THE RELATIONSHIP BETWEEN THE BENDER GESTALT TEST AND THE FROSTIG DEVELOPMENTAL TEST OF VISUAL PERCEPTION

Thesis Approved:

Kenneth D. Sedwick
Thesis Adviser

Dwaine J. Gould

D N Durham
Dean of the Graduate College

PREFACE

I would like to acknowledge the assistance of Drs. D. Ray, K. Sandvold, and D. Tyrrell, whose advice and criticism have been most helpful. In addition, a special thanks goes to my wife, Beverly who helped conceptualize the problem.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

The Problem

The purpose of this thesis is to explore the relationship between two clinical instruments, the Bender Visual Motor Gestalt Test, using Koppitz's Developmental Scoring System, and the Marianne Frostig Developmental Test of Visual Perception (Frostig, Lefever, and Whittlesey, 1964).

Review of the Literature

The Bender Visual Motor Gestalt Test (hereafter referred to as the Bender) is the third or fourth most frequently used psychological instrument in clinical diagnosis (Schulburg and Tolor, 1961). In a sample of experienced clinicians, they found only 5% who never used this test; four of every five clinicians sampled thought the Bender had "some or great value for clinical assessments" - either as a source of tentative personality hypotheses to be verified by other test data and behavioral observations, or for differential diagnosis or research projects.

History

The Bender Gestalt Test was developed in 1932 by Loretta Bender (1932). The test consists of nine geometric designs adopted from Max Wertheimer (1923). In the original administration of the test,

Wertheimer had subjects describe what they saw. When Bender first used the test, she had subjects copy the figures, thus the test became a visual-motor task.

Visual vs. Motor Factors. Evaluations of Bender Gestalt Test protocols pose interesting problems. Are the figures perceived accurately, but distorted in reproduction; or are the figures misperceived and accurately reproduced with the distorted perceptions? Two studies have investigated this problem.

McPherson and Pepin(1955) used two different methods of motor manipulation to ascertain whether performance on the Bender Gestalt Test is influenced primarily by motor or by perceptual factors. They compared the accuracy of designs reproduced in the conventional way with those constructed by placing pieces of felt on a felt board. They concluded that motor factors were not determining factors.

In the second study, Niebuhur and Cohen (1956) used a psychiatric hospital population. They employed separate memory and copy conditions. Motor factors were of importance only as the psychopathology of the subjects increased; holding psychopathology constant, motor factors were not of great importance.

In both of these studies, the nature of the relationship between perceptual and motor efficiency remained ambivalent. The fact that adult populations were used in both studies restricts the applicability of the results to children, especially in view of the developmental nature of perception and motor coordination. Secondly, the use of a hospital population prevents generalization to a normal population. Third, although the subjects had organic brain damage, they did not necessarily constitute a homogeneous group; thus, individual differences

in psychopathology may account for the obtained results. Thus, the issue of perceptual vs. motor factors remains unsettled.

Kephart (1960) tried to by-pass the whole issue. He suggests that perceptual and motor abilities are so interrelated, that any dichotomy of the two is artificial.

Getman (1965) goes beyond Kephart. He suggests that perceptual and motor abilities are so interrelated and inseparable that the words "visual-motor" ought not to be hyphenated. He speaks, instead, of visuamotor factors. However, both authors fail to supply empirical evidence to support their viewpoints.

Developmental Studies

Since development plays a crucial role in the reproduction of the Bender Gestalt Test figures, it has been studied in two different ways. One approach relies on a correlational methodology; investigators dealing primarily with other problems have incidentally reported on the relationship between their subjects' ages and performance on the Bender Gestalt Test. The second, and more intensive approach, depends on specific investigations of the changing nature of test functioning in a growing child and attempts to develop norms for different age levels.

Correlational Studies. Four studies have reported a correlation between age and Bender functioning. In three, the relationship was reported to be significant, while in one it was not.

Pascal and Suttell (1951) studied forty-six normal children, ranging in age from six-three to nine-three years. A product moment r of -.58 was found between the children's age and raw Pascal-Suttel Bender Gestalt score. Pascal (1952) cited an unpublished thesis by Suttell,

in which protocols were obtained from 500 normal children, ranging in age from six to fourteen years. The relationship between age and Bender scores was found to be curvilinear; the eta correlation was .63.

Armstrong and Houck (1960) studied 49 boys and 49 girls, ranging in age from six to twelve years, at a child guidance center. The Pascal-Suttell scores decreased steadily with age, and a correlation of -.60 was reported.

Zolik (1958) used a sample of 43 delinquents and 43 nondelinquents. The mean age for each group was sixteen. Correlations of .04 and .01 were between Bender performance and age obtained for the delinquent and nondelinquent groups, respectively.

In summary, the correlational studies cited seem to suggest that age is significantly related to the quality of Bender reproductions, for young subjects under age 16.

Maturational Studies. Maturational studies designed to investigate the specific changes associated with development of perceptual-motor ability in young children have been conducted. In these studies, the sophistication of the experimental methodology has varied greatly.

Early studies of the Bender Gestalt Test focused on the nature of the motoric movements made by young children. Schilder (1934) noted that the first scribblings represent motor play for the pleasure of motor expression; they are performed through large arm movements. Gradually, children tend to begin to inhibit their arm movements.

Bender (1933, 1938) described the visual-motor patterns of children two-six to four years old. She agreed with the findings of Schilder (1934). In the above studies by Bender involving children four through seven years of age, she found rapid growth in ability to

distinguish forms, between these ages. Bender thought that there was a constant interplay between motoric and sensory features, and these features were inseparable.

Bender (1952) considered the Bender Gestalt Tests designs valuable as a maturational test of visual-motor performance for subjects between the ages of four and twelve. By the age of eleven, all of the figures are reproduced satisfactorily; adults show no basic improvements in the design of the figures, but only in judging size and distance.

Koppitz (1960) tested 1,055 school children between the ages of five and ten years. She developed a scoring system for the Bender Test whereby scores decreased as the subjects got older. Between the ages of five and seven, there is a marked drop in scores; the decrease in scores then becomes more gradual and levels off at about the age of nine when most subjects draw the designs correctly. Koppitz contended that until the age of eight, the Bender Gestalt Test discriminates both those with outstanding and those with immature visual-motor abilities. After age 8, only poor and satisfactory performance can be discriminated.

