AN ENQUIRY INTO THE DETERMINANTS OF A DIFFERENTIATION BETWEEN ELDERLY "ORGANIC" AND "NON-ORGANIC" PSYCHIATRIC PATIENTS ON THE BENDER GESTALT TEST*

By

M. B. SHAPIRO, M.A., Ph.D.

JACK FIELD, M.A.

and

F. POST, M.B., M.R.C.P., D.P.M.

Institute of Psychiatry,
The Maudsley Hospital and the Bethlem Royal Hospital, London

I. Purpose

In a previous study, Shapiro, Post, Lofving and Inglis (10) found that a modified version of the Bender Visual-Motor Gestalt Test differentiated, at a high level of confidence, three groups of elderly psychiatric patients: brain-damaged patients, functionals and a group of doubtful diagnosis. This level of confidence was far higher than that of any of the other 24 tests used. Furthermore, a number of recent studies have shown that the Bender test differentiated Organic from non-Organic subjects (2, 4, 6, 7, 8). In our own Department Yates (13) found significant differences on measures of the reproductions of other designs between Organics and Functionals. In view of these findings we decided to investigate our own results further. Our general purpose was to measure and control some of the variables appearing relevant in the performance of the task and in this way make possible testable explanations for the results.

To achieve this purpose we had to overcome the following difficulty: The differentiation obtained by the Bender Gestalt Test was based on the use of subjective scores allocated according to a rating-scale. The scores for the individual reproduction were summed to give an aggregate score for each subject. It was this combined score which differentiated between the groups in our first study. It was impossible to say exactly what these aggregate scores were measuring and therefore what aspects of the drawing performances required explanation. For this reason we decided that the first step of our investigation must be to objectify the method of scoring which was used in the first instance. This paper is concerned with a description of the attempt and its outcome.

II. Subjects

The 89 patients involved in this study are described in more detail elsewhere (10). Their ages ranged from 60–92, with a mean age of 68·6 years. They were arranged into three diagnostic groups, 16 Organics, 58 Functionals and 15 "Doubtfuls", according to objective psychiatric criteria which were developed especially for our purpose (10). These groups of patients were

* The work reported in this paper was made possible by a grant from the Research Fund, made available from the endowment by the Board of Governors of the Bethlem Royal Hospital and the Maudsley Hospital.

defined respectively as suffering from "Psychiatric illnesses definitely associated with cerebral pathology", "Psychiatric illnesses not associated with cerebral pathology", and "Psychiatric illnesses associated with very mild or dubious organic involvement". Care had been taken to avoid contamination of diagnosis both by knowledge of test results, and by basing any part of the diagnosis on what the tests were supposed to measure.

III. Results of the Initial Study

The Bender Visual Motor Gestalt Test consists of nine simple designs presented successively, after the subject has been instructed to copy them with a pencil on a sheet of paper in front of him. In our first study the test had been modified in certain ways. Designs 1, 5 and 7 were excluded. Design 2 was simplified and Design 3 enlarged, to enable patients with slightly defective vision to do the test. The score for each reproduction was derived from a 5-point rating scale. The criteria for the rating scale were based on Bender's suggestions (1). Examples of the criteria used for one of the designs are given in Table II.

As was stated above, an Analysis of Variance of the results had shown that the Bender Test differentiated the diagnostic groups at a higher level of confidence than any of the other 24 tests used in that study ($P < \cdot 1$ per cent. $F = 8 \cdot 82$). It is important to note that neither the Wechsler weighted scores (four sub-tests) or age differentiated the diagnostic groups. A correlational analysis of all the differentiating tests indicated that the Bender test had little in common with the other tests, except the variance held in common with the Wechsler scores (5).

IV. Selection of Matched Groups

The preliminary analysis was carried out on two matched groups consisting of 15 Organics and 15 Functionals. The 15 Organics were in fact all the Organics used in the initial study, but one, whose record had been lost. The 15 Functionals were selected from our original Functionals by matching them as closely as possible with the Organics for age, Wechsler weighted scores and sex. The remaining 43 Functionals and 15 Doubtfuls were retained for cross-validation.

