for Vocabulary (N = 216), Block Design (N = 172), and Auditory Memory (N = 217). Underlined values not significantly different from zero ($p \geqslant 0.05$).

*TABLE 8. Partial Correlations: WRMT and AVI Battery**

	Vocabulary scores[a] partialled out			Block Design scores[b] partialled out			Both vocabulary and[b] block design scores partialled out		
	VV	AA	AVI	VV	AA	AVI	VV	AA	AVI
LI	0.27	0.12	0.16	0.26	0.07	0.20	0.21	0.03	0.16
WI	0.26	0.07	0.33	0.28	0.11	0.42	0.20	0.03	0.36
WA	0.32	0.06	0.28	0.32	0.09	0.39	0.24	−0.00	0.30
WC	0.29	0.01	0.30	0.31	0.08	0.39	0.21	−0.03	0.30
PC	0.27	0.13	0.33	0.29	0.20	0.45	0.19	0.11	0.37
Total	0.33	0.09	0.34	0.34	0.12	0.44	0.25	0.03	0.37

*Underlined values not significantly different from zero ($p \geqslant 0.05$).
[a] N = 216.
[b] N = 172.

3.4.3. Factor Analysis. The primary purpose of this analysis was to extract meaningful factors to submit to twin analysis. The Varimax rotated factor matrix for the three-factor solution of the combined grades sample is shown in Table 9. This was the solution with the most readily interpretable factors. Factor 1 is clearly a general reading achievement factor with high loadings for all five of the tests from the WRMT. That factor was also a more than 10% contributor to the variances of the WISC-R Vocabulary test and the auditory-visual integration task. Factor 2 seems to represent auditory memory, with high to moderate loadings for the ITPA Auditory Sequential Memory subtest and the auditory comparison and auditory-visual integration tasks. The third factor has high loadings for the visual comparison task and the WISC-R Block Design test; that factor was also a more than 10% contributor to the variances of the WISC-R Vocabulary test and the auditory-visual integration task. Factor 3 appears to represent a visual-spatial type ability. Note that each of the three derived factors was a more than 10% contributor to the variance of the auditory-visual integration task, with all three factor loadings being moderate in magnitude. Communalities for the variables which represent amount of variance accounted for, are also listed in Table 9. Appendix E presents the results for the factor analyses using the separate grade samples.

3.5. Twin Analyses

Twin analysis was performed for each of the variables in the children's test battery as well as for the three factors derived from the analysis reported immediately above. The results for the total twin sample and for the sample restricted to like-sexed twins are reported in Tables 10 and 11, respectively. The means and intraclass correlation coefficients for the MZ and DZ twins are presented in addition to the conventional test for the presence of genetic variance, which tests the significance of the ratio of the within pair mean squares (MS_{WDZ}/MS_{WMZ}). With a significant F-ratio, one may reject the null hypothesis that the amount of genetic variance presents equals zero. Recall that adjusted scores were used in these analyses, so the means reported were calculated from the adjusted scores and differ from those previously reported in Table 4.

The programs used for these analyses provide several statistics for testing assumptions of the twin model [14, 15]; the information presented in the tables indicates when the

TABLE 9. Factor Analysis of Children's Test Battery Varimax Rotated Factor Matrix for Combined-Grades Sample (N From 172 to 218)

	Factors[a]			
	1	2	3	Communality
Vocabulary	0.483	0.268	0.377	0.447
Block Design	0.135	0.101	0.827	0.713
Auditory Memory	0.111	0.782	0.124	0.639
Visual Comparison	0.307	0.035	0.697	0.582
Auditory Comparison	0.034	0.618	0.252	0.446
Auditory-Visual Integration	0.322	0.517	0.463	0.585
WRMT				
Letter Identification	0.701	0.096	0.022	0.501
Word Identification	0.924	0.087	0.183	0.894
Word Attack	0.829	0.081	0.270	0.767
Word Comprehension	0.842	0.106	0.312	0.817
Passage Comprehension	0.894	0.138	0.246	0.878

[a] Underlined values represent a contribution by that factor to the variance of that test of at least 10% ($\sqrt{0.10} = 0.316$).

assumption of the equivalence of the MZ and DZ means was judged suspect [16]. Tables in Appendix F provide information regarding the testing of two other assumptions: equivalence of total variance in the MZ and DZ samples; and equality of environmental covariance between the MZ and DZ samples.

3.5.1. Woodcock Reading Mastery Tests. Conventional twin analysis for each of the five individual tests and the Composite Index of Total Reading Achievement revealed evidence of genetic variance. The intraclass correlation coefficients indicated substantial familial resemblance for this battery of tests, with the MZ co-twins being more similar than the DZ co-twins. The F-ratio of the within-pair mean squares was significant in each case. The results for the total twin sample and for that restricted to like-sexed twins were quite similar.

3.5.2. Wechsler Tests. Intraclass correlations revealed substantial familial resemblance for the MZ twins on both tasks, but only for the Vocabulary test in the DZ sample. The nonsignificance of the intraclass correlation coefficient for the DZ twins on Block Design may be due to sampling fluctuation in this small sample since other data do not lead one to suspect the assumption of the equality of environmental covariances (see Appendix F). The F-ratio of the within-pair mean squares was significant in each case, suggesting the presence of genetic variance, but with some reservation in the Block Design test.

3.5.3. ITPA Auditory Sequential Memory. For this auditory memory task, evidence of genetic variance was present in the total twin sample but not in the analysis of the like-sexed twins alone. In neither analysis did intraclass correlation coefficients for the MZ and DZ twins significantly differ; however, the F-ratio of the within-pair mean squares was significant in the total twin sample.

3.5.4. Auditory-Visual Integration Test Battery. For the visual comparison task, the MZ and DZ co-twins showed some similarity. However, the MZ co-twins were only slightly more similar than the DZ co-twins, and the null hypothesis of no genetic variance could not be rejected.

*TABLE 10. Twin Analyses, All Twins Included**

	Means		Intraclass correlation coefficients (r_I)		Ratio of within-pair mean squares
	MZ	DZ	MZ	DZ	(MS_{WDZ}/MS_{WMZ})
WRMT (%iles)					
LI	74.3	72.1	0.787	0.460[b]	3.40
WI	63.8	64.5	0.914	0.636[b]	4.93
WA	60.8	62.0	0.870	0.365[b]	3.85
WC	51.7	53.1	0.804	0.662	1.70
PC	64.9	66.1	0.786	0.623	2.40
Total	68.7	68.9	0.934	0.592[b]	7.26
WISC-R (Scaled Scores)					
Vocabulary	7.78	8.20	0.696	0.326[b]	2.48
Block Design	10.8	10.0	0.727	0.048[b]	3.70
ITPA					
Auditory Memory	28.7	29.7	0.700	0.581	1.74
AVI Battery					
VV	8.18	8.18	0.423	0.320	1.12[d]
AA	5.97	5.48[a]	0.169	0.118	
AVI	9.15	9.71	0.552	0.195[b]	2.09[c]
Factors					
1	50.3	49.1	0.895	0.578[b]	4.80
2	49.9	50.5	0.649	0.430	2.32[c]
3	49.4	50.9	0.571	0.296	1.66

*$N_{MZ} = 57$ and $N_{DZ} = 51$ except in the following cases: for Vocabulary and the factors, $N_{DZ} = 50$; for Block Design, $N_{MZ} = 44$ and $N_{DZ} = 41$; and for Auditory Memory, $N_{MZ} = 56$. Solid underlining indicates value not significantly different from zero ($p \geqslant 0.10$); dashed underlining indicates value not significantly different from zero ($0.05 \leqslant p \leqslant 0.10$).

[a] MZ and DZ means significantly differ ($p < 0.05$).

[b] MZ and DZ r_I's significantly differ ($p < 0.05$).

[c] Refer to text for explanatory or qualifying comments.

[d] Ratio not significantly different from one ($p > 0.05$).

In the total twin sample, this analysis of the auditory comparison task suffered from several problems, most probably related to the low reliability of the task: the means and total variances for the MZ and DZ samples differed significantly, and the intraclass correlation coefficients were not significantly different from zero in either the MZ or DZ samples. With these multiple problems, it is unwise to proceed to a test for the presence of genetic variance.

For the auditory-visual integration task, the intraclass correlation coefficient for the MZ co-twins was moderate and significantly different from zero while that for the DZ co-twins was marginally significant ($p = 0.08$); the like-sexed DZ co-twins showed little similarity. The test for genetic variance was significant, indicating the presence of genetic variance. The marginal degree of similarity of the DZ co-twins may lead one to suspect possible misleading evidence of genetic variance due to violation of the assumption of equality of environmental covariances [15]. However, the inference of the possible presence of genetic variance remains worthy of discussion.

*TABLE 11. Twin Analyses, Like-Sexed Twins Included**

	Means		Intraclass correlation coefficients (r_I)		Ratio of within-pair mean squares
	MZ	DZ	MZ	DZ	(MS_{WDZ}/MS_{WMZ})
WRMT (percentiles)					
LI	74.3	70.2	0.787	0.558[a]	2.79
WI	63.8	64.7	0.914	0.645[a]	4.63
WA	60.8	60.8	0.807	0.387[a]	3.69
WC	51.7	55.0	0.804	0.659	1.74
PC	64.9	66.4	0.786	0.603	2.42
Total	68.7	68.9	0.934	0.620[a]	6.72
WISC-R (Scaled Scores)					
Vocabulary	7.8	8.3	0.696	0.368[a]	2.82
Block Design	10.8	10.1	0.727	−0.001[a]	3.96
ITPA					
Auditory Memory	28.7	28.7	0.700	0.637	1.42[b]
AVI Battery					
VV	8.18	8.09	0.423	0.253	1.10[b]
AA	5.97	5.58	0.169	0.175	
AVI	9.15	10.2	0.552	0.075[a]	2.29[c]
Factors					
1	50.3	49.5	0.895	0.651[a]	4.34
2	49.9	50.1	0.649	0.502	1.75
3	49.4	50.2	0.571	0.266	1.65

*N_{MZ} = 57 and N_{DZ} = 31 except in the following cases: for Block Design, N_{MZ} = 44 and N_{DZ} = 27; and for Auditory Memory, N_{MZ} = 56. Solid underlining indicates value not significantly different from zero ($p \geq 0.10$); dashed underlining indicates value not significantly different from zero ($0.05 \leq p \leq 0.10$).

[a] MZ and DZ r_I's significantly differ ($p < 0.05$).

[b] Ratio not significantly different from one ($p > 0.05$).

[c] Refer to text for explanatory comments.

3.5.5. Factors. Analysis for factor 1, which represents general reading achievement, revealed substantial familial resemblance with MZ co-twins being more similar than the DZ co-twins. The F-test of the within-pair mean squares was significant and indicated the presence of genetic variance.

For factor 2, which represents the auditory memory component of this test battery, the MZ co-twins were somewhat more similar than the DZ co-twins, and the conventional test for genetic variance was significant. However, in the total twin sample, the equivalence of total variances assumption was not tenable (see Appendix F); thus, the among-component estimate of genetic variance may be more appropriate [14]; the test of that estimate was not significant. Data from the like-sexed twin sample did not indicate a violation of that assumption, however, and the analysis indicated the presence of genetic variance.

The results for factor 3, which probably represents some type of visual-spatial ability revealed that, although co-twin resemblance was only moderate, genetic variance was a contributor to the total variation. In the like-sexed twin sample, however, the intraclass correlation coefficient was only marginally significantly different from zero ($p = 0.07$), which may lead one to question the equality of environmental covariances assumption (see Appendix F). The results were clearly parallel in the two analyses, however.

3.6. Relationships of Environmental Measures to Woodcock Reading Mastery Tests

The primary purpose of this study was to examine sources of the genetic variance contributing to reading achievement differences. To supplement that purpose, several environmental variables assessed via questionnaires were included to explore, in a preliminary manner, environmental contributions to individual variation in reading achievement. The results for the individual questionnaires follow.

3.6.1. Attitudes Toward Education. Both parents, when available, were asked to complete the questionnaire; data were available on 69 families for both parents and for 2 additional fathers and 10 additional mothers. Intercorrelations for the three factors are reported in Appendix C for individuals and between parents.

