

379
N81
NO. 5413

BENDER-GESTALT EMOTIONAL INDICATORS AND
ACTING-OUT BEHAVIOR IN YOUNG CHILDREN

THESIS

Presented to the Graduate Council of the
North Texas State University in Partial
Fulfillment of the Requirements

For the Degree of

MASTER OF SCIENCE

By

Donald Everett Trahan, B. S.

Denton, Texas

August, 1977

AA.

Trahan, Donald Everett, Bender-Gestalt Emotional Indicators and Acting-Out Behavior in Young Children. Master of Science (General Experimental Psychology), August, 1977, 31 pp., 4 tables, references, 30 titles.

This study was designed to investigate the relationship between 15 emotional indicators on the Bender-Gestalt Test and acting-out behavior in young children. The subjects were 93 children ranging in age from 5 to 12 years. Each was administered the Bender. A measure of each subject's overt acting-out behavior was then obtained by having teachers rate each student on a Behavioral Rating Scale. Subjects' records were then divided into groups on the basis of both sex and age. Results indicated that neither the total number of Bender indicators nor any of the individual Bender indicators were significantly correlated with total scores on the rating scale. Use of the Bender as a projective device to measure acting-out behavior was seriously questioned.

TABLE OF CONTENTS

	Page
LIST OF TABLES	iv
Thesis	
Introduction	1
Method	11
Subjects	
Procedure	
Results	15
Discussion	20
Appendix	24
References	28

LIST OF TABLES

Table		Page
1.	Relationship Between Number of Bender Acting-Out Indicators and Scores on Rating Scale	15
2.	Correlations Between Individual Bender Acting-Out Indicators and Scores on the Rating Scale	17
3.	Sex Differences in Number of Bender Acting-Out Indicators and Scores on Rating Scale	18
4.	Relationship Between Number of Bender Acting-Out Indicators and Scores on Rating Scale Combining Sexes	19

BENDER-GESTALT EMOTIONAL INDICATORS AND ACTING-OUT BEHAVIOR IN YOUNG CHILDREN

Since the introduction of the Bender Visual Motor Gestalt Test by Laretta Bender (1938), professionals in psychology and education have seen an amazing increase in its use in a number of divergent settings. The Bender, which was originally designed to provide a measure of perceptual motor development, is now being applied to problems in prediction of school achievement, assessment of mental functions, detection of organic disorders, prediction and diagnosis of learning problems, and screening for emotional disturbances. The current popularity of the instrument is highlighted in a recent survey (Lubin, Wallis, & Paine, 1971) which indicates that the Bender ranks third (behind the Wechsler Adult Intelligence Scale and Rorschach) in popularity among psychologists, and is the test most often used with the majority of their clients.

The present study focuses upon the utilization of the Bender as an indicator of maladaptive acting-out behavior in young children. For this investigation, acting-out behavior will be operationally defined by teachers' ratings of each child on a Behavioral Rating Scale (see Appendix A).

A review of the current literature reveals that the history of the Bender as an indicator of emotional disturbance

is a long one. Numerous investigators have attempted to find significant correlations between patterns of reproduction on the Bender and the presence of emotional maladjustment (Brannigan & Benowitz, 1975; Byrd, 1956; Clawson, 1959; Eber, 1958; Elliot, 1968; Gobetz, 1953; Handler & McIntosh, 1971; Hutt, 1969; Hutt & Briskin, 1960; Kai, 1972; Kitay, 1950; Koppitz, 1963; Mogin, 1966; Naches, 1967; Pascal & Suttell, 1951; Simpson, 1958; Tripp, 1957; Tucker & Spielberg, 1958; Watkins & Watkins, 1975; Zolik, 1958). Other authors have contributed reviews and theoretical considerations (Brown, 1965; Clawson, 1970; Koppitz, 1975; Learner, 1972). However, despite the large number of studies, additional scrutiny will reveal that very limited data are available which would allow clinicians to make diagnostic statements regarding specific patterns of maladaptive behavior from indicators on the Bender. In many of the preliminary studies, heterogeneous groups of neurotics and other "maladjusted" individuals were administered the Bender. Their records were scored for a number of indicators thought to be related to the various types of disorders and were then compared with the records of a group of "normal" controls. Statistical analysis then determined whether or not there were differences in the frequencies of the indicators between the two groups. The weakness of such an approach lies in the usefulness of the interpretation which can be drawn from the results. Even if it were found that the frequency of occurrence of a given indicator was significantly higher

in the heterogenous group of neurotics, it would be difficult to pinpoint the source of such differences within the mixed group.