We succeeded in matching perfectly for Sex, but for Age and Intelligence there remained unavoidable differences within pairs. As Table I shows, Organics are brighter than Functionals, the difference being significant at the <5 per cent. level. The Organics are older than the Functionals although the difference fails to reach the 5 per cent. level of significance. If anything, the group differences should result in it being more difficult to obtain large differences between the Organics and Functionals on the Bender test, as the Bender was significantly correlated with the prorated Wechsler scores in the second study (5).

Table I

Differences Remaining between Matched Groups

	Intelligence (Wechsler prorated weighted scores for 4 sub-tests)			Age (years)		
N = 15	Mean	Range	SD	Mean	Range	SD
Organics	54·80	15–127·5	30·50	71·00	62–92	8·76
Functionals	52·53	25–125	30·20	68·07	61–81	5·26
Correlated 2-tailed t test*	A = ·265 P< ·05			A = ·301 P< 10%		

* $t = N - 1/AN - 1$. Sandler (9).

https://doi.org/10.1192/bjp.103.431.364 Published online by Cambridge University Press

Using the rating-scale scores the Bender test now differentiated the matched Organics and Functionals at beyond the 5 per cent. level of confidence (A = ·265, 14df). The complete groups had been differentiated at beyond the ·1 per cent. level, the difference in the level of confidence being due to equating the two groups on the matched variables. We decided that any new measure would have to differentiate the means of the matched groups at beyond the 5 per cent. level of confidence at least, as did the rating-scale scores.

V. Procedure

Having examined a sufficient number of different measures some would differ significantly by chance alone. By definition one out of any 100 randomly-selected measures will differentiate two groups, identical in fact, at the 1 per cent. level of confidence. Any measures which differentiated between our two matched groups at an acceptable level of significance would, therefore, have to be cross-validated on another sample. For this purpose the 43 Functionals who remained after the selection of 15 to match the Organics could no longer be regarded as a strictly random sample of Functionals admitted to the Geriatric Unit over a given period. Such objection cannot be made against the 15 Doubtfuls, who were patients who fulfilled neither the organic nor functional psychiatric criteria. Being a mixed group their mean on a newly-found differentiating measure could be predicted to fall between the means of the matched samples. In any case, the best guess would be that, given two significantly different means, a third mean would fall between them. However, we could in addition, assume that the Doubtfuls contained a higher proportion of Functionals than of Organics as was the case in the whole sample of admissions. We therefore predicted that on a differentiating measure the mean score of the Doubtfuls would differentiate the Organics at a higher level of confidence than the Functionals.

A second independent contribution to validation was supplied by the 43 remaining unmatched Functionals. They are less similar to the Organics on the matched variables than the matched Functionals. It follows, therefore, that a measure which differentiates the matched Functionals from the Organics should also differentiate the remaining Functionals, and, if anything, at a higher level of significance.

Thirdly, evidence for the validity of an objective measure was available from correlations within each diagnostic group, and from these groups combined. Any measure which was truly an objectification of the original subjective scale would have to correlate significantly with that scale.

Fourthly, the original scores should not retain, independently of the objective measure, any part of their variance which differentiates significantly between Organics and Functionals. This could be checked by an Analysis of Covariance.

VI. The Objectification of Subjective Criteria

(i) *Procedure*

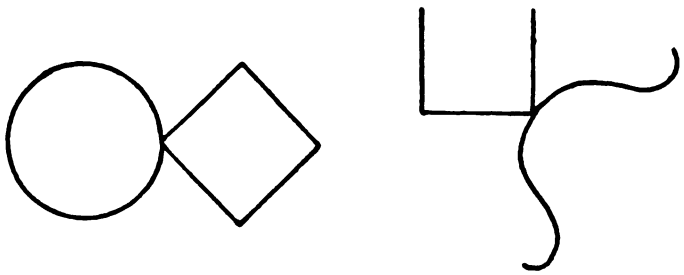
Our first attempt at objectification consisted of trying to express the criteria corresponding to each rating-score into clearly defined objective measures. For example, in Design A (Fig. 1) a score of 5 was awarded for a "good square and good circle. Especially diagonals should be about equal. They must still touch." Our attempt at objectification of these criteria was to measure the diagonals of the square and find a ratio by dividing the length of the longer

by that of the shorter. About 30 such measures, based on 57 measurements for each record were taken. Similar measures were grouped together in five classifications. The measure in our example was classified with similar ratios from other designs as a Proportionality measure (Table II).