Correlations of ATE with the reading achievement measures are shown in Table 12. To obtain these correlations, the co-twins were treated as individuals with the questionnaire data duplicated for the co-twins. Using this method permits the within family variation in the reading tests (same family environment ratings) to be included in the correlations, which will tend to produce a lower correlation than a single-child-parent or average-child-parent analysis. Of course, tests of significance had to be adjusted to account for the lack of independence of the questionnaire data. The conventional t-test was used with the degrees of freedom conservatively adjusted to the number of twin pairs minus 2. Using this method, none of the correlation coefficients were significant ($p < 0.05$).

3.6.2. Moos' Family Environment Scale. Both parents, when available, were asked to complete this questionnaire also; data were available in 67 families for both parents for 1 additional father and 11 additional mothers. The FES subscale intercorrelations are reported in Appendix C.

The correlations of the reading test scores with the FES mother, father, and average parent subscale scores and the incongruence score are shown in Table 13. The same strategy was used as in analysis of the ATE; both co-twins were included with the family environment data duplicated. Once again, the usual t-test was used with the degrees of freedom adjusted to the number of twin pairs minus 2. An interesting pattern was obvious: subscale 6, Intellectual-Cultural Orientation, correlated significantly ($p < 0.05$) with all but the Letter Identification test in the father and the average-parent samples; the correlations for this scale were slightly lower and nonsignificant in the mother sample. Using the average-parent values, from 4% to 9% of the variance in the reading tests could be accounted for by FES subscale 6.

3.6.3. Reading Instruction Approaches. Correlations of teachers' ratings with the WRMT are shown in Table 14; results using ranks rather than actual ratings were quite similar. A positive correlation in the table may be interpreted in the usual way: the higher the rating, the higher the reading achievement score. At the 5% level, only 2 of the 54 coefficients shown were significant, which was approximately equal to the expected number significant ($p < 0.05$) due to chance. It is unlikely, then, that these correlations were meaningful. On the outside chance that a certain combination of these approaches may have been related to reading achievement, stepwise regression was performed. No new information was revealed.

TABLE 12. Correlations of ATE With WRMT

ATE		WRMT (percentiles)					
		LI	WI	WA	WC	PC	Total
Father	I	0.06	0.08	0.02	0.01	0.05	0.05
(N = 142)	II	−0.13	−0.03	−0.03	−0.07	−0.03	−0.03
	III	−0.01	−0.05	−0.01	−0.03	−0.01	−0.02
Mother	I	0.05	0.09	0.12	0.09	0.06	0.09
(N = 158)	II	−0.17	−0.16	−0.13	−0.15	−0.18	−0.17
	III	−0.06	0.03	0.05	0.02	0.05	0.02
Average	I	0.05	0.11	0.07	0.06	0.06	0.08
parent	II	−0.19	−0.12	−0.12	−0.16	−0.14	−0.14
(N = 138)	III	−0.06	−0.05	−0.03	−0.03	−0.03	−0.05

4. DISCUSSION

4.1. Reading Models

The practical significance of this research project depends, in part, on the relevance of the design to theoretical models of reading acquisition. The variable of primary interest, auditory-visual integration, was not chosen in accordance with a strong commitment to any specific reading model; rather, AVI was chosen with the idea that decoding the written word to its verbal counterpart is an important and necessary skill for successful acquisition of reading skill and that AVI plays an important role in successful decoding. No attempt has been made to approach the question of whether this is an age-dependent relationship, such that AVI is important for reading acquisition but not so important once decoding skills are mastered. Grade level variation in this study was not sufficient for such analyses since children at both grade levels examined still seemed to be learning word analysis skills. With this noncommitment to a specific model of reading acquisition in mind, the results are discussed.

4.2. Constituent Factors in Reading Achievement

4.2.1. Intercorrelations of Children's Test Battery. A number of investigators have reported results for the relationships between reading achievement, general intellectual ability, auditory memory, and tasks similar to the ones in the auditory-visual integration test battery used in this study. Only a few have reported results for the first- and second-grade range, however [3, 6, 8, 10, 21, 25, 28, 32, 39, 54, 56].

Consider first the correlation of an auditory-visual integration task with a reading achievement measure (Table 7). The correlations of the AVI task with the five reading tests and the composite reading score from the WRMT ranged from 0.26 to 0.48, with Letter Identification being the only test of this battery with a correlation below 0.40. Table 15 summarizes the information presented in the introduction regarding the relationship between AVI skill and reading achievement in first- and second-grade children from previous studies. Those correlations were generally positive and significant and consistent with the results of the present study. The AVI task used in this study was a slight modification of the task used by Kahn and Birch [25]; the results of the present study closely resembled those of Kahn and Birch. Information regarding analyses of the present sample separated according to grade level may be found in Appendix E. The

*TABLE 13. Correlation of FES With WRMT**

		WRMT (percentiles)					
FES		LI	WI	WA	WC	PC	Total
Father	1	0.13	0.13	0.13	0.10	0.15	0.13
(N = 136)	2	0.17	0.23	0.14	0.17	0.21	0.20
	3	−0.13	−0.04	−0.09	−0.04	−0.01	−0.06
	4	0.05	−0.04	−0.06	−0.07	−0.03	−0.05
	5	0.05	−0.00	0.03	0.03	0.06	0.03
	6	0.19	0.23	0.26	0.24	0.23	0.25
	7	0.03	0.03	0.02	−0.02	0.00	0.01
	8	0.10	0.11	0.21	0.10	0.09	0.13
	9	0.07	0.08	0.04	0.12	0.09	0.09
	10	0.01	0.01	0.05	0.04	0.04	0.05
Mother	1	−0.01	0.14	0.05	0.17	0.10	0.12
(N = 156)	2	−0.01	0.09	0.05	0.12	0.08	0.09
	3	−0.11	−0.16	−0.05	−0.21	−0.17	−0.16
	4	−0.01	−0.01	0.06	0.06	0.02	0.03
	5	−0.14	−0.18	−0.13	−0.16	−0.16	−0.17
	6	0.13	0.16	0.16	0.22	0.18	0.18
	7	−0.04	−0.02	−0.02	−0.00	−0.02	−0.02
	8	−0.03	−0.01	−0.00	−0.00	−0.02	−0.01
	9	0.01	0.10	0.03	0.12	0.10	0.10
	10	0.03	−0.01	0.02	−0.03	−0.01	0.02
Average	1	0.09	0.21	0.15	0.22	0.20	0.21
Parent	2	0.08	0.19	0.08	0.17	0.14	0.16
(N = 134)	3	−0.12	−0.13	−0.09	−0.16	−0.11	−0.15
	4	−0.00	−0.06	−0.05	−0.06	−0.05	−0.05
	5	−0.04	−0.09	−0.05	−0.07	−0.04	−0.07
	6	0.22	0.27	0.28	0.30	0.28	0.29
	7	−0.02	0.03	−0.01	−0.00	−0.01	0.01
	8	0.04	0.05	0.16	0.05	0.06	0.08
	9	0.05	0.10	0.05	0.14	0.12	0.12
	10	0.06	−0.00	0.04	−0.00	0.03	0.04
Incongruence		−0.11	0.07	−0.11	−0.05	−0.05	−0.08

*Underlined values significantly different from zero ($p < 0.05$).

TABLE 14. Correlation of WRMT With Teachers' Ratings of First-Grade Reading Instruction Approaches (N = 101)*

	LI%	WI%	WA%	WC%	PC%	Total %
Directed reading	0.05	0.08	−0.11	0.09	0.07	0.09
Directed reading-thinking	−0.04	−0.02	−0.08	−0.05	−0.09	−0.02
Language experience	−0.05	−0.09	−0.04	−0.08	−0.06	−0.06
Individualized reading	−0.02	0.23	0.11	0.14	0.15	0.17
Linguistic	−0.00	0.16	−0.01	0.15	0.13	0.12
Intensive phonics	0.01	0.08	0.13	0.05	0.12	0.10
Changed alphabet	−0.13	0.07	−0.04	0.01	−0.03	−0.01
Programmed instruction	0.00	0.06	−0.04	0.03	0.12	0.01
Computer-assisted instruction	−0.15	−0.14	−0.21	−0.17	−0.08	−0.18

*Underlined values significantly different from zero ($p < 0.05$).

*TABLE 15. Correlation of AVI Tasks With Reading Achievement**

Study	Reading test(s)	AVI task	Grade(s)	Correlation coefficients
Birch and Belmont, 1965	Metropolitan Readiness Test (grade 1)	10 items (multiple choice type response)	1	0.70
	Stanford Achievement Test (grade 2)		2	0.42
Muehl and Kremenak, 1966	Metropolitan Achievement Tests Reading Subtest	6 items (same-different response)	1	0.52
Kahn and Birch, 1968	Metropolitan Reading Achievement Battery Word Knowledge (WK) Reading Comprehension (RC)	20 items (adapted from Birch & Belmont, 1964)	2	*WK* 0.37 *RC* 0.44
Reilly, 1971	Gates-MacGinitie Reading Test Vocabulary (RV) Comprehension (RC) Total Reading (RT)	20 items (adapted from Birch & Belmont, 1964)	1	*RV* 0.23 *RC* 0.14 *RT* 0.21
			2	*RV* 0.65 *RC* 0.71 *RT* 0.70
Blackman and Burger, 1972	Metropolitan Achievement Tests Word Knowledge	10 items (from Birch & Belmont, 1964)	1	0.11
Whiton et al, 1975	Primary Reading Profiles Word Knowledge (WK) Word Attack (WA) Comprehension (RC) Total (RT)	12 items (from Birch & Belmont, 1964, 1965)	1	*WK* 0.36 *WA* 0.28 *RC* 0.39 *RT* 0.44

*Underlined values not significantly different from zero (p ⩾ 0.05).

results of this study, then, support the notion suggested by previous studies: auditory-visual integration skill has a positive and significant relationship to reading achievement at the first- and second-grade level.

Several studies that reported correlation coefficients for the relationship between reading achievement and general intellectual ability are presented in Table 16. Results from three of the five studies are very similar to those in the present research: the correlation coefficients were significantly different from zero and moderate in magnitude. Findings of the two remaining studies were unusual in that one revealed essentially zero correlation between general intelligence and reading achievement and the other reported a rather high correlation between the two variables, 0.74. Although this study does not include one general "IQ" measure, the results from the Vocabulary and Block Design tests of the WISC-R support the conclusion one can deduce from the literature: there exists a positive and significant relationship between reading achievement and general intellectual ability.

Birch and Belmont [5] reported the correlation of AVI with IQ in first- and second-grade samples, which were 0.56 and 0.42, respectively. The results of the present study are consistent with these and support the idea of a positive and significant relationship between AVI and intellectual ability, although the literature data are obviously limited.

Barker [3] and Hare [21] included measures of auditory memory in their test batteries. Barker reported results for a group of second- through fifth-grade children; correlation coefficients for two reading achievement tasks with memory were 0.48 and 0.46. Hare used the Auditory Sequential Memory task from the ITPA, which was also used in this

*TABLE 16. Correlation of General Intellectual Ability With Reading Achievement**

Study	Reading test(s)	IQ measures	Grade(s)	Correlation coefficients
Bryan, 1964	California Achievement Test Reading Vocabulary (RV) Reading Comprehension (RC)	Kuhlman-Anderson Intelligence Test	1	0.34 (RV) 0.32 (RC)
			2	0.74 (RV) 0.48 (RC)
Birch and Belmont, 1965	Metropolitan Readiness Test (grade 1)	Otis Quick-Scoring Tests of Mental Ability	1	0.56
	Stanford Achievement Test (grade 2)		2	0.53
Machowsky and Meyers, 1975	California Achievement Test Reading	California Test of Mental Maturity	1	0.29
Warren et al, 1975	Cooperative Primary Test	Lorge-Thorndike Intelligence Test	2	0.52
Hare, 1977	Wide Range Achievement Test (reading) Stanford Achievement Test (paragraph meaning)	Lorge-Thorndike Intelligence Test	2	0.00 (WRAT) 0.11 (SAT)

*Underlined values not significantly different from zero ($p \geqslant 0.05$).

study; the correlation coefficients for a second-grade sample were reported to be 0.33 and 0.24 for two reading measures. The results from this study are not consistent with those in the literature and, because of a lack of data for this age group, no definite conclusions can be drawn.

