5-1969

An Investigation of the Relationship Between the Bender-Gestalt, Draw-a-Man, and Wechsler Preschool and Primary Scale of Intelligence

G. Edward Allen
Utah State University

AN INVESTIGATION OF THE RELATIONSHIP BETWEEN THE BENDER-GESTALT, DRAW-A-MAN, AND WECHSLER PRESCHOOL AND PRIMARY SCALE OF INTELLIGENCE

by

G. Edward Allen, Jr.

A thesis submitted in partial fulfillment
of the requirements for the degree

of

MASTER OF SCIENCE

in

Psychology

Approved:

UTAH STATE UNIVERSITY
Logan, Utah

1969

ACKNOWLEDGMENTS

The author wishes to gratefully acknowledge the guidance and support of his committee: Dr. Glendon Casto, Dr. Heber Sharp, and Dr. David Stone. As most graduate students eventually realize, thesis committees are sincerely interested in assisting the student. This committee was no exception. While the author is solely responsible for any experimental shortcomings which are present in this study, any credit must be jointly shared by the committee.

A special debt of gratitude is due Dr. Philip Langer. His encouragement and advice have proven invaluable to the author throughout his graduate studies.

Finally, my wife Vicky should be publicly acknowledged. Without her loyal support, both financial and moral, this degree would not have been possible.

G. Edward Allen, Jr.

TABLE OF CONTENTS

LIST OF TABLES

ABSTRACT

An Investigation of the Relationship Between

the Bender-Gestalt, Draw-a-Man, and

Wechsler Preschool and Primary

Scale of Intelligence

by

G. Edward Allen, Jr., Master of Science

Utah State University, 1969

Major Professor: Glendon Casto
Department: Psychology

This study investigated the relationship between the Bender-Gestalt, Draw-a-Man, and the Wexler Preschool and Primary Scale of Intelligence. Twenty-two children enrolled in the Logan, Utah Head-Start Program comprised the sample. Product-moment correlations indicated a significant relationship between these instruments.

The following tentative conclusions were drawn:

1. The Wexler Preschool and Primary Scale of Intelligence bears a relationship to the Draw-a-Man and Bender-Gestalt tests similar to that between the Wechsler Intelligence Scale for Children and these instruments.

2. The Bender-Gestalt test, using the Koppitz scoring system, and the Draw-a-Man test can serve a similar checking function with the Wexler Preschool and Primary Scale of Intelligence as they do with the Wechsler Intelligence Scale for Children.

Methodological shortcomings prohibit over-generalization of these findings. The results, however, are seen as indicative of the promise of these instruments, and further investigation was advocated.

(39 pages)

INTRODUCTION

Psychologists have long been concerned with the need for a valid, yet relatively rapidly scored intelligence test. Such tests have been sought to provide: (1) a general estimate of intelligence, or (2) a supplemental validity check when used with more elaborate intelligence testing instruments. Consequently, tests such as the Goodenough Draw-a-Man Test and more recently, the Bender-Gestalt Test have been frequently employed in such a supplemental role with the Wechsler Intelligence Scale for Children (WISC) and the Stanford-Binet Intelligence Test (Anastasi, 1961; Koppitz, 1964).

For many years, the WISC and the Stanford-Binet have been accepted as the standard testing instruments for young (below age six) children. The theoretical and practical shortcomings of these tests at this age level, however, are well-known (Anastasi, 1961). Other preschool and infant scales exist, but their applicability and use are curtailed by defects in standardization, reliability, appeal, and feasibility of I.Q. conversions (Cronbach, 1960).

Wechsler (1967) has developed a downward extension of the WISC, known as the Wechsler Preschool and Primary Scale of Intelligence (WPPSI). This instrument is specifically designed to measure intelligence in children four to six and one-half years of age. The WPPSI is based on the assumption that the four to six year old possesses the ability to think for himself and learn from mistakes to the degree which language development and experience permits. A further assumption is that these potentialities "may be systematically appraised through an appropriate battery of tests" (Wechsler, 1967, p. 1). The WPPSI is intended to provide such a battery of tests.

Because of the recency of the WPPSI (1967), virtually no comparative data exist between it and other measures of intelligence. This study will compare the WPPSI profiles of Head-Start children with their scores on the Bender-Gestalt Test (using the Koppitz scoring system), and the Goodenough Draw-a-Man Test.

REVIEW OF LITERATURE

As noted earlier, there are virtually no published data available on the WPPSI, other than that in Wechsler's testing manual. Consequently, this review will concentrate on the Goodenough Draw-a-Man Test (DAM) and the Bender-Gestalt Test with preschool and/or primary school populations. In addition, the Head-Start Program will be briefly discussed. Finally, the WPPSI will be discussed at the conclusion of this section.

The Head-Start Program

The Head-Start Program was initiated in 1964 as part of the Economic Opportunity Act. This program is designed to assist communities in establishing schools for culturally deprived preschool (under age six) children. Admittance into this program is based solely on family income level. For example, a family of four with an annual income at or below $3,200.00 would theoretically qualify for this program, regardless of its cultural standard (Meyer, 1965). It is assumed that restricting admission to those below a predetermined economic level will automatically include the majority of those designated as culturally deprived. This results in cultural level entirely being a function of the financial situation existing at the moment.