TABLE II

First Analysis of Rating-Scores for Design A

Original Rating Score	Original Scoring Criteria	Objective Measures	Classification
1	Scribble or drawing which has no resemblance to test design		
2	Two loops incompletely closed; they may touch, overlap or have considerable space between them. Or one closed loop resembling either a square or a circle		
3	Two closed loops in horizontal plane, the second modified to resemble a square and the first modified to resemble a circle; they may touch or have some space between them	Deviation from horizontal best-fit line	Horizontal orientation
4	A fairly good circle and a fairly good square and square orientated on the diagonal; they should touch and size of two reproductions should be about equal	Deviation from vertical of vertical diagonal. Distance between parts	Vertical orientation Separation of parts
5	Good square and good circle. Especially diagonals should be about equal. They must still touch	Ratio of diagonals of square	Proportionality



Design A Design 4

FIG. 1.—The two Bender Gestalt Test Designs used for the Diagonals and Angles Measures.

(ii) *Results*

Histograms were drawn for each measure and for the similar measures combined under their separate classifications. There was a marked trend in the histograms for Organics to copy the designs less accurately on all the measures than their matched Functionals, but in no case did the means of the two groups differ as significantly as the means of the subjective rating-scale scores.

4A

https://doi.org/10.1192/bjp.103.431.364 Published online by Cambridge University Press

It is not suggested that the measures selected are the only or the best objectification of the subjective criteria. The task proved difficult and laborious, but our findings were not promising enough to pursue further.

(iii) *Conclusion*

One reason for the lack of differentiation suggested itself to us. Some reproductions could not be measured at all because of their poor quality, although they could be scored according to the subjective criteria. The reproductions, as their quality falls, diverge more and more from the original designs. The more they diverge, the less amenable do they become to measurement by a common objective standard. For example, on Design A a score of 1 is given when there is only a scribble, but a measurement on Proportionality can only be made if it is possible to draw diagonals in the square. This would not be possible till the design is good enough to deserve a high subjective rating. Thus, compared to the rating-scores, our measures appeared seriously blunted at the lower level of quality of the reproduction of the designs.

VII. "Unmeasurability" as a Score

(i) *Procedure*

To check the correctness of this view we set up an "Unmeasurability" criterion. Confining ourselves to the most clearly definable of our measures (Vertical and Horizontal Orientation, Proportionality and Area), the score of 1 was given for each measure per reproduction which could be made. A score of 2 was given in each case where the parts of the reproduction required for any measure were clearly missing but could be identified from what was reproduced, i.e. the rest of the drawing was well enough done to indicate that the part in question was clearly missing. A score of 3 was given in each case where the parts of the reproduction for any measure were not identifiable.

(ii) *Results*

The results were consistent with our explanation. The rating scale scores could quantify the least accurate copies. The objective measures we had used could not.

"Unmeasurability" identified the same 4 Organics in the same order outside the overlap as the rating-scores. Of the 15 matched pairs, 9 Organics and 14 Functionals gained the same minimum score on the "Unmeasurability" criterion; that is, all the measures could be made on all of their reproductions.

No conventional statistic could aptly be used to reject the null hypothesis for this distribution. But if we think of these 30 cases as ranked, the probability of finding the same 4 cases in the same order at the same end of the scale as with the subjective scores is not less than 1 in 27^4. We can therefore reject the explanation attributing the results to chance at beyond the ·001 per cent. level of confidence.

"Unmeasurability" was not as effective as the original subjective scoring in the total sample. Here the subjective method placed 10 of the 58 Functionals outside the overlap but "Unmeasurability" identified none. All of the Doubtful group produced perfect "Unmeasurability" scores and so cross-validation of this finding was not possible.

