1989

Examining the Bender Recall as a Test of Visual Memory

Bernadette F. Evans
University of Rhode Island

Follow this and additional works at: <https://digitalcommons.uri.edu/theses>

Recommended Citation

Evans, Bernadette F., "Examining the Bender Recall as a Test of Visual Memory" (1989). *Open Access Master's Theses*. Paper 1683.
<https://digitalcommons.uri.edu/theses/1683>

This Thesis is brought to you for free and open access by DigitalCommons@URI. It has been accepted for inclusion in Open Access Master's Theses by an authorized administrator of DigitalCommons@URI. For more information, please contact digitalcommons-group@uri.edu.

**EXAMINING THE BENDER RECALL
AS A TEST OF VISUAL MEMORY**

BERNADETTE F. EVANS

**A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
MASTERS OF ART
IN
PSYCHOLOGY**

UNIVERSITY OF RHODE ISLAND

1989

ABSTRACT

A scoring system was devised to measure the Bender Visual-motor Gestalt recall segment. Three dependent measures were analyzed: number of retrieved designs; "quality recall" such that recalled figures were compared to a child's original Bender drawings; and "average quality per design" in which the average degree of distortion per recalled design was calculated. Seven, nine, and eleven year old reading disabled and nondisabled children's memory performance was compared. It was hypothesized that a complex visual memory task, such as the Bender recall, and analysis of errors or distortions in memory would significantly discriminate between the two reading and among the three age groups. A Multivariate Analysis of Variance indicated significant main effects for age ($F = 5.69$, $df = (6, 224)$, $p < .001$) and for the two reading groups ($F = 6.21$, $df = (3, 112)$, $p < .001$). A discriminant analysis revealed each of the three dependent variables significantly discriminated between reading groups, with "quality recall" accounting for the largest portion of the variance (13%). Also, a trend analysis revealed significant linear trends across age for each reading group and on each dependent measure. T-tests indicated reading disabled subjects lagged

behind nondisabled readers in support of a developmental lag hypothesis. It was concluded that the Bender recall segment may be an adequate measure of visual memory and that distortion in recall may be worth further exploration with reading disabled children as well as with other memory deficient populations.

ACKNOWLEDGMENTS

I am especially grateful to Dr. Allan Berman, my major professor, for sharing his knowledge, experience, and time to enhance my professional growth. Similarly, I am grateful to Dr. Janet Kulberg and Dr. Susan Brady for their creative ideas and valuable insights. Lastly, I thank Dr. Joseph Rossi for dedicating his time to assist me with statistical analysis, and Dr. Dick Nelson for giving of his time and providing objective feedback.

I also want to give special thanks to Richard Lloyd of the Warwick School System for inspiring this project and assisting me with data collection. Gratitude is also extended to other Warwick school personnel -- Rosalie Fairman, Tom Barry, and Jeff Sharkey, for allowing me access to their files. I am also grateful to Warren School personnel -- James Hoebbel, Lou Perella, Charles Morris, Ken Sergeant, Dot Shea, teachers, parents, and children, for their various contributions to this research project. Another special thanks must be extended to Feliciano Figueriedo who dedicated her time and thoughts to aid me in my acquisition of subjects. Lastly, I am grateful to Sarah Cilano for her tireless effort in scoring the many Bender protocols.

TABLE OF CONTENTS

ABSTRACT	ii
ACKNOWLEDGEMENTS	iv
LIST OF TABLES	vii
LIST OF FIGURES	viii
INTRODUCTION	1
BACKGROUND	3
Statement of Problem	14
METHOD	16
Subjects	16
Instruments	18
Scoring	21
Procedure	22
RESULTS	26

DISCUSSION	37
REFERENCES	43
APPENDICES	51
BIBLIOGRAPHY	55

LIST OF TABLES

Table 1.	Means, Standard Deviations, and Range of Full Scale IQ and Reading Achievement T-Scores, and Ratio of Females to Males Across Age Groups.	18
Table 2.	Means and Standard Deviations for Three Recall Measures Across Two Reading Groups and Three Age Groups.	26
Table 3.	F-Ratio's for each Dependent Measure Across Three Age Groups.	29
Table 4.	T-Scores for each of Three Age Groups, Between Reading Disabled and Nondisabled Readers, for each Dependent Measure.	33
Table 5.	Summary Table of F-Ratio's for Linear Functions on each Measure, Across Age for Reading Groups.	34
Table 6.	Percent of Subjects Accurately Classified According to Age, Quality Recall, and Number Recalled.	36

LIST OF FIGURES

- Figure 1. Comparison of trend across age groups for each reading group on quality recall. 30
- Figure 2. Comparison of trend across ages and between reading groups on number recalled. 31
- Figure 3. Comparison of trend across ages and between reading groups for average quality per design. 32

INTRODUCTION

The present study was designed to examine the Bender Visual-Motor Gestalt test (Bender) recall segment. A scoring system was devised that incorporated the traditional number recalled and a new variable which measures "distortion" of recall. Primarily, a child's initial perception and eventual recall of the Bender designs was contrasted. This study then examined the difference between the traditional and "distortion" scoring methods in discriminating between reading disabled and nondisabled reading groups and three age groups.

In addition to proposing an alternative scoring system with the Bender visual recall task, its use in discriminating reading disabled from nondisabled readers was examined. Previous studies contrasting these two reading groups on visual working memory have provided equivocal results (cf., Katz, Shankweiler, & Liberman, 1981; Vellutino, Smith, Steger, & Kaman, 1975). The tasks used in these studies have generally lacked in complexity, either in administration or in their scoring method. This study, therefore, also aimed to examine the Bender recall's

discriminating power by incorporation of a more complex administration.

Furthermore, this study evaluated the recall measures' sensitivity to developmental differences across three age groups and between two reading groups.

BACKGROUND

Many diagnostic tools are utilized by professionals in performing psychological evaluations. One such tool, frequently used in a standard assessment procedure, is the Bender Visual-Motor Gestalt test (Sattler, 1982). Its purpose has ranged from utilization as a projective test to providing an evaluation of visual-motor perception. The Bender employs the use of geometric designs, some complex, due to their incorporation of two related images. Immediately following the administration of the Bender, it is common to ask the client to recall as many figures as he/she can (the Bender recall). Some psychologists view the recall data as indicative of one's memory capacity or ability to attend to the initial material. Since this information is often used in diagnosing such disorders as learning disabilities or attention deficits, this information is potentially very valuable.

Presently, however, the presentation and interpretation of the recall segment are not standardized. No normative data exist. Researchers have not adequately answered the question of whether the recall segment is a valid diagnostic measure. Of the minimal studies available, most are flawed. Heterogeneous samples are commonly used, eliminating replication or generalization to similar groups. Often subjects are chosen globally representing "psychological

referrals" (Hutton, 1966; Armentrout, 1976; Finch, Spirito, Garrison, & Marshall, 1983). Only one study examined specific school-related referrals (Finch, Spirito, Garrison, & Marshall, 1983) and it included subjects with broad "academic and/or behavior problems."

Most research examining memory for designs over the past two to three decades has focused on two other visual retention tests -- Graham and Kendall's Memory-for-Designs (cf., Carroll, 1972; Grundvig & Needham, 1970; Leton, 1962; Lyle, 1968; Marsh & Hirsch, 1982; Walters, 1961) and the Benton Visual Retention Test (cf., Brooks, 1975; Marsh & Hirsch, 1982; Vellutino, Steger, & Kandel, 1972). Upon examination, both tests possess problems either with their method of scoring or in presentation, interfering with an accurate assessment of visual retention. The Graham-Kendall Memory-for-Designs, for instance, does not seem to adequately measure recall or memory (Lyle, 1968). Its method of scoring includes a designation of zero for those figures not recalled and those produced satisfactorily (Graham & Kendall, 1960). Thus, Graham and Kendall equated a nearly perfect recall with zero recall of a design. They failed to consider the importance of recalling a "satisfactory" design.