Warren et al [54] and Whiton et al [56] both included an auditory comparison task and the latter also included a visual (spatial) comparison task in their test batteries. Whiton et al [56] tested second-semester, first-grade children and reported positive, nonsignificant correlations between their auditory comparison task and reading achievement measures; the visual comparison task correlated significantly with the total reading and reading comprehension measures (0.29 and 0.25, respectively), but the coefficients for the other reading measures were not significant although similar in magnitude (0.22 and 0.24). Warren et al [54] included second-semester, first- and second-grade children in their study; the correlations of the auditory comparison task with the reading measure were 0.32 and 0.50 (first and second grade, respectively). In the present study, the correlations between auditory comparison and reading achievement were generally low but significant. The differences in the results between studies are most likely due to differences in the ways in which this type of task was administered or difficulty of the items in the test. The visual comparison task in the present study also correlates moderately and significantly with reading achievement.

In summary, the results from this study are consistent with those of previously reported research regarding the interrelationships between reading achievement, auditory-visual integration, and general intellectual ability: a positive and significant relationship exists between these three variables.

4.2.2. Partial Correlations of AVI with Reading Achievement. Because of the interrelationships just described between reading achievement, general intellectual ability, and auditory-visual integration, one could argue that the reason reading achievement and AVI exhibit positive, significant correlation coefficients is their covariance with general intellectual ability. This theory of IQ covariance may be tested by partialling out IQ estimates and then examining the remaining partial correlations between reading achievement and AVI. The results of the present study revealed that, even after partialling out the WISC-R test scores, the correlations between reading achievement and AVI remained significant. This kind of data has not previously been reported for children in the first and second grades. Kahn and Birch [25] reported similar results for grades three through six. These data substantiate the claim that the relationship between reading achievement and auditory-visual integration is independent of general intellectual ability, even at a beginning reading level.

4.2.3. Factor Analyses of the Children's Test Battery. To provide more insight into the interrelationships between the tasks in the children's test battery, factor analysis was performed. To my knowledge, this type of analysis has not been reported for a similar set of variables.[5] These analyses revealed an auditory memory factor separate from the reading achievement factor; auditory memory has been proposed as a skill that "mediates" the significant relationship between auditory-visual integration and reading achievement. For these data, this does not appear to be the case. The variance for auditory-visual integration task tended to be split between the three extracted factors; in this sample, 10% or more of the variance for the AVI task was accounted for by each of the three factors. It is also interesting that a meaningful visual-spatial factor could be extracted, since this skill is also necessary for success on the AVI task. These data, then, support the idea that the integration aspect of this AVI task (auditory-visual and/or temporal-spatial integration) is related to reading achievement.

4.2.4. Environmental Variables. Analysis of the three factors from the ATE questionnaire yielded no significant correlations with reading achievement, which was similar to the results reported by Garfinkle [20] for her previously described test battery. In the present study, however, the pattern of correlations suggested that ATE factor II (Parental Participation) may be related to reading achievement. It is interesting that two of the four items assigned to this factor deal with the importance of reading as a pasttime and the parents' participation in reading to the twins. This suggests that parents' positive attitudes toward reading may promote, to a limited extent, achievement in reading.

The correlation matrix for the FES was a bit more revealing. A consistent pattern of a positive (and usually significant) relationship between Intellectual-Cultural Orientation (subscale 6) and reading achievement emerged, accounting for approximately 4–9% of the variance on each reading task. This finding was not inconsistent with those of Garfinkle [20]. In her study, Moos' Intellectual-Cultural Orientation subscale correlated positively and significantly with a mathematical concepts battery and the Peabody Picture Vocabulary Test (0.16 and 0.20, respectively). The findings of this study, then, are consistent with the idea that exposure to and emphasis on intellectual and cultural activities will promote, to a limited extent, achievement in reading.

More striking results might have been found for these environmental measures, especially for the parental questionnaires, if the sample had been more representative (eg, regarding socioeconomic status, education); it is difficult to know what effect restriction of the variation in SES and education has on questionnaires such as these. Garfinkle [20]

---

[5] Blackman and Burger [8] reported factor analysis results for a set of reading readiness variables and subsequent reading achievement. This included an auditory-visual integration task, but many other tasks were not at all similar to ones used in this study so a comparison will not be attempted.

reported correlations of 0.47 and 0.46 for Moos' Intellectual-Cultural Orientation subscale with father's and mother's education, respectively. No family data were available in this study to evaluate such a relationship, but it is possible that the positive relationship between the Intellectual-Cultural Orientation subscale of the Moos' FES and reading achievement as assessed in the present study may be mediated by a common relationship with parent's education or socioeconomic status.

The correlation matrix for the teacher's ratings of first-grade reading instruction approaches and the children's reading achievement scores revealed no consistent relationship. This is in agreement with the ideas expressed in teachers' textbooks [12], that successful acquisition of reading skills may depend more upon the skill of the teacher than the actual reading instruction methods used.

In summary, these data support the notion that family attitudes may affect achievement in school, although the effect in this sample was rather small. This study, obviously, included only a subsample of the environmental factors that could affect reading achievement and the results, in that respect, are rather encouraging.

4.3. Genetic Contributions to Variation in the Children's Test Battery

4.3.1. Reading Achievement. Previous twin investigations provide ample evidence of a genetic contribution to variation in reading achievement (Table 3); in fact, the twin pair intraclass correlation coefficients are quite similar among these studies. This is rather remarkable considering the wide age range reported (8–20 years), the two cultures sampled, and the varied measures of reading achievement used. Several studies included various grade levels of children, but did not report results separately according to grade (or age); this is rather regrettable, but was probably due to small sample sizes. None of these studies reported results separately for beginning grade school children (first and second graders); in fact, it is unlikely that any of the studies included children just completing the first grade. The results from this study for the first- and second-grade combined sample were quite comparable to those in the literature, compatible with substantial familial aggregation and the presence of genetic variance. The DZ intraclass correlations were greater than expected on the basis of a genetic hypothesis and the similarity of the MZ co-twins (except for the Word Attack Test); this was also true for the literature reports reviewed and is probably due to assortative mating, if not for reading achievement per se, for some related variable(s) such as occupation or educational level.

4.3.2. WISC-R Tests. The Vocabulary and Block Design tests of the WISC-R were administered, in part, to be able to assess representativeness of the sample. The types of tasks represented by these tests usually exhibit significant genetic variance [40, 41]. The results for the Vocabulary test were not unreasonable; both MZ and DZ co-twins exhibited significant similarity and results of the twin analysis suggested the presence of genetic variance. The results for the Block Design test were not so clear cut. The DZ co-twins showed very little similarity. For some unknown reason, this study seemed to have a greater than average number of markedly discordant DZ pairs: 6 of 42 pairs differed by more than two standard deviations (> 6 scaled score units). Wilson (personal communication) noted a similar pattern of intraclass correlations for the WISC Block Design test in a sample of twins tested at age 8 years; however, data available on an additional 55 sets of DZ twins who were administered the WISC-R at 8 years of age revealed a significant intraclass correlation coefficient (0.38) for the Block Design test. Thus, the results of the present study for WISC-R Vocabulary were rather typical, while those for Block Design were rather worrisome in consideration of a genetic model but perhaps not so unusual for the age group tested.

4.3.3. Auditory Memory. The results of twin studies of auditory memory seem to be somewhat dependent on the measure used: Strandskov [48] and Vandenberg [50] both administered the Primary Mental Abilities Test to adolescent twins and reported nonsignificant F-ratios for the within-pair mean squares. Block [9], Pezzullo et al [38], and Vandenberg [50] administered the WISC Digit Span test (or a similar, but modified, version) to adolescent and preadolescent twins, and all reported significant F-ratios for the within-pair mean squares. This study, of course, involved much younger children but the ITPA Auditory Sequential Memory is similar to the WISC Digit Span; the WISC test includes both digits forward and digits backward while the ITPA measure involves only digits forward. The results of this twin analysis supported the findings of significant genetic variance reported in the studies using the WISC measure. The similarity of the DZ co-twins was greater than expected on the basis of the MZ intraclass correlations and the presence of genetic variance. Such inflation usually suggests assortative mating, but this explanation is rather difficult to imagine for auditory memory; auditory memory exhibited only low correlations with reading achievement and the WISC-R Vocabulary test, which are susceptible to such effects.

4.3.4. Auditory-Visual Integration Test Battery. These twin analyses represent the first of their kind. As previously mentioned, the results for the visual comparison task revealed no evidence of significant genetic variance; the same was true for the auditory comparison task, which suffered from low reliability. The results for the auditory-visual integration task were much more interesting and suggested a genetic contribution to the variation in the AVI task. It is unknown whether this is a reflection of the integration skills required or merely of related variables which also exhibit genetic variance (eg, auditory memory, general intellectual ability). Nonetheless, it is notable that the variation in this lower level skill contains a heritable component.

4.3.5. Factors Derived from the Children's Test Battery. The results from the analyses of factors 1 and 2, suggestive of the presence of genetic variance, were not surprising since these factors represent reading achievement and auditory memory, respectively, in the present sample, which have already been shown to exhibit heritable variation. The analysis of factor 3, which seemed to represent some type of visual-spatial ability, also suggested the presence of genetic variance but with no apparent assortative mating. This was typical of the pattern described for most Block Design analyses [40]. In view of the results obtained for the Block Design test alone in this study, it was interesting that this factor (on which Block Design loaded highly along with Visual Comparison) revealed a more typical pattern of genetic variance.

4.3.6. Summary. The present study supports evidence from the literature that a heritable component contributes to individual variation in reading achievement, auditory memory, visual-spatial ability, and verbal IQ, even during the beginning school years. This study presents the first reported evidence of the contribution of genetic variance to auditory-visual integration, a "lower-level" skill related to reading achievement.

5. CONCLUSIONS

Previous twin studies provide evidence for genetic contributions to individual differences in reading achievement, but the nature of these effects is uncertain. The main thrust of this project was to investigate possible lower-level sources of this heritable variation, particularly auditory-visual integration, in young readers. The following conclusions may be drawn:

1. Individual variation in reading achievement is influenced by genetic factors in first- and second-grade twin children, which has been observed in older children also.
2. In this sample, a statistically significant relationship exists between reading achievement and auditory-visual integration, independent of general intellectual ability.
3. The integration aspect of auditory-visual integration task (auditory-visual and/or spatial-temporal) was the aspect of this skill related to reading achievement rather than auditory memory or visual-spatial ability, which also contribute to success on the AVI task.
4. Individual variation in auditory-visual integration seems to be influenced by genetic factors in first- and second-grade twin children, with the reservation that the intraclass correlation coefficient for the DZ twins is only marginally significant in this sample.
5. Because AVI was found to be related to other skills exhibiting genetic variance, it is unclear whether the aspects of the AVI task related to reading achievement alone produced the evidence of heritable variation.
6. The only environmental measure that showed a significant correlation with reading achievement in these young children was intellectual-cultural orientation of the family.

This study substantiates the results from previous twin studies with older children that genetic factors play an important role in individual variation in reading achievement. Since reading is such a complex behavior, studies of more basic skills or underlying factors may help in elucidating the sources of heritable variation. However, a substantial amount of the variance cannot be accounted for by genetic factors and, therefore, research on specific environmental influences will also be important in elucidation of sources of individual variation. Perhaps by studying the etiology of individual differences within the normal range, we will learn more about abnormal variation.

ACKNOWLEDGMENTS

I would like to express my appreciation to my advisor and friend, Dr. Richard Rose, for his guidance and support during my training at Indiana University. I am grateful for his encouragement during my research and the writing of this dissertation. I consider myself fortunate to have been one of his students.

I am grateful to Drs. Joe Christian, Michael Conneally, and Leo Fay for their helpful criticisms and comments in the preparation of this manuscript and to Dr. James Norton for his expert statistical advice and his suggestions for this manuscript. I am also grateful to Dr. Pao-lo Yu for his "emergency" statistical consultations.

This project could not have been completed without the interest and cooperation of a special group of people: twins and their parents. I am grateful to them and to the parents who permitted me to test their singleton children during the development of my test battery.