In conclusion, the maturational studies seem to be in close agreement. There is a generally observed random scribbling in very young children, scribbling which gradually becomes more refined. The ability to reproduce form rapidly increases between the ages of four and seven. Between the ages of nine and eleven, all of the designs can be reproduced successfully and the relevant developmental processes seem completed.

Children's Scoring System

The earliest method of evaluating a child's Bender Test protocol relied on the developmental table presented by Bender (1938, 1946). It permits the clinician to compare his subjects' protocols with the reproductions Bender obtained in her sample. The clinician can then decide which age level his subject's protocol most closely approximates. This decision is basically subjective. Because Bender's scoring system is so subjective in nature, little research has been conducted to determine its reliability and validity.

Keller (1955) was the first to try to develop a quantitative scoring system intended exclusively for children. However, all of his subjects were mentally retarded children, so that the general utility of his system is not known.

Keogh and Smith (1961) tested a kindergarten population. They obtained Bender protocols and scores them using the Pascal-Suttell system (1951) as well as their own. They developed a five point rating scale designed to yield a single total score based on "discriminated degrees of quality of production". The rating scale ranges from a score of one, assigned when the reproduction is "not recognizable" to a score of five, assigned when "all parts of the design are present and recognizable in good form". The scale was found to have an intra-judge reliability of .86, an inter-judge reliability of .83, and an .80 correlation with the Pascal-Suttel raw score.

Koppitz' scoring system (1960) is of major interest and is used to evaluate the Bender protocols in this study. Her normative population consisted of 1055 school children, representing 44 entire classes in 11 different schools, located in rural, small town, surburban, and urban

areas. The sample included both Negro and white children, whose ages ranged from five to ten-five years. Protocols were scored by Koppitz with no knowledge of subjects' age, sex, and grade placement. The Koppitz' scoring system was devised specifically for use with children. It had been cross-validated in two previous studies (Koppitz, 1958a, 1958b). The Revised Scoring System (see Appendix A) includes only items which consistently differentiate between above and below average students in the first and second grades. It consists of 30 scoring items, none of which are mutually exclusive. Each item is scores as 1 when present and a 0 when absent. A child may obtain a maximum score of 30. The scoring is for errors only. Thus, a high score indicates poor test performance, while a low score indicates high test performance.

Miller, Loewenfeld, Linder and Turner (1962) made an inter-observer reliability investigation of the Koppitz Developmental Bender Scoring System. A sample of 30 Bender Gestalt Tests was drawn from a 100 cases at a guidance center and scored independently by five raters. The results showed 15 intercorrelations ranged from .88 to .96 with a mean correlation among raters of .93. Koppitz participated by correspondence in this study. The authors concluded that the Koppitz Scoring System can be learned and applied with consistent agreement.

Koppitz (1964, p. 13) notes that the split-half method and the alternate form method are not appropriate reliability measures. She suggests instead a test-retest method. Using the test-retest method with an interval of four months, Kendall's Rank Correlational Coefficients ranged from .547 to .659.

Koppitz (1958b) conducted two studies exploring the relationship between the Bender Gestalt Test and the WISC. She sampled 90 children who

had been referred to a school psychologist: 20 first graders, 20 second graders, 25 third graders, and 25 fourth graders. The subjects were divided into two groups. One group included the first and second graders, and the other group consisted of the third and fourth graders. Separate chi-squares were computed for WISC Full Scale IQ, Verbal IQ and Performance IQ, comparing the number of subjects with above and below average Bender performance to those with above and below average WISC IQ's. It was found that the WISC Full Scale IQ was more closely related to the Bender performance of third and fourth graders than to the Bender performance of first and second graders. The divergence in correlations between the first and second grade on the one hand, and the third and fourth grades on the other was explained by Koppitz as resulting from the considerable maturation in visual-motor perception that occurs between six and seven years. In addition, the normative curve for the Developmental Bender Scoring System displayed a marked irregularity between the six and a half and seven year levels (Koppitz, 1964, p. 34). The verbal IQ was significantly related to the Bender Gestalt Test of first but not second graders. The Performance IQ was significantly related to Bender Gestalt Test performance for both groups. These results should be interpreted cautiously as (Koppitz's) Revised Scoring System was not used.

In a second study (Koppitz, 1964, p. 46) Bender protocols were correlated wither with IQ scores on the Stanford-Binet Intelligence Scale, Form L or with IQ scores on the Wechsler Intelligence Scale for children. Two-hundred and thirty-nine children were tested. These children were seen for emotional or learning problems by a school psychologist or at a guidance center. The children ranged in age from

5 to 10 years. The subjects were divided into two groups: the first consisted of 176 children with IQ's ranging from 75 to 149. The second contained 63 retarded children with an IQ range from 40 to 74. Correlations were reported for the first group by age levels. All correlations were statistically significant ranging from -.79 to -.48. For the retarded group, the correlation was -.44. It is interesting to note that the highest correlation was among five year olds. Koppitz interprets this to suggest that the Bender Gestalt Test may be a good screening instrument for children entering school.

Tests of Readiness

A number of studies have attempted to explore experimentally that the Bender Test is a good screening device. Koppitz, Mardis and Stephens (1961) explored the relationship between the Lee-Clark Reading Readiness Test and the Metropolitan Readiness Test, as well as both of these tests with achievement (Metropolitan Achievement Test), and with the Bender Gestalt Test. Correlations ranged from -.21 to -.67 for the Bender Gestalt Test and the Readiness Tests. It was also found that the Bender Gestalt Test correlated as well with achievement as did the Readiness Tests.

Summary - Bender Gestalt Test

In conclusion, it is to be noted that there is an absence of validation data, except, possibly, for the normative sample using age level as an approach to validity. However, no data is available to support the conclusion that the Bender Gestalt Test, using Koppitz' scoring system is an indicator of visual-motor ability beyond the visual-motor

performance required of the test. It is also to be noted that studies correlating intelligence with Bender functioning relied on clinical populations. One may question, then, the applicability of obtained results to normal populations.