The Bender-Gestalt Test

The Bender-Gestalt Test is among the most widely-used clinical tests (Sundberg, 1961). Bender developed this test primarily as a means of qualitatively determining the presence of brain damage or psychological disorders (Anastasi, 1961). Although a number of scoring systems have been devised for this instrument, they have largely been designed for use with adult psychiatric patients or with institutionalized retarded children. Few, if any, are intended for use with young children of normal intelligence (Koppitz, 1964).

However, Bender (1938) pointed out years ago that copying Gestalt figures tends to reflect the maturation level of visual-motor perception. The degree of visual-motor maturity is closely related to language ability and other functions associated with childhood intelligence. Wewetzer (1959) found significant relationships between Binet-Norden I.Q. scores and Bender performance for both brain-damaged and normal children. Similarly, Aaronson (1957) found a positive relationship between Bender recall scores and Porteus Maze I.Q. scores in children.

The Koppitz Scoring System

Koppitz (1964) devised a developmental scoring system that attempts to differentiate distortions reflecting immaturity or perceptual malfunctioning from those reflecting emotional factors and attitudes in children. Normative data for this system were derived from 1104 public school children between the ages of 5 and 11.

Koppitz (1964) assumed that age and visual-motor maturation were the determining factors in the use of the Bender Test as a measure of intelligence. She correlated scores from the Developmental Bender Scoring System with I.Q. scores at various age levels. All correlations between Bender scores and WISC and Stanford-Binet I.Q.'s at all ages were statistically significant at the .01 level.

In another study Koppitz (1958) explored the relationship between Bender scores and the WISC. For the first and second grades, separate

Verbal and Performance I.Q. scores correlated significantly with the Bender at the .02 level. The relationship between Bender scores and Full Scale I.Q. scores were significant at the .10 level of confidence. Koppitz (p. 50) concluded: "The Bender Test can be used with some confidence as a short nonverbal intelligence test for young children, particularly for screening purposes."

Thweatt (1963, p. 217) investigated the validity of the Koppitz scoring technique as a predictor of learning disabilities. He concluded: "The results indicate that the Bender-Gestalt Test with Koppitz' scoring system can predict with accuracy future reading problems regardless of the causal factors of the reading disability."

Plenk and Jones (1967, p. 367) examined the reliability of Koppitz' scoring system, emphasizing its utility with the three to four year age group. Their reason for conducting this study was that "there is little or no information concerning the useability of this test with younger age groups, except for that reported in Koppitz." Their results showed that Koppitz' scoring system was not applicable to the three to four year age group. Below five years, Koppitz scores were positively rather than negatively correlated with I.Q. Miller et al. (1963) reported reliability coefficients ranging from .88 to .96 for the Koppitz scoring system.

Keough (1965) explored the Bender Test as a predictive and diagnostic indicator for reading performance. Her results suggested that the Bender was related to an over-all pattern of school achievement. High Bender performance was a more reliable predictor of school achievement than was low performance. There was little observable relationship between poor Bender scores and school achievement.

Chang and Chang (1967) studied the relation between visual-motor skills and reading achievement in superior primary grade students. They reported correlations between Bender scores and WISC scores of .50 and .46, respectively. The authors concluded that the Bender Test was promising as a technique for indicating reading level and warranted further investigation in this capacity.

Keough and Smith (1967) investigated the relationship between visual-motor ability, as measured by the Bender Test, and school achievement. Their data indicated that the Bender Test was an effective screening device for school achievement. It was most effective, however, with younger groups (five to six years old). Past the third grade, Bender scores were unreliable as a screening technique.

In summary, the available literature comparing the Bender Test and the WISC (found largely in Koppitz, 1964) reported correlations ranging from .40 to .79.

The Goodenough Draw-a-Man Test

In 1926, Goodenough published a Draw-a-Man Test which she hoped would provide a quick-scoring and easily administered estimate of intelligence. The degree of her success is indicated by Sundberg's (1961) report that the DAM ranked third in frequency of use. Correlations between .41 and .80 have been consistently reported between the DAM and other measures of intelligence (Anastasi, 1961).

Shipp and Louden (1964) explored the relationship between the DAM and predicted first grade achievement. They concluded that the DAM was of some value as a predictor of general achievement in the first grade. The authors suggested its use as a quick-scoring screening device.

Dunn (1967a), investigating the validity of the DAM, found a correlation of .64 between the WISC and the DAM. In a later study, Dunn (1967b) reported the following correlations between the WISC and the DAM: .77 with Full-Scale score, .73 with the Verbal Scale, and .75 with the Performance Scale.

Vane and Kessler (1967, p. 52) studied the long-term reliability and validity of the DAM. They found an average correlation of .56 with the Stanford-Binet, when measured four times over a period of 10 years. They concluded: "The test [DAM] is valuable as a quick estimate of intelligence."

Datta (1967) investigated the effect of impoverished home conditions on attained DAM scores. Her sample consisted of 965 Head-Start participants. "Impoverished conditions" were defined as those conditions necessary to gain admittance into the Head-Start program (Meyer, 1965). She found that the Head-Start group had substantially lower performances on the DAM than did the normative control group.