A very special thanks is deserved by the typist of this dissertation, Debbie Crismon, for her superb typing of this manuscript and her patience and speed during preliminary drafts.

Immeasurable thanks go to many friends and colleagues for transmitting their enthusiasm for learning and offering their guidance throughout my years at Indiana University. I am especially grateful to Jo Boughman, who helped introduce me to the world of behavior genetics, to Marilyn Dumont-Driscoll, with whom I grew in behavior genetics, and to Dan Driscoll, Vicki Veigl, Stuart Schwartz, and Pam and Dave Weaver for their special friendships and their support.

Finally, my family deserves much credit and many thanks for their unwavering support and faith in me.

REFERENCES

1. Badian NA (1974): Auditory-visual integration, auditory memory and reading behavior. EdD dissertation, Boston University School of Education.

2. Badian NA (1977): Auditory-visual integration, auditory memory, and reading in retarded and adequate readers. J Learn Disabil 10:108–114.
3. Barker BM (1976): Interrelationships of perceptual modality, short-term memory and reading achievement. Percept Mot Skills 43:771–774.
4. Birch HG (1962): Dyslexia and the maturation of visual function. In J Money (ed): "Reading Disability: Progress and Research Needs in Dyslexia" Baltimore: Johns Hopkins Press.
5. Birch HG, Belmont L (1964): Auditory-visual integration in normal and retarded readers. Am J Orthopsychiatry 34:852–861.
6. Birch HG, Belmont L (1965): Auditory-visual integration, intelligence and reading ability in school children. Percept Mot Skills 20:295–305.
7. Blank M, Bridger WH (1966): Perceptual abilities and conceptual deficiencies in retarded readers. Proc Am Psychopathol Assoc 56:401–412.
8. Blankman LS, Burger AL (1972): Psychological factors related to early reading behavior of EMR and nonretarded children. Am J Ment Defic 77:212–229.
9. Block J (1968): Hereditary components in the performance of twins on the WAIS. In Vandenberg SG (ed): "Progress in Human Behavior Genetics" Baltimore: The Johns Hopkins Press.
10. Bryan QR (1964): Relative importance of intelligence and visual perception in predicting reading achievement. Calif J Educ Res 15:44–48.
11. Bryant PE (1975): Cross-modal development and reading. In Duane D, Rawson MB (eds): "Reading, Perception and Language." Baltimore: York Press, Inc.
12. Burns PC, Roe BD (1976): "Teaching Reading in Today's Elementary Schools." Chicago: Rand McNally College Publishing Company.
13. Calfee RC (1975): Memory and cognitive skills in reading acquisition. In Duane DD, Rawson MB (eds): "Reading, Perception and Language." Baltimore: York Press, Inc.
14. Christian JC, Kang KW, Norton JA Jr (1974): Choice of an estimate of genetic variance from twin data. Am J Hum Genet 26:154–161.
15. Christian JC, Feinleib M, Norton JA Jr (1975): Statistical analysis of genetic variance in twins. Am J Hum Genet 27:807.
16. Christian JC, Norton JA Jr (1977): A proposed test of the difference between the means of monozygotic and dizygotic twins. Acta Genet Med Gemellol 26:49–53.
17. Cooley WW, Lohnes PR (1971): "Multivariate Data Analysis" New York: John Wiley & Sons, Inc.
18. DeFries JC, Singer SM, Foch TT, Lewitter FI (1978): Familial nature of reading disability. Br J Psychiatry 132:361–367.
19. Doehring DG (1968): "Patterns of Impairment in Specific Reading Disability." Indiana University Publications: Science Series, No. 23 Bloomington: Indiana University Press.
20. Garfinkle AS (1981): Genetic and environmental influences on the development of Piagetian logico-mathematical concepts and other specific cognitive abilities: A twin study. Acta Genet Med Gemellol 30:000–000.
21. Hare BA (1977): Perceptual deficits are not a cue to reading problems in 2nd grade. Reading Teacher 30:624–628.
22. Husen T (1959): "Psychological Twin Research. I. A Methodological Study." Stockholm: Almqvist & Wiksell.
23. Husen T (1960): Abilities of twins. Scand J Psychol 1:125–135.
24. Jarvik LF (1975): Human intelligence: Sex differences. Acta Genet Med Gemellol 24:189–211.
25. Kahn D, Birch HG (1968): Development of auditory-visual integration and reading achievement. Percept Mot Skills 27:459–468.
26. Kirk SA, McCarthy JJ, Kirk WD (1968): "Examiner's Manual: Illinois Test of Psycholinguistic Abilities," revised edition Urbana, Illinois: University of Illinois Press.
27. Maccoby EE, Jacklin CN (1974): "The Psychology of Sex Differences" Stanford, California: Stanford University Press.
28. Machowsky H, Meyers J (1975): Auditory discrimination, intelligence, and reading achievement at grade 1. Percept Mot Skills 40:363–368.
29. Margolis H (1977): Auditory perceptual test performance and the reflection-impulsivity dimension. J Learn Disabil 10:164–172.
30. Matheny AP Jr, Dolan AB (1974): A twin study of genetic influences in reading achievement. J Learn Disabil 7:99–102.
31. Moos RH (1974): "Preliminary Manual for Family Environment Scale, Work Environment Scale, and Group Environment Scale." Palo Alto, California: Consulting Psychologists Press, Inc.

32. Muehl S, Kremenak S (1966): Ability to match information within and between auditory and visual sense modalities and subsequent reading achievement. J Educ Psych 57:230–239.
33. National Center for Health Statistics (1971): Intellectual development of children as measured by the Wechsler Intelligence Scale, United States. "Vital and Health Statistics" PHS Pub No. 1,000, Series 11, No. 107 Public Health Service. Washington: US Government Printing Office.
34. Newman HH, Freeman FN, Holzinger KJ (1937): "Twins: A Study of Heredity and Environment" Chicago: University of Chicago Press.
35. Ounsted C, Taylor DC (1972): "Gender Differences: Their Ontogeny and Significance" London: Churchill Livingstone.
36. Paraskevopoulos JN, Kirk SA (1969): "The Development and Psychometric Characteristics of the Revised Illinois Test of Psycholinguistic Abilities." Urbana, Illinois: University of Illinois Press.
37. Pelham WE, Ross AO (1977): Selective attention in children with reading problems: A developmental study of incidental learning. J Abnorm Child Psychol 5:1–8.
38. Pezzullo TR, Thorsen EE, Madaus GF (1972): The heritability of Jensen's Level I and II and divergent thinking. Am Educ Res J 9:539–546.
39. Reilly DH (1971): Auditory-visual integration, sex and reading achievement. J Educ Psychol 62:482–486.
40. Rose RJ, Harris EL, Christian JC, Nance WE (1979): Genetic variance in nonverbal intelligence: Data from the kinships of identical twins. Science 205:1153–1155.
41. Rose RJ, Boughman JA, Corey LA, Nance WE, Christian JC, Kang KW (1980): Data from kinships of monozygotic twins indicate maternal effects on verbal intelligence. Nature 283:375–377.
42. Rourke BP (1976): Reading retardation in children: Developmental lag or deficit? In Knights RM, Bakker DJ (eds): "The Neuropsychology of Learning Disorders: Theoretical Approaches." Baltimore: University Park Press.
43. Samuels SJ, Anderson, RH (1973): Visual recognition memory, paired-associate learning, and reading achievement. J Educ Psychol 65:160–167.
44. Samuels SJ, Turnure JE (1974): Attention and reading achievement in first-grade boys and girls. J Educ Psychol 66:29–32.
45. Smith SM, Penrose LS (1954–55): Monozygotic and dizygotic twin diagnosis. Ann Hum Genet 19:273–289.
46. Specht DA (1978): "Users Guide to Subprogram Reliability and Repeated Measures Analysis of Variance." Wrubel Computing Center Report #352.
47. Specht DA, Hohlen M (1978): "SPSS Subprogram Reliability: Item and Scale Analysis." Wrubel Computing Center Report #351.
48. Strandskov H (1955): Some aspects of the genetics and evaluation of man's behavioral characteristics. Eugenics Quarterly 2:152–161.
49. Torgensen J (1975): "Problems and Prospects in the Study of Learning Disabilities." Chicago: The University of Chicago Press.
50. Vandenberg SG (1962): The hereditary abilities study: Hereditary components in a psychological test battery. Am J Hum Genet 14:220–237.
51. Vande Voort L, Senf GM (1973): Audiovisual integration in retarded readers. J Learn Disabil 6:170–179.
52. Vande-Voort L, Senf GM, Benton AL (1972): Development of audio-visual integration in normal and retarded readers. Child Dev 43:1260–1272.
53. Vernon MD (1971): "Reading and Its Difficulties: A Psychological Study." London: Cambridge University Press.
54. Warren DH, Anooshian LJ, Widawski MH (1975): Measures of visual-auditory integration and their relations to reading achievement in early grades. Percept Mot Skills 41:615–630.
55. Wechsler D (1974): "Manual for the Wechsler Intelligence Scale for Children-Revised." New York City: The Psychological Corporation.
56. Whiton MB, Singer DL, Cook H (1975): Sensory integration skills as predictors of reading acquisition. J Read Behav 7:79–89.
57. Wictorin M (1952): Bidrag till raknefardighetens psykolog. En tvilling undersokning. Goteborg: PhD thesis. Cited in Husen T (1959): "Psychological Twin Research: A Methodological Study." Stockholm: Almqvist & Wiksells.
58. Woodcock RW (1973): Woodcock Reading Mastery Tests Manual. Circle Pines, Minnesota: American Guidance Service, Inc.

Correspondence: Dr. Emily L. Harris, Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA 15261.

Appendix A: Zygosity Determination and Distribution According to Sex and Grade

Appendix A contains additional information regarding methods of zygosity determination and the distribution of zygosity type according to sex and grade.

Zygosity of the like-sexed twin pairs was determined through various combinations of the following: similarity of physical appearance (general appearance, eye color, hair color, height, weight), genetic markers in blood, urine and saliva (ABO, Rh, MNS, Kell, Duffy, P, Kidd, Hp, AcP, PGM, AKl, ADA, Gc, Amy_1, Amy_2, the salivary proteins: Pr, Pa, Pb, and Db, and Secretor), and dermatoglyphics as well as placental information when available and comments by parents and twins about "confusibility." The zygosity of the 20 unlike-sexed twin pairs was based on the sex difference. Table A1 shows the distribution of the zygosity determination methods and zygosity distribution for the 88 like-sexed twin pairs. Genetic markers (primarily in the blood) were used as the major criteria for zygosity determination in 59 of the 88 like-sexed twin pairs; discordance for at least one marker constituted the criterion for a dizygotic classification. (If a pair was discordant for only one marker, that marker was rerun.) A minimum of 7 and a maximum of 18 markers were examined for each pair; the median number was 13. For the pairs that are concordant on all the markers examined, a probability of dizygosity (or monozygosity) may be calculated. The calculations by Smith and Penrose [1955] for individual blood groups were used as the basis for the calculations reported here; seven markers were included in both the Smith and Penrose article and this study: ABO, Rh, MNS, P, Kell, Duffy and Kidd. The maximum probability of like-sexed co-twins being dizygotic with concordance for these seven blood groups (given no parental information) is approximately 8%. The actual range in this sample was from less than 0.1% to 7.3%; in some cases, information on one or both parents was available and was included in the probability calculations.

In eight cases, placental information (from a pathology report) was diagnostic of monozygosity. Seven of these twin pairs also had blood typing done; all results were consistent with the classification of monozygotic twins.

Dermatoglyphics were used as the primary method of zygosity determination when blood typing was not available or when placental information was not diagnostic; this included nine cases. Seven were classified as monozygotic; the probability of dizygosity based on dermatoglyphic differences alone was less than or equal to 15%, and physical appearance of the co-twins was very similar in these cases. The two sets classified as dizygotic were also dissimilar in appearance.

Due to the lack of other data, 13 pairs had to be classified according to physical appearance similarities or differences (eg, eye color, hair color, curliness of hair, general facial appearance). Five twin pairs were obviously dissimilar and were classified as dizygotic; seven sets were extremely similar and still mistaken for their co-twin by friends and, on occasion, by close family members and were classified as monozygotic. One set remained unclassified and was not included in the twin analyses.