The Marianne Frostig Developmental Test of Visual Perception

The Marianne Frostig Developmental Test of Visual Perception (Frostig et al., 1961) and hereafter referred to as the Frostig is relatively new test. It was developed at the Marianne Frostig School of Educational Therapy. Unlike the Bender Gestalt Test, it has not yet found its way into wide use.

History

The Developmental Test of Visual Perception stemmed from clinical observation of children with brain damage and/or learning disabilities at the Marianne Frostig School of Educational Therapy. Problems in visual perception were noted frequently, specifically in five areas: eye-hand coordination, figure-ground perception, form constancy, position in space, and spatial relationships (Frostig et al., 1961). The Developmental Test of Visual Perception was constructed to attempt to explore these five areas. The test consists of five subtests.

1. Eye-hand coordination. The child's task here is to draw straight and curved lines between increasingly narrow boundaries.
2. Figure-ground perception. Here the child is asked to discriminate figures between intersecting figures.

3. Perception of Form Constancy. The task here is to detect squares and circles and distinguish them from other forms on the page.
4. Perception of position in space. This test requires the child to detect a reversed or rotated figure in a sequence.
5. Perception of spatial relationships. The task here is to copy patterns by linking dots.

An attempt was made to discriminate between perceptual and motor abilities. Subtest 1 requires "simple motor skill" and subtest 5 requires copying ability. Subtests 2, 3, 4, require only perceptual recognition.

The units of test measurement consist of perceptual age level, scale scores, and perceptual quotients. The perceptual age level for each subtest was defined in terms of the average performance of children in age groups from three to nine years. Scale scores are the perceptual ages divided by the chronological ages, multiplied by 100, and adjusted to the nearest whole number. The Perceptual Quotient (hereafter referred to as the PQ) is a deviation score obtained from the sum of subtest scales scores, corrected for age variation. The PQ has a median of 100, a lower quartile of 90 for each age group. Other percentile PQs are somewhat consistent with IQ values of the WISC. The five perceptual age levels and the scale scores (one for each of the subtests) indicate the child's development in each visual perceptual ability. The authors believe that a perceptual age of 6 to 6½ years on each of the subtests is a prerequisite for classroom learning. A PQ of 90 is thought to be required for learning to read (Frostig et al., 1961). This statement was based on the failure of six children to learn to read.

Thus scale scores below the child's age as well as PQ below 90 indicates to these authors that children should have remedial perceptual training. It should be noted in this connection, that the similarity between the PQ and the IQ has a disadvantage, if the PQ is interpreted to indicate a fixed level of ability. However, the authors point out that the PQ is a better prognostic indicator than the perceptual age level.

Standardization

The Frostig was standardized in 1961 on a sample of 2,100 unselected nursery school and public school children between the ages of three and nine years. All children were from schools in southern California. Information as to the socio-economic status of each child was not available, although the authors' thought the sample represented children from low-middle to upper-middle socio-economic strata. No Negro children were included. It appears that children from the lower class or so-called culturally disadvantaged children were not represented.

The Developmental Test of Visual Perception may be administered either individually or in a group. Scoring is objective, with the manual giving clear examples (Frostig, 1964).

Reliability

Test-Retest Reliability. Frostig, Lefever, and Whittlesey (1961), using a sample of 50 children with learning disabilities, reported a Pearson's correlation of .98 between test and retest total scores with an interval of three weeks.

In her study (Frostig, et al., 1961) the Frostig Test was administered to two groups of first graders, 35 in each group, and two groups of second graders, 37 in each group. The interval between test and retest was two weeks. The reliability coefficient was .80 for the entire sample. Subtest scale scores (test-retest) correlations ranged from .42 to .80.

Frostig et al. (1964) reported another test-retest reliability study, one not conducted by a psychologist. The interval was two weeks. Pearson's coefficients for both the kindergarten and first grade groups were .69. Subtest reliability correlations ranged from r .29 to r .69.

Split-Half Reliability. Split-half reliability was calculated on four age groups, five to nine years old, using a total sample of 1,459 children. Correlations ranged from .78 to .89. Correlations for children age's five to six were highest and children age eight to nine were lowest. Subtest reliability correlations ranged from .48 to .96 (Frostig et al., 1964).

Validity

Frostig et al., (1964) reported correlations between PQ and teachers' ratings of classroom adjustment, motor coordination, and intellectual functioning. Three hundred and seventy-four children were involved, all kindergarten. The Pearsons r's were respectively .441, .502, and .497.

They also reported a correlation between PQ and intelligence as measured by the Goodenough Test. Correlations were: kindergarten group, .460, first grade, .318, second grade, .366. It is difficult to interpret these results, due to the low reliability (r is .77) of the

Goodenough IQs (Goodenough, 1955).

Corah and Powell (1963) conducted a factor analysis on the Frostig Test with 40 nursery school children. Ten factors were intercorrelated: age, sex, the Ghent Overlapping Geometric Figures Test (a measure of form discrimination), the Full-Range Picture Vocabulary Test, Form B (an intelligence test), and the five subtests of the Developmental Test of Visual Perception. Four factors emerged: Factor I general intelligence, Factor II sex differences, Factor III age, Factor IV a perceptual variable related to the discrimination of forms. Using an oblique rotation, the authors thought that two major factors would account for the most variance: general intelligence and developmental changes in perception. They further suggested that the Full-Range Picture Vocabulary Test may be a good measure of the latter. The mean IQ for the group was 106.32; the mean PQ was 106.90. However, the correlation between the two scores was .386.

The above study may be criticized for the smallness of size of the sample and for the homogeneity of ages. In addition, Silverstein (1965) noted that Corah and Powell included a number of variables in addition to the subtests. He suggests that this may increase the communality of the subtests, at the expense of their specificity.

Silverstein (1965) conducted a factorial study on the five subtests of the test. Three common factors were extracted for the kindergarten children and two for the first and second graders. A general factor, perceptual development or intelligence was the only one of major importance. As the author noted that no objective criterion exists for determining the specificity of a variable, relative to its communality, he thought the results supported consistently the hypothesis that the

five subtests are independent.