Koppitz et al. (1959), using a sample of 143 first graders, explored the potential of the Bender Test and Human Figure Drawings in predicting school achievement. These instruments were administered after the first six weeks of school, then correlated with an achievement test given at the end of a seven-month period. The authors found a multiple correlation of .65 between Bender scores, Human Figure Drawings, and school achievement.

The Wechsler Preschool and Primary Scale of Intelligence

The only data currently available relating the WPPSI to other intelligence tests is found in the WPPSI Test Manual (Wechsler, 1967). A study was conducted to investigate the relationship between the WPPSI and three individually administered intelligence tests: the Stanford-Binet Intelligence Scale, the Peabody Picture Vocabulary Test, and the Pictorial Test of Intelligence.

The subjects were 98 children ranging from 60 to 73 months of age. Coefficients of correlation between WPPSI scores and the Stanford-Binet were .76 (Verbal I.Q.), .56 (Performance I.Q.), and .75 (Full Scale I.Q.). The Peabody Picture Vocabulary correlated .57 (Verbal I.Q.), .44 (Performance I.Q.), and .58 (Full-Scale I.Q.) with the WPPSI. Correlations of .53 (Verbal I.Q.), .60 (Performance I.Q.), and .64 (Full Scale I.Q.) were found between the Pictorial Test of Intelligence and the WPPSI. Wechsler (1967, p. 35) concluded: "Although the correlations between the WPPSI and the three other scales indicate positive relationships among them, these coefficients are not so high as to suggest that the scales are interchangeable."

PROBLEM

Based on the literature cited above, we can assume that both the Bender-Gestalt Test and the Draw-a-Man Test have some value as quick-scoring estimates of intelligence. As previously noted, however, there is a paucity of validating data on the WPPSI as well as on the Head-Start group as a sample. This study will try to partially fill this void by examining aspects of the concurrent validity of the WPPSI. Specifically, this thesis will attempt to answer the following questions:

1. Is there a significant relationship between the Bender-Gestalt Test, using the Koppitz Scoring System, and the WPPSI?

2. Is there a significant relationship between the Goodenough Draw-a-Man Test and the WPPSI?

3. Do the reported correlations for a Head-Start sample differ from those found between the aforementioned instruments and the WISC using a "normal" population sample?

PROCEDURE

Sample

The subjects were 22 children enrolled in the 1967 Head-Start program at the Wilson Elementary School in Logan, Utah. The sample included approximately half of the total members of the class. These were chosen on the basis of availability during the testing period. The sample included 12 boys and 10 girls, ranging in age from four and one-half to six years. (Only one child was under the age of five at the time of testing.) There was no systematic demographic analysis of the sample. However, reports from the instructors indicated that their cultural environments were diverse, ranging from parents attending Utah State University to those on county relief.

The WPPSI was administered by the researcher and one other trained tester during the Head Start program at the Wilson school. The program terminated before the Bender Test and the DAM could be administered. Consequently, the remaining tests were administered at random intervals over the ensuing three months. The periods of testing were contingent upon the availability of the subjects and the scheduling of the respective schools in which they were currently enrolled. Three trained testers (including the researcher) gave the Bender Test and the DAM at the school where each child was currently enrolled. The WPPSI was scored by a single trained individual, as were the Bender and the DAM tests. This was done in an attempt to reduce scoring variability. Each test was administered and scored in strict accordance with the procedures established by Goodenough (DAM), Koppitz (Bender), and Wechsler (WPPSI). As a reliability check, questionable DAM and Bender protocols were scored by others. Close agreement was found, although interscorer reliability was not computed.

The WPPSI

The Wechsler Preschool and Primary Scale of Intelligence (Wechsler, 1967) was developed to more adequately appraise the abilities of the preschool child. The WPPSI is based on the same theoretical and methodological foundations as the WISC. Like the WISC, the WPPSI is divided into Verbal and Performance Test Scales. These scales consist of a battery of subtests which are intended as measures of different abilities.

Eight of the 11 subtests on the WPPSI " . . . provide the same measures as the WISC, and may be seen as continuous with the WISC" (Wechsler, 1967, p. 7). These include Information, Vocabulary, Arithmetic, Similarities, Comprehension, Picture Completion, Mazes, and Block Design. The remaining three subtests (Sentences, Animal House, and Geometric Design) are unique to the WPPSI and will be described. The Animal House subtest " . . . requires the child to associate sign with symbol and may be considered as a measure of learning ability" (p. 11). This subtest corresponds to the Coding task found in the WISC. Geometric Design is included " . . . because previous studies indicated that the young child's ability to reproduce geometric figures correlates quite well with other measures of intelligence" (p. 11). This subtest is designed to measure abilities dependent on perceptual and visual-motor organization. The Sentences scale corresponds to the WISC subtest of Digit Span. It is intended to measure recall, and to some degree, vocabulary development. Descriptions of each subtest are presented in Appendix A.

The WPPSI standardization sample consisted of 1200 children taken from groups " . . . considered to be representative of the United States population of children aged four through six and one-half years" (p. 13). Where applicable, the split-half technique was used to determine reliability coefficients. The test-retest method was used on speeded tests. The reliability coefficients on the three primary I.Q. scores (Verbal, Performance, and Full Scale) ranged from .84 to .94.