In summary, then, the chance of misclassification is small. The greatest risk is that of classifying a dizygotic pair as monozygotic; it is unlikely that any MZ pairs have been classified as DZ.

The distribution of zygosity classification according to sex and grade is presented in Table A2.

Appendix B: WISC-R Test-Scale Correlations and Reliability Estimates for Children's Test Battery

Appendix B contains the WISC-R test-scale correlations for the Vocabulary and Block Design tests and the previously reported reliability estimates for the tasks in the children's test battery, except the AVI test battery. Estimates of internal consistency, based on the twin data available from the present study, are reported for the AVI battery.

All of the tests administered for this study involve some degree of error in measurement. Through reliability estimates, one attempts to measure, on a population level, how accurately an individual's score on a test

TABLE A1. Distributions of Primary Method of Determining Zygosity and Zygosity for Like-Sexed Twin Pairs

Placental information	Blood, urine saliva		Dermatoglyphics		Appearance, confusibility		Total	
MZ	MZ	DZ	MZ	DZ	MZ	DZ	MZ	DZ
8	35	24	7	2	7	5	57	31

TABLE A2. Sex, Zygosity, and Grade Distribution

	MZ		Like-sexed	DZ		Unknown
	Male	Female	Male	Female	Unlike-sexed	Female
Both grade 1[a]	13	22	9	10	12	1
Both Grade 2[b]	8	13	4	8	7	—
Grade 1, grade 2	—	1	—	—	1	—

[a]Actual grade range = 1.7–2.0.
[b]Actual grade range = 2.7–2.9.

estimates his or her true score. If there is no error of measurement, the reliability coefficient will be one; if all of the variation in the observed scores is due to errors of measurement, the reliability coefficient will be zero. The reliability of most tests lies somewhere in between.

WECHSLER INTELLIGENCE SCALE FOR CHILDREN-REVISED

Table B1 presents the test-scale correlation coefficients for the Vocabulary and Block Design tests. These two tests were chosen from the entire WISC-R because of their relatively high correlations with their respective scales and the Full Scale estimate of intellectual ability across the age range included in the present research. Also, reliability estimates from the normative study were relatively high for the Vocabulary and Block Design tests, as seen in Table B2.

WOODCOCK READING MASTERY TESTS

The most remarkable reliability data from this test battery are the internal consistency measures and the test-retest alternate-form reliabilities reported for the Woodcock Reading Mastery Tests, presented in Table B3. Certainly one advantageous characteristic of this battery is the lack of opportunity for successful guessing. This reading achievement measure is certainly more reliable than most, if not all, of the measures used in previous twin studies. When the reliability estimates of the WRMT are compared with those of another well-standardized, individually administered task, the WISC-R, only the combined test scores (Verbal, Performance, and Full Scales) have reliabilities comparable to those of the individual reading achievement tests (except Letter Identification, which has lower reliabilities in both grades).

ITPA AUDITORY SEQUENTIAL MEMORY

The reliability estimates for the auditory memory task are presented in Table B4. The estimates were relatively high and comparable to those reported for the two WISC-R tests used in the present study.

*TABLE B1. WISC-R Correlations**

	6½ years			7½ years			8½ years		
	Verbal scale	Performance scale	Full scale	Verbal scale	Performance scale	Full scale	Verbal scale	Performance scale	Full scale
Vocabulary	0.66^a	0.53	0.65^a	0.72^a	0.51	0.67^a	0.71^a	0.51	0.66^a
Block Design	0.59	0.66^a	0.67^a	0.58	0.66^a	0.68^a	0.50	0.62^a	0.61^a

*From WISC-R Manual [55: pp 36–38].

[a] After correction for contamination (when a test was correlated with a scale of which it was a contributing member).

*TABLE B2. WISC-R Reliability Estimates**

	Split-half			Test-retest
Age (in years)	6½	7½	8½	6½–7½
Vocabulary	0.74	0.70	0.86	0.68
Block Design	0.80	0.82	0.85	0.78

*From WISC-R Manual [55: pgs 28, 32].

*TABLE B3. Woodcock Reading Mastery Tests Reliability Estimates**

	Split-half reliabilities (corrected by Spearman-Brown formula)		Test-retest alternate form reliabilities
	Pre-A^a grade 1.9	A grade 2.9	grade 2.9
Letter Identification	0.55	0.79	0.84
Word Identification	0.99	0.99	0.94
Word Attack	0.95	0.97	0.90
Word Comprehension	0.96	0.88	0.90
Passage Comprehension	0.96	0.95	0.88
Total Reading	0.99	0.99	0.97

*From WRMT Manual [58 pgs 57,58].

[a] Prepublication form of WRMT—very similar to the final Form A.

*TABLE B4. ITPA Auditory Sequential Memory Reliability Estimates**

	Internal consistency	Test-retest
Ages 6–7/7–1 years	0.81 (0.85)	
7–7/8–1 years	0.90 (0.92)	
Age 8 years		0.86 (0.89)

*Paraskevopoulos and Kirk [36]; estimates calculated from sample with restricted intelligence range; values in parentheses represent estimates corrected for that restricted range.

TABLE B5. Auditory-Visual Integration Test Battery Reliability Estimates

	Visual comparison		Auditory comparison		Auditory-visual integration	
	KR-20[a]	Split-half[b]	KR-20[a]	Split-half[b]	KR-20[a]	Split-half[b]
Grade 1 (n = 136)	0.634	0.679	0.067	0.134	0.668	0.657
Grade 2 (n = 82)	0.620	0.737	0.393	0.381	0.795	0.791
Total sample (n = 218)	0.649	0.708	0.245	0.278	0.760	0.749

[a] Kuder-Richardson Formula 20.

[b] With Spearman-Brown correction; split was odd-even.

AUDITORY-VISUAL INTEGRATION TEST BATTERY

Reliability estimates for the auditory-visual integration test battery are reported in Table B5. In the present study, only one score was available for each individual on a test, so estimates of internal consistency were calculated: Kuder-Richardson 20, which is Cronbach's alpha applied to dichotomous data, and the Spearman-Brown correction of the split-half coefficient. The split-half coefficient is the correlation between two parts of a test, odd items and even items in the present study; from this correlation, the Spearman-Brown correction predicts the reliability of the whole test (the two parts combined). Cronbach's alpha is based on analysis of variance and is calculated by comparing the average item variance with the variance of the sum over all the items [46]. Results are reported for the separate grades and the combined sample. All coefficients were significantly different from zero ($p<0.01$), except those for the auditory comparison task in the grade one sample.

The internal consistencies for the visual comparison and auditory-visual integration tasks were comparable to those for the Vocabulary and Block Design tests of the WISC-R. The low reliabilities for the auditory comparison task, especially in the first-grade sample, reflects the difficulty of this task for this age range and the design of the task which permits, and even encourages, guessing when the child does not know the correct answer. Whiton et al [56] reported internal consistencies for a similar battery of tasks that each consisted of 12 items and were administered to second-semester, first-grade males. The reliabilities (Kuder-Richardson Formula 21) were similar in magnitude to those reported in the present study, except in the case of the auditory comparison task. The coefficients for the visual comparison, auditory comparison, auditory-visual, and visual-auditory tasks were 0.63, 0.68, 0.60, and 0.65, respectively. It is difficult to speculate on the origin of the difference in the reliabilities of the auditory comparison task used by Whiton et al [56] and the one used in the present study: the administration of the tasks seemed very similar and the ages were comparable (using the present study's grade one sample); the items used in the present research probably increased in difficulty more quickly.

One might be tempted to exclude the auditory comparison from further analyses on the basis of the nonsignificant coefficients in the first-grade sample. However, that variable was retained for comparative purposes in view of the significant, albeit marginal, coefficients in the second-grade sample; but the results should be interpreted with great caution, in view of the very low reliabilities.

Appendix C: Environmental Questionnaires

Appendix C contains a copy of the questionnaire completed by the children's first- and second-grade teachers regarding methods of reading instruction used. Results of the factor analysis of the Attitudes Toward Education (ATE) questionnaire are also included, as well as the intraparent and interparent correlations for the subscales or factors of the Moos' Family Environment Scale (FES) and the ATE.

ATTITUDES TOWARD EDUCATION QUESTIONNAIRE

Factor Analysis.

The ATE questionnaire data were submitted for factor analysis in an attempt to replicate the findings of Garfinkle [20]. The Varimax rotated factor matrix is shown in Table C1. In general, the patterns of the factor loadings were fairly comparable; asterisked items in Table 1 represent items that had a similar pattern of factor loadings to those reported by Garfinkle. The underlinings indicate to which factor each item was assigned by Garfinkle.

Factor Intercorrelations.

Table C2 presents factor intercorrelations for the mothers and fathers separately as well as the mother-father factor correlations. The correlation coefficients between factors II and III were moderate and significantly different from zero in the mother and father samples. In the mother-father correlations, only the coefficients for the same-factor relationships were significantly different from zero.

MOOS' FAMILY ENVIRONMENT SCALE

The FES subscale intercorrelations are presented in Table C3. The sample was divided by sex; mothers results are shown in the lower left portion of the table and results for fathers are in the upper right portion. Comparison of these two patterns of correlations revealed substantial similarity with few significant deviations. The mother-

TABLE C1. Factor Analysis of ATE (N = 154)

	Factors		
Item	I	II	III
1.*	0.01	0.46	0.16
2.*	0.01	0.12	0.74
3.*	−0.00	0.31	0.20
4.*	0.69	0.03	0.01
5.	−0.05	0.35	−0.27
6.*	−0.11	0.37	0.36
7.*	0.69	−0.09	−0.12
8.*	−0.01	0.07	0.52
9.*	0.07	0.62	0.06
10.*	−0.22	0.58	−0.03
11.*	0.73	−0.02	0.13
12.*	0.17	−0.17	0.70
13.	−0.03	−0.61	0.01
14.	0.05	0.34	0.48
15.*	0.72	−0.01	0.11

*Results of factor analysis consistent with that of Garfinkle [20] for asterisked items; underlined factor loadings indicate to which factor that item was assigned in the study of Garfinkle.

father subscale correlations are shown in Table C4. The correlations on the main diagonal revealed that the parents' perceptions of the environment assessed by this questionnaire were somewhat similar. The off-diagonal correlations were similar in pattern to those reported in the previous table for subscale correlations of individuals.

Brief Descriptions of Reading Instruction Approaches for Use With the Enclosed Questionnaire:

1. Basal reader approach
   a. Directed reading approach (DRA)—The DRA is a strategy utilized by teachers who follow the lesson plans or manuals usually included with a basal reader series. This includes directed silent and oral reading of a story, skill building activities, and follow-up practice and enrichment activities.
   b. Directed reading-thinking activity (DRTA)—The DRTA is a more general strategy for encouraging the children to think as they read either the basal reader or content area selections. The children are encouraged to develop questions or hypotheses about the material and then evaluate them after reading the material.
2. Language experience approach—The language experience approach uses the experiences of the children as a basis for reading materials through the experience chart. The teacher transcribes stories related by the children onto paper for the children to read.
3. Individualized reading approach—The individualized reading approach is designed to foster independent reading and includes self-selection of reading material, self-pacing, and much independent work. Skills are developed with the help of the teacher as needed.
4. Linguistic approaches—In linguistic approaches, beginning readers generally are presented material in

TEACHER'S QUESTIONNAIRE

Child's Name ________________________ Grade ____________

Please rate the contribution of the following approaches used in the child's reading program for the above-designated grade (descriptions available on the following page). For each approach, please circle the appropriate number.