Summary

In conclusion, we suggest that further studies in validation are needed. Classroom teachers' ratings of pupil adjustment, motor coordination, and intellectual functioning is not to be taken as evidence that this is a test of perceptual development.

Statement of the Problem

As stated above, the purpose of this study is to explore the relationship between the Bender Gestalt Test, using the Koppitz Scoring System, and the Developmental Test of Visual Perception. Both these tests appear to enjoy some face validity, in that the studies reviewed indicated that visual-motor ability is required. In part, this study is an attempt to determine the presence of concurrent validity (APA, 1957); if two tests porport to measure the same variables, one should expect some type of relationship between them.

As noted in the review of the literature, specific socio-economic differences were not reported. Also, there is an absence of data on lower class or culturally deprived children. In this study, we shall investigate children from lower class families, as well as, families from other socio-economic levels, and shall attempt to control for three other variables: age, sex, and intelligence.

CHAPTER II

PROCEDURE

This was a correlation study conducted in the fall of 1966 in Lawton, Oklahoma. Subjects were randomly selected from the elementary school system in Lawton. Eighty-nine subjects were involved. All were given the Bender Gestalt Test and the Developmental Test of Visual Perception; a sample was given the Peabody Picture Vocabulary Test - a measure of intelligence. As noted above, four variables were statistically controlled: age, sex, intelligence, and socio-economic status. A correlational method was used to determine the relationship between PQ and Bender scores.

Methods

Subjects

Of eighteen elementary schools in Lawton, six were reported by principals to be predominantly middle class, while four were reported to be predominantly lower class. A school from each of these two socio-economic classifications was randomly selected. Within each of these two schools, two classes, one first and one second grade, was randomly selected. Subjects for the study consisted of 89 children, 46 males and 43 females.

Age. We divided our subjects into four age groups:

TABLE I

SUBJECTS BY AGE GROUP AND SEX

Age	N	Male	Female
6.0 - 6.5	29	13	16
6.6 - 6.11	19	11	8
7.0 - 7.5	23	11	12
7.6 - 7.11	18	8	10
Total	89	46	43

Subjects were grouped into socio-economic classes on the basis of a condensed classification of Reiss' socio-economic criteria (Reiss et al., 1961). Three socio-economic categories were employed: lower class, defined as those children whose families received public assistance or whose fathers were laborers; middle class, defined as those children whose fathers were skilled factory workers, retail clerks, skilled workmen, salesmen, and military personnel of noncommissioned rank; upper middle class, defined as those children whose fathers were commissioned officers, professional men, and merchants. Socio-economic status was determined for each child by use of the pre-enrollment cards and information obtained from the school teacher. The lower class group consisted 21 subjects; the middle class group of 45 subjects, and the upper middle class of 30 subjects.

Intelligence. Intelligence was measured using the Peabody Picture Vocabulary Test (Dunn, 1965). From the original sample of 89, 44 subjects were randomly selected and administered this test. IQ ranged from 67 to 128. The children were divided into three groups on the basis of IQ: Low Group, n - 13, the IQ ranged from 67 to 89, mean IQ

83.307, SD = 6.34. Middle Group, n = 21, the IQ ranged from 90 to 109, mean IQ 100.286, SD = 5.15. High Group, n = 10, the IQ ranged from 111 to 128, mean IQ 118.6, SD = 4.27.

Test Administration

The Developmental Test of Visual Perception. This test was administered in groups to each class (Frostig, 1964). The investigator, who was specially trained in the administration of this test, his assistant, and the classroom teacher participated in the administration of the test. Special attention was paid the first graders, to prevent their becoming fatigued. Rest and stretch periods at fifteen minute intervals were given throughout the administration of the test, in accordance with the manual (Frostig et al., 1961).

The Bender Gestalt Test. The Bender Gestalt Test was administered by group method, using each class as a group. Keogh and Smith (1961) used a group method of administration and found no difference when compared to individual administration. Transpariencies were made from the Bender Gestalt Test Cards and were projected on a screen with an overhead projector.

Peabody Picture Vocabulary Test. The investigator administered the Peabody Picture Vocabulary Test to 44 children, individually. The children were randomly selected from the original group.

Time Span. Each group was given all three tests within a two-day period. Each test was administered as a whole in one period.

Scoring of Tests

The Peabody Picture Vocabulary Test and the Developmental Test of Visual Perception are both objectively scored. Each child was assigned a number and then all of his test protocols received that number.

The Bender-Gestalt protocols were scored independently by three raters. Raters A and B scored 89 protocols. The inter-rater reliability coefficient was .93 (p .001). Thirty-five protocols were randomly selected and scored by rater C. Correlations between A and B and C were .95, .92 respectively. All correlation were significant (p .01).

Statistical Analysis

The Bender Gestalt Test scores were transformed to standard scores, by dividing the standard deviation (appropriate for the age level) into the Bender Raw Score. The Standard Score was then added to the number 10 in order to avoid negative numbers. A high Bender scale score indicates poor performance, while a low standard score indicates good performance.

Pearson product-moment correlations were used to determine the relationships between Bender Gestalt Test Scores and PQs, between Bender Scores and IQs, and between PQs and IQs. Correlations with the Bender Gestalt Test are always negative, due to the fact that the Bender Gestalt Test is scored for errors.

CHAPTER III

RESULTS

The overall Pearson's correlation between PQ and Bender scores was $r = -.63$, (p .001).

TABLE II

CORRELATIONS BETWEEN THE BENDER GESTALT TEST AND THE DEVELOPMENTAL TEST OF VISUAL PERCEPTION BY SEX DIFFERENCES AND BY SEX DIFFERENCES WITHIN EACH AGE GROUP

Age Group	Sex	N	Bender and Frostig
6.0 - 6.5	M	16	-.86 ***
6.0 - 6.5	F	13	-.76 **
6.0 - 6.5	M & F	29	-.83 ***
6.6 - 6.11	M	11	-.38 a
6.6 - 6.11	F	8	-.00 a
6.6 - 6.11	M & F	19	-.29 *
7.0 - 7.5	M	11	-.81 **
7.0 - 7.5	F	12	-.62 *
7.0 - 7.5	M & F	23	-.72 ***
7.6 - 7.11	M	8	-.82 **
7.6 - 7.11	F	10	-.11 a
7.6 - 7.11	M & F	18	-.69 **
Total	M	46	-.59 ***
Total	F	43	-.69 ***
Total	M & F	89	-.63 ***

a Not significant
* Significant at the .05 level
** Significant at the .01 level
*** Significant at the .001 level

Sex Differences

The overall correlation for males between PQ and Bender scores was $r = -.59$ (p .001); for females, it was $r = -.69$ (p .001). The correlation coefficients were significantly different from each other (p .0001).