The Developmental Bender Scoring System

The Developmental Bender Scoring System " . . . consists of 30 mutually exclusive scoring items which are scored as either present or absent. All scorings are added into a composite score" (Koppitz, 1964, p. 12). The scoring items in the Koppitz system were validated against school achievement as measured by the Metropolitan Achievement Test. The subjects for the item analysis were 165 first and second grade school children. These subjects were selected to represent a socio-economic cross section of the population.

Test score reliability was determined by the test-retest method. Correlation coefficients ranged from .55 to .66; all were significant at the .001 level. Scorer reliability coefficients ranged from .88 to .96. Detailed definitions of each scoring item are presented in Appendix B.

The Goodenough Draw-a-Man Test

The Goodenough Draw-a-Man places emphasis on the child's accuracy of observation and level of conceptual thinking (Goodenough, 1926). Credit is given for including individual body parts, detail, perspective, and proportion. A total of 73 scorable items are included and are based on age differentiation and similar factors. Anastasi (1961) surveyed the literature dealing with the reliability of the DAM. She reported reliability coefficients ranging from .68 to .89. Because of the widespread administration of the DAM, no further description will be included here. A description of the scoring criteria is presented in Appendix C.

RESULTS AND DISCUSSION

Analysis of the data involved two phases: (1) the child's scores on each test were converted to standard scores, and (2) product-moment correlation coefficients were computed between the three primary scores on the WPPSI (Full-Scale, Verbal, Performance) and Bender and DAM scores. The results are presented in Table 1. (A table of raw data is presented in Appendix D.)

Table 1. Correlation coefficients between the WPPSI and the Bender-Gestalt and Draw-a-Man

WPPSI	Bender-Gestalt	Draw-a-Man
Verbal	.535**	.539*
Performance	.688*	.610*
Full-Scale	.664*	.624*

*Significant at 1 percent level.
**Significant at 5 percent level.

As previously noted, this study was designed to determine the relationship between the WPPSI and the Bender Test; and between the WPPSI and the DAM. The first question was: Is there a significant relationship between the Bender-Gestalt Test, using the Koppitz scoring system, and the WPPSI? In this study, the Bender correlated .535 with the WPPSI Verbal Scale ($p < .05$), .688 with Performance Scale ($p < .01$), and .644 with Full Scale score ($p < .01$).

The second question asked was: Is there a significant relationship between the Goodenough Draw-a-Man Test and the WPPSI? The following correlations were obtained between the WPPSI and the DAM: .539 ($p < .01$) with the Verbal Scale, .610 ($p < .01$) with the Performance Scale, and .624 ($p < .01$) with Full Scale.

The third question investigated was: Do the reported correlations for a Head-Start sample differ from those found between the aforementioned instruments and the WISC? The correlations reported in this study are comparable with those cited in the review of literature comparing these tests.

The results indicate a significant relationship exists between the Bender Test and the WPPSI. These results parallel those found between the WISC and the Bender Test (Koppitz, 1964). She reported chi-square coefficients of 2.1 ($p > .10$) for Verbal Scale, 8.1 ($p < .01$) for Performance Scale, and 4.4 ($p < .05$) for Full Scale scores.

Although all correlations reported in this study were significant, the coefficient between the WPPSI Verbal Scale and the Bender Test was somewhat lower than the others. It is plausible that the lower coefficient was caused by the greater emphasis placed on logical reasoning, factual information, and social understanding by tests of verbal intelligence. Koppitz (1964) pointed out that none of these factors bear a clear relationship to the copying of Gestalt figures. She concluded, however, that WISC Performance and Full-Scale scores are closely related to Bender Test performance. Apparently this conclusion can be extended to WPPSI Performance and Full Scale scores and Bender Test performance.

The results presented in this study tend to confirm the presence of a significant relationship between the DAM and the WPPSI. All correlations between these instruments were significant at the .01 level of confidence. These data seem to support Anastasi's (1961) conclusion that performance on the DAM is, to some degree, representative of general intelligence.

This investigation presented data to indicate that the instruments used are equally valid for a Head-Start sample or a normal population. The magnitude of the correlations reported in this study compare closely with those presented in the previously cited literature between the WISC, DAM, and Bender Test for a normative sample (for example, Anastasi, 1961; Wechsler, 1967; and Koppitz, 1964).

Analysis of the data yielded one other interesting finding. The Bender scores correlated with the DAM at the .01 level of significance (.659). This indicates a close relationship between factors measured on these instruments.

CONCLUSIONS AND IMPLICATIONS

From the data reported in this investigation, it can be concluded that:

1. The WPPSI seems to bear a relationship to the Bender-Gestalt and Draw-a-Man similar to that between the WISC and these instruments. This indicates that the WPPSI relates to preschool intelligence (four to six and one-half years) in much the same way as the WISC relates to childhood intelligence (seven to fourteen years), at least as measured by the Bender Test and DAM.

2. The Bender-Gestalt Test, using Koppitz' scoring system, and the Goodenough Draw-a-Man Test may serve a similar function with the WPPSI as they do with the WISC; for example, as supplemental validity checks or as quick-scoring estimates of preschool intelligence.

Because of the limited size and lack of a systematic demographic analysis of the sample, it would be unwise to over-generalize these findings. Other uncontrolled variables included a lack of systematic test conditions and presentation. Consequently, no conclusion can be drawn with respect to the Head-Start population, and those conclusions drawn above must be regarded as tentative until further data are presented.