VIII. The Measurement of an "Obvious Characteristic" of the Drawings

(i) *Procedure*

It was clearly necessary to make another attempt to objectify the scoring. It seemed to us that in view of the nature of the subjective criteria, many attempts at analysis could be made and the differentiating parts of the drawings might still remain undiscovered. We therefore decided to alter our line of approach. Instead of a deductive analysis based on the criteria of the subjective scores we set up, on *a priori* considerations, conditions with which our measure would be expected to comply. These were:

(*a*) that the aspect concerned must be obvious, likely to catch the scorer's eye and hence influence his rating;

(*b*) that the measure must be applicable to the records of every subject so that we would have to omit as few cases as possible;

(*c*) that the particular characteristic must be easily and unequivocally measurable.

The characteristic that complied most completely with these conditions was a ratio of diagonals in the square of Design A and the open square of Design 4 (Fig. 1). Firstly, the ratio of diagonals is a measure on which quantitative differences are obvious and therefore likely to influence ratings. Secondly, there was only one Organic where the Diagonals Measure was impossible on both copies, and two Organics where it was impossible on one or other of the copies. For purposes of statistical analysis we omitted the one Functional who was matched with the Organic who produced two unmeasurable designs. In the case of the two Organics who produced only one measurable copy each we used the measure on that one reproduction. In all cases where the two copies were used an average was arrived at. The third condition was also met; the diagonals were easily and unequivocally measurable. None of the other designs fulfilled the three conditions.

(ii) *Results*

The significance of the difference between the means of the Diagonals Measure of the matched Organics and Functionals, was found to be beyond the 5 per cent. level of confidence (A = ·247, 13 df) and in the expected direction.

It was now necessary to cross-validate and further investigate this, the new measure.

Firstly, the records of the 15 Doubtfuls and 43 unmatched Functionals were measured and the ratio of diagonals derived. A one-tailed test for uncorrelated means was used, the direction of the difference having been correctly predicted. The means of the Doubtfuls and Organics differed significantly at beyond the 2·5 per cent. level of confidence. The difference between the means of all the Functionals combined and the Doubtfuls was significant at beyond the 5 per cent. level. The difference between the Doubtfuls and Organics is at a higher level of confidence than the difference between Doubtfuls and Functionals, in accordance with the prediction.

Secondly, the means of the unmatched Functionals and Organics differed significantly at beyond the ·05 per cent. level and the means of the combined matched and unmatched Functionals differed significantly from the mean of the Organics also at beyond the ·05 per cent. level of confidence. The differences and other relevant information are given in Table III.

TABLE III

Differences found between the Groups by the Diagonals Measure

Groups compared (1)	(2)	Group (1) M	SD	N	Group (2) M	SD	N	Diff.	A or t Ratio	P %	df
Organics	Matched Functionals	1·266	·150	14	1·158	·144	14	·108	·247	<5	13
Organics	Doubtfuls	1·266	·150	14	1·164	·080	15	·102	2·304	<2·5	27
All Functionals	Doubtfuls	1·131	·097	58	1·164	·080	15	·033	1·212	<5	71
Unmatched Functionals	Organics	1·118	·074	43	1·266	·150	14	·148	4·933	< ·05	55
All Functionals	Organics	1·131	·097	58	1·266	·150	14	·135	4·156	< ·05	70

Inhomogeneity of variance was allowed for by the Cochran-Cox Method (11).

Thirdly, significant correlations were found in each diagnostic group and in all the groups combined between the Diagonals Measure and the original subjective scores as shown in Table IV.

The reliability of the Diagonals Measure was investigated. Two of our colleagues measured the diagonals for the matched Organics and Functionals. One of them did this for the Organics and the other for the Functionals. The Diagonals Measure derived from their measurements correlated ·92 with the original Diagonals Measure.