With respect to the Benton Visual Retention Test, Form A, each stimulus card is presented for ten seconds and then recall of that one figure is tested. An updated edition

allows for greater than ten seconds after presentation of the image before requesting the client to recall the design. This increase in time results in improved differentiation of clinical groups (Brooks, 1975). The Visual Retention Test, however, does not consider the initial perception of the stimulus figure by examining a drawn copy produced by the client beforehand. Perceptual problems could create a recall profile that appears incorrect. This would skew the rater's interpretation of the recall segment (Hanawalt, 1959). Also, measuring recall after each individual design is less taxing on memory processes than if serial recall is required. Recall after a series of designs increases the likelihood of differentiating between clinical/nonclinical groups (cf., Cummings & Faw, 1976; Rogers & Swenson, 1975; Sipe & Engle, 1986; Torgeson & Houck, 1980; Vellutino, Smith, Steger, & Kaman, 1975).

Since the Bender is the second most commonly used measure in diagnostic batteries, it might be efficient to implement a recall segment if it offered diagnostic information. Drawbacks of other memory for design tests could also be eliminated when measuring the Bender recall. For example, one might consider the initial perception of designs when scoring the recall segment.

The Bender recall has the potential to tap such variables as visual short and long term memory capacity (Rogers & Swenson, 1975; Tolor, 1956), visual-perceptual

distortions occurring during encoding or retrieval phase of memory, attention span, etc. Considering this, the school referred population is ideal to study. Children with learning disabilities and attention-deficit hyperactivity disorders experience difficulties with memory and attention-based tasks. Reading disabled children, in particular, are believed to suffer from memory deficits (Brady, 1986; Doehring, 1985; Liberman, Mann, Shankweiler, & Werfelman, 1982; Lyle, 1968; Lyle & Goyen, 1975; Vellutino, 1987; Vellutino, Steger, Kaman & DeSetto, 1975).

Various theoretical models exist purporting the causative and correlative nature of this memory deficit. Some researchers postulate that reading disabled children exhibit a phonological coding deficit in working or short term memory (Brady, 1986; Liberman, Mann, Shankweiler, & Werfelman, 1982; Torgeson & Houck, 1980; Vellutino, 1979), while others propose deficiencies in visual memory recall (Carroll, 1972; Lyle, 1968; Lyle & Goyen, 1975). Complicating the situation, some researchers suggest reading disabilities can be subtyped (Doehring, 1985; Boder, 1973) including phonological coding and visual memory deficits as subtypes (Boder, 1973; Hynd & Cohen, 1983).

Those espousing the phonological coding hypothesis do provide convincing evidence to support their views. In order for these researchers to advocate a phonological coding versus visual deficit, the studies have employed

comparative techniques. Results indicated significant differences between good and poor readers on tasks incorporating words or items reliant on verbal mediation.

In contrast, results from visual memory tasks (those using geometric figures or nonsense/foreign words) have been equivocal. Overall, it appears that the degree of complexity in administration and demands on recall varied within and across studies. Weaknesses were particularly noted in studies concluding reading disabled children did not evidence a visual working memory deficit. For instance, some studies contrasting phonological coding and visual processing in working memory of good and poor readers, would present an item briefly and subsequently require subjects to recall the item within ten to fifteen seconds (Katz, Shankweiler, & Liberman, 1981; Vellutino, Smith, Steger, & Kaman, 1975; Vellutino, Steger, Kaman, & DeSetto, 1975). Limiting recall to isolated items versus a sequence of items is less complex and, in turn, less demanding.

The visual memory task per se also appears less complex in most studies. Vellutino, Steger and Kandel (1972) utilized diamond and triangle shapes in their study. These visual stimuli require less cognitive processing than would an unfamiliar, abstract geometric design. Studies measuring visual memory skills also varied by incorporating a recognition versus free recall paradigm (Katz, Shankweiler, & Liberman, 1981; Liberman, Mann, Shankweiler, & Werfelman,

1982). Again, a visual recognition task is less taxing on attending and memory processes than free recall per se. Furthermore, some researchers inappropriately concluded that good and poor readers did not differ in visual memory performance based on data evidencing a floor effect (Katz, Shankweiler, & Liberman, 1981; Vellutino, Steger, Kaman, & DeSetto, 1975) or ceiling effect (Vellutino, Smith, Steger, & Kaman, 1975).

Issues such as the aforementioned need further consideration before concluding that reading disabled children do not exhibit a visual working memory deficit. The Bender's unfamiliar, abstract designs, serial administration and ten second delay prior to recall compensate for noted weaknesses of these other studies.

There are, however, researchers who continue to support a visual memory deficit hypothesis with reading disabled children. Those studies that support a visual memory deficit incorporated nonsense/verbally noncodable geometric figures to distinguish poor from good readers. (Visually presented words or letters were not used since they lent themselves to being verbally coded.) In particular, Lyle (1968) and Carroll (1972) emphasized the importance of measuring distortions or errors in recall in order to discriminate reading groups.

Lyle (1968) conducted a study to examine good versus poor readers' performance on Graham and Kendall's memory-

for-designs test (MFD). He used two scoring methods: (a) Graham and Kendall's standard method and (b) one in which number of lines missing, added or distorted was counted. Lyle found that both scoring methods significantly differentiated the groups, with Graham and Kendall's method surfacing as slightly more sensitive. Lyle concluded that: (1) poor readers had a visual memory deficit and; (2) "distortions in remembering figures from memory" predominated the poor versus good readers' recall (via Graham and Kendall's scoring method).

Similarly, in 1972, Carroll examined visual memory performance of neurologically and nonneurologically impaired children, measured on the Visual Memory Scale (VMS) (recognition task of simple to complex geometric designs). Of significance to Carroll's study was the significant correlation between reading readiness ($r = -.39$) and, separately, reading achievement of first graders with the VMS ($r = -.25$). Specifically, the more memory errors committed, the lower the achievement and readiness scores obtained by the children.

Authors of the above-mentioned studies both refer to distortions or errors in memory recall or recognition as significant discriminators of poor and good readers. As mentioned earlier, when scoring the recall, a comparison between recall and initial perception ("distortion" measure) might provide valuable additional diagnostic information.

This issue is worthy of consideration since it has been neglected in the above-mentioned and most other studies examining the reading disabled population, and memory studies in general.

Researchers neglect another issue, that of visual long term memory in relation to reading disabilities. Any serial recall of four to seven items, excluding rehearsal strategies or chunking devices (maintaining information in short term memory), begins to tap long term memory (in accordance with Atkinson and Shiffrin's stage model of memory) (Best, 1986). Vellutino, et al. (1975) briefly refer to this issue. When second grade poor readers were required to recall four or more configurations or designs, a significant divergence between them and good readers surfaced. Recalling four items, Vellutino stated, extends beyond short term memory capacity. Vellutino, et al. did not address this further.

Age is another important variable when examining memory retention. Most researchers concur that memory changes with age. Young children, for example -- eight years old and under -- usually recall fewer items than those over eight years. Various theories explaining memory development abound. Most theories center on: (a) the child's increasing use of strategies (e.g., rehearsing and/or organizing information) (Engle & Marshall, 1983; Hulme, Thomson, Muir, & Lawrence, 1984); (b) the child's knowledge

base increasing with age (Huttenlocher & Burke, 1976); and (c) cognitive changes involving a child's eventually learning to make inferences and think abstractly (Howe, 1983). Theories presently advocating efficiency in processing (semi-incorporating the three former theories) are gaining the most support (Bjorklund, 1985; Case, Kurland, & Goldberg, 1982; Chechile, et al., 1981; Chechile & Richman, 1982). Chechile (1981) found that retrieval versus recognition of information increases systematically from first through sixth grade. In later experiments, Chechile and Richman (1982) concluded that this phenomenon was due to the increasing development of semantic memory. The more expansive the knowledge base in long term memory, the more efficiently the individual is able to process information. Bjorklund (1985) added that language is another important variable affecting information processing. Case et al. (1982) added yet another dimension, emphasizing that reducing operating (or processing) space in working memory provided more space for storage in working memory. (Case, et al. assumed working memory occupied a constant space and that less operating space was necessitated when processing familiar material. Therefore, more space would be available for storage.) Thus, the noted increase in information retention from early childhood to adolescence reflects the important roles of increased knowledge and conceptual understanding. There are others who may disagree

with this theory, continuing to advocate rehearsal (Bauer, 1977, 1979; Engle et al., 1983; Hulme, et al., 1984) or organizational strategies [Flavell, 1985 (cited in Goodman and Haith, 1987)] as the prime variables differentiating memory performance across ages. (On an aside, Bjorklund believed that such strategies are important but only during adolescence and beyond.)