It can be assumed, however, that these instruments are promising as measures of preschool intelligence, and further investigation is warranted.

SUMMARY

This study investigated the relationship between the Bender-Gestalt, Draw-a-Man, and the WPPSI. Twenty-two children enrolled in the Logan, Utah Head-Start Program comprised the sample. Product-moment correlations indicated a significant relationship between these instruments.

The following tentative conclusions were drawn:

1. The WPPSI bears a relationship to the DAM and Bender Test similar to that between the WISC and these instruments.

2. The Bender-Gestalt, using the Koppitz scoring system, and the Draw-a-Man can serve a similar checking function with the WPPSI as they do with the WISC.

Methodological shortcomings prohibit over-generalization of these findings. The results, however, are seen as indicative of the promise of these instruments, and further investigation was advocated.

LITERATURE CITED

Aaronson, B. S., 1957, The Porteus Mazes and Bender-Gestalt recall. *Journal of Clinical Psychology* 13, 186-187.

Anastasi, Anne, 1961, *Psychological testing.* New York: Macmillan.

Bender, Loretta, 1938, A visual motor Gestalt test and its clinical use. *The American Orthopsychiatric Association Research Monograph,* No. 3.

Chang, Thomas and Chang, Vivian, 1967, Relation of visual-motor skills and reading achievement in primary-grade pupils of superior ability. *Perceptual and Motor Skills,* 24 (1), 51-53.

Cronbach, Lee J, 1960, *Essentials of psychological testing.* New York: Harper and Row.

Datta, Lois-Ellin, 1967, Draw-a-Person test as a measure of intelligence in preschool children from very low income families. *Perceptual and Motor Skills,* 24 (1), 346-350.

Dunn, James A., 1967a, Note on the relation of the Harris Draw-A-Woman to WISC I.Q.'s. *Perceptual and Motor Skills,* 24 (1), 316.

Dunn, James A., 1967b, Validity coefficients for the new Harris-Goodenough Draw-a-Man test. *Perceptual and Motor Skills,* 24 (6), 299-301.

Goodenough, Florence, 1926, *Measurement of intelligence by drawings.* Tarryton-on-Hudson, New York: World Book Company.

Hutton, Jerry B., 1966, Bender recall of children 2-5 related to age and intelligence. *Perceptual and Motor Skills,* 23 (1), 34.

Keough, B. K., 1965, The Bender-Gestalt as a predictive and diagnostic test of reading performance. *Journal of Consulting Psychology,* 29, 83-84.

Keough, B. K. and Smith, C., 1967, Visuo-motor ability for school achievement. *Perceptual and Motor Skills,* 25, 101-110.

Koppitz, Elizabeth, 1964, *The Bender-Gestalt Test with Children.* New York: Grune and Stratton.

Koppitz, Elizabeth, 1958, Relationships between the Bender-Gestalt Test and the Wechsler Intelligence Scale for Children. *Journal of Clinical Psychology,* 14, 413-416.

Koppitz, E. M., Sullivan, J., Blyth, D., and Shelton, J., 1959, Prediction of first grade school achievement with the Bender-Gestalt Test and Human Figure Drawings. *Journal of Clinical Psychology,* 15, 164-168.

Meyer, Alberta L., 1965, Project Head-Start. *Childhood Education*, 41, 432-432.

Miller, L. C., Lowenfeld, R., Linder, R., and Turner, J., 1963, Reliability of Koppitz' scoring system for the Bender-Gestalt. *Journal of Clinical Psychology*, 19 (2), 211.

Plenk, A. and Jones, J., 1967, An examination of the Bender Gestalt performance of three- and four-year-olds and its relationship to Koppitz' scoring system. *Journal of Clinical Psychology*, 23 (3), 367-370.

Shipp, D. and Loudon, Mary, 1964, The Draw-a-Man Test and achievement in the first grade. *Journal of Educational Research*, 57 (10), 518-521.

Sundberg, N. D., 1961, The practice of psychological testing in clinical services in the United States. *American Psychologist*, 16, 79-83.

Thweatt, R. C., 1963, Prediction of school learning disabilities through the use of the Bender-Gestalt Test: A validation study of Koppitz' scoring technique. *Journal of Clinical Psychology*, 19 (2), 216-217.

Vane, Julia R. and Kessler, Rosalyn T., 1967, The Goodenough Draw-a-Man Test: Long-term reliability and validity. *Perceptual and Motor Skills*, 24 (1), 51-53.

Wechsler, David, 1967, *Manual for the Preschool and Primary Scale of Intelligence*. New York: The Psychological Corporation.

Wewetzer, K. H., 1964, *Das hirngeschadigte kind*, cited in Koppitz, E. *The Bender-Gestalt test with children*. New York: Grune & Stratton.