IX. THE NATURE OF THIS "OBVIOUS CHARACTERISTIC"

(i) *Discussion*

We had now almost completed albeit in an unexpected manner, the first step of our investigation. We had found an objective measure which appeared to do the same kind of job as the original subjective scale. Nevertheless, the

TABLE IV

Product Moment Correlations between Diagonals Measure and Subjective Rating-scores for the Groups

Group	N	r	P %
Organics*	12	·74	<1
Matched Functionals	15	·66	<1
Unmatched Functionals	43	·41	<1
All Functionals	58	·54	< ·1
Doubtfuls	15	·56	<5
All Cases	85	·61	< ·1

* Only cases where both copies were measurable are included.

Diagonals Measure still did not appear to be a satisfactory means of investigating the mechanisms affecting the copying of designs by different kinds of psychiatric patients. It was not a direct measure of the outcome of any part of the active performance of making a drawing. The subject does not *draw* the diagonals in the sense that he draws, say, the lines making up his copy. We therefore asked ourselves what parts of the drawing performance would contribute to the ratio of the diagonals. It seemed that there were at least two such determinants; those resulting in the lengths of the sides, and those resulting in the angles between the sides. An investigation of these two aspects of the drawing was, therefore, carried out.

(ii) *Procedure*

The deviation from a right angle of each of the 8 angles in the copies of the two designs was measured in degrees and summed regardless of sign. The result was called the Angles Measure. The Sides Measure consisted of the sum

of the lengths of the three longest sides divided by the length of the shortest for each of the two copies. For these measures the "open" square (Fig. 1) was regarded as being a closed figure with 4 sides. If one of these measures correlated significantly higher with the Diagonals Measure than the other one it would be the more important determinant of the Diagonals Measure.

As in the case of the Diagonals Measure there was one Organic (the same individual) whose reproduction could not be measured. Both he and the matched Functional were omitted from the analysis. Two Organics (one of whom was the same as in the Diagonals study) only produced one copy which could be measured. Their measures were arrived at by doubling the single measures.

(iii) *Results*

The Angles Measure differentiated between matched Organics and Functionals at beyond the 5 per cent. level of confidence (A = ·227, 13 df). None of the angles taken separately came near to differentiating the groups. The Sides Measure showed no tendency to differentiate the two groups.

The correlation between the Angles Measure and the Diagonals Measure for the matched Organics and Functionals taken together was ·93, and separately ·85 and ·92 respectively ($P<1$ per cent.). The correlation between the Sides Measure and the Diagonals Measure for the matched groups taken together was ·15 and separately was ·14 in both groups, which is not significantly different from zero.

The correlations of the Angles Measure with the Diagonals Measure are therefore consistently and significantly higher than the correlation of the Sides Measure with Diagonals Measure. This suggests that it is the Angles Measure rather than the Sides Measure that is a determinant of the Diagonals Measure.

The validity of the Angles Measure was now further examined.

An Analysis of Variance was carried out using the Organics, matched and unmatched Functionals and Doubtfuls. It seemed unnecessary to carry out the cross-validating procedure independently for each pair of groups as had been done for the Diagonals Measures, as we are now dealing with a correlate of that measure. The Analysis of Variance would be shorter and provide us with estimates of error variance based on the total sample. The F ratio was 8·42, significant at beyond the ·1 per cent. level (the F ratio found for the original subjective criteria was 8·82). The Within Variance with 83 degree of freedom was used for the one-tailed t tests, the direction of the differences having been predicted. The means of the Organics and Doubtfuls differed significantly at beyond the ·5 per cent. level, the unmatched Functionals and Doubtfuls at beyond the 5 per cent. level and the unmatched Functionals and Organics at beyond the ·05 per cent. level (Table V). As with the Diagonals Measure there was a greater tendency for the means of the Doubtfuls to fall nearer to that of the Functionals than to that of the Organics, as had been predicted.