Despite the paucity of evidence linking specific indicators on the Bender to specific patterns of emotional maladjustment, the tendency for clinicians to use the test in this manner continues to increase. It appears that in many instances clinicians and researchers are making rather specific hypotheses based on their own clinical experience and the limited data available, then are conducting studies using non-specific, heterogeneous groups, and finally are mistakenly assuming that results of the studies provide support for their original positions.

The situation quite recently has improved to a limited degree. A number of authors have approached the problem by utilizing homogeneous groups of individuals in an attempt to link specific indicators to more delimited categories of behavior (Brannigan et al., 1975; Clawson, 1959; Handler et al., 1971; Naches, 1967). The remainder of this review will thus focus on the attempts to identify specific indicators useful in predicting acting-out behavior. Although a thorough search will produce a list of almost two dozen response patterns which have been suggested as indicators of acting out, there are 15 which appear with some consistency.

1. Careless overwork or heavily reinforced lines has consistently appeared in records of acting-out individuals

(Brannigan et al., 1975; Brown, 1965; Handler et al., 1971; Hutt, 1969; Koppitz, 1963, 1975). The individual may impulsively redraw a design or part of it with heavy lines without taking care or even erasing the former design. The person usually draws the lines very rapidly as if slashing at the paper. This indicator is said to reflect an underlying impulsivity, hostility, and aggressiveness.

2. Progressively increasing size on Figures 1, 2, and 3 is another indicator which consistently has been reported to be correlated with acting-out behavior (Kai, 1972; Koppitz, 1963, 1975; Naches, 1967). In committing this error, the individual progressively increases the size of either the dots or circles within the first three designs so that the last ones are at least three times as large as the first ones. Koppitz (1963, 1975) attributes these patterns to explosiveness and low frustration tolerance.

3. Collision of designs has been related to hostile and aggressive behavior. Brown (1965) suggests that such overlapping of designs reflects an indifference to the life-space boundaries of other people as well as a disregard for their rights and comfort. Although a number of investigators have sought to validate this hypothesis (Brannigan et al., 1975; Brown, 1965; Byrd, 1956; Hutt, 1969; Mogin, 1966; Tripp, 1957), the results have been equivocal. Most of the studies utilized heterogeneous groups of neurotics rather than groups of acting-out individuals, thus making the results somewhat unclear.

Brannigan et al. (1975) did attempt to correlate scores on the Bender with specific patterns of acting out as measured by a behavioral rating scale. However, they found the presence of collision was not significantly correlated with acting out as measured by the rating scale.

4. Expansion 1, defined as the use of two or more sheets of paper, is one of the better-established indicators of acting out (Brown, 1965; Clawson, 1970; Koppitz, 1963, 1975; Naches, 1967). Expansion of this type is assumed to reflect precarious inner controls (Brown, 1965) and impulsivity (Koppitz, 1963, 1975). Although presence of this indicator appears to differentiate between normal and emotionally disturbed groups, the base rate of occurrence is quite low; and efforts to link its presence specifically to records of acting-out individuals have had limited success.

5. Expansion 2 refers to the pattern of spacing figures far apart on the page. This indicator is scored if the space between two successive drawings is more than half the size of the relevant axis of the preceding figure (Hutt, 1969). The presence of this indicator is assumed to reflect weak inner controls, hostility, and aggressiveness. Although a number of authors have contributed statements regarding the theoretical basis for such a hypothesis (Brown, 1965; Clawson, 1970; Hutt, 1969), there has been only a limited amount of actual data. Brannigan et al. (1975) and Byrd (1956)

attempted to verify this relationship but found that expansion of this type was not significantly correlated with acting out.

6. Expansion 3 refers to a pattern in which there is a progressively increasing size in the reproductions of the stimulus figures. This indicator is scored when there is an increment in the size of the drawings over at least six figures. Such a pattern is said to be associated with irritability, loss of control, and acting out impulsively (Brown, 1965; Clawson, 1970; Hutt, 1969). A number of authors have contributed empirical data which tends to support the hypothesis (Brannigan et al., 1975; Clawson, 1959).

7. Large size of drawings is another response pattern purported to be able to differentiate between records of maladjusted and normal individuals (Clawson, 1970; Elliot, 1968; Hutt, 1969; Kai, 1972; Koppitz, 1963, 1975; Naches, 1967). Although several definitions of this indicator have been offered, the one most consistently reported defines large size as being present when five or more figures show increase of the vertical or horizontal axis by more than one-fourth of the corresponding axis of the stimulus figure. It is interesting to note that few of these studies report actual data; and, with one exception, those which do report data have utilized heterogeneous groups of neurotics or psychotics for comparison with normal control groups. However, Naches (1967) did find that large size was significantly related to specific acting-out behavior.