Others, who study memory as it relates to reading performance (Benton, 1962; Rourke, 1975), postulated that visual-perceptual-spatial skills are important at earlier stages of learning to read. However, "language and formal operational thought become increasingly more relevant at . . . advanced stages of reading development" (Rourke, 1975, p.917). If these authors are correct, one might find good and poor readers diverging on visual memory tasks at earlier ages but converging on nonverbal memory tasks as reading becomes less reliant on visual cues. Vellutino, Smith, Steger and Kaman (1975) supported this conclusion. They found second grade poor readers performed more poorly than good readers on visual memory tasks but sixth grade good and poor readers did not differ. They concluded sixth grade poor readers relied on orthographic knowledge, thus, enhanced their performance.

The developmental lag hypothesis has also been used to explain such results. Bauer (1977), Bryant and Impey (1986), and Tarver, Hallahan, Kauffman, and Ball (1976),

proponents of this hypothesis, believed that reading disabled children lag behind rather than exhibit an abnormal reading pattern. All three studies, however, provided different explanations for the observed lag. Bauer conjectured that the lag is due to delayed use of conscious rehearsal strategies in working memory. Bryant and Impey suggested that phonological competency differentiates good from poor readers. This theory supports the phonological coding deficit hypothesis in understanding working memory. Tarver et al. hypothesized that the developmental lag between reading groups is due to differences in selective attention. Whichever theorist is correct, all would agree that children with reading disabilities should experience more difficulty retaining information obtained through reading than nondisabled readers. Because of the aforementioned deficiencies in studies measuring visual memory, a similar lag could be found in this study. Therefore, complex memory tests, even visual ones, should reveal significant differences between readers at any age. The lag hypothesis is in opposition to Vellutino's (1975) or Rourke's (1975) hypothesis which suggests that poor readers' performance on memory tasks eventually "catches up" to that of good readers.

Those advocating a "failure in development" (Seymour, 1986; Temple, 1987) would not expect reading disabled children and nondisabled readers eventually to perform at an

equivalent level. Rather, reading disabled children would fail to improve in their memory capabilities across age. Seymour (1986), for instance, found reaction time measures on memory tests between good and poor readers to deviate significantly. Reaction time measures increased with age for good readers; this was not the case for poor readers. Similarly, Roder and Chechile (cited in Brainerd and Pressley, 1985) hypothesized that storage capacity in memory varies as a function of reading ability but not of age. If reading disabled children experience failure (versus a lag) in memory development, one would not necessarily expect their ability to retain short-term information to improve with age per se.

Statement of Problem

This study set out to: (1) devise a standardized system of administering and scoring the Bender recall; and (2) explore its usefulness in discriminating poor and good readers. If the Bender recall is proven beneficial in differentiating poor and good readers, various hypotheses regarding reading disabilities could be explored: (1) the nature of the proposed memory deficit and (2) the nature of the developmental trend. It is therefore, predicted that:

(1) The Bender recall, administered as a composite of nine designs, will significantly differentiate reading disabled children from nondisabled readers. It is also predicted that a measure of distortion or error in recall

will differentiate the groups, exclusive of a standard number count.

(2) Irrespective of reading ability, memory performance on the Bender recall will significantly improve from seven through eleven years of age. The significant improvement across age is consistent with the commonly-documented developmental trend. Also, reading disabled childrens' performance will lag behind that of nondisabled readers.

METHOD

Subjects

Sixty reading disabled and sixty good readers, equally distributed in seven, nine and eleven year old age groups, were chosen for study from two local suburban school systems in Rhode Island. All subjects included in the three age groups ranged from either 7 years 0 months to 7 years 11 months, 9 years 0 months to 9 years 11 months, or 11 years 0 months to 11 years 11 months. Mean and standard deviation scores were determined within each reading and age group to verify subjects were evenly distributed by age between the two reading groups. A record search was conducted in one school district to locate previously classified reading disabled children fitting the following selective criteria. All reading disabled children's full scale IQ scores ranged between 80 and 120. Reading disabled subjects were required to evidence a T-score discrepancy of eight points or greater between Total Reading Achievement scores and full scale IQ. Reading disabled subjects did not receive any special services at the time of diagnosis. Children with dual diagnoses of behavior or emotional disorders, or with a physical handicap, were excluded.

The T-score discrepancy formula was employed to classify a reading disability. An eight T-point discrepancy is considered significant at the .05 level in detecting

learning disabilities (Hanna, Dyck, & Holen, 1979). A discrepancy model is an acceptably applied index, based on federal regulations, Public Law 94-142 (U.S.D.H.E.W., 1977) in identifying learning disabilities.

The comparison group of nondisabled readers was selected from a separate school system, matched demographically with the other middle class income school district. Students' records were devoid of behavioral, emotional or learning problems. These records were confirmed by the classroom teachers' verbal reports. The comparison group's full scale IQ also ranged between 80 and 120. Those subjects with eight or greater T-point discrepancies (determined through testing) or Total Reading achievement scores at or below the thirtieth percentile, were excluded. Approximately 110 students underwent testing. This group provided an adequate representation for selection of sixty subjects who fit the inclusionary criteria.

Table 1 presents the means, standard deviations, and ranges of subjects' full scale IQ scores, reading achievement T-scores, and ratio of males to females across age groups.

Table 1.

Means, Standard Deviations, and Range of Full Scale IQ and Reading Achievement T-Scores, and Ratio of Females to Males Across Age Groups^a

		Reading Disabled			Nondisabled Reader		
Ages		7	9	11	7	9	11
FSIQ	M	105	102	99	106	109	106
	SD	10	9	10	10	8	9
	Range	82-118	87-118	83-117	80-119	92-120	96-120
Reading T-scores	M	36	38	35	56	59	56
	SD	8	6	11	7	8	7
	%ile	8th	13th	7th	73rd	83rd	73rd
	Range	14-45	25-47	1-48	45-70	45-73	45-73
Ratio F		.50	.30	.25	.60	.70	.75
to M		.50	.70	.75	.40	.30	.25

^an=20 per age group

Instruments

The instruments used were the Wechsler Intelligence Scale for Children-Revised (WISC-R) (Wechsler, 1974), Metropolitan Achievement Test (MAT) (Durost, Bixler, Wrightstone, Prescott, & Balow, 1986) and Bender Visual-Motor Gestalt initial administration (Bender) and recall

measure (Bender recall). The WISC-R and MAT were used to assess discrepancies between ability and achievement and to ensure subjects fell within the preset 80 to 120 IQ range (± 1.5 standard deviation). The WISC-R is a widely used and accepted means of assessing general intelligence. It renders a Full Scale IQ consisting of a verbal and performance component. These components contain six subtests each measuring a range of skills relevant to an ability index. Its validity and reliability have been well established. Test-retest reliability estimates remain high with respect to learning disabled children. Smith and Rogers (1978), for example, examined the test-retest performance of 160 learning disabled children over a six month interval. Reliability correlations were significant at .79 for Full Scale and .82 for verbal and performance scales.

A short form version of the WISC-R was used to measure the comparison group's IQ and in matching reading disabled/nondisabled reader samples. The short form consisted of similarities, vocabulary, picture arrangement, and block design. This tetrad correlates .947 with the ten subtests of the WISC-R, indicating its validity as an IQ measure (Sattler, 1982).