APPENDIXES

Appendix A

Wechsler Preschool and Primary Scale of Intelligence

Subtests

Verbal Scale	*Description*
Information	Samples remote memory, experience, and cultural background.
Vocabulary	Samples word knowledge from experience, and experimentive information.
Arithmetic[a]	Measures basic quantitative concepts without involving the explicit use of numbers.
Similarities[a]	Measures verbal concept-formation by presenting simple analogies rather than emphasizing the concept of "similar."
Comprehension	Measures practical knowledge and common sense.
Sentences (supplementary test)[b]	Measures recall and language development, corresponds to "Digit Span" found in WISC.
Performance Scale	
Animal House[b]	Requires the association of sign with symbol and is considered a measure of learning ability; corresponds to "coding" found in WISC.
Picture completion	Measures visual concentration.
Mazes[a]	Measures ability to plan ahead and show foresight; test begins with unidirectional horizontal mazes and has a new scoring system.

Geometric design[b]	Included as a general indicator of intelligence, measures abilities dependent on perceptual and visual-motor organization.
Block design[a]	Measures visual-motor coordination; blocks are larger and designs simpler than on the WISC.

[a]Subtests modified from the WISC.
[b]Subtests developed especially for the WPPSI.

Appendix B

Instructions for the Administration and Scoring of the Bender Test

Koppitz (1964, p. 15-32)

Seat the child comfortably at an uncluttered table on which two sheets of paper, size 8-1/2" by 11", and a #2 pencil with an eraser have been placed. After rapport has been established show the stack of Bender cards (Bender, 1946) to the child and say: "I have nine cards here with designs on them for you to copy. Here's the first one. Now go ahead and make one just like it." After the child has adjusted the position of the paper to suit himself, place the first Bender card, Figure A, at the top of the blank paper in front of the child. No comments are made while observations and notes are made on the child's test behavior. There is no time limit for this test. When a child has finished drawing a figure, the card with the stimulus design is removed and the next card is put in front of him and so on. All nine cards are presented in this fashion in orderly sequence.

If a child asks questions concerning the number of dots or the size of drawings, etc., he should be given a noncommital answer like: "Make it look as much like the picture on the card as you can." He should be neither encouraged nor discouraged from erasing or making several attempts at drawing a design. It has been found practical to discourage the counting of dots on Figure 5 since this requires much time and adds little new information. The children who count dots on Figure 5 also tend to count dots and circles of Figures 1, 2, and 3. When a child begins counting dots on Figure 5 the examiner may say: "You do not have to count those dots; just make it look like the picture." If the child still persists in counting the dots, it then takes on diagnostic significance. The indications are that the child is most likely quite perfectionistic or compulsive. If the child has filled most of the sheet of paper and turns it sideways to fit Figure 8 into the remaining space, this should be noted on the protocol as this is not considered to be a rotation of design.

Each child is permitted to use as much or as little paper as he desires. If he asks for more than the two sheets of paper provided, he should be given additional paper without comment. Even though the test has no time limit, it is helpful to keep a record of the time needed to complete the test, as an extremely short or unusually long period is diagnostically significant.

Care should be taken that the Bender Test is presented at the beginning of the testing session when the child is well rested, as a fatigued child will not perform optimally. If it is felt that a child has been rather hasty in the execution of the test or if maximum performance has not been obtained, he may be asked to repeat the drawing of a Bender figure on another sheet of paper. If additional testing for maximum achievement seems indicated, a notation to this effect should be made on the protocol.

All Bender scoring items are scored as one or zero, that is, as "present" or "absent." Only clearcut deviations are scored. In case of doubt, an item is not scored. Since the Scoring System is designed for young children with as yet immature fine motor control, minor deviations are ignored. All scoring points are added into a composite score upon which the normative data are based.

Figure A

1. Distortion of Shape
   a) Square or circle or both are excessively flattened or misshapen; one end of circle or square is twice as long as the other one. If two sides of square do not meet at point of junction with circle, then the shape of the square is evaluated as *if* the two sides did meet. Extra or missing angles (in case of doubt do not score).
   b) Disproportion between size of square and circle; one is twice as large as the other one.

2. Rotation
   Rotation of figure or any part of it by 45° or more; rotation of stimulus card even if then copied correctly in rotated position.

3. Integration
   Failure to join circle and square; curve and adjacent corner of square more than 1/8" apart; this applies also to overlap.

Figure 1

4. Distortion of Shape
   Five or more dots converted into circles; enlarged dots or partially filled circles *not* considered circles for scoring of this item--in case of doubt do not score; dashes *not* scored.

5. Rotation
   Rotation of figure by 45° or more; rotation of stimulus card even if then copied correctly as shown on rotated card.

6. Perseveration
   More than 15 dots in a row.

Figure 2

7. Rotation
   Rotation of figure by 45° or more; rotation of stimulus card even if then copied correctly as shown on rotated card.

8. Integration
   One or two rows of circles omitted; row of dots of Figure 1 used as third row for Figure 2; four or more circles in the majority of columns; row of circles added.

9. Perseveration
   More than 14 columns of circles in a row.

Figure 3

10. Distortion of Shape
    Five or more dots converted into circles; enlarged dots or partially filled-in circles *not* considered circles for this scoring item--in case of doubt do not score; dashes *not* scored.

11. Rotation
    Rotation of axis of figure by 45° or more; rotation of stimulus card even if then copied correctly as shown on rotated card.

12. Integration
    a) Shape of design lost; failure to increase each successive row of dots; shape of arrow head not recognizable or reversed; conglomeration of dots; single row of dots; blunting or incorrect number of dots not scored.
    b) Continuous line instead of row of dots; line may be substituted for dots or may be addition to dots.