	None	Little	Moderate	Strong	Very Strong
Basal reader approach					
Directed reading approach	1	2	3	4	5
Directed reading-thinking activity	1	2	3	4	5
Language experience approach	1	2	3	4	5
Individualized reading approach	1	2	3	4	5
Linguistic approach	1	2	3	4	5
Intensive phonics approach	1	2	3	4	5
Changed alphabet approach	1	2	3	4	5
Systems approach					
Programmed instruction	1	2	3	4	5
Computer-assisted instruction	1	2	3	4	5
Other (please describe below)	1	2	3	4	5

The person completing this questionnaire was or is (please check one):

______ child's teacher for above-designated grade
______ other teacher
______ reading specialist
______ principal
______ other (please specify below)

*TABLE C3. FES Subscale Intercorrelations**

		Above diagonal: Father (N = 68) 1	2	3	4	5	6	7	8	9	10
Below	1		0.44	−0.29	0.19	−0.03	0.36	0.30	0.15	0.34	−0.19
diagonal:	2	0.32		−0.00	0.39	−0.08	0.37	0.12	0.05	0.05	−0.55
mother	3	−0.50	−0.20		−0.07	0.25	−0.11	−0.02	−0.19	−0.28	0.21
(N = 78)	4	0.13	0.23	−0.08		0.04	0.22	0.25	−0.22	0.02	−0.40
	5	−0.00	−0.20	0.18	0.01		0.11	0.17	−0.17	0.16	0.24
	6	0.24	0.29	−0.10	0.14	0.06		0.32	0.11	0.36	−0.09
	7	0.36	0.27	−0.14	0.31	0.19	0.39		−0.07	0.09	−0.07
	8	0.10	−0.09	−0.21	−0.15	0.18	0.01	−0.25		0.26	0.13
	9	0.47	−0.02	−0.34	0.05	0.26	0.08	0.13	0.29		0.19
	10	0.03	−0.36	0.07	−0.14	0.48	−0.15	−0.08	0.14	0.24	

*Underlined values significantly different from zero ($p < 0.05$).

*TABLE C4. FES Intrafamily Subscale Correlations (N = 67)**

		Father 1	2	3	4	5	6	7	8	9	10
Mother	1	0.49	0.34	−0.21	0.13	−0.04	0.25	0.09	0.15	0.30	−0.13
	2	0.05	0.50	−0.04	−0.01	−0.25	0.32	0.03	0.06	−0.16	−0.45
	3	−0.25	−0.06	0.59	−0.20	0.13	−0.13	−0.09	−0.20	−0.19	0.02
	4	−0.06	0.28	−0.04	0.35	−0.04	0.10	0.11	−0.21	−0.03	−0.38
	5	−0.11	0.03	0.29	0.09	0.52	0.14	0.14	−0.14	0.13	0.12
	6	0.22	0.33	−0.09	0.13	−0.02	0.69	0.25	0.10	0.15	−0.12
	7	0.41	0.25	−0.07	0.23	0.16	0.29	0.69	−0.16	0.21	−0.18
	8	0.05	−0.18	−0.09	−0.26	0.04	0.08	−0.25	0.78	0.17	0.12
	9	0.33	0.09	−0.23	0.07	0.25	0.31	0.17	0.27	0.76	0.17
	10	−0.03	−0.24	0.10	−0.08	0.41	−0.19	−0.04	0.14	0.32	0.50

*Underlined values significantly different from zero ($p < 0.05$).

*TABLE C2. ATE Intercorrelations**

		Mother			Father	
		I	II	III	I	II
Father	I	0.42	0.05	−0.13		
	II	−0.17	0.37	0.20	−0.01	
	III	−0.16	0.12	0.59	0.03	0.38
Mother	I		0.03	0.11		
	II			0.25		

*N = 71 for fathers and N = 79 for mothers; N = 69 for mother-father correlations. Underlined values significantly different from zero ($p < 0.05$).

which each letter has only one sound. Word attack skills are taught by presenting words that vary only by a single letter in their spellings; however, the sounds of the letters are not isolated from the word.

5. Intensive phonics approaches—Phonics approaches emphasize attacking words through the sounds of their individual letters. Either a sight vocabulary may first be developed with phonic analysis of these words leading to generalizations for other words or sounds of individual letters may first be taught with subsequent blending to produce words.
6. Changed alphabet approaches—The purpose of this approach is to obtain a one-to-one sound-symbol relationship. This may be accomplished through adding symbols, deleting letters, modifying existing letters, or printing the traditional alphabet in various colors.
7. Systems approaches
   a. Programmed instruction—This type of approach is designed to be self-instructional, not requiring direct teacher supervision. This program presents the material in small, sequential steps with a response required by the child. The child is immediately informed of the correctness or incorrectness of the response and can move at his/her own pace.
   b. Computer-assisted instruction—In this approach, the programmed instruction is administered through the use of a computer.

Appendix D: Adjustments: Rationale and Details of the Analyses

Appendix D contains, in more detail, results of the stepwise regression analyses and the rationale for the decision of which independent variable(s) to include in the adjustment equation for each dependent variable.

RESULTS AND RATIONALE

Stepwise regression, with grade, age, sex, and order of administration as the independent variables[1], revealed that for the two comparison tasks (auditory and visual comparison) only grade level entered the regression equations with significant F-values ($p<0.05$); and for the auditory-visual integration task, grade and sex ($N=218$) entered into the regression equation with significant F-values ($p<0.05$). As expected, the children in the second grade, in general, performed better than children in the first grade on all of these tasks. In the AVI task, the

---

[1] The independent variables were represented in the following ways: grade in tenths of a school year; age in years, months, and days converted to years and decimal fractions of years; for sex, males were designated by the number 1 and females by 2; and order of administration, according to whether the reading achievement tests were administered before (1) or after (2) the auditory-visual integration test battery.

boys tended to score higher than the girls within each grade. The F-value for sex was just significant at $p = 0.05$. Although Reilly [39] reported no sex differences for this type of task in first- and second-grade children, sex was included in the final adjustment equation because of other sex differences found in this study and reported later in this section, namely in the WISC-R tests.

Similar analysis of the auditory memory task (ITPA Auditory Sequential Memory) raw scores ($N = 217$) revealed that adjustment for grade level would be sufficient (using a 5% level of significance). Once again, second-grade children scored higher, on the average, than first graders.

The scaled scores (according to age in 4 month bands) from the Wechsler tests were examined for adequacy of standardization for this sample. For Block Design ($N = 172$), only sex entered the regression equation with a significant F-value ($p < 0.05$); boys performed better than girls. The relationships in the Vocabulary test ($N = 216$) were a bit more complex. Sex, grade, and age all entered into the regression equation with significant F-values ($p < 0.05$). As in Block Design, boys scored higher than girls within each grade and age level. Grade and age adjustments were in opposite directions; within a given grade, the younger children achieved higher scaled scores than the older children and at a certain age, children in the higher grade likewise achieved higher scaled scores. This seems to imply that children at a given grade level had similar Vocabulary performance, so that the younger ones, referred to the scaled score table for their younger age, were given higher scaled scores, while the older ones, with the same Vocabulary performance but referred to the scaled score table for their older age, were given correspondingly lower scaled scores. Analysis of the available raw scores for the Vocabulary test ($N = 210$) revealed that the opposing effects of grade level and age in the standardized scores were probably due to the standardization by age, when grade was more important in this sample. When raw scores were analyzed, the regression coefficients were positive for both age and grade, as expected, and negative for sex (boys' scores greater than girls' scores).

The pattern of sex differences found for the WISC-R tests in the present study has been reported previously for this age group by the National Center for Health Statistics [33]. In that Health Examination Survey, over 7,000 six- to eleven-year-old children were administered an extensive battery of tests, including the WISC Vocabulary and Block Design tests. Boys on the average consistently scored higher than girls throughout the age range tested for both of the WISC tests.

Analyses of the WRMT revealed that the standardization (by grade in tenths of a school year) was adequate for this sample (using a 5% significance level) except in the Word Attack test. Similar to the situation in the WISC-R Vocabulary test, grade and age entered into the regression equation at significant F-values with opposite signs on their coefficients. To determine whether this might be a result of standardization procedures, the Mastery Scores of the reading tests were examined in the same manner. As one would expect, grade was a significant factor for adjustment in the tests. Age also entered into the Word Attack regression equation at a significant level but with a negative sign. So, in a given grade the younger children achieved higher Mastery Scores on the Word Attack test; and at a given age, the higher grade level children achieved higher Mastery Scores. This effect could be the result of children being placed in a grade lower than their age would predict, ie, children being held back in a grade because of failure and/or children starting school a year later than expected because of immaturity, so that the oldest children in the grade are most likely achieving at an average level or below while the younger children will be achieving at a wider range of levels (from below average to above average); this may result in a negative age effect within a grade. One would be more concerned if this phenomenon were observed consistently in the analyses of the WRMT. However, since only the Word Attack test of the WRMT exhibited such an effect, the decision was made not to include adjustments for this test; the analysis of the Mastery Scores showed that grade, indeed, had the greatest effect of the control variables.

The regression coefficients utilized for adjustment may be found in Table D1.

DISCUSSION

The stepwise regression results for adjustment of the children's tasks for sex, grade, age, and order of administration were rather unremarkable except for the lack of a significant effect of sex in the reading tests (WRMT). The generally held opinion is that boys are usually slower at developing reading skills than girls and, if one examines the reading disability literature, it is evident that reading disability is much more prevalent in boys than in girls [18,24,27,35]. One explanation for the lack of sex differences in the present study might have been that the Woodcock Reading Mastery Tests were designed to eliminate sex biases. However, the description of test development in the Manual [58] does not suggest this and, indeed, separate sex norms for the total reading index presented in the Manual [58: pp 110–111] indicate that boys performed less well than girls in the normative sample. However, other investigators have reported no sex differences in reading achievement in this age group [21,54].

TABLE D1. Adjustments: Stepwise Regression Results

Dependent variable	Independent variable	Coefficient	F to remove[a]
WISC-R (scaled scores)			
Vocabulary	Sex	− 1.48	17.58
(N = 216)	Grade	1.67	10.86
	Age	− 1.04	5.49
Block Design	Sex	− 1.26	8.02
(N = 172)			
ITPA (Raw Scores)			
Auditory Sequential Memory (N = 217)	Grade	7.22	37.24
AVI Battery (N = 218)			
Visual Comparison	Grade	.832	12.93
Auditory Comparison	Grade	.889	13.85
Auditory-Visual	Sex	− 1.03	3.84
Integration	Grade	3.54	43.24

[a] Terminology of BMD2R Stepwise Regression program.

The major purpose of these analyses was to be able to adjust for any extraneous factors that would complicate interpretation of the results in an unwarranted fashion (eg, inflation of intraclass correlation coefficients due to identical ages of cotwins or inflation of the correlation coefficient for two variables due to covariance with age). An interesting sidelight is the lack of a significant sex effect (as determined by stepwise regression) in the reading achievement tests.

Appendix E: Comparison of First- and Second-Grade Samples

Appendix E contains information regarding the interrelationships of the tasks in the children's test battery with the sample partitioned into first- and second-grade groups.

A NOTE ON THE SAMPLE

Children in this sample ranged from grade 1.7 to grade 2.9. Because of the time of the testing, the distribution of grades was bimodal with a distinct break: one group included grade 1.7–2.0 and the other group 2.7–2.9. These two groups will be referred to as "grade one" and "grade two," respectively. It should be noted then that these labels refer to the grade most recently completed or nearly completed.

MEANS

The means and standard deviations for the grade-one and grade-two samples are shown in Table E1. Note that, for the nonstandardized tests, the means for grade two sample were higher than those for the grade one sample. None of the other means of the two grades were significantly different from one another ($p<0.05$).

INTERCORRELATIONS

To examine the relationships between the tests administered to the children, co-twins were treated as individuals. This obviously, is not the ideal procedure since the sample was not made up of random independent observations,

TABLE E1. Children's Test Battery: Means and Standard Deviations

	Grade 1[a]		Grade 2[b]	
	$\bar{x}$	SD	$\bar{x}$	SD
Auditory-visual integration test battery				
Visual comparison	8.1	1.8	8.9[c]	1.4
Auditory comparison	5.7	1.6	6.5[c]	1.8
Auditory-visual integration	9.6	3.6	12.7[c]	4.3
Auditory Memory (ITPA)	28.4	7.2 (135)	35.3[c]	9.9
WISC-R tests (scaled scores)				
Vocabulary	9.0	2.7	9.7	2.6 (80)
Block Design	10.8	3.0 (123)	11.1	2.5 (49)
WRMT (percentiles)				
Letter Identification	74.7	17.7	70.7	28.2
Word Identification	63.7	31.0	63.6	26.2
Word Attack	59.7	29.9	63.9	28.3
Word Comprehension	53.0	29.9	50.1	21.8
Passage Comprehension	66.9	23.4	62.2	24.4
Total	67.5	31.3	69.7	29.1

[a] N = 136 unless otherwise noted in table inside parentheses.
[b] N = 82 unless otherwise noted in table inside parentheses.
[c] Grade-two mean significantly greater than grade-one mean ($p < 0.05$).

but this method was deemed adequate for superficial examination of the interrelationships of these variables. Table E2 lists the correlations; grade one correlations are in the lower left portion of the Table and grade two correlations are in the upper right portion.