The MAT is a nationally-standardized achievement test consisting of six levels: Primer, Primary I, Primary II, Elementary, Intermediate, and Advanced. The levels yield a

standard score convertible to T-scores. Psychometrically, the MAT is viewed as a sound and reliable instrument. Its content is well suited for measuring achievement (Anastasi, 1982). The internal reliability estimate for Total Reading on all levels except the Primer is .96 (Kuder-Richardson). The Primer's (mid-first grade) internal reliability coefficient is .93. Studies have been conducted to determine its stability when testing learning disabled subjects (cf., Smith & Roger, 1978; Zingale, Smith, & Doeckki, 1980). Both Smith and Roger (1978) and Zingale et al. (1980) found the MAT maintained adequate temporal stability. Zingale et al.'s study measured retest reliability after one month with 82 subjects ranging in age from 6.2 to 13.2 years. Coefficients of the MAT's temporal stability in Total Reading ranged from .82 (Primary level) to .97 (elementary level). The authors concluded that the MAT yielded a reliable index of achievement with learning disabled children. The use of the MAT and WISC-R, therefore, seems appropriate in determining T-point discrepancies and IQ/reading achievement ranges.

The other instrument essential to this study was the Bender. The Bender, as previously mentioned, consists of geometric designs (see Appendix A). The test's psychometric properties vary in relation to the scoring method used. For this study, Elizabeth Koppitz' (1965) Development Bender Scoring System (a common scoring system) was applied to the

initial Bender protocol and recall measure. Koppitz examined "gross motor errors such as distortions, rotation of designs, integration problems, and perseveration" (Koppitz, 1965, p. 5). The reliability and validity is well documented. Kendall's rank correlation for test-retest reliability estimates range from .597 to .659 and is significant at $p < .001$ after four months. Interrater reliability estimates, assessed via a Pearson product-moment correlation, range from .88 to .96. The Koppitz scoring method also has relatively high predictive validity in reference to achievement test scores, $-.53$ to $-.75$, to detecting brain injuries and mental retardation (Koppitz, 1965). (A negative correlation was attained since the Bender scores for errors.) The recall segment can be measured in various ways: number recalled, number recalled with an adjustment for quality relative to initial drawings, and an average quality of each recalled design.

Scoring

Each method of scoring was explored with prime interest in the measure examining number recalled with a comparison between each design recalled and the subjects' initial perception of the Bender stimulus. First, all nine designs were scored on the recall protocol as either absent or present, providing number recalled. Second, those designs present were compared to the initial figure drawn by the subject versus the Bender stimulus figure. This procedure

ensured that the subject's perception of the image, as indicated on his/her first drawing, was taken into account. (See Appendix B) Simultaneously, and relatively important, this procedure may provide an index of distortion occurring in memory. If the subject recalled a perfect design as indicated on the original stimulus card or if the subject's recalled design was the same as the initial drawing, the figure was rated "3." A "2" was designated if the recalled design evidenced a rotation of 45 degrees or more or was poorly integrated relative to the initial drawing, as defined by Koppitz. Also, those designs in which Koppitz scores for perseveration received a "2" if perseveration occurred relative to the original drawing. A recalled design received a "1" if distorted such that the shape of the design was lost, unless referred to under Koppitz' "integration" definition. (See Appendix B) (For example, figure number two [see Appendix A] is scored as poorly integrated if circles are converted to dots; they are not distorted. Therefore, the design is ranked "2.") Third, the average quality per design was determined by dividing the number recalled into the second dependent measure. (See Appendix C for Koppitz' definitions of perseveration, distortion, and rotation.)

Procedure

A record search was conducted in order to locate reading disabled subjects, as previously defined, across

each age group. WISC-R's and MAT's were administered within two months of each other by qualified professionals -- two school psychologists and trained special education teachers, respectively -- primarily between 1982 through 1987. The two school psychologists also administered the Bender segment and Bender recall in a standardized manner when administering the WISC-R.

Nondisabled readers were selected from grade levels (second, fourth, and sixth) corresponding to age groups seven, nine and eleven years. Selected subjects fit the inclusionary criteria. Written parental consent was obtained before testing began. Signed letters were returned indicating "yes, I agree to let my child participate in your project" or "no, I prefer my child not participate." (See Appendix D) Testing occurred in the same nondistracting room for each age group. Children chosen from the same classroom were tested individually in the same day to lessen the likelihood of testing procedure leakage. Subjects were also instructed not to inform their classmates as to the nature of the testing.

School records provided MAT Total Reading achievement scores, which were administered within one to two months, prior to, or following, the administration of the WISC-R and Bender. The experimenter administered the shortened version of the WISC-R and Bender, in compliance with standardized instructions, in the order stated, within an approximate

thirty minute time period. Ten seconds elapsed between the initial Bender and recall segment as dictated by the procedure used with the reading disabled subjects. Ten seconds allowed enough time for the stimulus card and completed protocol to be removed, a clean sheet of paper to be placed in front of the subject, and brief standardized instructions stated. For example, the examiner stated, "Now, on this piece of paper, draw all the designs you can remember. Take your time and think about it." Since the subjects were not expecting to recall the designs, conscious rehearsal of stimuli was minimized. Upon subject's statement of completion, the experimenter inquired once, "Are you sure you can't remember any more?" No further prompting followed.

The initial phase of the Bender was timed. An advantage/disadvantage posited by more or less exposure to the stimulus figures was not permitted. Subjects whose time in minutes fell beyond ± 1.5 standard deviation from their sample's mean time were eliminated. Because outliers would be discarded on the basis of time, more than twenty subjects per age group were tested (nondisabled readers) or had data collected on them (reading disabled). Approximately thirty to thirty-five subjects per age group were examined

initially. Subjects within the nondisabled reader group were also eliminated if a ≥ 8 T-point discrepancy resulted.

Upon completion of the data collection phase, the experimenter blindly rated the initial Bender and recall segment. Another school psychologist, unfamiliar with the purpose of the study, also blindly rated 60% of the protocols in order to assess interrater reliability. The correlations attained ranged from .86 (seven-year-old nondisabled readers) to .99 (nine-year-old reading disabled), with the mean reliability correlation at .94.

RESULTS

This study generated three dependent variables: (a) "quality of recall" (a measure including number recalled and "distortion"); (b) number recalled; and (c) average quality per design ("distortion" per design). These variables were examined between two reading groups, reading disabled and nondisabled readers, and across three age groups, seven, nine, and eleven years old. Table 2 presents a summary of means and standard deviations of each group on all three measures. (Also see Figures 1, 2, and 3.)

Table 2.

Means and Standard Deviations for Three Recall Measures
Across Two Reading Groups and Three Age Groups

		7		9		11	
		RD	NonRD	RD	NonRD	RD	NonRD
Recall Measures							
Quality Recall	<u>M</u>	8.4	11.2	11.4	14.0	13.1	15.6
	<u>SD</u>	2.8	3.4	3.3	3.6	4.4	3.0
Number Recalled	<u>M</u>	4.3	5.1	5.2	5.9	5.7	6.0
	<u>SD</u>	1.0	1.1	1.3	1.2	1.4	1.1
Average Quality	<u>M</u>	1.9	2.2	2.2	2.3	2.2	2.6
	<u>SD</u>	.53	.41	.46	.28	.43	.34

The analysis most appropriate to implement with three variates, two levels of reading groups, and three levels of age, was a multivariate analysis of variance (Harris, 1985). A violation of the Assumption of Homogeneity of Covariance Matrices was observed as indicated by Box's test. The Box's test, however, is considered overly powerful yielding extremely large degrees of freedom (Harris, 1985). At the least, considering that the significance attained in this study was $p < .001$, the results would still prove significant exceeding the .05 level. An overall main effect for age proved significant, $F = 5.69$, $df = (6, 224)$, $p < .001$. A significant main effect was also obtained in comparing reading disabled and nondisabled readers, $F = 6.21$, $df = (3, 112)$, $p < .001$. No interaction between reading and age groups was found ($F = .91$, $df = (6, 224)$, $p > .4$).

