SIGNIFICANCE OF PSYCHOLOGICAL TEST RESULTS OF EXOGENOUS AND ENDOGENOUS CHILDREN *

By Z. Pauline Hoakley, M.A., and Helen A. Frazeur, M.A.

Wayne County Training School, Northville, Michigan

IN THE field of medicine and psychiatry, disease entities have their syndromes which are followed with sufficient fidelity to make a diagnosis possible in the majority of cases. As clinical psychology matures, the psychologist dreams of the possibility of observing and isolating diagnostic patterns in the mental reactions of his subjects. Evidence of this is seen in recent years in the number of articles in the literature reporting relationships between test results which differentiate between the adjusted and maladjusted, the psychotic and the non-psychotic, the feeble-minded and the normal. Strauss and Werner have made numerous contributions differentiating between the exogenous or brain-injured mentally defective child and the endogenous or familial type with respect to perception, visuo-motor activity (6, 8) and thought organization (7, 9). The object of this study is to determine whether or not differences between the two last named groups are reflected in their performance on Form L of the revised Binet (10).

From the files of the Wayne County Training School were taken every exogenous, white, male moron under 16 years of age for whom there was a like endogenous case, who differed from him by not more than 6 months in chronological and mental age, and by not more than 5 points in I.Q. The diagnosis, in each case, had been made by the neuropsychiatrist according to the following criteria: (5)

For the brain crippled,

(1) Evidence of pre-natal, natal, or post-natal injury to the brain, either of a traumatic nature or due to an inflammatory process;
(2) Presence of neurological signs of brain lesion; and
(3) No evidence of mental deficiency in the members of the immediate family.

For the familial type,

(1) Absence of significant factors in the developmental history, according to the case record;
(2) Presence of mental deficiency among the members of the immediate family; and
(3) Absence of neurological signs.

In the selection of the subjects, all cases of mixed or questionable etiology were excluded, as were all whose condition was complicated by motor disturbances or epidemic post-encephalitis. After these eliminations had been made, 18 pairs remained. A comparison of

* From the Wayne County Training School at Northville, Michigan, Robert H. Haskell, M.D., Medical Superintendent.

the two groups with respect to the range in chronological and mental ages and I.Q. is given in the following table:

TABLE I

	Exogenous	Endogenous
Chronological Age		
Range	9–2 to 15–1	9–4 to 15–1
Q Range	10–5 to 13–7	10–9 to 13–10
Mean	12–7½	12–4
Mental Age		
Range	6–2 to 10–6	6–2 to 10–10
Q Range	6–8 to 9–0	6–8 to 9–6
Mean	7–6	7–9
I.Q.		
Range	50–74	50–75
Q Range	60–69	63–70
Mean	65	67.5

From the scores of the pairs on the various items under consideration, the t was computed to determine the significance of the difference in the means. The level of confidence was established by reference to Fisher's table of t (2).

The first comparison of the items was on the basis of the Terman-Merrill scoring: that is, the number right or the number passed or failed. Aside from an occasional half credit, Terman and Merrill take no account of the degree of badness of the failure or of the relative superiority of the success. It is possible that a more refined scoring might uncover differences which were obscured by this all or none method. Consequently, various criteria were set up for a more refined scoring of those items which were previously rated as plus or minus. On these bases, the diamond, the designs from memory, and the paper cutting tests were scored individually by four judges:* three experienced clinical psychologists and a research fellow, trained in psychology.

The test blanks, opened to the item to be scored and designated only by number, were presented to the judges in random order so there was no clue to the name or the classification of the subject. When the scoring was completed, the ratings of the four judges were compared and, if the value assigned by one judge differed from that of the others by more than one point, re-examination was made in order to be sure that the score represented the judgment of the scorer. If one subject of a pair did not have an item, his mate was also omitted from that comparison.

Criteria for the diamond:

Score 1. A perfect performance.
2. Four good angles. Sides parallel but not necessarily equal. Angles not necessarily opposite. A square.
3. Three good angles with no "ears." Four sides.
4. Only 2 good angles. Two may have "ears." Four sides.
5. Malformed, as triangles, etc. Control attempted.
6. Not to be recognized as an attempt.

If one subject of a pair drew but two diamonds and his mate drew three, the third was not scored. The final score was the sum of the mean scores given by the four judges on each diamond, divided by the number of diamonds considered.

Criteria for design from memory (a):

Score 1. Drawings receiving full credit according to Terman-Merrill.
2. Drawings receiving half credit according to Terman-Merrill.