Age Factors

Correlation coefficients between Bender scores and PQ for the age groups ranged from -.00 to -.86. The only sex difference within age groups to reach statistical significance among correlations was for the age group 7.6 - 7.11.

Intellectual Factors

For those subjects who received IQ scores, the following correlation coefficients were noted.

TABLE III

CORRELATIONS BETWEEN THE BENDER GESTALT TEST, THE DEVELOPMENTAL TEST OF VISUAL PERCEPTION, AND THE PEABODY PICTURE VOCABULARY TEST

Sex		PQ and Bender	IQ and Bender	PQ and IQ
M	22	-.62 **	-.57 **	.60 **
F	22	-.81 ***	-.51 *	.48 *
Total	24	-.74 ***	-.55 ***	.56 ***

* Significant at the .02 level
** Significant at the .01 level
*** Significant at the .001 level

The correlation between IQ and PQ r - .56. The correlation between IQ and PQ Bender scores r = .55. The correlation between the PQ and the Bender scores was r = -.74. A partial correlation coefficient of r = .63 was obtained between PQ and Bender scores, holding IQ constant. All correlations were significant at the .001 level.

When correlation coefficients were compared for statistical differences by Fisher's t (males vs. famales) no sex differences were noted.

TABLE IV

CORRELATIONS BETWEEN THE DEVELOPMENTAL TEST OF VISUAL PERCEPTION AND THE BENDER GESTALT TEST FOR THREE DIFFERENT INTELLECTUAL GROUPS

Intellectual Group	N	PQ and Bender
Low Group (IQ 67 - 89)	13	-.78 *
Middle Group (IQ 90 - 109)	21	-.71 **
High Group (IQ 110 - 128)	10	-.23 a

a Not significant
* Significant at the .05 level
** Significant at the .01 level

When subjects were divided into three groups based on intellectual factors, the following correlations were noted. The correlation for the Low Group between the Bender and the PQ was r = .78. The correlation for the Middle Group was r = .71. The correlation for the High Group was not significant.

Socio-economic Status

In comparing the mean performance of the three socio-economic groups by t-tests, it was noted that the Lower Groups was significantly lower

in PQ and Bender scores than the other two socio-economic groups.

TABLE V

COMPARISON OF MEAN PERFORMANCE ON THE BENDER GESTALT TEST, THE DEVELOPMENTAL TEST OF VISUAL PERCEPTION (PQ), AND THE PEABODY PICTURE VOCABULARY TEST (IQ) BY SOCIO-ECONOMIC LEVEL

Socio-economic Status	N (IQ)	Mean IQ	N	Mean PQ	Mean Bender
Lower	13	93.166	21	92.476	11.0308
Middle	18	99.722 a	45	100.844 **	10.1560
Lower	13	93.166	21	92.476	11.0308
Upper middle	13	105.346 *	30	105.133 ***	10.0038 ***
Middle	18	99.722	45	100.844	10.0560
Upper middle	13	105.346 a	30	105.133 a	10.0038 a

a Difference between means not significant
* Difference between means significant at the .05 level
** Difference between means significant at the .02 level
*** Difference between means significant at the .01 level

There were no significant differences between the Upper-middle and Middle Groups on Bender performance, intelligence, and PQ.

TABLE VI

CORRELATIONS BETWEEN THE BENDER GESTALT TEST AND THE DEVELOPMENTAL TEST OF VISUAL PERCEPTION BY SOCIO-ECONOMIC AND SEX DIFFERENCES

Socio-economic Group	N	Bender and Frostig
Lower Group		
Male	14	-.55 *
Female	7	-.15 a
Total	21	-.54 ***
Middle Group		
Male	24	-.71 ***
Female	21	-.66 ***
Total	45	-.68 ***
Upper Middle Group		
Male	13	-.68 **
Female	17	-.55 *
Total	30	-.60 ***

a Not significant
* Significant at the .05 level
** Significant at the .02 level
*** Significant at the .001 level

Correlations between PQ and Bender scores for the three socio-economic groups ranged from -.54 to -.68. All correlations were significantly different from zero.

When correlations between PQ and Bender scores were computed for males and females within socio-economic groups no statistical significant differences were noted.

CHAPTER IV

DISCUSSION

The overall correlation between the Bender Gestalt Test scores and the PQ indicates a substantial relationship. Although caution should be used in interpreting a correlation coefficient, this study does seem to indicate that the Developmental Test of Visual Perception and the Bender Gestalt Test are measuring, to some degree, the same behavior. The fact that the correlation coefficients are far from perfect is not surprising when one compares the internal characteristics of each test and their low reliability. The Frostig contains 57 items, giving a possible total raw score of 77 and takes between 30 and 45 minutes to administer. The Bender Gestalt Test contains 9 figures with a total score of 30 and takes about 3 minutes to administer. Another source of variation may be motoric. Motoric ability is needed in the Bender Test, while as Frostig noted (1962) such motor behavior is believed to be held to a minimum in her test.

Although the Frostig does contain a larger sample of behavior, this does not indicate that the Test is a more valid test than the Bender Gestalt. As noted in previous discussion, validity studies are scarce for both tests. There is certainly no clear indication from this study that the Bender is any less valid than the Frostig Test. However, this study does suggest that both tests might be used together to enhance the testing of perceptual development especially in view of

their low reliability. The Frostig Test purports to be able to break down perceptual ability into five specific factors. Although more factorial studies are needed, such a breakdown could aid the clinician in his diagnosis and recommendation.