Each age group differed significantly on all recall measures. Table 3 provides a summary of F-ratios and level of significance for each dependent measure across the three age groups and collapsed across reading groups. As follow-up, a Tukey test was conducted to test for significance (Harris, 1985) between each age group within each reading group on the three dependent measures. On "quality recall," reading disabled subjects performed significantly different ($p < .01$) between seven and eleven years of age but not significantly different between seven and nine years of age or nine and eleven years of age. The results obtained for

nondisabled readers were similar to the reading disabled subjects. Nondisabled readers performed significantly different between seven and eleven years of age ($p < .01$) but not with the other two respective comparisons. (See Figure 1) On "number recalled," reading disabled subjects again indicated a significant difference ($p < .05$) between seven and eleven years of age but not with the other two pairwise comparisons. The Tukey test revealed no significant difference with any age comparisons for the nondisabled readers on "number recalled." (See Figure 2) With the "average quality per design" measure, reading disabled subjects revealed no significant difference between seven and nine, seven and eleven, or nine and eleven years of age. In contrast, nondisabled readers revealed a significant difference between seven and eleven years of age ($p < .05$). (See Figure 3) Also, results indicated reading disabled subjects performed more poorly than their counterparts on all recall measures.

Table 3.

F-Ratios for each Dependent Measure Across Three Age Groups^a

Quality Recall	16.98***
Number Recalled	8.95***
Average Quality	7.22***

^adf=(2,114)

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

Three T-tests (Gravetter, 1975) were conducted within each age group to examine the degree of difference between reading disabled children and the comparison group, on each dependent measure. Table 4 provides a summary of T-scores and level of significance. Reading groups significantly differed on all three dependent measures in the seven-year-old group. The reading disabled and nondisabled readers in the nine-year-old group were significantly discriminated on the quality recall and number recalled measure ($p<.001$), but results were nonsignificant for average quality per design. Within the eleven-year-old group, reading groups were again differentiated by the quality recall measure ($p<.05$) and, this time, by average quality ($p<.001$), whereas number recalled proved nonsignificant.

Figure 1. Comparison of trend across age groups for each reading group on quality recall.

Figure 2. Comparison of trend across age groups between reading groups on number recalled.

Figure 3. Comparison of trend across age groups and between reading groups for average quality per design.

Table 4.

T-Scores for each of Three Age Groups, Between
Reading Disabled and Nondisabled Readers, for each
Dependent Measure^a

Age	7	9	11
Quality Recall	4.30***	7.96***	2.05*
Number Recalled	5.76***	10.00 ***	.75
Average Quality	2.28*	.66	4.22***

^adf=(1,38)

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

Correlations were also calculated between the three dependent measures. Since quality recall encompasses both number recalled and distortion measure (as determined by the average quality variable), moderate to high intercorrelations were obtained between the quality recall measure with number recalled and distortion measure (.82 and .57, respectively). The correlation between average quality and number recalled was .07.

The function across ages for each reading group, on each measure, was examined with a trend analysis. All functions proved significant for a linear trend except number recalled for nondisabled readers. Table 5 provides a

summary of F-ratios and level of significance for linear trends.

Table 5.

Summary Table of F-Ratio's for Linear Functions on each Measure, Across Age for Reading Groups^a

	Quality Recall	# Recall	Avg. Quality
Reading Disabled	17.5049***	12.4757***	5.2*
Nondisabled	15.6750***	1.6	6.779*

^adf=(1,114)

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

Effect size (Harris, 1985) was estimated across reading and, separately, across age groups. Effect size for quality recall, for reading and age groups, was .10 and .19, respectively. Effect size for the average quality measure for reading groups was .07, and age groups was .09. Lastly, effect size for number recalled across reading groups was .04 and across age groups was .11.

As further follow-up, a discriminant analysis was conducted with age, quality recall, number recalled, and average quality as predictor variables in discriminating reading disabled from nondisabled groups. Assumptions of Normality and Equality of Covariance Matrices were not

violated. Age was entered first. Of the remaining predictors, quality recall accounted for a larger portion of the variance (13%), $F(2,117) = 8.579$, $p < .001$, with a standardized weight of 1.59. Number recalled was the third variable placed into the discriminant function, $F(3,116) = 6.37$, $p < .001$, with a standardized weight of $-.60$. The variance accounted for by quality recall and number recalled lessened the likelihood of average quality being included in the discriminant function. The total variance accounted for by these predictor variables was approximately 14%. However, structural coefficients revealed a significant correlation of average quality, as well as the other two predictors, with the discriminant function (average quality .70; quality recall, .80; and number recalled .53).

Table 6, a classification table, summarizes the percentages of subjects accurately discriminated by the function. The significance of these results can be calculated by applying a z-score formula developed by Huberty (1984). In this instance, a z-score of 3.469 was

attained, suggesting an overall 66% classification rate was significant at $p < .001$.

Table 6.

Percent of Subjects Accurately Classified According to Age, Quality Recall, and Number Recalled

Actual Group	Number of Cases	Predicted Group Membership	
		RD	NonRD
Reading Disabled	60	40 66.7%	20 33.3%
Nondisabled Readers	60	21 35.0%	39 65.0%

DISCUSSION

As hypothesized, reading disabled and nondisabled readers were differentiated by their visual memory performance on the Bender recall. These findings support Lyle (1968) and Carroll's (1972) results which suggest that reading disabled children do perform more poorly on visual recall tasks than do nondisabled. In contrast, the significant findings in this study do not lend support to Katz, et al. (1981) or Vellutino et al.'s (1972; 1975) assertions that good and poor readers do not differ on visual working memory tasks. Their conclusion was based on studies that utilized verbally noncodable geometric figures similar to, but less complex, than the Bender designs. The present authors viewed Katz and Velluntino's findings as potentially due to a floor effect. Given the Bender recall task required disabled and nondisabled readers to recall a sequence of more complex geometric figures, it is possible the floor effect was eliminated and, thus, partially explain the significant results that were obtained. This conclusion supports Cummings and Faw (1976), Sipe and Engle (1986), and Vellutino et al. (1975), for example, who suggested free recall, after a series of designs, is apt to differentiate clinical/nonclinical groups more so than a recognition task or free recall after exposure to one design. Therefore, the

Bender recall, administered in this manner proved to be a useful diagnostic tool with this population.

Besides discriminating between reading groups, the Bender recall was a sensitive measure in detecting developmental differences between age groups with all three dependent measures. This obtained developmental trend corresponds to most studies that find an overall developmental increase in recall (cf., Chechile, 1981; Engle & Marshall, 1983; Hulme, et al., 1984). Each group exhibited a significant linear trend across age for quality recall and average quality. Reading disabled subjects also evidenced a significant linear trend across number recalled and nondisabled readers approximated one. Also, both reading groups performed significantly different between seven and eleven years of age on the quality recall measure, revealing a parallel developmental trend. This similar performance with quality recall suggested that each group shared some similar memory process. This study, however, was not designed to determine what that similarity was for example, the use of conscious rehearsal strategies (Hulme, et al., 1984) or "efficient processing" (Case et al., 1982).

Aside from the parallel trend, the reading groups differed in their performance on number recalled and average quality measures. For example, the reading disabled group demonstrated a significant difference with number recalled from seven to eleven years of age, whereas nondisabled

readers did not indicate a significant developmental increase with number recalled. In contrast, reading disabled children did not reveal a significant decline with distortions in recall from seven to eleven years of age as did the nondisabled readers.

The present authors attributed this divergence in performance to the types of scoring methods used. Thus, upon closer examination, number recalled and average quality per design possibly measured two different constructs (e.g., span versus distortion, respectively) since they correlated minimally with one another. To further support this notion, it appeared that these two variables discriminated reading groups differently at each age level. For instance, number recalled was a better discriminator of reading groups at seven years of age whereas the distortion measure was a greater discriminator of readers at eleven years of age. Also, reading disabled childrens' performance on average quality per design seemed to stabilize, whereas nondisabled readers' performance continued to ascend. These various results suggested the presence of two different processes.

The three different scoring measures were further examined to assess their relative contributions in discriminating groups. These variables, combined, correctly classified a significant percentage of the children into appropriate reading groups. Also, the structural coefficient of average quality per design correlated more

with the discriminant function than number recalled. This supports Lyle's (1968) study in which measures of distortion from memory significantly differentiated good from poor readers. Lyle concluded that distortion in recall provided valuable information, aside from number recalled, per se. Carroll (1972) also found a significant correlation between errors in visual recall and reading achievement levels of young children. Therefore, it is suggested these two scoring methods may be useful, independent, indices of visual memory ability.