Figure 4

13. Rotation
    Rotation of figure or part of it by 45° or more; rotation of stimulus card even if then copied correctly as shown on rotated card.

14. Integration
    Curve and adjacent corner more than 1/8" apart, this applies also to overlap; curve touches both corners.

Figure 5

15. Distortion of Shape
    Five or more dots converted into circles; enlarged dots or partially filled circles are *not* scored; dashes are *not* scored.

16. Rotation
    Rotation of total figure by 45° or more; rotation of extension, e.g. extension points toward left side or extension begins left of center dot of arc; rotation is only scored once if arc and extension are both rotated independently of each other.

17. Integration
    a) Shape of design is lost; conglomeration of dots; straight line or circle of dots instead of arc; extension cuts through arc; square or point instead of arc is *not* scored.
    b) Continuous line instead of dots in either arc or extension or both.

Figure 6

18. Distortion of Shape
    a) Three or more distinct angles substituted for curves (in case of doubt do not score).
    b) No curve at all in one or both lines; straight line.

19. Integration
    Two lines not crossing or crossing at the extreme end of one or both lines; two wavy lines interwoven.

20. Perseveration
    Six or more complete sinusoidal curves in either direction.

Figure 7

21. Distortion of Shape
    a) Disproportion between size of two hexagons; one must be at least twice as large as the other one.
    b) Hexagons are excessively misshapen; extra or missing angles in one or both hexagons.

22. Rotation
    Rotation of figure or any part of it by 45^{o} or more; rotation of stimulus card even if then copied correctly as shown on rotated card.

23. Integration
    Hexagons do not overlap or overlap excessively; that is, one hexagon completely penetrates through the other one.

Figure 8

24. Distortion of Shape
    Hexagon or diamond excessively misshapen; extra or missing angles; diamond omitted.

25. Rotation
    Rotation of figure by 45^{o} or more; rotation of stimulus card even if then copied correctly as shown on rotated card (turning of paper in order to make most economical use of paper *not* scored and should be noted on the protocol).

Appendix C

Scoring Criteria for Goodenough

Draw-a-Man Test

Harris-Goodenough Scoring System

"On this paper I want you to make a picture of a *man*. Make the *very best* picture you can. *Take your time and work carefully.* I want to see if you can *do as well* as other boys and girls. *Try very hard* and see what a *good* picture you can make."

HEAD: Not features alone.

LEGS: Two in full face. One or two in profile.

ARMS: Two in full face. One or two in profile. Not fingers alone unless definite space between base of fingers and point of attachment to the body.

TRUNK: Straight line or two dimensional. Combined with head if features take up only the upper half or if there is a crossline between.

LENGTH OF TRUNK GREATER THAN WIDTH: Not if a single line. Not if length and width are equal. Measure at points of greatest length and width.

SHOULDERS: Broadening and rounding of trunk at this point in full face. Not square, rectangular, or elliptical trunks. Expansion of chest in profile.

ARMS AND LEGS ATTACHED TO TRUNK: Not if no trunk. Any point on trunk. Arms attached to neck or junction of head and trunk.

LEGS ATTACHED TO TRUNK AND ARMS ATTACHED TO TRUNK AT CORRECT POINTS: Exactly at shoulders or where shoulders should be.

NECK: Distinct from head and trunk.

NECK CONTINUOUS WITH HEAD, TRUNK, OR BOTH.

EYES: One or two. Any kind.

NOSE: Any kind.

MOUTH: Any kind.

NOSE AND MOUTH TWO DIMENSION AND TWO LIPS: Nose not straight line, dot, two dots, circle or square. Mouth must have a line or other separation between the two lips which in turn are in two dimension. In profile nose must be distinct from forehead and upper lip. Mouth must show separate modeling of two lips, or mouth line continuous with face outline.

NOSTRILS: Any kind. Two dots. In profile bottom nose outline extends inward across upper lip outline.

HAIR: Any kind.

HAIR: NO MORE THAN CIRCUMFERANCE AND BETTER THAN A SCRIBBLE AND NON TRANSPARENT: Not if head outline shows through hair.

CLOTHING: Any kind, Hat. Belt. Row of buttons. Series of vertical or horizontal lines on trunk.

TWO ARTICLES OF CLOTHING NON-TRANSPARENT: Concealing what they are supposed to cover. Not hat flush with head. Not buttons alone.

FOUR OR MORE ARTICLES OF CLOTHING: Definitely indicated. Must be among hat, shoes, coat, collar, shirt, necktie, belt, suspenders, or trousers. Shoes must have laces, toe cap, or double line for sole. Not heel alone. Coat or shirt must have sleeves, pockets, lapels, or shading by spots or stripes. Not buttons alone. Collar not merely neck insert. Not lapels.

COSTUME COMPLETE WITHOUT INCONGRUITIES: A definite recognizable costume. Complete in all essentials. Not confusing costumes. Sleeves, trousers, and shoes must always be shown. Also hat, collar, and tie if usually part of costume.

FINGERS: Any kind. On both hands if shown.

CORRECT NUMBER OF FINGERS: Five. On both hands if shown.

DETAIL OF FINGERS: Two dimension. Length greater than width. Span not greater than 180°. On both hands if shown. Must have correct number.