TABLE V

Differences found between the Groups by the Angles Measure by an Analysis of Variance (df=83)

Groups compared (1)	(2)	Group (1) M	SD	N	Group (2) M	SD	N	Diff.	t	P
										%
Organics	Doubtfuls	113·43	50·92	14	76·87	34·15	15	36·56	2·73	< ·5
Unmatched Functionals	Doubtfuls	56·86	26·54	43	76·87	34·15	15	20·01	1·73	<5
Unmatched Functionals	Organics	56·86	26·54	43	113·43	50·92	14	56·57	4·99	< ·05

The Angles Measure was correlated with the rating-scale scores and significant correlations found in all groups except for the Doubtfuls (Table VI).

https://doi.org/10.1192/bjp.103.431.364 Published online by Cambridge University Press

TABLE VI

Product Moment Correlations between Angles Measures and Subjective Rating-scores for the Groups

Group	N	r	P %
Organics	12	·80	< ·1
Matched Functionals	15	·75	< ·1
Unmatched Functionals	43	·41	< 1
All Functionals	58	·58	< ·1
Doubtfuls	15	·44	<10
All cases	85	·63	< ·1

Although we have shown that the Angles Measure differentiates the diagnostic groups at the same level of confidence as the rating-scale scores and correlates significantly with them, we had still to find out whether the subjective rating-scores retained a significant differentiating power independently of the objective Angles Measure. An Analysis of Covariance was carried out, holding constant the amount of each score predictable from the Angles Measure and then comparing 58 Functionals with 12 Organics (measurable on both reproductions). The t ratio equalled 0·035 which is very far from reaching an acceptable level of significance (Table VII).

TABLE VII

Difference between Organics and All Functionals, on the Subjective Rating-scores, holding constant the Amount Predictable from the Angles Measure

	M	SD	N
Organics	−·2650	2·48	12
All Functionals	+·0284	2·67	58
Difference	·2934	t = ·035	70

(iv) *The Emergence of a New Problem*

One point remains to be mentioned concerning Designs A and 4 from which the Diagonals and Angles Measures were derived. The measures on these two designs were combined because they appeared to be alike. However, the ratios of diagonals for Design A correlate only ·167 with the ratios of the diagonals of Design 4 (P>10 per cent., N = 85). Similarly, the deviation from right angles for Design A correlate ·272 with that of Design 4. This correlation, although significant at beyond the 2 per cent. level is still very low.

We have not yet begun an investigation of these low correlations. We can only note that Billingslea (3, pp. 10–11) also found that different measures of what appeared to him to be the same features in copies of the Bender designs failed to correlate. One possible explanation is that these measures are of very small parts of performance and therefore have only low reliability. It is also possible that the Diagonals and Angles Measures are determined differently in each of these designs.

Further investigation shows that the measures on each design make a significant contribution to the variance of the scores which they objectify. Taking the whole sample (N = 86) on Design A and 4 respectively, the ratios of Diagonals correlate ·313 (P<1 per cent.) and ·549 (P<·1 per cent.) with the original subjective scores, and the deviations from right angles correlate

·499 ($P < \cdot 1$ per cent.) and ·562 ($P \cdot < 1$ per cent.). It is clear that each design is contributing something to the subjective scores but doing so for different reasons.

The fact remains, however, that the way in which a subject copies the angles in a diamond-orientated square has little relation with the way in which he copies the angles in an open-ended square which is square-orientated. This finding, if confirmed, might provide a valuable lead for further investigation.

IX. Conclusions

One of the main outcomes of the study is the fact that we did not succeed in directly objectifying the original subjective scoring criteria. There are three possible explanations of this failure. In the first place it might be due to our own personal inability to deduce the correct forms of measurement from the definitions. When we, with our more experienced eyes, now re-examine our deductions we are inclined to think that there is some truth in this explanation.

In the second place the aggregate subjective score is composed of a number of scores corresponding to different parts of the criteria. Each score might make a small and independent contribution to the differentiating power of the aggregate score. This is supported by the fact that each of the 57 measures which we tried out provided means which were in the expected direction.

In the third place the real reasons why a scorer gives a certain rating to a drawing may not be the same as the apparent reason. That there is something in this argument is indicated by our results. For example, the "Unmeasurability" criterion identified only Organics. This implies that brain-damaged patients tend to produce drawings that are obviously distorted or incomplete. The subjective rater would here have an easy source of valid judgments which do not require a set of relatively complex subjective scoring criteria. Similarly, obvious sources of distortion, such as the disproportionality of the diagonals of a simple design like a square, would provide another source of valid judgments which do not require a set of relatively complex subjective scoring criteria. It might be argued that those who insist on using complex subjective criteria may, in fact, be ensuring that their judgments are determined only by the simplest aspects of human behaviour.