The results of this study also revealed a developmental lag between reading groups as predicted. Reading disabled children recalled less and exhibited more distortions per design than nondisabled readers within each age group. The reading disabled children did not "catch up" to the nondisabled readers at any point, as Vellutino (1975) proposed. Rather, the results supported Bauer (1977), Bryant and Impey (1986), and Tarver et al.'s (1976) developmental lag hypothesis but not necessarily their respective theoretical explanations.

Therefore, for practical application (i.e., individual psychological testing), the Bender recall has the potential to offer valuable diagnostic information -- developmental, distortion in recall, and memory span. Of the three scoring methods, quality recall may be the most appropriate measure to utilize. Number and average quality provide a limited

range of data points. Number recalled ranges from "0" to "9." Average quality ranges from "0" to "3." Both ranges restrict an examiner's ability to determine an adequate versus inadequate recall profile on an individual basis. The quality recall measure, incorporating these two distinct variables, provides a range from "0" to "27." Also, quality recall accounts for most of the variance in the discriminant function and, therefore, seemingly the preferable measure.

Future research should focus on improving the scaling of the recall segment and examination of other relevant populations. Considering the exploratory nature of this study, the Bender recall's temporal stability must be established. Replication of this study, with a built-in cross-validation, is needed. Further comparisons of other memory deficient and nondeficient populations is necessary to examine the extent of the recall's utility. A developmental comparison extension of both reading groups into adolescence would enhance our understanding of the similarities/differences between good and poor readers on memory tests. For instance, it may be interesting to explore whether reading disabled children ever approximate good readers in their visual memory performances.

Further research should focus on comparisons of recall on phonological memory-based tasks and recall of sequences of complex visual stimuli. This procedure would correct for weaknesses noted in studies that solely advocate for a

phonological coding deficit hypothesis in poor readers' memory performance (cf., Brady, 1986; Liberman, et al., 1982; Vellutino, 1979). An extension of this study might consist of looking at other predictive measures that would increase the likelihood of discriminating good from poor readers and incorporating them into a classification equation.

Other future research may be directed towards examination of distortions in recall. Studies of this nature may provide information on higher order storage/retrieval processes in both disabled and nondisabled populations. Continuation with this aspect of memory may make a significant impact on theory and study of memory and become a valuable tool in neuropsychological evaluations.

REFERENCES

- Anastasi, A. (1982). Psychological testing. New York: Macmillan.
- Armentrout, J. A. (1976). Bender Gestalt recall: Memory measure or intelligence estimate? Journal of Clinical Psychology, 32, 832-834.
- Bauer, R. H. (1977). Memory processes in children with learning disabilities: Evidence for deficient rehearsal. Journal of Experimental Child Psychology, 24, 415-530.
- Bauer, R. H. (1979). Memory, acquisition, and category clustering in learning disabled children. Journal of Experimental Child Psychology, 27, 365-383.
- Benton, A. L. (1975). Developmental dyslexia: Neurological aspects. In W. J. Friedlander (Ed.). Advances in neurology, 7. New York: Raven.
- Best, J. (1986). Cognitive Psychology. St. Paul, MN: West.
- Boder, E. (1973). Developmental dyslexia: A diagnostic approach based on three atypical reading-spelling patterns. Developmental Medicine and Child Neurology, 15, 63-87.
- Brady, S. (1986). Short-term memory, phonological processes and reading ability. Annals of Dyslexia, 36, 138-153.

- Brainerd, C. J., & Pressley, M. (Eds.). (1985). Basic processes in memory development. New York: Springer-Verlag.
- Brook, R. M. (1975). Visual Retention test: Local norms and impact of short-term memory. Perceptual and Motor Skills, 40, 967-970.
- Bryant, P., & Impey, L. (1986). The similarities between normal readers and developmental and acquired dyslexics. Cognition, 24, 121-137.
- Bjorklund, D. F. (1985). The role of conceptual knowledge in the development of organization in children's memory. In C. J. Brainerd & M. Pressley (Eds.). Basic processes in memory development. New York: Springer-Verlag.
- Carroll, J. L. (1972). A visual memory scale (VMS) designed to measure short-term visual recognition memory in 5- and 6-year-old children. Psychology in the Schools, 9, 152-158.
- Case, R., Kurland, D. M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. Journal of Experimental Child Psychology, 33, 386-404.
- Chechile, R. A., & Richman, C. L. (1982). The interaction of semantic memory with storage and retrieval processes. Developmental Review, 2, 237-250.

- Chechile, R. A., Richman, C. L., Topinka, C., & Ehrensbeck, K. (1981). A developmental study of the storage and retrieval of information. Child Development, 52, 251-259.
- Cummings, E. M., & Faw, T. T. (1976). Short-term memory and equivalence judgments in normal and retarded readers. Child Development, 47, 286-289.
- Doehring, D. G. (1985). Reading disability subtypes: Interaction of reading and nonreading deficits. In B. P. Rourke (Ed.). Neuropsychology of learning disabilities, (pp. 133-146). New York: Guilford.
- Durost, W., Bixler, H., Wrightstone, J., Prescott, G., & Balow, I. (1986). Metropolitan Achievement Test. New York: Harcourt Brace Jovanovich.
- Engle, R., & Marshall, K. (1983). Do developmental changes in digit span result from acquisition strategies? Journal of Experimental Child Psychology, 36, 429-436.
- Finch, A. J., Spirito, A., Garrison, S., & Marshall, P. (1983). Developmental differences in Bender-gestalt recall of children with learning and behavior problems. Perceptual and Motor Skills, 56, 87-90.
- Goodman, G. S., & Haith, M. M. (1987). Memory development and neurophysiology: Accomplishments and limitations. Child Development, 58, 713-717.

- Graham, F. K., & Kendall, B. S. (1960). Memory-for-designs test: Revised general manual. Perceptual and Motor Skills, 11, 147-188.
- Gravetter, F. (1975). Statistics for psychology. Lexington, MA: Xerox Individualized Publishing.
- Grundig, J. (1970). Comparison of different scoring techniques of visual perceptual memory. Journal of Clinical Psychology, 26, 353-357.
- Hanawalt, N. G. (1959). Review of the Benton Visual Retention test. In O. K. Buros (Ed.). The fifth mental measurements yearbook, (pp. 536-537). Highland Park, NJ: Gryphon.
- Hanna, G., Dyck, N., & Holen, M. (1979). Objective analysis of achievement-aptitude discrepancies in LD classification. Learning Disability Quarterly, 2, 32-38.
- Harris, R. (1985). A primer of multivariate statistics, second edition. Orlando, FL: Academic Press.
- Howe, M. (1983). Introduction to the psychology of memory. New York: Harper & Row.
- Huberty, C. (1984). Issues in the use and interpretation of discriminant analysis. Psychological Bulletin, 95, 156-171.

- Hulme, C., Thomson, N., Muir, C., & Lawrence, A. (1984). Speech rate and the development of short-term memory span. Journal of Experimental Child Psychology, 38, 241-253.
- Huttenlocher, J., & Burke, D. (1976). Why does memory span increase with age? Cognitive Psychology, 8, 1-31.
- Hutton, J. (1966). Bender recall of children as related to age and intelligence. Perceptual and Motor Skills, 23, 34.
- Hynd, G., & Cohen, M. (1983). Dyslexia: Neuropsychological theory, research, and clinical differentiation. New York: Grune & Stratton.
- Katz, R. B., Shankweiler, D., & Liberman, I. Y. (1981). Memory for item order and phonetic recoding in the beginning reader. Journal of Experimental Child Psychology, 32, 474-484.
- Koppitz, E. M. (1965). The Bender Gestalt test for young children. New York: Grune & Stratton.
- Liberman, I. Y., Mann, V., Shankweiler, D., & Werfelman, M. (1982). Children's memory of recurring linguistic and nonlinguistic material in relation to reading ability. Cortex, 18, 367-375.
- Lyle, J. G. (1968). Performance of retarded readers on Memory-for-designs test. Perceptual and Motor Skills, 26, 851-854.