Consider first the reading tests alone. These intercorrelations were all significant ($p<0.01$) and of a rather high magnitude (ranging from 0.39 to 0.90 with most values above 0.60). The patterns of correlations were comparable in the grade one and grade two samples. Woodcock [58: p 64] reported similar results for a grade 1.9 sample.

Turning to the WISC-R score correlations, the Vocabulary and Block Design tests both correlated at least moderately with the reading tests in the grade-one sample. This was also true of the Vocabulary test in the second-grade sample. However, in the second graders, the Block Design and the noncomprehension reading test scores correlations were not significantly different from zero ($p\geq0.05$). This may have been due, in part, to a reduced sample size ($N=49$) for this series of correlations. Results from previous studies, presented in Table 16, revealed, in most cases, moderate correlations between measures of reading achievement and general intellectual ability in first- and second-grade children.

The two groups also differed in some of the WISC-R score correlations with the remaining variables (Memory, VV, AA, and AVI). One difference was in the auditory comparison task (AA); the second graders showed significant, moderate correlations while the first graders' correlations were nonsignificant. This was probably due to a high frequency of guessing on the AA task and its very low reliability in the first-grade sample. Correlations with memory were low or insignificant in both grades. The auditory-visual integration task (AVI) correelated at a moderate level (0.38 to 0.41) with the WISC-R scores in both grade groups. For the visual comparison task (VV), the patterns also differed. In the first-grade sample, both correlations were significantly different from zero and moderate; in contrast, results from the second grade sample revealed a correlation significantly different from zero only with Block Design, not with Vocabulary.

Now consider the intercorrelations of memory and the auditory-visual integration test battery and their relationships with the reading measures. There was an obvious difference between the grade groups in the pattern of correlations; in the first-grade sample, the auditory comparison task correlated significantly with the AVI task while the same was true of the auditory memory task in the second-grade sample. Again, the low or near zero correlations for AA in the first-grade sample may have been due, in part, to a high rate of guessing in this task and its low reliability (see Appendix B). Neither guessing nor a low ceiling can explain the findings in the memory task, however. The slight difference in magnitude between the grades in the visual comparison

*TABLE E2. Intercorrelations**

		Above diagonal: grade two^b											
		WRMT							Block				
		LI (%)	WI (%)	WA (%)	WC (%)	PC (%)	Total (%)	Vocab.	Design	Memory	VV	AA	AVI
Below diagonal: grade one^a	LI (%)		0.61	0.57	0.39	0.58	0.71	0.24	−0.05	0.05	0.22	0.31	0.31
	WI (%)	0.59		0.89	0.77	0.89	0.96	0.49	0.25	0.11	0.31	0.27	0.54
	WA (%)	0.51	0.79		0.71	0.83	0.93	0.41	0.21	−0.00	0.41	0.31	0.47
	WC (%)	0.60	0.88	0.75		0.78	0.81	0.61	0.43	0.20	0.32	0.22	0.45
	PC (%)	0.59	0.90	0.74	0.89		0.92	0.58	0.31	0.08	0.35	0.35	0.50
	Total (%)	0.67	0.96	0.85	0.94	0.93		0.50	0.25	0.10	0.37	0.33	0.53
	Vocabulary	0.30	0.44	0.53	0.54	0.54	0.50		0.48	0.18	0.17	0.32	0.38
	Block Design	0.34	0.32	0.39	0.40	0.37	0.40	0.25		0.25	0.37	0.37	0.39
	Memory	0.18	0.16	0.27	0.18	0.18	0.21	0.19	0.06		−0.00	0.15	0.22
	VV	0.45	0.41	0.45	0.46	0.45	0.48	0.45	0.41	0.20		0.17	0.25
	AA	0.04	0.10	0.06	0.09	0.14	0.10	0.16	0.12	0.06	0.15		0.40
	AVI	0.20	0.41	0.39	0.46	0.46	0.45	0.38	0.41	0.28	0.40	0.28	

*Underlined values are not significantly different from zero ($p \geqslant 0.05$).

[a] N = 136 except for Block Design (N = 123) and memory (N = 135).

[b] N = 82 except for Vocabulary (N = 80) and Block Design (N = 49).

correlations was also notable; the decreased magnitude in the second-grade may have been due, in part, to a low ceiling. The patterns for the auditory-visual integration task were similar in the two grades except for a reversing of the role of VV and AA. Partial explanations of low ceiling and a high rate of guessing have already been discussed and are applicable to that finding.

PARTIAL CORRELATIONS

Tables E3 and E4 present partial correlations of reading achievement with the AVI battery when the estimates of general intellectual ability (WISC-R Vocabulary and/or Block Design) were partialled out.

Except for the Letter Identification test, the reading tests still exhibited positive and moderate correlations with the AVI task. In the first-grade sample, the AA correlations were not significantly different from zero (probably due to the low reliability of AA) and the VV correlations were low to moderate and significantly different from zero. The situation in the second-grade sample was less clear. When Vocabulary scores were partialled out, the VV correlations were significantly different from zero (except for Letter Identification), and the AA correlations were mixed, with Word Identification and Word Comprehension exhibiting nonsignificant correlations. When Block Design scores were partialled out, the correlations for both the VV and AA tasks were low and mainly nonsignificant. When both Vocabulary and Block Design were partialled out, the AA correlations appeared to decrease (all now nonsignificant) and the VV correlations were intermediate in magnitude between those obtained partialling out Vocabulary and those obtained partialling out Block Design (two were significant with $p<0.05$). Note that the partial correlations of AVI with the reading tests were larger in magnitude in the second-grade sample than in the first-grade sample.

FACTOR ANALYSES

Grade-One Sample.

The three-factor solution for this sample was rather difficult to interpret, but may be found in Table E5. The four-factor solution was more readily interpretable and is shown in Table E6. The first factor was, as in the total sample, a reading achievement factor, with high loadings for all of the five tests from the WRMT. Factor 1 was a contributor of at least 10% to the variances of the WISC-R Vocabulary test and the visual comparison task. Factor 2 was a more specific factor and was entirely due to the only significant correlation coefficient for AA in grade one (see Table E2); the auditory comparison task had a very high loading and the auditory-visual integration task, a moderate loading. The high loading of AA on this factor (0.90) is rather troublesome in light of that task's very low reliability in the grade-one sample. That anomalous finding may be due to the manner in which missing data was handled by the factor analysis program (cf 2.3.4). Factor 3 was also a specific factor and represented auditory memory with the only high or moderate loading variable being the ITPA Auditory Sequential Memory test. The fourth factor probably represented the same visual-spatial ability represented by factor 3 in the combined sample, mainly involving the Block Design test and the

*TABLE E3. Partial Correlations: Grade One**

	Vocabulary scores partialled out[a]			Block design scores partialled out[b]			Vocabulary and block design scores both partialled out[b]		
	VV	AA	AVI	VV	AA	AVI	VV	AA	AVI
LI (%)	0.38	0.03	0.17	0.36	0.05	0.18	0.35	0.03	0.10
WI (%)	0.28	−0.02	0.23	0.34	0.03	0.33	0.26	−0.03	0.23
WA (%)	0.26	−0.06	0.22	0.34	0.00	0.33	0.22	−0.09	0.21
WC (%)	0.31	−0.04	0.28	0.37	0.02	0.37	0.28	−0.06	0.24
PC (%)	0.29	−0.03	0.30	0.36	0.10	0.41	0.26	0.03	0.29
Total (%)	0.34	−0.02	0.28	0.40	0.03	0.37	0.31	−0.04	0.25

*Underlined values not significantly different from zero ($p \geqslant 0.05$).

[a] N = 136.

[b] N = 123.

Visual Comparison task. That factor also was a contributor of at least 10% to the variances of the auditory-visual integration task and the Letter Identification test of the WRMT.

Grade-Two Sample.

The Varimax rotated factor matrix for the three-factor solution seemed to be the solution that was most clearly interpretable and is shown in Table E7. The first factor was, once again, primarily a reading achievement factor with the five reading tests from the WRMT exhibiting high loadings; the WISC-R Vocabulary test and the auditory-visual integration task both showed moderate-level loadings. The visual-spatial factor surfaced in this sample in factor 3, with high loadings for the WISC-R Block Design test and the visual comparison task. The second factor was not quite as clear as the other two factors, but obviously involved auditory memory evidenced by the high to moderate loadings on the ITPA Auditory Sequential Memory test, the auditory comparison task, and the auditory-visual integration task. Extraction of a fourth factor did not clarify the issue (see Table E8).

*TABLE E4. Partial Correlations: Grade Two**

	Vocabulary scores partialled out[a]			Block design scores partialled out[b]			Vocabulary and block design scores both partialled out[b]		
	VV	AA	AVI	VV	AA	AVI	VV	AA	AVI
LI (%)	0.20	0.25	0.13	0.13	0.18	0.26	0.16	0.13	0.23
WI (%)	0.28	0.16	0.37	0.16	0.22	0.51	0.21	0.15	0.48
WA (%)	0.39	0.25	0.36	0.30	0.27	0.51	0.36	0.22	0.49
WC (%)	0.29	0.10	0.24	0.17	0.13	0.33	0.22	0.06	0.29
PC (%)	0.32	0.27	0.30	0.14	0.31	0.46	0.21	0.25	0.43
Total (%)	0.35	0.25	0.35	0.24	0.25	0.50	0.30	0.18	0.47

*Underlined values not significantly different from zero (p ⩾ 0.05).
[a] N = 80.
[b] N = 49.

TABLE E5. Factor Analysis of Children's Test Battery Varimax Rotated Factor Matrix for Grade-One Sample Three-Factor Solution (N from 123 to 136)

	Factors[a]			
	1	2	3	Communality
Vocabulary	0.549	0.310	0.161	0.423
Block Design	0.383	0.457	0.120	0.370
Auditory Memory	0.082	0.056	0.938	0.889
Visual Comparison	0.462	0.379	0.380	0.501
Auditory Comparison	−0.082	0.829	−0.095	0.703
Auditory-Visual Integration	0.328	0.645	0.310	0.620
WRMT				
Letter Identification	0.713	0.004	0.151	0.531
Word Identification	0.929	0.130	−0.008	0.879
Word Attack	0.826	0.119	0.223	0.746
Word Comprehension	0.914	0.179	0.082	0.874
Passage Comprehension	0.909	0.208	0.055	0.872

[a] Underlined values represent a contribution by that factor to the variance of that test of at least 10%.

TABLE E6. Factor Analysis of Children's Test Battery Varimax Rotated Factor Matrix for Grade-One Sample Four-Factor Solution (N from 123 to 136)

	Factors[a]				
	1	2	3	4	Communality
Vocabulary	0.591	0.296	0.188	0.101	0.483
Block Design	0.183	0.078	−0.068	0.883	0.823
Auditory Memory	0.112	0.039	0.957	0.074	0.936
Visual Comparison	0.342	0.109	0.262	0.648	0.617
Auditory Comparison	0.019	0.901	−0.019	0.042	0.814
Auditory-Visual Integration	0.307	0.506	0.276	0.440	0.620
WRMT					
Letter Identification	0.624	−0.184	0.067	0.385	0.575
Word Identification	0.938	0.069	−0.008	0.146	0.905
Word Attack	0.814	0.022	0.203	0.225	0.754
Word Comprehension	0.902	0.079	0.062	0.242	0.882
Passage Comprehension	0.914	0.131	0.049	0.200	0.894

[a] Underlined values represent a contribution by that factor to the variance of that test of at least 10%.