8. Dashes substituted for circles is another response pattern assumed to be related to impulsive acting-out tendencies. This indicator is scored when at least half of the circles in Figure 2 are replaced with dashes at least 1/16-inch long. The paucity of data supporting this assumed relationship is alarming. Koppitz (1963, 1975) includes this indicator among her list of emotional indicators, yet she failed to find a significant relationship between its presence and maladjustment in her own data. Brannigan et al. (1975) also reported nonsignificant correlations between this indicator and acting out in adolescents. Zolik (1958) similarly found that the indicator did not differentiate between delinquent and non-delinquent groups. On the otherhand, Handler et al. (1971) did find an association between this response pattern and aggressiveness. Despite the continued reference to this indicator as a sign of impulsivity, aggressivenesss, or acting out, there is a dearth of evidence to justify this position.

9. Second attempt, another of the purported indicators of maladjustment, is scored when the drawing of a design or part of it is spontaneously abandoned before or after it has been completed and a new drawing of the design is made (Koppitz, 1963). Data providing support for the relationship between second attempt and acting out are sparce. Koppitz (1975) suggests that this indicator is related to impulsive behavior; however, in her original research (Koppitz, 1963) utilizing a heterogeneous group of emotionally disturbed children, she

found that is was related to maladjustment only in children of age 8 to 10 years. Mogin (1966) found that the presence of second attempt differentiated between groups of maladjusted and normal children; however, there was no attempt to link the presence of the indicator to specific acting-out tendencies. Zolik (1958) attempted to discriminate between delinquent and nondelinquent adolescents on the basis of second attempt. He found significant differences between the groups only for Figure 4. Handler et al. (1971) did find that aggressive children showed significantly more second attempts than either withdrawn or control children. However, the overall evidence demonstrating that the presence of second attempt is significantly correlated with any characteristic related to acting-out behavior is quite poor.

10. Dashes substituted for dots is another response pattern on the Bender reported to be correlated positively with emotional maladjustment (Brannigan et al., 1975; Brown, 1965; Hutt, 1969; Learner, 1972; Zolik, 1959). This indicator is scored when at least two of the dots in any of Figures 1, 3, or 5 are replaced by dashes at least 1/16-inch long. Despite the frequency with which this indicator is reported, there is a very limited amount of supportive data. Brown (1965) reports that on the basis of his clinical experience he has found this substitution to be a reliable indicator of impulsivity and indifference to control imperatives. Hutt (1969) similarly suggests that the substitution of dashes for dots is related

to maladjustment; however, he presents no data of his own. Zolik (1959) did attempt to relate the presence of this indicator to delinquency. He found a significantly higher occurrence of this indicator in the delinquent group only on Figure 3. Brannigan et al. (1975) found nonsignificant negative correlations between the presence of this response pattern and antisocial acting-out tendencies in adolescents.

11. Circles for dots is yet another substitution reported to be related to acting-out behavior. Presence of this indicator is scored when circles replace at least two dots in any of Figures 1, 3, or 5. Brown (1965) suggests that such distortion of dots has high validity as an indicator of severe tension and control difficulty in individuals who have strong acting-out impulses. Some support for the dot-distortion/acting-out relationship is provided by Tripp (1957) and Zolik (1959). However, in perhaps the best research designed to investigate this hypothesis, Brannigan et al. (1975) found nonsignificant correlations between circle-for-dot substitutions and a number of acting-out behaviors.

12. Sharp angles on Figure 6 has been suggested by Brown (1965) to indicate difficulty in holding aggressive drives under restraint. He reported that, on the basis of his clinical experience, even a single spike or angle in an otherwise flattened copy warrants alertness to the danger of acting out. There appear to be few, if any, statistical data to support this assumption.

13. Integration difficulty of Figures 3 and 7 has also been reported by Brown (1965) as being related to acting out. A marked disassociation of the arrow point in Figure 3 has reportedly been observed in records of individuals who are apt to lose control over their aggressive impulses. Brown (1965) also reports that Figure 7 suggests acting-out potentials when the leaning figure seems to be pushing against the vertical one in such a way as to give the impression that the latter is tilting to the right. However, no data are presented to support this hypothesis.

14. Boundary violation is considered present in the record if part of a figure fails to stop at the margin of the paper and "runs beyond this natural boundary" (Brown, 1965). This indicator reportedly suggests difficulty in maintaining emotional restraint.

15. Exaggerated curve on Figures 4, 5, and 6 has been reported to be related to impulsiveness, poor adjustment, and acting out (Brannigan et al., 1975; Byrd, 1956; Clawson, 1959; Hutt, 1969). The following scoring criteria were developed by Clawson (1959). In Figure 4, the base-altitude ratio of the curve must be within the range of 1:1 to 3:1. For Figure 5, the base-altitude must be within the range of 1:1 to 9:5. For Figure 6, the base-altitude ratios for the horizontal curve must be within the limits of 2:1 to 5:1, while for the vertical curve they must fall within the limits of $1\frac{1}{2}$:1 to 6:1. If these ratios are exceeded, exaggerated curve is scored.