Age Factors

The relationship between the Bender Gestalt and the Frostig is a substantial one between the ages of 6.0 and 8.0, with the exception of age 6.6, 7.11. The investigator is unable to explain the lack of relationship for this age group. However, it is interesting to note that Koppitz (1964, p. 30) found the greatest variation at this age level in Bender Test performance. She attributed this to a maturational spurt. Although males tend to display a higher relationship between Bender Gestalt and Grostig performance at certain age levels, females on an overall basis, show a significantly higher relationship. This may be interpreted as indicating that females are more consistent that males at these age levels. This is not to imply that there is a sex difference in performance. It is also interesting to note that the highest relationship for both sexes was between the age 6-6.5. Such relationships are not that surprising, if it is remembered that both tests correlate highly with Readiness Tests that are given to that age group. Possibly such perceptual factors as measured by these tests are consistent within this age group. It can be hypothesized that perception is a dominant ability at this age level and declines as cognitive ability becomes more dominant. It would be interesting to correlate both the Bender Gestalt and the Frostig with measures of intelligence for each specific age group.

Intellectual Factors

Both the Bender Gestalt and the Frostig Tests are moderately related to intelligence. This seems to give support to the hypothesis that perceptual factors are related to intelligence (or vice versa) of that both are related to some third factor. The fact that Bender scores and Frostig Scores are substantially related to each other, as well as to intelligence, was demonstrated by partialling out intelligence. The relationship between the Bender and Frostig was substantial and consistent over the below and average levels of intelligence. However, this was not true for the above average level of intelligence. Possibly the small sample in this range may account for the lack of a significant relationship. Additional research on the relationship between perception and intelligence on the one hand and achievement on the other hand is recommended. Specifically, it might be interesting to explore the relationship between low perception scores (Bender and Frostig) and achievement of students with varying levels of intelligence. Possibly children with superior intellectual ability are not that dependent on perceptual factors for reading achievement.

Socio-Economic Factors

The substantial relationship between the Bender Gestalt and Frostig was consistent over various socio-economic levels. The only non-significant relationship noted was for the females in the lower class group. This was an extremely small sample and this might suggest under representation.

What is interesting to note is the consistently poor performance of the lower class group in intelligence and perception. As noted

previously, children from the lower class have not been adequately represented in the standardization of these clinical instruments. However, such results suggest that children from lower class homes are at a perceptual disadvantage, as well as an intellectual one. Such findings suggest a need for cross-validation. If such results were confirmed, there would be useful implications for perceptual readiness training in the Head Start Programs.

CHAPTER V

SUMMARY

The relationship between the Bender-Gestalt Test using Koppitz scoring system, and the Frostig Developmental Test of Visual Perception was investigated. Correlations between these two tests indicated a substantial relationship. This relationship was consistent over different socio-economic levels, age and sex differences, and independent of intelligence.

Since both tests do not enjoy high reliability, it was suggested that both tests may complement each other in clinical diagnosis. The consistently poor performance of the lower-socio-economic group in intelligence and perception suggests needs for further research in the role of perception, as measured by the Bender and Frostig, and school achievement.

BIBLIOGRAPHY

American Psychological Association. Technical Recommendations for Psychological Tests and Diagnostic Techniques. Washington: American Psychological Association, 1957.

Armstrong, R.G. & P.A. Hauck. Correlates of the Bender-Gestalt scores in children. J. Psychol. Study, 11:15, 3-158.

Bender, Lauretta. Principles of Gestalt in copies form in mentally defective and schizophrenic persons. Arch. Neurol. & Psychiat., 1932, 27:661, 686.

Bender, Lauretta. Disturbances in visuomotor Gestalt function in organic brain disease associated with sensory aphasia. Arch. Neurol. & Psychiat., 1933, 30, 514-537.

Bender, Lauretta. A visual motor Gestalt test and its clinical use. Amer. Orthopsychiat. Ass., Res. Monogr., 1938, No. 3.

Bender, Lauretta. Instructions for the use of visual motor Gestalt test. N.Y. Amer. Orthopsychiat. Assoc., 1946.

Bender, Lauretta. Child Psychiatric Techniques: Diagnostic and Therapeutic Approach to Normal and Abnormal Development Through Patterned, Expressive, and Group Behavior. Springfield, Thomas, 1952.

Corah, N.L. & B.J. Powell. A factor analytic study of the Frostig Developmental Test of visual perception. Percept. mot. Skills, 1963, 16, 59-63.

Frostig, M., LeFever, D., Welty, D. & J.B. Whittlesey. The Marianne Frostig Developmental Test of Visual Perception. 1961 Standardization, First edition, Consulting Psychologists Press, Palo Alto, California.

Frostig, M., LeFever, D.W. & J.R.B. Whittlesey. The Marianne Frostig Developmental Test of Visual Perception. Third edition, Consulting Psychologists Press, Palo Alto, California, 1964.

Getman, G.N. The visuomotor complex in the acquisition of learning skills. Learning Disorders, Vol. 1. Ed by J. Hellmuth. Bernie Straub & J. Hellmuch, Seattle, Washington, 1965.

Keller, J.E. The use of the Bender-Gestalt maturation level scoring system with mentally handicapped children. Amer. of Orthopsychiat., 1955, 25:563-573..

Keogh, Barbara K. & E. Carol Smith. Group techniques and proposed scoring system for the Bender-Gestalt Test with children. J. clin. Psychol., 1961, 17:172-175..

Keport, N.C. The Slow Learner in the Classroom. Charles E. Merrill Books, Inc. Columbus, Ohio, 1960.

Koppitz, Elizabeth M. The Bender Gestalt Test for Young Children. N.Y. Grune & Shatton, 1964.

Koppitz, E.M., Mardis, V. & T. Stephens. A note on screening school beginners with the Bender Gestalt Test. J. Educ. Psychol., 52:80-81.

Koppitz, Elizabeth M. The Bender-Gestalt Test for children: a normative study. J. clin. Psychol., 1960, 16:432-435.

Koppitz, Elizabeth M. Relationships between the Bender-Gestalt Test and Wechsler Intelligence Test for children. J. clin. Psychol., 1958b, 14:413-416.

Koppitz, Elizabeth M. The Bender-Gestalt Test and learning disturbances in young children. J. clin. Psychol., 1958a, 14:292-295.