* The authors wish to express their appreciation to Mrs. Ruth Melcher Patterson, Ph.D., and Miss Laura Lehtinen, M.A., for their assistance in scoring.

3. No credit according to Terman-Merrill except example 26.*
4. Recognized as an attempt. Example 26.
5. Not to be recognized as an attempt.

Criteria for design from memory (b):

Score 1. Drawings receiving full credit according to Terman-Merrill.
2. Drawings receiving half credit according to Terman-Merrill.
3. Lines too far from corners and Terman-Merrill's no credit example 17.**
4. Box in perspective and the remainder of Terman-Merrill's no credit.
5. Not to be recognized as an attempt.

The final score in each of the drawings from memory was the mean of the scores given by the four judges.

The refined scoring for the paper cutting tests differed from the above pattern in that the score on each was cumulative. The maximum score was 8 and 11 for the first and second designs respectively. One point was given for each of the following:

(a) Subject uses printed frame as the paper.
(b) Subject uses printed frame as open sheet of paper. (If seen as open for one portrayed element, give credit, although seen as folded for the second.)
(c) For each element portrayed.
(d) For each element correctly portrayed as open.
(e) For each element correctly portrayed and in the correct position.
(f) Subtract one point for the portrayal of additional elements.

Of the above tests, all of which may be called perceptual and visuo-motor in type, only the diamond, when scored according to Terman-Merrill, resulted in a significant score. In this, t was —3.06, significant below the 1 per cent level of confidence. But this, in itself, cannot be used as a diagnostic tool since, of the fifteen pairs having the test, nine made equal scores.

Comparisons were made with respect to the verbal items: vocabulary, abstract words, absurdities and similarities. The first three were scored only according to Terman-Merrill. In all, save the vocabulary, no word or situation was used which had not been presented to both subjects of the pair. In each case the score was the number right. For the comparison of our groups with similarities, subtests VII, 2 and XI, 6 were combined, making one score.

The qualitative scores for the similarities were taken from the Martinson-Strauss evaluation of responses to the 1916 Binet (4), where the item was common to both studies. For the remainder of the items, VII, 2 c and d, and XI, 6 a, d and e, letters were assigned to responses in accord with those of the Martinson-Strauss standards. That is, "a_1" was given to the response which compared the named objects by classification as, for example, wood and coal are fuel; "a," to comparisons in terms of use as, they both burn; "b," to inferior but correct responses as, both are hard; "c," to nonsensical answers; "d," to answers which contradict the general or specific facts involved in the problem; "g," to answers based on a misunderstanding of the question; and "i" to the response that the subject does not know.

Martinson and Strauss did not attempt to determine the degree of superiority of one response over another but only

* Terman and Merrill, Measuring Intelligence, p. 251.
** Terman and Merrill, *ibid.*, p. 253.

that "a_1" was superior to "a"; "a" to "b"; and "b" to "d" and "g." "C" was the poorest of all, and "i" was superior to "d," "g" and "c," in that it was more often given by normal children.

In order to submit our scores to statistical analysis, it was necessary to ascribe a numerical value to these letters. Following the order of merit indicated by Martinson and Strauss, we arbitrarily assigned the following credits: $a_1=4$; $a=3$; $b=2$; $i=1$; d and $g=0$; $c=-.1$. The final rating was the mean score for the number of items used. Only those were included which had been presented to both members of the pair.

No significant difference was found in the verbal items.

Contact with brain-crippled mentally defective children leads to the conclusion that they have a short attention span. In the Binet, digital memory is probably the simplest measure of the ability to attend and to recall. Consequently, we have made comparisons of the performance of our two groups on digits in both normal and reversed order.

The scoring for all comparisons with degits was according to Terman-Merrill. In Plan A, in both forward and reversed series, ability to repeat digits falling below the basal was assumed. In Plan B, no ability was assumed that was not demonstrated. This materially reduced the size of our groups since we do not routinely give digits below the basal age. There were eight pairs who established their lower limit with digits in normal order, and three with re-