If these 15 indicators are, in fact, valid correlates of acting-out behavior as measured by teachers' perceptions in a classroom setting, then there should be a strong relationship between the Bender score and the observed acting-out behavior. For the purposes of this study, it will be assumed that the educational classroom provides a setting divergent enough to allow the expression of numerous types of acting out. It follows that Bender acting-out scores should correlate positively with observations of acting-out behavior in that setting.

Method

Subjects

There were 47 females and 46 males in the sample of 93 subjects selected for the study. Of these, 59 children comprising the entire first five grades in an urban private school participated. In addition, 21 subjects were randomly selected from a rural elementary school, while 13 subjects were selected from an urban kindergarten class. The age ranges of the groups were 5 to 8 ($\underline{n} = 50$) and 8 to 12 ($\underline{n} = 43$).

Procedure

Each subject was administered the Bender Visual Motor Gestalt Test (Bender) according to the standardized procedure for individual administration (Bender, 1946). For each subject, the test was introduced in the following manner: "I am going to show you a series of nine cards. Each card has a design on it. I would like for you to copy each of the

drawings as best you can." Subjects were allowed access to as many sheets of paper as they desired to use. Questions directed to the examiner during the administration of the test were answered as nondirectively as possible. The statement, "Do it however you like," was used whenever appropriate.

Each subject's record was then scored by the examiner and a second rater for the presence or absence of each of the 15 proposed indicators. High interrater reliability for the Bender scoring system was demonstrated, $\chi^2(1) = 154.45$, $p < .001$. The total number of scored indicators which appeared in each record was then used to represent the subject's acting-out score as reflected in the Bender. In each case, the score obtained by the examiner's ratings was used for this indice. Complete directions for scoring each indicator are included (see Appendix B).

The second phase of the study involved obtaining a measure of each subject's acting-out behavior in a classroom setting. Teachers of each student were given copies of a Behavioral Rating Scale developed specifically for this study (see Appendix A). The rating scale contained 10 items. Two items from the scale contributed to each of the following categories of behavior: (a) aggressiveness, (b) unethical actions, (c) defiance-resistance, (d) poor emotional control, and (e) inability to delay. The latter four categories were suggested by Brannigan et al. (1975). Teachers were also given a direction sheet containing information about the

scoring of each category of behavior. Teachers were instructed to rate each subject on all items of the rating scale. Each behavior represented by a scale item was given a score ranging from 1 to 6, with each score representing a stated frequency of occurrence. The scores for each item were arranged so that a score of 1 represented a low rate of the behavior being considered, while a score of 6 represented a high rate of occurrence. The total score for the rating scale was then obtained by summing the subject's scores from the 10 individual items. This total score was then used to represent the subject's rate of overt acting-out behavior in the classroom as perceived by the teacher.

For 58 of the subjects, ratings from two teachers on the rating scale were available. An estimate of the interrater reliability was obtained by calculating Spearman's rho. Excellent scoring reliability was demonstrated for this sample, $\rho(56) = .80$, $p < .001$.

After all records had been collected and scores from both the Bender and rating scale were available, the subjects' data were divided into four categories. Male and female records were evaluated separately since there appears to be adequate evidence that sex differences exist in the frequency of expression of acting-out behavior (Werry & Quay, 1971). Also, the records of children of ages 5 to 8 years were evaluated separately from those of ages 8 to 12 years. Koppitz (1963)

found small differences in the frequency of Bender emotional indicators between younger and older children.

The final design included four categories: (a) females 5 to 8 years old, (b) males 5 to 8 years old, (c) females 8 to 12 years old, and (d) males 8 to 12 years old. For each category, a median split technique was used to divide the subjects' records into groups of high and low acting out based on the total number of Bender acting-out indicators. Similarly, records for each category were divided by a median split into high and low acting-out groups based on total scores from the rating scale. Thus, for each category a 2 X 2 table was produced.

For categories (a) and (b), chi-square was calculated in order to determine if there was a significant relationship between the total number of acting-out indicators on the Bender and the actual occurrence of acting-out behavior in the classroom as measured by teachers' evaluations on the rating scale. For categories (c) and (d), the exact probabilities were derived using the method suggested by McNemar (1969). For these categories, the expected cell frequencies were too small to allow use of the chi-square statistic.

Finally, in order to provide a measure of the degree to which each individual acting-out indicator was related to observed acting-out behavior, the point-biserial correlations between each Bender indicator and total scores on the rating

scale were obtained (Nunnally, 1967). Each correlation was then tested for significance by utilizing the t -ratio method suggested by McNemar (1969).