THUMB: One lateral digit definitely shorter than any of the others. Or angle between it and index finger twice or more as great as between any other two digits. Or point of attachment to hand distinctly nearer the wrist than other fingers. On both hands if shown.

ARM JOINT: Elbow or shoulder. Elbow must show an abrupt bend about middle of arm. Not a curve. One arm sufficient. Shoulder must show arms at the side. Distinct curve at point of attachment to body. One arm need not be at side if there is a logical reason for it not being there.

LEG JOINT: Knee or hip. Knee must show an abrupt bend about the middle of leg. Not a curve. Can show narrowing of leg at this point.

For hip, inner lines of two legs meet at point of junction with the body.

PROPORTION - HEAD: Area of head not more than one-half or less than one-tenth of trunk.

PROPORTION - ARMS: As long or slightly longer than trunk. Not reaching knees. Width less than trunk.

PROPORTION - LEGS: As long or longer than trunk. Not greater than twice as long. Width less than trunk.

PROPORTION - FEET: Feet and legs must be shown. Length of foot greater than height. Length of foot not more than one-third or less than one-tenth of total leg.

PROPORTION - TWO DIMENSION: Both arms and legs shown in two dimension. Hands and feet neet not be.

HEEL SHOWN: Any kind.

MOTOR COORDINATION - LINES A: All lines reasonably firm. Not marked tendency to overlap or gap at points of junction. "Sketchy" type of drawings are acceptable.

MOTOR COORDINATION - LINES B: All lines firmly drawn. Correct joining.

MOTOR COORDINATION - HEAD OUTLINE: No obvious unintentional irregularities in outline. Not if head is crude circle or ellipse.

MOTOR COORDINATION - TRUNK OUTLINE: No obvious unintentional irregularities in outline. Not if trunk is crude circle or ellipse.

MOTOR COORDINATION - ARMS AND LEGS: No obvious unintentional irregularities in outline. Without tendency to narrow at point of junction with body. Both arms and legs must be in two dimension.

MOTOR COORDINATION - FEATURES: Eyes, nose, and mouth two dimension. *Full face*, eyes equidistant from nose and corners of mouth. Nose above center of mouth and equidistant from corners of mouth. Two sides of mouth alike and mouth at right angles to axis of head. *Profile*, distance from center of eye to back of head twice or more as great as center of eye to outer edge of nose. Nose in proportion to head and other features and forms an obtuse angle with forehead. Mouth in proportion to head and other features.

EARS: Two in full face. One in profile. Any kind.

EARS - CORRECT PROPORTION AND POSITION: Height greater than width. Placed in middle two-thirds of head. Shall extend toward back of head. Must have dot for aural canal in profile.

EYE DETAIL: Brow or lashes. Any kind.

EYE DETAIL: Pupil. Not a dot with curved line above. In both eyes if shown.

EYE DETAIL: Width greater than height. In both eyes if shown.

EYE DETAIL GLANCE: Face must be profile. Glance distinctly straight ahead from face.

CHIN AND FOREHEAD: *Full face,* eyes and mouth must be present. Sufficient space above eyes and below mouth to represent forehead and chin. Not if outline not present to separate chin from neck. *Profile,* eyes and mouth not necessary if outline of face shows clearly limits of forehead and chin.

PROJECTION OF CHIN: Full face must have a curved line below lip.

PROFILE A: Head, trunk, and feet must be shown in profile without error. Entire drawing contains no more than one of the following errors: bodily transparency, legs not in profile, arms attached to outline of back and extending foreward.

PROFILE B: Figure must be shown in true profile, without error or bodily transparency.

Appendix D

Individual Raw Scores

Table 2. Comparisons between WPPSI, Bender-Gestalt, and Draw-a-Man Scores

Student	WPPSI I.Q. Full-scale score	Bender-Gestalt raw score	Draw-a-Man raw score
1	104	9	15
2	116	12	14
3	119	10	12
4	96	14	10
5	96	16	12
6	99	15	8
7	86	21	8
8	115	13	11
9	101	14	9
10	84	19	6
11	101	13	8
12	114	10	19
13	92	12	9
14	70	14	7
15	93	15	11
16	96	15	16
17	107	9	17
18	109	9	23
19	76	20	12
20	101	8	15
21	122	12	19
22	99	10	16

VITA

G. Edward Allen, Jr.

Candidate for the Degree of

Master of Science

Thesis: An Investigation of the Relationship Between the Bender-Gestalt, Draw-a-Man, and Wechsler Preschool and Primary Scale of Intelligence

Major Field: Psychology

Biographical Information:

Personal Data: Born February 22, 1943, son of Mr. and Mrs. George E. Allen; married Vicky Peterson December 18, 1964.

Education: Received the Bachelor of Science degree from Utah State University, with a major in psychology, in 1965; completed requirements for the Master of Science degree, specializing in psychology, at Utah State University in 1968.

Professional Experience: 1968 to present, research intern at Far West Laboratory for Educational Research and Development; 1967-68, instructor, Utah State Extension Service; 1966-68, teaching assistant, Utah State University; 1966, survey aide, Community Action Program.

Organizations and Distinctions: 1966-67, member and president of Psi Chi (psychology honorary fraternity); 1962-63, academic tuition scholarship.