TABLE II

BINET ITEMS WITH TERMAN-MERRILL AND REFINED SCORING. RESULTS OF GROUP COMPARISONS

Items	Scoring Method	No. of Pairs	Level of Confidence	Inferior Group
Diamonds	T.M.	15	Below 1%	Exog.
"	Refined	15	5–10%	Exog.
Designs from Memory:				
Design A	T.M.	17	2–5%	Exog.
"	Refined	17	10–20%	Exog.
Design B	T.M.	17	5–10%	Exog.
"	Refined	17	5–10%	Exog.
Paper Cutting:				
Form A	T.M.	17	70–80%	Endog.
"	Refined	17	80–90%	Exog.
Paper Cutting:				
Form B	T.M.	17	**	
"	Refined	17	60–70%	Exog.
Vocabulary	T.M.	18	70–80%	Endog.
Abstract Words	T.M.	11	20–30%	Exog.
Similarities: Two and Three Things	T.M.	17	10–20%	Exog.
" " " "	Refined	17	30–40%	Exog.
Verbal Absurdities	T.M.	18	40–50%	Endog.
Digits Forward	Plan A	18	30–40%	Endog.
" "	Plan B	8	*	
Digits Reversed	Plan A	18	2–5%	Exog.
" "	Plan B	3	*	

** All but 1 child (endog.) failed.
* Mean difference: zero.

versed. In both comparisons of Plan B the mean difference in score was zero.

Plan A, with reversed digits, was the only item of this group to give a score approaching significance. That this is meaningless becomes apparent when we consider that the rating for one or both members of all pairs, save the three named above, was based on the assumption that the child could reverse three digits.

Inspection of Table II reveals the fact

TABLE III

Binet Items with Terman-Merrill and Refined Scoring. Results of Sub-Group Comparisons

Items	Group	Scoring Method	No. of Pairs	Level of Confidence	Inferior Group
Diamonds	M.A. 6–0 to 8–0	T.M.	10	5–10%	Exog.
"	M.A. 8–1 to 10–10	"	5	5–10%	Exog.
"	M.A. 6–0 to 8–0	Refined	10	20–30%	Exog.
"	M.A. 8–1 to 10–10	"	5	5–10%	Exog.
"	I.Q. 50–65	"	9	1–2%	Exog.
"	I.Q. 65–75	"	6	40–50%	Exog.
Designs from Memory:					
Design A	M.A. 9–0 to 10–10	T.M.	5	5–10%	Exog.
"	I.Q. 50–65	"	8	40–50%	Exog.
"	I.Q. 65–75	"	9	5–10%	Exog.
"	M.A. 9–0 to 10–10	Refined	5	20%	Exog.
"	I.Q. 50–65	"	8	50–60%	Exog.
"	I.Q. 65–75	"	9	20–30%	Exog.
Design B	M.A. 9–0 to 10–10	T.M.	5	10–20%	Exog.
"	I.Q. 50–65	"	8	50–60%	Endog.
"	I.Q. 65–75	"	9	Below 1%	Exog.
"	M.A. 9–0 to 10–10	Refined	5	1–2%	Exog.
"	I.Q. 50–65	"	8	50–60%	Endog.
"	I.Q. 65–75	"	9	Below 1%	Exog.
Paper Cutting:					
Form A	M.A. 9–0 to 10–10	T.M.	5	50–60%	Exog.
"	I.Q. 65–75	"	9	*	
"	M.A. 9–0 to 10–10	Refined	5	90–100%	Endog.
"	I.Q. 65–75	"	9	70–80%	Exog.
Form B	M.A. 9–0 to 10–10	"	5	70–80%	Endog.
"	I.Q. 50–65	"	8	*	
"	I.Q. 65–75	"	9	60–70%	Exog.
Vocabulary	I.Q. 50–65	T.M.	9	*	
"	I.Q. 65–75	"	9	50–60%	Endog.
Abstract Words	M.A. 9–0 to 10–10	"	4	20–30%	Exog.
" "	I.Q. 65–75	"	6	40–50%	Exog.
Similarities: Two and Three Things	M.A. 7–0 to 10–10	"	10	30–40%	Exog.
" "	I.Q. 65–75	"	8	50–60%	Exog.
" "	M.A. 7–0 to 10–10	Refined	10	30–40%	Exog.
" "	I.Q. 50–65	"	9	30–40%	Exog.
" "	I.Q. 65–75	"	8	50–60%	Exog.
Verbal Absurdities	M.A. 8–0 to 10–10	T.M.	7	70–80%	Endog.
" "	I.Q. 50–65	"	9	60–70%	Endog.
" "	I.Q. 65–75	"	9	40–50%	Endog.
Digits Forward	I.Q. 50–65	Plan A	9	70–80%	Exog.
" "	I.Q. 65–75	"	9	2–5%	Endog.
Digits Reversed	M.A. 9–0 to 10–10	"	5	60–70%	Exog.
" "	I.Q. 50–65	"	9	30–40%	Exog.
" "	I.Q. 65–75	"	9	5–10%	Exog.