McPherson & Pepin McPherson, Marion, W. & Lauretta A. Pein. Consistency of reproductions of Bender Gestalt designs. J. clin. Psychol., 1955, 11:163-166.

Moller, L.C., Lowenfield, R., Linder, R., and J. Turner. Reliability of Koppitz scoring system for the Bender Gestalt. J. clin. Psychol., 1963, 19:2111.

Miller, L.C., Lowenfield, R., Linder, R., & J. Turner. Reliability of Koppitz scoring system for the Bender Gestalt. J. clin. Psychol., 1962, 19:2111.

Niebuhur, H. Jr. & D. Cohen. The effect of psychopathology on visual discrimination. J. abnorm. soc. Psychol., 1956, 53:173-177.

Pascal, G.R. Gestalt functions: The Bender-Gestalt, mosaic and word test. In D. Brower and L.E. Abt (Eds), Progress in Clinical Psychology, N.Y., Grune & Stratton, 1952.

Pascal, G.R. & Barbara J. Suttell. The Bender-Gestalt Test: Its Quantification and Validity for Adults. N.Y., Grune & Stratton, 1951.

Reiss, A.J., Duncan, O.D., Hatt, P.K. & North,L Occupation and Social Status. Free Press of Glencoe, N.Y., 1961.

Schilder, P. Psychic disturbances after head injuries. Amer. J. Psychiat., 1934, 91:155-188.

Schulburg, H.C. & A. Tolor. The use of the Bender-Gestalt in clinical practice. J. proj. Tech., 1961, 25:347-351.

Silverstein, A.B. Variance comments in the developmental test of visual perception. Percept. mot. Skills, 1965, 20:973-976.

Wertheirmer, M. Studies in the theory of Gestalt psychology. Psychol. Forsch., 1923, 14:301-350.

Zolike, E.S. A comparison of the Bender-Gestalt reproductions of delinquents and non-delinquents. J. clin. Psychol., 1958, 14:24-26.

APPENDIX A

KOPPITZ STANDARDIZATION

DISTRIBUTION OF BENDER SCORES BY AGE

Age Group	N	Mean Bender Score	S.D	± I S.D.
5-0 to 5-5	81	13.6	3.61	10.0 to 17.2
5-6 to 5-11	128	9.8	3.72	6.1 to 13.5
6-0 to 6-5	155	8.4	4.12	4.3 to 12.5
6-6 to 6-11	180	6.4	3.76	2.6 to 10.2
7-0 to 7-5	156	4.8	3.61	1.2 to 8.4
7-6 to 7-11	110	4.7	3.34	1.4 to 8.0
8-0 to 8-5	62	3.7	3.60	.1 to 7.3
8-6 to 8-11	60	2.5	3.03	.0 to 5.5
9-0 to 9-5	59	1.7	1.82	.0 to 3.5
9-6 to 9-11	45	1.4	1.43	.0 to 2.8
10-0 to 10-5	19	1.5	1.31	.0 to 2.8

APPENDIX B

AGE 6.0 to 6.5 BENDER AND PQ SCORES

No.	Raw Bender	Bender Standard Score	Standard Score +10	Frostig PQ
3	16	1.847	11.847	87
4	12	.874	10.874	79
5	7	-.339	9.661	98
7	12	.874	10.874	74
9	13	1.116	11.116	82
11	7	-.339	9.661	104
12	6	-.583	9.417	113
13	3	-1.310	8.690	116
16	4	-1.067	8.933	119
24	15	1.601	11.601	70
25	2	-1.553	8.447	118
27	9	.339	10.339	102
28	8	-.097	9.903	103
29	6	-.583	9.417	106
30	5	-.825	9.175	116
31	10	.388	10.388	100
32	11	.631	10.631	102
102	16	1.847	11.847	87
106	7	-.339	9.661	110
108	12	.874	10.874	87
111	9	.339	10.339	102
112	16	1.847	11.847	82
113	6	-.583	9.417	104
114	7	-.339	9.661	98
121	10	.388	10.388	116
122	8	-.097	9.903	108
125	18	2.204	12.204	89
129	16	1.847	11.847	84
208	7	-.339	9.661	113

APPENDIX C

AGE 6.6 - 6.11 BENDER AND PQ SCORES

No.	Raw Bender	Bender Standard Score	Standard Score +10	Frostig PQ
2	6	-.106	9.894	102
10	7	.159	10.159	105
1	8	.425	10.425	85
21	9	.691	10.691	116
22	11	1.223	11.223	96
23	6	-.106	9.894	103
26	7	.159	10.159	100
103	8	.425	10.425	119
104	1	-1.436	8.564	119
105	16	2.553	12.553	116
110	4	-.638	9.362	111
115	15	2.287	12.287	76
120	6	-.106	9.894	94
123	8	.425	10.425	96
124	8	.425	10.425	100
127	9	.691	10.641	106
205	5	-.159	9.841	103
205a	6	-.106	9.894	96
220	8	.425	10.425	94

APPENDIX D

AGE 7.0 - 7.5 BENDER AND PQ SCORES

No.	Raw Bender	Bender Standard Score	Standard Score +10	Frostig PQ
14	11	1.717	11.717	88
15	8	.886	10.886	103
40	6	.332	10.332	113
42	5	.055	10.055	116
43	8	.886	10.886	87
44	5	.055	10.055	113
48	4	-.221	9.779	94
50	6	.332	10.332	93
62	1	-1.052	8.948	121
64	3	-.498	9.502	122
66	7	.609	10.609	92
67	7	.609	10.609	116
69	9	1.163	11.163	114
70	3	-.498	9.502	113
71	4	-.221	9.779	125
100	10	1.440	11.440	95
107	7	.609	10.609	80
126	15	2.825	12.825	78
202	4	-.221	9.779	103
206	2	-.775	9.225	110
207	2	-.775	9.225	122
221	10	1.440	11.440	87
224	8	.886	10.886	83