Results

Analysis of the data presented in Table 1 indicates that significant relationships between the total number of Bender acting-out indicators and total scores on the rating scale were not found for any of the four groups. For males 5 to 8 years old, the pattern of scores obtained in this study could have occurred by chance, $\chi^2(1) = .371$, $p > .50$. Similarly, the

Table 1
Relationship Between Number of Bender
Acting-Out Indicators and Scores on Rating Scale

Gender of Group	Age	**Bender Indicators	**Rating Scale Scores		Chi- Square	p
			Low	High		
Male	5 to 8	Low	8	4	.371	p > .50
		High	5	5		
Female	5 to 8	Low	7	5	.051	p > .70
		High	7	6		
Male	8 to 12	Low	9	3	*	p = .071
		High	3	6		
Female	8 to 12	Low	8	4	*	p = .361
		High	5	5		

*The small n for these categories made it necessary to calculate the exact probability (McNemar, 1969).

**Low and High refer to scores below and above median.

data for females of ages 5 to 8 fail to support the hypothesis that acting-out behavior can be predicted by the number of acting-out indicators on the Bender, $\chi^2(1) = .051$, $p > .70$. In both the male and female 8 to 12 groups, calculation of the exact probability of those patterns of scores again revealed no relationship between the number of acting-out indicators on Bender records and actual classroom behaviors as measured by the rating scale. The exact probabilities of the pattern of scores for those groups were .071 and .361, respectively.

The relationship between the presence or absence of each individual indicator and total scores on the rating scale is shown in Table 2. Unfortunately, for 23 of the 60 relationships, the frequency of occurrence of the Bender indicator was so low that correlational analysis was not appropriate. For those cases in which the indicator appeared only once or not at all in a given group, the point-biserial correlations were not reported. The correlation coefficients for the remaining 37 categories are listed in Table 2. Analysis of this data failed to support the hypothesis that acting-out behavior can be predicted by the presence of any of the 15 indicators. Not a single correlation coefficient reached even the .05 level of significance.

The entirely negative results obtained here were somewhat surprising in view of some of the research cited earlier. However, a closer look at the data revealed important sex differences in both the number of observed Bender-Gestalt

Table 2
Correlations Between Individual Bender Acting-Out
Indicators and Scores on the Rating Scale

Bender-Gestalt Indicator	Males		Females	
	Age 5 to 8	Age 8 to 12	Age 5 to 8	Age 8 to 12
Careless Overwork	.110	-.110	NC	NC
Progressive Increase In Size	-.066	NC	-.183	NC
Collision	.179	.290	.251	.097
Expansion 1	.119	.097	.120	NC
Expansion 2	-.050	-.299	.059	.063
Expansion 3	NC	NC	NC	NC
Large Size	-.117	NC	NC	NC
Dashes for Circles	-.132	NC	NC	NC
Second Attempt	-.091	.146	-.211	NC
Dashes for Dots	.174	NC	-.067	NC
Circles for Dots	-.053	-.045	-.149	.258
Sharp Angles on Fig. 6	.135	.160	.065	NC
Integration Difficulty	.161	-.125	.130	-.125
Boundary Violation	NC	NC	NC	NC
Exaggerated Curve	.171	.101	.137	-.280
	<u>n</u> = 25	<u>n</u> = 21	<u>n</u> = 25	<u>n</u> = 22

Note: Correlation coefficients were not calculated for items marked NC. In such cases the indicator appeared on fewer than 10% of the subjects' records.

acting-out indicators and total scores on the rating scale (see Table 3). For subjects 8 to 12 years old, males scored significantly higher than females on both the number of observed acting-out indicators and scores on the rating scale. Significant differences were not observed between males and females in the 5- to 8-year age group. It is believed that these differences might partially explain the significant correlations reported in earlier studies in which males and females were combined into single groups. This would be

Table 3
Sex Differences in Number of Bender
Acting-Out Indicators and Scores on Rating Scale

Age Range		Bender Acting-Out Indicators			<u>t</u>
		Mean	S.D.	<u>n</u>	
5 to 8	Male	4.88	2.19	25	1.05
	Female	4.32	1.52	25	
8 to 12	Male	3.43	1.43	21	3.45*
	Female	2.14	.99	22	
Rating Scale Scores					
5 to 8	Male	21.84	8.08	25	1.51
	Female	18.60	7.06	25	
8 to 12	Male	23.48	9.42	21	3.08*
	Female	16.18	5.75	22	

*p < .001

particularly true for those studies in which chi-square or a similar statistic was used. Since males tend to score higher on both measures of acting out, at least for the older group, their scores would form a cluster in one cell of the distribution while the scores of females, which were lower on both measures, would form a cluster in another cell of the distribution.