* Mean difference: zero.

that only the diamond, when scored according to Terman-Merrill, gave a level of confidence below 2 per cent. It seemed possible that the significance of some items might be obscured by the diversity of mental age and I.Q., particularly since some of the items under consideration occur well above the mean mental age of the groups. Consequently, it appeared advisable to restrict the group by taking, as the lower limit, the mental age at which the item occurs since those of lesser mental age could not be expected to perform well, irrespective of their classification. Also, a narrowed I.Q. range might yield significant findings. For results of these comparisons, see Table III.

The only addition to the list of significant differences is memory design (b), drawn by subjects with I.Q.'s 65 to 75, regardless of the method of scoring.

Strauss and Werner (6), in their article concerning the mental organization of the brain-injured mentally deficient child, point to a striking difference in the manner of performance between endogenous and exogenous groups with the marble board, although, objectively, the two may reach results equally correct. Can this be shown with the items of the Binet?

Two of the outstanding errors in drawing diamonds are the inverted angles and the angles which are determined, not by the natural meeting of the sides but by a turn in the line resulting in a loop or an "ear." Exogenous children have been observed to produce this kind of drawing. Reviewing the diamonds produced by the subjects of this study, we found that 21 "ears" or inverted angles had been drawn by the brain-crippled as compared with 7 by the familial. The difference is not significant and dividing the group into more homogeneous mental age and I.Q. subgroups does not increase the reliability of the difference. (See Table IV.)

The drawings of the paper cutting were scrutinized for the portrayal of additional elements. But 2 of the 30 subjects made additions. Both were exogenous. One, with a mental age of 6 years 9 months, had drawn 10 holes but no creases, apparently unable to inhibit his tendency to perseverate.

With the memory designs, comparisons were made between the whole groups and between the subgroups to determine whether one type of child drew too poorly for recognition more often than the other. No significant difference was found for any group with either design.

Two comparisons, each using the whole group, were made of the quality of the associations by similarity: one, to test if one group was superior to the other in the number of "a" and "a_1" responses; and the second, to discover whether one group, more than the other, reported that he did not know instead of making erroneous answers. Neither proved significant.

The results of the qualitative differances are summarized in Table IV.

Our final comparison has to do with the range of the Binet. Range is defined as the number of years lying between the first basal and the last zero age. A basal is that year level at which all six sub-tests are passed; a zero, where all six are failed. Our testing procedure is to establish a basal age below the earliest failure and, at the upper end,

TABLE IV

Results of Group Comparisons with Respect to Quality of Responses

Item	Group	No. of Pairs	Level of Confidence	Inferior Group
Angles of Diamonds				
Inverted and "ears"	Whole	15	10–20%	Exog.
" " "	M.A. 7–0 to 10–10	8	10–20%	Exog.
" " "	I.Q. 50–65	9	30–40%	Exog.
" " "	I.Q. 65–75	6	40–50%	Exog.
Paper Cutting, a and b Additional Elements	Whole	17	10–20%	Exog.
Designs from Memory				
Design A				
Attempt unrecognizable	Whole	17	30–40%	Endog.
" "	M.A. 6–0 to 9–0	12	10–20%	Endog.
" "	M.A. 9–1 to 10–10	5	60–70%	Exog.
Design B				
Attempt unrecognizable	Whole	16	•	
" "	M.A. 9–1 to 10–10	5	10–20%	Endog.
Similarities: Two and Three Things				
"a" and "a_1"	Whole	17	70–80%	Exog.
"d" and "g" versus "i"	Whole	14	20–30%	Exog.

• Mean difference: Zero.

two consecutive zero ages unless the first zero is at or past the subject's life age. As an example, if a basal age has been found at year level IX but the child has failed to meet the VIII year vocabulary requirement, a second basal is established below the VIII year level. If the subject passes no test at year XII but has already passed the paper cutting test at XIII, the test is continued through XIV and the average adult level unless the subject is in or below his fourteenth year.

Contrary to our expectation, our endogenous children more frequently had the longer range and the multiple zero age. Neither comparison resulted in a significant difference, t falling between the 30 and 40 per cent levels of confidence in both cases. Only one child, exogenous in type, had a multiple basal.

Discussion

In this study we have attempted to determine whether brain-crippled mentally defective male children respond to items of the Binet in a manner, or with a degree of accuracy, different from male children whose deficiency is of familial origin when they are closely matched in a one to one relationship with respect to chronological age, mental age, and I.Q. Furthermore, by subjecting the scores of the two groups to statistical measures adapted to small samples, we have sought to learn whether or not observed differences were sufficiently reliable to establish the given subtest as a psychological tool for differentiating between the two groups of children.