TABLE E7. Factor Analysis of Children's Test Battery Varimax Rotated Factor Matrix for Grade Two Sample Three-Factor Solution (N from 49 to 82)

	Factors[a]			
	1	2	3	Communality
Vocabulary	0.390	0.528	0.309	0.526
Block Design	−0.014	0.613	0.687	0.847
Auditory Memory	−0.052	0.776	−0.226	0.656
Visual Comparison	0.237	−0.059	0.796	0.694
Auditory Comparison	0.255	0.524	0.137	0.358
Auditory-Visual Integration	0.435	0.523	0.222	0.512
WRMT				
Letter Identification	0.799	0.031	−0.161	0.665
Word Identification	0.907	0.211	0.178	0.899
Word Attack	0.880	0.079	0.286	0.862
Word Comprehension	0.687	0.363	0.342	0.720
Passage Comprehension	0.870	0.237	0.250	0.876

[a] Underlined values represent a contribution by that factor to the variance of that test of at least 10%.

TABLE E8. Factor Analysis of Children's Test Battery Varimax Rotated Factor Matrix for Grade-Two Sample Four-Factor Solution (N from 49 to 82)

	Factors[a]				
	1	2	3	4	Communality
Vocabulary	0.526	0.450	0.373	0.020	0.618
Block Design	0.115	0.856	0.240	0.213	0.848
Auditory Memory	0.028	0.110	0.797	0.173	0.677
Visual Comparison	0.254	0.636	−0.475	0.198	0.734
Auditory Comparison	0.126	0.193	0.110	0.868	0.819
Auditory-Visual Integration	0.424	0.296	0.212	0.500	0.563
WRMT					
Letter Identification	0.672	−0.302	−.0110	0.448	0.757
Word Identification	0.934	0.106	0.035	0.165	0.913
Word Attack	0.882	0.146	−0.153	0.201	0.862
Word Comprehension	0.814	0.376	0.204	−0.031	0.847
Passage Comprehension	0.904	0.185	0.025	0.184	0.886

[a] Underlined values represent a contribution by that factor to the variance of that test of at least 10%.

Appendix F: Assumptions of the Twin Model

Appendix F contains a brief discussion of some of the assumptions of the twin model and analysis of variance for the variables in the present study as well as tests of the validity of two assumptions.

Traditional twin analysis techniques involve the comparison of the within-pair mean squares of MZ and like-sexed DZ twins. This model requires several assumptions for testing for the presence of genetic variance: 1) total variance of MZ twins equals that of DZ twins ($\sigma^2_{MZ} = \sigma^2_{DZ}$); 2) the covariance among environmental effects within-pairs of MZ twins equals that within-pairs of DZ twins ($C_{MZ} = C_{DZ}$); and 3) covariance between genetic and environmental effects on the same individual equals the covariance between genetic effects on one member of a twin pair and environmental effects on the other member of that twin pair ($\sigma_{GE} = \sigma^*_{GE}$) [14].

The equality of total variances assumption may readily be tested and, if found to be invalid, may indicate a difference in the environmental variance components' contributions to the total variances of MZ and DZ twins. If such a difference is detected, an estimate of genetic variance unbiased by differing environmental variances may still be obtained using the among component estimate [14]. A more serious concern is that of equal environmental covariances in MZ and DZ twins. It has often been argued that, because they look more alike, MZ co-twins are treated more alike than DZ co-twins; this may or may not be true. The major problem with this possible inequality is that it will bias all estimates of genetic variance from this model [14]. A test to evaluate the possibility that apparently significant genetic variance is in fact due to MZ environmental covariance being greater than DZ environmental covariance is discussed in Christian et al [15].

Tables F1 and F2 present the statistics for testing assumptions 1 and 2 mentioned above, as well as the test for the significance of the among component estimate of genetic variance, which is not biased by unequal environmental variances for the MZ and DZ twin samples, when appropriate; Table F1 provides information for the entire twin sample and Table F2 includes information for the sample restricted to like-sexed twin pairs.

*TABLE F1. Twin Analyses: Additional Information, All Twins Included**

	Analysis of variance				Ratio of sum of mean squares		Test to exclude $C_{MZ} > C_{DZ}$	
	MZ		DZ					
	MS_A	MS_W	MS_A	MS_W	F′	p	MS_{ADZ}/MS_{WDZ}	p
WRMT (percentiles)								
LI	776.2	92.5	850.3	314.9	1.34	0.21	2.70	<0.01
WI	1509.6	68.1	1507.0	335.8	1.17	0.54	4.49	<0.01
WA	1461.8	156.0	1291.3	601.0	1.17	0.50	2.15	<0.01
WC	1328.2	144.4	1204.8	244.9	1.02	0.95	4.92	<0.01
PC	867.2	103.8	1071.6	249.2	1.36	0.20	4.30	<0.01
Total	1659.3	56.2	1594.9	408.5	1.17	0.54	3.90	<0.01
WISC-R (Scaled scores)								
Vocabulary	10.2	1.82	8.89	4.52	1.12	0.62	1.97	<0.01
Block Design	13.3	2.10	8.57	7.78	1.06	0.82	1.10	0.38
ITPA								
Auditory Memory	105.6	18.6	122.4	32.4	1.25	0.35	3.78	<0.01
AVI battery								
VV	3.90	1.58	3.45	1.78	1.05	0.82	1.94	0.01
AA	2.67	1.90	3.75	2.96	1.46	0.05[a]	1.27	0.20
AVI	20.7	5.97	18.6	12.5	1.16	0.47	1.49	0.08
Factors								
1	175.2	9.67	173.6	46.4	1.19	0.48	3.74	<0.01
2	137.2	29.2	170.2	67.8	1.43	0.11[a]	2.51	<0.01
3	137.2	37.4	114.4	62.1	1.01	0.96	1.84	0.02

*MS_A = among-pair mean square; MS_W = within-pair mean square; p = probability that ratio listed just previously equals 1; C = environmental covariance.

[a] Assumption of equality of total variance not tenable; among component estimate of genetic variance perhaps more appropriate than within pair estimate for testing for presence of genetic variance [14] in these cases, null hypothesis ($\sigma_G^2 = 0$) could not be rejected.

*TABLE F2. Twin Analyses: Additional Information, Like-Sexed Twins Included**

	Analysis of variance				Ratio of sum of mean squares			
	MZ		DZ				Test to exclude $C_{MZ} > C_{DZ}$	
	MS_A	MS_W	MS_A	MS_W	F′	p	MS_{ADZ}/MS_{WDZ}	p
WRMT (percentiles)								
LI	776.2	92.5	909.9	258.1	1.34	0.26	3.53	< 0.01
WI	1509.6	68.1	1463.1	315.5	1.13	0.66	4.64	< 0.01
WA	1461.8	156.0	1303.9	575.7	1.16	0.56	2.67	0.01
WC	1328.2	144.4	1228.0	252.0	1.00	0.97	4.87	< 0.01
PC	867.2	103.8	1016.1	251.5	1.31	0.32	4.04	< 0.01
Total	1659.3	56.2	1612.7	377.9	1.16	0.59	4.27	< 0.01
WISC-R (Scaled scores)								
Vocabulary	10.2	1.82	11.1	5.15	1.36	0.22	2.17	0.02
Block Design	13.3	2.10	8.29	8.31	1.08	0.78	.997	0.50
ITPA								
Auditory Memory	105.6	18.6	118.7	26.4	1.17	0.55	4.50	< 0.01
AVI battery								
VV	3.90	1.58	2.92	1.74	1.18	0.51	1.68	0.08
AA	2.67	1.90	2.78	1.95	1.03	0.87	1.42	0.17
AVI	20.7	5.97	15.9	13.65	1.11	0.66	1.16	0.34
Factors								
1	175.2	9.67	198.8	42.0	1.30	0.34	4.73	<0.01
2	137.2	29.2	154.3	51.2	1.23	0.40	3.01	< 0.01
3	137.2	37.4	106.4	61.7	1.04	0.89	1.72	0.07

*MS_A = among-pair mean square; MS_W = within-pair mean square; p = probability that ratio listed just previously equals 1; C = environmental covariance.

LIST OF TABLES

1 Difference between Good or Normal Readers and Subnormal Readers in Measures of Auditory Memory, Auditory and Visual Discrimination, Auditory-Visual Integration, and Spatial-Temporal Processing ....
2 Relationships of Auditory and Visual Memory, Auditory and Visual Discrimination, Auditory-Visual Integration, and Spatial-Temporal Processing to Normal Variation in Reading Achievement ......
3 Twin Studies of Reading Achievement ....
4 Means and Standard Deviations for Children's Test Battery ....
5 WRMT—Raw Scores: Means and Standard Deviations ....
6 WISC-R Scaled Scores: Means and Standard Deviations ....
7 Intercorrelations of Children's Test Battery ....
8 Partial Correlations: WRMT and AVI Battery ....
9 Factor Analysis of Children's Test Battery: Varimax Rotated Factor Matrix for Combined Grades Sample ....
10 Twin Analyses: All Twins Included ....
11 Twin Analyses: Like-Sexed Twins Included ....
12 Correlation of ATE with WRMT ....
13 Correlation of FES with WRMT ....
14 Correlation of WRMT with Teachers' Ratings of First-Grade Reading Instruction Approaches ....
15 Correlation of AVI Tasks with Reading Achievement ....
16 Correlation of General Intellectual Ability with Reading Achievement ....

LIST OF APPENDICES

A Zygosity Determination and Distribution According to Sex and Grade . . . .
B WISC-R Test-Scale Correlations and Reliability Estimates for Children's Test Battery . . . .
C Environmental Questionnaires . . . .
D Adjustments: Rationale and Details of the Analyses . . . .
E Comparison of First- and Second-Grade Samples . . . .
F Assumptions of the Twin Model . . . .

ABBREVIATIONS USED IN TEXT AND TABLES FOR TESTS IN PROTOCOL

Auditory-visual integration test battery	AVI Battery
Visual comparison	VV
Auditory comparison	AA
Auditory-visual integration	AVI
Illinois Test of Psycholinguistic Abilities	ITPA
Auditory Sequential Memory	Memory
Wechsler Intelligence Scale for Children-Revised	WISC-R
Vocabulary	Vocab
Block Design	Blocks
Woodcock Reading Mastery Tests	WRMT
Letter Identification	LI
Word Identification	WI
Word Attack	WA
Word Comprehension	WC
Passage Comprehension	PC
Index of Total Reading	Total
Attitudes Toward Education	ATE
Moos' Family Environment Scale	FES

SUBJECT INDEX

Adjustments
Attitudes Toward Education
  correlation with WRMT
  description
  factor analysis
  factor intercorrelations
  reliability
Auditory discrimination
Auditory memory
Auditory Sequential Memory (from ITPA)
  description
  mean and standard deviation
  twin analysis
Auditory-visual integration
  test battery
    description
    mean and standard deviation
    reliability
    twin analysis

Block Design (from WISC-R)
  description
  mean and standard deviation
  twin analysis
Factor analysis
  combined grade sample
    twin analysis
  method
  separated by grade
Familial environmental factors
  Attitudes Toward Education
  Family Environment Scale
Family Environment Scale
  correlation with WRMT
  description
  reliability
  subscale intercorrelations
    intrafamily
Intercorrelation
  Attitudes Toward Education
  calculation of
  children's test battery
  Family Environment Scale
Reading instruction methods questionnaire
  correlation with WRMT
  description
Reliability
  Attitudes Toward Education
  Auditory Sequential Memory (from ITPA)
  auditory-visual integration test battery
  Block Design (from WISC-R)
  calculation of
  Family Environment Scale
  Vocabulary (from WISC-R)
  Woodcock Reading Mastery Tests
Sample description
Spatial-temporal processing
Testing procedures
Twin analysis
  entire twin sample
  like-sexed twin sample
  method
  tests of assumptions
Twin studies
  auditory memory
  reading achievement
  WISC-R tests
Visual discrimination

Visual memory
Vocabulary (from WISC-R)
  description
  mean and standard deviation
  twin analysis
Woodcock Reading Mastery Tests
  correlation with
    Attitudes Toward Education
    auditory discrimination
    auditory memory
    auditory-visual integration
    Family Environment Scale
    reading instruction approaches
    visual discrimination
    WISC-R tests
  description
  mean and standard deviation
  partial correlation with
    auditory discrimination
    auditory-visual integration
    visual discrimination
  reliability
  twin analysis
Zygosity determination