- Lyle, J. G., & Goyen, J. (1975). Effects of speed of exposure and difficulty of discrimination on visual recognition of retarded readers. Journal of Abnormal Psychology, 8, 613-616.
- Marsh, G., & Hirsch, S. (1982). Effectiveness of two tests of visual retention. Journal of Clinical Psychology, 38, 115-118.
- Rogers, D. L., & Swenson, W. M. (1975). Bender-gestalt recall as a measure of memory versus distractibility. Perceptual and Motor Skills, 40, 919-922.
- Rourke, B. P. (1975). Brain-behavior relationships in children with learning disabilities. American Psychologist, 911-918.
- Sattler, J. M. (1982). Assessment of children's intelligence and special abilities. Boston, MA: Allyn and Bacon, Inc.
- Seymour, P. (1986). Cognitive analysis of dyslexia. New York: Routledge & Kegan Paul.
- Sipe, S., & Engle, R. W. (1986). Echoic memory processes in good and poor readers. Journal of Experimental Psychology: Learning, Memory and Cognition, 12, 402-412.
- Smith, M. D., & Rogers, M. (1978). Reliability of standardized instruments when used with learning disabled children. Learning Disability Quarterly, 1, 23-31.

- Tarver, S., Hallahan, D., Kauffman, J., & Ball, D. (1976). Verbal rehearsal and selective attention in children with learning disabilities: A developmental lag. Journal of Experimental Child Psychology, 22, 375-385.
- Temple, C. M. (1987). The nature of normality, the deviance of dyslexia and the recognition of rhyme: A reply to Bryant and Impey (1986). Cognition, 27, 103-108.
- Tolor, A. (1956). A comparison of the Bender-gestalt test and the digit span test as measures of recall. Journal of Consulting Psychology, 20, 305-309.
- Torgeson, J. K., & Houck, D. G. (1980). Processing deficiencies of learning disabled children who perform poorly on the digit span test. Journal of Educational Psychology, 72, 141-160.
- U. S. Department of Health, Education, and Welfare. (1977). Assistance to states for education of handicapped children. Federal Register, 42, 65083.
- Vellutino, F. R. (1979). Dyslexia: Theory and research. Cambridge, MA: MIT Press.
- Vellutino, F. R. (1987). Dyslexia. Scientific American, 256(3), 34-41.
- Vellutino, F. R., Smith, H., Steger, J. A., & Kaman, M. (1975). Reading disability: Age differences and the perceptual-deficit hypothesis. Child Development, 46, 487-493.

- Vellutino, F. R., Steger, J., & Kandel, G. (1972). Reading disability: An investigation of the perceptual deficit hypothesis. Cortex, 8, 106-118.
- Vellutino, F. R., Steger, J. A., Kaman, M., & DeSetto, L. (1975). Visual form perception in deficient and normal readers as a function of age and orthographic-linguistic familiarity. Cortex, 11, 22-30.
- Walters, C. (1961). Reading ability and visual-motor function in second grade children. Perceptual and Motor Skills, 13, 370.
- Wechsler, D. (1974). Manual for the Wechsler Intelligence Scale for Children, Revised. New York: Psychological Corporation.
- Zingale, S., Smith, M., & Doeckci, P. (1980). Temporal stability of the Metropolitan Achievement test when used with learning disabled children. Learning Disability Quarterly, 3, 84-86.

APPENDIX A

Designs and types of errors scored by Koppitz accordingly.

Plate 1. Nine Figures of the Bender Gestalt Test, adapted from Wertheimer. (Reproduced from Plate 1 of Research Monograph No. 3, "A Visual Motor Gestalt Test and its Clinical Use," published by the American Orthopsychiatric Association in 1938. Copyright, the American Orthopsychiatric Association, Inc. Reproduced by permission.)

- Sample A. Distortion; Rotation; Integration
 Figure 1. Distortion; Rotation; Perseveration
 Figure 2. Rotation; Integration; Perseveration
 Figure 3. Distortion; Integration; Rotation
 Figure 4. Rotation; Integration
 Figure 5. Distortion; Rotation; Integration
 Figure 6. Distortion; Integration; Perseveration
 Figure 7. Distortion; Rotation; Integration
 Figure 8. Distortion; Rotation

APPENDIX B

SAMPLE OF PROTOCOL WITH SCORING

ORIGINAL

RECALL

Scoring

Figure A =	0	
1 =	2	rotation
2 =	0	
3 =	0	
4 =	2	rotation
5 =	3	same as perceived
6 =	2	rotation
7 =	2	rotation
8 =	3	same

APPENDIX C

KOPPITZ' DEFINITIONS

Koppitz defines distortion, rotation, integration, and perseveration in detail for each of the nine designs.

General definitions, however, for scoring a design as such are provided below:

Distortion: Shape is "excessively misshapen," for example, circles have points, squares have extra or missing angles, or dots are converted into circles. Also, the size is disproportioned, for example, one shape is twice as large as the other.

Rotation: The total figure or part of it is rotated by 45 degrees or more.

Integration: Shapes are not "integrated." For example, a circle and square are not joined or overlap more than 1/8 inch; one or two rows of circles are omitted or added; sometimes the shape is lost due to poor integration; or two lines (e.g., Figure 6) do not cross or do so at extreme ends of the other line.

Perseveration: An "excessive" number of dots or curves are included in the design.

APPENDIX D

WARREN PUBLIC SCHOOLS

P. O. Box 404
Warren, Rhode Island 02885

Superintendent's Office (401) 245-1150
Business Office 245-4687
Warren High School 245-2010
Mary V. Quirk School 245-3392
Hugh Cole School 245-1460

, 1987

Main Street School 245-6739
Child Street School 245-2391
Special Services Dept. 245-7649
245-1990

Dear Parent(s):

We are instituting a project within the Warren School System where your child's participation would be of great value to us. We would like to compare children without learning disabilities, such as yours, to those with learning disabilities. Unfortunately, there are children who suffer from learning disabilities, such as in reading or math, who require special services throughout their school years. Research is needed to further our understanding of these children's problems to improve their chances of overcoming such disabilities. The current project would compare your child's performance on a simple test of drawing designs with that of a child who experiences a learning disability. We will also administer a brief test measuring general aptitude. If you permit your child to participate, and he or she is selected, only 30 minutes of your child's time would be required. Testing would be performed in May or June during school hours and at a time convenient for the teacher and your child. (This study is conducted under the supervision of Dr. Allan Berman of the University of Rhode Island.)

Participants and results will be kept strictly anonymous. Results will be used for research purposes only. You may withdraw at any time. If you decide not to participate this will not be reflected in your child's school records.

Your assistance is greatly appreciated and needed. You will be making an important contribution. If there are any questions, feel free to contact Bernadette Evans at 245-1990.

Please check one of the responses below, indicating your preference, and have your son or daughter return this letter no later than _____. Thank you very much!

Sincerely,

Bernadette Evans
Bernadette Evans, Psychologist Intern

Keith H. Sargent
Keith H. Sargent, Special Education Director

L. D. Perella
Dr. Perella, Principal

— YES, I AGREE TO LET MY CHILD PARTICIPATE
IN YOUR PROJECT. I have informed my child
that he/she may be selected for this project
and that he/she may withdraw at any time.

Parent's Signature, Date

Child's Signature

— No, I prefer my child not participate.

BIBLIOGRAPHY

- Anastasi, A. (1982). Psychological testing. New York: Macmillan.
- Armentrout, J. A. (1976). Bender Gestalt recall: Memory measure or intelligence estimate? Journal of Clinical Psychology, 32, 832-834.
- Bauer, R. H. (1977). Memory processes in children with learning disabilities: Evidence for deficient rehearsal. Journal of Experimental Child Psychology, 24, 415-530.
- Bauer, R. H. (1979). Memory, acquisition, and category clustering in learning disabled children. Journal of Experimental Child Psychology, 27, 365-383.
- Benton, A. L. (1975). Developmental dyslexia: Neurological aspects. In W. J. Friedlander (Ed.). Advances in neurology, 7. New York: Raven.
- Best, J. (1986). Cognitive Psychology. St. Paul, MN: West.
- Boder, E. (1973). Developmental dyslexia: A diagnostic approach based on three atypical reading-spelling patterns. Developmental Medicine and Child Neurology, 15, 63-87.
- Brady, S. (1986). Short-term memory, phonological processes and reading ability. Annals of Dyslexia, 36, 138-153.