We have found that the only statistically reliable differences occurred in the perceptual or visuo-motor field. That

the exogenous child is handicapped in this area has been shown by the studies of Strauss and Werner, and by Lord and Wood (3). Possibly we should not expect greater differences when we consider the number of factors operating against the possibility of their existence.

In the first place, our groups are very small. Significant differences might be uncovered if they were sufficiently large to allow for subdividing on the basis of quality. That is, there is a possibility that, for a given mental age or I.Q., there are degrees of poorness which the endogenous would not show. Also, greater differences might be found if the groups were more homogenous. We know there are differences in type of brain damage: in areas having received injury, in the extent of the injury, and in the degree of restitution. Neither can we be certain that our familial group is unitary in character. Consequently, until we have a sufficiently large number of cases to divide into still sizable, homogenous groups, and until our knowledge is sufficient to allow us to isolate types with greater strictness, we probably should expect different reactions within groups and exceptions to established trends.

Again, our negative results may have their origin in the test as well as in the group submitted to it. It may be that the Binet technique is too controlled to catch, in the mentally defective child, the nuances, the disinhibitions, and the flickerings of normalcy which, in a clinical situation, lead us to suspect the presence of brain damage. This possibility is strengthened by a comparison of our findings with those of Doll, Phelps and Melcher (1) who, in 1932, attempted to discover "some general principles concerning the psychology of birth injured feebleminded children" * from the results on the 1916 Binet.

Their subjects, 11 in number, showed a great disparity in chronological age, mental age, and location of lesion. Also, they were of both sex. The one factor common to all, save one, was gross motor disability, a factor which set them off sharply from the subjects of this study. In spite of these differences, and in spite of the smallness of both study groups, the trends brought out in the two studies are in the same direction in five of the eight subtests common to both. In only two instances could we say the results wholly disagreed. However, in the light of our statistical findings, we should not read into this agreement more than is there. It is true that this study has revealed some interesting trends or tendencies, and it is true that some of our results are in agreement with those of another study made in a wholly different manner and with subjects whose only consistent likeness to ours was brain damage. But, it is also true that our statistical procedure has not revealed any significant number of statistically reliable differences.

This study is by no means conclusive and, as we have just said, the results are preponderantly negative. We have presented it as a warning to psychologists to use diagnostic patterns with extreme caution until they are thoroughly authenticated, and even then to keep

* Doll, Phelps, and Melcher. Mental Deficiency Due to Birth Injuries. Page 103.

in mind that the case under consideration may be the exception to the rule.

BIBLIOGRAPHY

1. DOLL, PHELPS, AND MELCHER. Mental Deficiency Due to Birth Injury. New York, N. Y.: The Macmillan Co., 1932.

2. LINDQUIST, E. F. Statistical Analysis in Educational Research. Boston, Mass.: Houghton Mifflin Co., 1940.

3. LORD, E., AND WOOD, L. Diagnostic Values in a Visuo-Motor Test. *Amer. J. Orthopsychiat.*, 1942, 12, 414–428.

4. MARTINSON, B., AND STRAUSS, A. A. A Method of Clinical Evaluation of the Responses to the Stanford-Binet Intelligence Test. *Amer. J. Ment. Def.*, 1941, 46, 48–59.

5. STRAUSS, A. A. Typology in Mental Deficiency. *Amer. J. Ment. Def.*, 1939, 44, 85–90.

6. STRAUSS, A. A., AND WERNER, H. The Mental Organization of the Brain-Injured Mentally Defective Child. *Amer. J. Psychiat.*, 1941, 97, 1194–1203.

7. STRAUSS, A. A., AND WERNER, H. Disorders of Conceptual Thinking in the Brain-Injured Child. *J. Nerv. Ment. Dis.*, 1942, 96, 153–172.

8. STRAUSS, A. A. Diagnosis and Education of the Cripple-Brained, Deficient Child. *J. Except. Child.*, 1943, 9, 163–168.

9. STRAUSS, A. A. Ways of Thinking in Brain-Crippled Children. *Amer. J. Psychiat.*, 1944, 100, 639–647.

10. TERMAN, L. M., AND MERRILL, M. A. Measuring Intelligence. Boston, Mass.: Houghton Mifflin Co., 1937.

70TH ANNUAL MEETING AT

HOTEL MT. ROYAL, MONTREAL, CANADA

OCTOBER 2ND, 3RD, 4TH AND 5TH, 1946