APPENDIX E

AGE 7.5 - 7.11 BENDER AND PQ SCORES

No.	Raw Bender	Bender Standard Score	Standard Score +10	Frostig PQ
8	3	-.508	9.492	110
41	3	-.508	9.492	108
46	3	-.508	9.492	108
47	9	1.287	11.287	81
52	6	.389	10.389	97
63	4	-.209	9.791	113
65	4	-.209	9.791	95
68	6	.389	10.389	104
201	6	.389	101389	87
203	5	.089	10.089	85
209	6	.389	10.389	95
210	6	.389	10.389	97
222	2	-.808	9.192	98
223	4	-.209	9.791	95
226	6	.389	10.389	95
228	4	-.209	9.791	121
229	2	-.808	9.192	118
230	3	-.508	9.492	114

APPENDIX F

IQ (PEABODY PICTURE VOCABULARY) BENDER, AND PQ SCORES

No.	Bender Standard Score	Frostig PQ	IQ	No.	Bender Standard Score	Frostig PQ	IQ
23	9.894	103	108	210	10.389	97	88
2	9.894	102	95	225	13.195	83	81
22	11.223	96	98	222	9.192	98	128
25	8.447	118	104	227	12.305	81	88
3	11.847	87	104	224	10.886	83	95
100	11.440	95	83	52	10.389	97	111
103	10.425	119	100	65	9.791	95	88
110	9.362	116	100	40	10.332	113	111
129	11.847	85	108	41	9.492	108	118
126	12.825	78	81	48	9.779	94	110
125	12.204	89	82	66	10.609	92	67
51	12.603	84	77	68	10.389	104	83
203	10.089	85	106	70	9.502	113	100
221	11.440	87	90	62	8.948	121	124
201	10.389	87	88	64	9.502	122	102
228	9.791	121	102	67	10.609	116	104
220	10.425	94	96				
205	9.841	103	91				
208	9.661	113	109				
205a	9.894	96	89				
207	9.225	122	122				
71	9.779	125	118				
63	9.791	113	88				
46	9.492	108	99				
44	10.055	113	122				
45	10.361	92	100				
42	10.055	116	116				
229	9.192	118	95				

APPENDIX G

LOWER-SOCIO-ECONOMIC GROUP
BENDER AND PQ SCORES

No.	Bender Standard Score	Frostig PQ
7	10.874	74
51	12.603	84
200	12.609	84
204	10.239	87
205	9.841	103
207	9.225	122
210	10.389	97
221	9.661	87
224	10.886	83
227	12.305	81
102	11.847	85
105	12.553	116
108	10.874	87
110	9.362	116
111	10.339	102
112	11.847	82
113	9.417	104
125	12.204	89
126	12.825	78
129	11.847	85
205a	9.894	96

APPENDIX H

UPPER-MIDDLE SOCIO-ECONOMIC GROUP
BENDER AND PQ SCORES

No.	Bender Standard Score	Frostig PQ	No.	Bender Standard Score	Frostig PQ
2	9.894	102	5	9.661	98
13	8.690	116	65	9.741	95
3	11.847	87			
16	8.933	119			
21	10.691	116			
25	8.447	118			
31	10.388	100			
42	10.055	116			
43	10.866	87			
44	10.055	113			
46	9.492	108			
47	11.287	81			
48	9.779	94			
50	10.332	93			
67	10.609	116			
69	11.163	114			
70	9.502	113			
71	9.779	125			
72	9.806	100			
30	9.175	116			
9	11.116	82			
10	10.159	105			
11	9.661	104			
12	9.417	113			
27	10.339	102			
68	10.389	104			
120	9.894	96			
62	8.948	121			

APPENDIX I

MIDDLE SOCIO-ECONOMIC GROUP
BENDER & PQ SCORES

No.	Bender Standard Scores	Frostig PQ	No.	Bender Standard Scores	Frostig PQ
1	10.425	85	216	10.389	95
			218	9.791	121
8	9.492	110	229	9.192	118
14	11.717	88	230	9.492	114
15	10.886	103	100	11.440	95
22	11.223	96	103	10.425	119
23	9.894	103	104	8.564	119
24	11.601	70	106	9.661	116
26	10.159	100	107	10.609	80
28	9.903	103			
32	10.631	102	114	9.661	98
40	10.332	113	115	12.287	76
41	9.442	108	121	10.388	116
45	10.361	92	122	9.903	108
52	10.389	97	123	10.425	96
			124	10.425	100
			127	10.691	106
63	9.791	113			
64	9.502	122	206	9.225	110
66	10.609	92	203	10.089	85
73	10.138	92	209	10.389	95
201	10.389	87	222	9.192	98
202	9.779	103	223	9.791	98
			220	10.425	94
208	9.661	113	225	13.195	83

APPENDIX J

MEANS AND STANDARD DEVIATIONS BY SEX AND SOCIO-ECONOMIC STATUS

SEX

	$\bar{X}$ PQ	SD	$\bar{X}$ Bender	SD
Male	100.847	12.67	10.254	.29
Female	101.325	13.84	10.133	.29

SOCIO-ECONOMIC STATUS

	$\bar{X}$ IQ	SD	$\bar{X}$ Bender	SD	$\bar{X}$ PQ	SD
Upper-Middle	105.346	12.05	10.004	.79	105.133	12.45
Middle	99.722	14.53	10.056	1.74	100.844	12.50
Lower	93.166	14.04	11.031	1.23	92.476	13.38

VITA

William J. O'Connor

Candidate for the Degree of

Master of Science

Thesis: THE RELATIONSHIP BETWEEN THE BENDER GESTALT TEST AND THE FROSTIG DEVELOPMENTAL TEST OF VISUAL PERCEPTION

Major Field: Psychology

Biographical:

Personal Data: Born in New York City, New York, March 22, 1940; married to Beverly Comman, 1965.

Education: Graduated from Martin Van Buren High School in 1957; received the Bachelor of Science degree from New York University, with a major in General Business, pursued graduate work in Psychology at Hofstra University, completed requirements for the Master of Science degree in May, 1967.

Professional experience: Employed by the Reading Center at Oklahoma State University as a psychometrician, June, 1965-May, 1967; worked for Commanche County Guidance Center as a psychologist, Summer, 1966; Psychological Consultant in Kingfisher County from December, 1966 to April, 1967.