Our main interest, however, lies in our other findings which are:

(i) Elderly brain-damaged psychiatric patients copy a square less accurately than do non-brain damaged elderly psychiatric patients.

(ii) This inaccuracy arises primarily in copying the right angles of the square.

(iii) Between at least two types of squares there is little correlation in the extent of the inaccuracy.

(iv) The comparative inaccuracy of brain-damaged patients is little affected by age and intelligence which did not differentiate the groups.

We are now subjecting these findings to a systematic experimental analysis to determine in what ways organic patients find it difficult to draw squares. The outcome of this work should enable us to formulate some precise and testable generalizations about the effects of brain damage in elderly patients.

Summary

1. The subjective rating-scale scores of the Bender Gestalt test-records of a year's intake of psychiatric patients over 60, which differentiated organic, functional and doubtful diagnostic groups at a high level of significance, are analysed.

2. Two objective measures, a Diagonals Measure and an Angles Measure, depending on only 2 of the 6 designs used were found. These differentiated the groups at the same high level of significance as the subjective scores.

3. The Angles Measure is shown to account for the differentiating power of the subjective scores.

4. Age, and Intelligence did not differentiate the diagnostic groups.

ACKNOWLEDGMENTS

Thanks are due to Dr. R. Beech and members of the Psychological Department, and Messrs. A. E. Maxwell and P. Slater of the Statistics Department for discussion and advice.

REFERENCES

1. BENDER, L., *A Visual-Motor Gestalt Test and its Clinical Use*, 1938, Research Monograph No. 3. New York: American Orthopsychiatric Association.
2. BENSBERG, G. J., "Performance of Brain Injured and Familial Mental Defectives on the Bender Gestalt Test", *J. Consult. Psychol.*, 1952, **16**, 61–5.
3. BILLINGSLEA, F. Y., "The Bender Gestalt Test: An Objective Scoring Method and Validation Data", *J. Clin. Psychol.*, 1948, **4**, 1, 1–27.
4. HANVIK, L. J., and ANDERSON, A. L., "The Effect of Focal Brain Lesions on Recall and on the Production of Rotation on the Bender Gestalt Test", *J. Consult. Psychol.*, 1950, **14**, 197–8.
5. INGLIS, J., SHAPIRO, M. B., and POST, F., " 'Memory Function' in Psychiatric Patients over 60. The Role of Memory Tests Discriminating between Functional and Organic Groups", *J. Ment. Sci.*, 1956, **102**, 428, 589–98.
6. KOTTENHOFF, H., "Bender Gestalt Test und die Untersuchung des Cerebalen Funktionswandels", *Psychol. Beitrage*, 1955, **2**, 1, 40–55.
7. MARK, J. C., and MORROW, R. S., "The Use of the Bender Gestalt Test in the Study of Brain Damage", *Amer. Psychol.*, 1955, **10**, 8, 323.
8. ROBINSON, N. M., "Bender Gestalt Performances of Schizophrenics and Paretics", *J. Clin. Psychol.*, 1953, **9**, 3, 219–23.
9. SANDLER, J., "A Test of the Significance of the Difference between the Means of Correlated Measures based on a Simplification of Student's t", *Brit. J. Psychol.*, 1955, **46**, 3, 225–6.
10. SHAPIRO, M. B., POST, F., LOFVING, B., and INGLIS, J., " 'Memory Function' in Psychiatric Patients over 60. Some Methodological and Diagnostic Implications", *J. Ment. Sci.*, 1956, **102**, 427, 233–46.
11. SNEDECOR, G. W., *Statistical Methods*, 1946, 4th ed., 83–4. Iowa Collegiate Press: Ames.
12. WECHSLER, D., *The Measurement of Adult Intelligence*, 1944, 3rd ed. Baltimore: The Williams and Wilkins Co.
13. YATES, A. J., "An Experimental Study of the Block Design Rotation Effect, with Special Reference to Brain Damage", 1955. Unpublished Ph.D. Thesis, University of London.