- Brainerd, C. J., & Pressley, M. (Eds.). (1985). Basic processes in memory development. New York: Springer-Verlag.
- Brook, R. M. (1975). Visual Retention test: Local norms and impact of short-term memory. Perceptual and Motor Skills, 40, 967-970.
- Bryant, P., & Impey, L. (1986). The similarities between normal readers and developmental and acquired dyslexics. Cognition, 24, 121-137.
- Bjorklund, D. F. (1985). The role of conceptual knowledge in the development of organization in children's memory. In C. J. Brainerd & M. Pressley (Eds.). Basic processes in memory development. New York: Springer-Verlag.
- Carroll, J. L. (1972). A visual memory scale (VMS) designed to measure short-term visual recognition memory in 5- and 6-year-old children. Psychology in the Schools, 9, 152-158.
- Case, R., Kurland, D. M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. Journal of Experimental Child Psychology, 33, 386-404.
- Chechile, R. A., & Richman, C. L. (1982). The interaction of semantic memory with storage and retrieval processes. Developmental Review, 2, 237-250.

- Chechile, R. A., Richman, C. L., Topinka, C., & Ehrensbeck, K. (1981). A developmental study of the storage and retrieval of information. Child Development, 52, 251-259.
- Cummings, E. M., & Faw, T. T. (1976). Short-term memory and equivalence judgments in normal and retarded readers. Child Development, 47, 286-289.
- Doehring, D. G. (1985). Reading disability subtypes: Interaction of reading and nonreading deficits. In B. P. Rourke (Ed.). Neuropsychology of learning disabilities, (pp. 133-146). New York: Guilford.
- Durost, W., Bixler, H., Wrightstone, J., Prescott, G., & Balow, I. (1986). Metropolitan Achievement Test. New York: Harcourt Brace Jovanovich.
- Engle, R., & Marshall, K. (1983). Do developmental changes in digit span result from acquisition strategies? Journal of Experimental Child Psychology, 36, 429-436.
- Finch, A. J., Spirito, A., Garrison, S., & Marshall, P. (1983). Developmental differences in Bender-gestalt recall of children with learning and behavior problems. Perceptual and Motor Skills, 56, 87-90.
- Goodman, G. S., & Haith, M. M. (1987). Memory development and neurophysiology: Accomplishments and limitations. Child Development, 58, 713-717.

- Graham, F. K., & Kendall, B. S. (1960). Memory-for-designs test: Revised general manual. Perceptual and Motor Skills, 11, 147-188.
- Gravetter, F. (1975). Statistics for psychology. Lexington, MA: Xerox Individualized Publishing.
- Grundig, J. (1970). Comparison of different scoring techniques of visual perceptual memory. Journal of Clinical Psychology, 26, 353-357.
- Hanawalt, N. G. (1959). Review of the Benton Visual Retention test. In O. K. Buros (Ed.). The fifth mental measurements yearbook, (pp. 536-537). Highland Park, NJ: Gryphon.
- Hanna, G., Dyck, N., & Holen, M. (1979). Objective analysis of achievement-aptitude discrepancies in LD classification. Learning Disability Quarterly, 2, 32-38.
- Harris, R. (1985). A primer of multivariate statistics, second edition. Orlando, FL: Academic Press.
- Howe, M. (1983). Introduction to the psychology of memory. New York: Harper & Row.
- Huberty, C. (1984). Issues in the use and interpretation of discriminant analysis. Psychological Bulletin, 95, 156-171.

- Hulme, C., Thomson, N., Muir, C., & Lawrence, A. (1984). Speech rate and the development of short-term memory span. Journal of Experimental Child Psychology, 38, 241-253.
- Huttenlocher, J., & Burke, D. (1976). Why does memory span increase with age? Cognitive Psychology, 8, 1-31.
- Hutton, J. (1966). Bender recall of children as related to age and intelligence. Perceptual and Motor Skills, 23, 34.
- Hynd, G., & Cohen, M. (1983). Dyslexia: Neuropsychological theory, research, and clinical differentiation. New York: Grune & Stratton.
- Katz, R. B., Shankweiler, D., & Liberman, I. Y. (1981). Memory for item order and phonetic recoding in the beginning reader. Journal of Experimental Child Psychology, 32, 474-484.
- Koppitz, E. M. (1965). The Bender Gestalt test for young children. New York: Grune & Stratton.
- Liberman, I. Y., Mann, V., Shankweiler, D., & Werfelman, M. (1982). Children's memory of recurring linguistic and nonlinguistic material in relation to reading ability. Cortex, 18, 367-375.
- Lyle, J. G. (1968). Performance of retarded readers on Memory-for-designs test. Perceptual and Motor Skills, 26, 851-854.

- Lyle, J. G., & Goyen, J. (1975). Effects of speed of exposure and difficulty of discrimination on visual recognition of retarded readers. Journal of Abnormal Psychology, 8, 613-616.
- Marsh, G., & Hirsch, S. (1982). Effectiveness of two tests of visual retention. Journal of Clinical Psychology, 38, 115-118.
- Rogers, D. L., & Swenson, W. M. (1975). Bender-gestalt recall as a measure of memory versus distractibility. Perceptual and Motor Skills, 40, 919-922.
- Rourke, B. P. (1975). Brain-behavior relationships in children with learning disabilities. American Psychologist, 911-918.
- Sattler, J. M. (1982). Assessment of children's intelligence and special abilities. Boston, MA: Allyn and Bacon, Inc.
- Seymour, P. (1986). Cognitive analysis of dyslexia. New York: Routledge & Kegan Paul.
- Sipe, S., & Engle, R. W. (1986). Echoic memory processes in good and poor readers. Journal of Experimental Psychology: Learning, Memory and Cognition, 12, 402-412.
- Smith, M. D., & Rogers, M. (1978). Reliability of standardized instruments when used with learning disabled children. Learning Disability Quarterly, 1, 23-31.

- Tarver, S., Hallahan, D., Kauffman, J., & Ball, D. (1976). Verbal rehearsal and selective attention in children with learning disabilities: A developmental lag. Journal of Experimental Child Psychology, 22, 375-385.
- Temple, C. M. (1987). The nature of normality, the deviance of dyslexia and the recognition of rhyme: A reply to Bryant and Impey (1986). Cognition, 27, 103-108.
- Tolor, A. (1956). A comparison of the Bender-gestalt test and the digit span test as measures of recall. Journal of Consulting Psychology, 20, 305-309.
- Torgeson, J. K., & Houck, D. G. (1980). Processing deficiencies of learning disabled children who perform poorly on the digit span test. Journal of Educational Psychology, 72, 141-160.
- U. S. Department of Health, Education, and Welfare. (1977). Assistance to states for education of handicapped children. Federal Register, 42, 65083.
- Vellutino, F. R. (1979). Dyslexia: Theory and research. Cambridge, MA: MIT Press.
- Vellutino, F. R. (1987). Dyslexia. Scientific American, 256(3), 34-41.
- Vellutino, F. R., Smith, H., Steger, J. A., & Kaman, M. (1975). Reading disability: Age differences and the perceptual-deficit hypothesis. Child Development, 46, 487-493.

- Vellutino, F. R., Steger, J., & Kandel, G. (1972). Reading disability: An investigation of the perceptual deficit hypothesis. Cortex, 8, 106-118.
- Vellutino, F. R., Steger, J. A., Kaman, M., & DeSetto, L. (1975). Visual form perception in deficient and normal readers as a function of age and orthographic-linguistic familiarity. Cortex, 11, 22-30.
- Walters, C. (1961). Reading ability and visual-motor function in second grade children. Perceptual and Motor Skills, 13, 370.
- Wechsler, D. (1974). Manual for the Wechsler Intelligence Scale for Children, Revised. New York: Psychological Corporation.
- Zingale, S., Smith, M., & Doeckci, P. (1980). Temporal stability of the Metropolitan Achievement test when used with learning disabled children. Learning Disability Quarterly, 3, 84-86.