Scores of both males and females were thus combined for each of the age groups as presented in Table 4. Results of this manipulation failed to support the hypothesis that the presence of acting-out indicators on the Bender-Gestalt is related to overt acting-out behavior. For the 5- to 8-year age group, the distribution of scores was exactly that which would be expected if there were no relationship between the two measures, $\chi^2(1) = 0.0$, $p = 1.00$. For subjects in the 8 to 12

Table 4

Relationship Between Number of Bender Acting-Out Indicators and Scores on Behavioral Rating Scale Combining Sexes

Age Group	n	*Bender Acting-Out Score	*Rating Scale Scores		Chi-Square	p
			Low	High		
5 to 8	50	Low	12	12	0.0	p = 1.00
		High	13	13		
8 to 12	43	Low	11	8	2.60	p > .10
		High	8	16		

*Low and High refer to scores below and above median.

group, scores likewise reflected no more than chance fluctuations, $\chi^2(1) = 2.60, p > .10$.

Discussion

The results of the present study failed to support the hypothesis that the 15 patterns of reproduction on the Bender Visual Motor Gestalt Test (Bender) were related to overt acting-out behavior as measured by teachers' classroom ratings. Neither the total number of scored indicators, nor the presence of any individual indicator, correlated significantly with scores on the Behavioral Rating Scale. The pattern of negative results was consistent for both males and females in each age group.

Although the present findings were somewhat inconsistent with conclusions drawn from past research (Byrd, 1956; Clawson, 1959; Koppitz, 1963), there were several important distinctions between the present study and previous ones. First, the present study focused upon the relationship between Bender indicators and rather specific types of overt acting-out behavior. Many of the previous studies have investigated only the differences in the frequency of occurrence of various indicators on records of a mixed group of neurotics and a group of "normals." Although there may have in fact been significantly higher frequencies of "acting-out indicators" in the records of neurotics in general, there have been very few attempts to determine whether or not the indicators were actually present in the records of those neurotics who tended to act out or

whether they were present in the records of other neurotics within the group. In view of this, it was difficult to determine whether the present results contradict those previously reported or whether such discrepancies reflect methodological differences.

This study was also methodologically different from previous research in that records of males and females were evaluated separately. This may be an important methodological improvement since there appear to be sex differences not only in the presence of Bender indicators but also in the rate of acting-out behavior. Data from the present study indicated that males tend to score higher on both of these measures. These findings were consistent with those reported by Koppitz (1963) and Werry & Quay (1971). Significant correlations reported in some previous research in which male and female records were combined may be reflecting only that males score higher on both indices, while females score lower on both. However, while this may seem logical, data from the present study failed to support the hypothesis that Bender indicators were significantly correlated with acting-out behavior, even for the combined groups.

Another very important consideration in conducting research on this issue is that a rather large amount of evidence exists which indicates that some of the Bender indicators purported to be related to acting out are directly related to the subject's level of perceptual motor development (Koppitz,

1963; Pascal & Suttell, 1951). If the ability to reproduce Bender designs without these indicators is related to perceptual motor maturity, then any attempts to determine the relationship between these indicators and acting-out behavior would be confounded by developmental differences. This is particularly true for younger children.

On the basis of these considerations and the data presently obtained, it appeared that none of the 15 Bender indicators utilized in this study were able to predict differences in overt acting-out behaviors as measured by teachers' ratings in a classroom setting. Although this conclusion varies somewhat from some of the general results in studies previously cited, it was quite consistent with the results of many other authors (Brannigan et al., 1975; Byrd, 1956; Koppitz, 1963; Naches, 1967; Zolik, 1958).

Overall, the methodology of the present study appeared adequate. The interrater reliability of both the Bender scoring system and the rating scale were well within acceptable limits. The high correlation between teachers' ratings of each student was particularly surprising in view of the widespread criticism of the effects of bias in teachers' ratings of students. In analyzing the data, further attempts were made to control for teacher bias. By utilizing median split techniques, subjects were divided into high and low groups on each measure. This procedure would tend to eliminate the

differences due to bias which may have occurred at either extreme on the two measures.

The most obvious weakness of the present study was the rather limited sample of subjects. Although subjects were drawn from three separate locations in an attempt to obtain a representative sample, over half attended an urban private school and probably represented an atypical group. Also, the sample size for each of the four categories studied tended to be rather small.

On the basis of this study, there were a number of implications for future research. First, it appeared that records of males and females should be evaluated separately. Second, it is suggested that levels of perceptual motor development be controlled. This might be done by focusing on narrower age ranges or by providing an independent measure of perceptual motor maturity. Finally, larger and more representative samples appear necessary in order to assure generality of the results.

Appendix A

Behavioral Rating Scale

- A. Overt Physical Aggression Against Another Person.
1. Less than once per month
 2. Once monthly to once weekly
 3. Two times per week
 4. Several times per week
 5. One to two times daily
 6. Three times daily or more
- B. Verbal Aggression Against Another Person.
1. Less than once per month
 2. Once monthly to once weekly
 3. Two times per week
 4. Several times per week
 5. One to two times daily
 6. Three times daily or more
- C. Destruction of Personal Property of Self or Others.
1. Less than once per month
 2. Once monthly to once weekly
 3. Two times per week
 4. Several times per week
 5. One to two times daily
 6. Three times daily or more
- D. Disobeying Direct Orders of Teacher.
1. Less than once per week
 2. One to two times per week
 3. Several times per week
 4. One to two times daily
 5. Three to four times daily
 6. Five times daily or more
- E. Violating Established School Rules.
1. Less than once per month
 2. Once monthly to once weekly
 3. Two times per week
 4. Several times per week
 5. One to two times daily
 6. Three times daily or more

F. Taking Property of Another Individual.

1. Less than once per month
2. Once monthly
3. Twice monthly to once weekly
4. Two times per week
5. Almost daily
6. Twice daily or more

G. Disrupting Classroom Activity.

1. Less than once per week
2. One to two times per week
3. Several times per week
4. One to two times daily
5. Three to four times daily
6. Five times daily or more

H. Use of Profane Language.

1. Less than once per week
2. One to two times per week
3. Several times per week
4. One to two times daily
5. Three to four times daily
6. Five times daily or more

I. Talks in Class Without Permission.

1. Less than once per week
2. One to four times per week
3. One to two times daily
4. Three to four times daily
5. Five to seven times daily
6. Eight times daily or more

J. Leaves Seat Without Permission.

1. Less than once per week
2. One to two times per week
3. Several times per week
4. One to two times daily
5. Three to four times daily
6. Five times daily or more

Appendix B

Bender-Gestalt Acting-Out Scoring Key

- ___ 1. Careless Overwork: A design or part of it is redrawn or drawn over with heavy lines, impulsive lines. If attempts are made to erase and redraw, the design is not scored.
- ___ 2. Progressively Increasing Size on Figures 1, 2, and 3: Either the dots or circles of Figures 1, 2, or 3 are increased in size such that the last ones are at least three times as large as the first ones.
- ___ 3. Collision: Any part of one design overlaps with part of another design.
- ___ 4. Expansion 1: Use of two or more sheets of paper (use of front and back of one sheet is also scored).
- ___ 5. Expansion 2: Figures are spaced far apart on page. If the vertical or horizontal distance between two successive drawings is more than half the size of the relevant axis of the preceding figure. Must occur 2 or more times to be scored.
- ___ 6. Expansion 3: Progressively increasing size of the stimulus figures. There is an increment in the size of the drawings over at least six figures.
- ___ 7. Large Size of Drawings: Five or more figures show an increase of the vertical or horizontal axis by more than one-fourth of the corresponding axis of the stimulus figure.
- ___ 8. Dashes Substituted for Circles: At least half of the circles in Figure 2 are replaced by dashes at least one-sixteenth inch long.
- ___ 9. Second Attempt: The design or part of it is spontaneously abandoned before or after it has been completed and a new drawing of the design is made. The two drawings must be made on different parts of the paper.
- ___ 10. Dashes Substituted for Dots: At least two of the dots in any of Figures 1, 3, or 5 are replaced by dashes at least one-sixteenth of an inch long.

- ___ 11. Circles for Dots: Circles replace at least two dots in any of Figures 1, 3, or 5.
- ___ 12. Sharp Angles on Figure 6: Sharp angles replace any of the smooth curves in Figure 6.
- ___ 13. Integration Difficulty: Scored when the arrow point in Figure 3 is markedly dissociated or if the left figure in 7 is "pushing against" the other so that the right figure is leaning to the right of vertical.
- ___ 14. Boundary Violation: Part of a figure runs off the edge of the paper.
- ___ 15. Exaggerated Curve: In Figure 4 the base-altitude ratio of the curve must be within the ratio of 1:1 to 3:1. For Figure 5, the base-altitude ratio must be within the range of 1:1 to 9:5. In Figure 6, the base-altitude ratio for the horizontal curve must be within the limits of 2:1 to 5:1, while for the vertical curve they must fall within the limits of $1\frac{1}{2}:\frac{1}{2}$ to 6:1. If the base-altitude ratio exceeds these values, this indicator is scored.

References

- Bender, L. A visual motor Gestalt test and its clinical use. New York: American Orthopsychiatric Association, 1938.
- Bender, L. Manual for instruction and test cards for visual motor Gestalt test. New York: American Orthopsychiatric Association, 1946.
- Brannigan, G. G., & Benowitz, M. L. Bender Gestalt signs and antisocial acting out tendencies in adolescents. Psychology in the Schools, 1975, 12(1), 15-17.
- Brown, F. The Bender-Gestalt and acting out. In L. Abt & S. Weissman (Eds.), Acting Out. New York: Grune & Stratton, 1965.
- Byrd, E. The clinical validity of the Bender-Gestalt Test with children: A developmental comparison of children in need of psychotherapy and children judged well-adjusted. Journal of Projective Techniques, 1956, 20, 127-136.
- Clawson, A. The Bender Visual Motor Gestalt Test as an index of emotional disturbance in children. Journal of Projective Techniques, 1959, 23, 198-206.
- Clawson, A. The Bender Visual Motor Gestalt Test for children: A manual (5th ed.). Los Angeles: Western Psychological Services, 1970.
- Eber, M. A Bender Gestalt validity study: The performance of mentally retarded children. (Doctoral dissertation, Florida State University, 1958). Dissertation Abstracts, 1958, 18, 296. (University Microfilms No. 58-433)

- Elliot, J. A validation study of the Koppitz and Pascal and Suttell systems with eleven through fourteen year old children. Unpublished master's thesis, University of Michigan, 1968.
- Gobetz, W. A quantification, standardization, and validation of the Bender-Gestalt Test on normal and neurotic adults. Psychological Monographs, 1953, 67 (6, Whole No. 356).
- Handler, L., & McIntosh, J. Predicting aggression and withdrawal in children with the Draw-a-Person and Bender-Gestalt. Journal of Personality Assessment, 1971, 35, 331-337.
- Hutt, M. The Hutt adaptation of the Bender-Gestalt Test (2nd ed.). New York: Grune & Stratton, 1969.
- Hutt, M., & Briskin, G. The clinical use of the revised Bender-Gestalt Test. New York: Grune & Stratton, 1960.
- Kai, T. An examination of the Koppitz Bender Gestalt Test (II): The correlation between each item of emotional indicators (EI) and emotional problems in younger children. Memoirs of the faculty of education, Kumamoto University, 1972, 20, section 2.
- Kitay, J. The Bender-Gestalt Test as a projective test. Journal of Clinical Psychology, 1950, 6, 170-174.
- Koppitz, E. The Bender Gestalt for young children. New York: Grune & Stratton, 1963.
- Koppitz, E. The Bender Gestalt for young children: Volume 2. New York: Grune & Stratton, 1975.

- Learner, E. The projective use of the Bender Gestalt.
Springfield, Illinois: Charles C. Thomas, 1972.
- Lubin, B., Wallis, R., & Paine, C. Patterns of psychological test usage in the United States: 1935-1969. Professional Psychology, 1971, 2, 1970-1974.
- McNemar, Q. Psychological Statistics (4th ed.). New York: John Wiley & Sons, 1969.
- Mogin, L. Administration and objective scoring of the Bender Gestalt Test in group screening of primary grade children for emotional maladjustment (Doctoral Dissertation, Rutgers State University, 1966). Dissertation Abstracts, 1966, 27, 6-A. (University Microfilm No. 66-12,076).
- Naches, A. The Bender Gestalt Test and acting out behavior in children. Unpublished doctoral dissertation, Colorado State University, 1967.
- Nunnally, J. Psychometric Theory. New York: McGraw-Hill, 1967.
- Pascal, G., & Suttell, B. The Bender-Gestalt Test. New York: Grune & Stratton, 1951.
- Simpson, W. A study of some factors of the Bender Gestalt reproduction of normal and disturbed children. (Doctoral dissertation, University of Oklahoma, 1958). Dissertation Abstracts, 1958, 19, 1120. (University Microfilm No. 58-3879).
- Tripp, A. Some graphomotor features of the Bender Visual Motor Gestalt Test in relation to delinquent and non-delinquent

- white adolescent males. Unpublished doctoral dissertation, New York University, 1957.
- Tucker, J., & Spielberg, M. Bender Gestalt Test correlates of emotional depression. Journal of Consulting Psychology, 1958, 22, 56.
- Watkins, J., & Watkins, D. Comparison of normal and emotionally disturbed children by the Plenk scoring system for the Bender Gestalt. Journal of Clinical Psychology, 1975, 31(1), 71-74.
- Werry, J., & Quay, H. The prevalence of behavior symptoms in younger elementary school children. American Journal of Orthopsychiatry, 1971, 4(1), 136-143.
- Zolik, E. A comparison of the Bender Gestalt reproductions of delinquents and non-delinquents. Journal of Clinical Psychology, 1958, 14, 24-26.