

Eastern Illinois University
The Keep

Masters Theses

Student Theses & Publications

1982

Validation of the Tactual Performance Test as an Organicity Screening Device

Stephen C. Lippold

Eastern Illinois University

This research is a product of the graduate program in [Psychology](#) at Eastern Illinois University. [Find out more](#) about the program.

Recommended Citation

Lippold, Stephen C., "Validation of the Tactual Performance Test as an Organicity Screening Device" (1982). *Masters Theses*. 2954.
<https://thekeep.eiu.edu/theses/2954>

This is brought to you for free and open access by the Student Theses & Publications at The Keep. It has been accepted for inclusion in Masters Theses by an authorized administrator of The Keep. For more information, please contact tabruns@eiu.edu.

THESIS REPRODUCTION CERTIFICATE

TO: Graduate Degree Candidates who have written formal theses.

SUBJECT: Permission to reproduce theses.

The University Library is receiving a number of requests from other institutions asking permission to reproduce dissertations for inclusion in their library holdings. Although no copyright laws are involved, we feel that professional courtesy demands that permission be obtained from the author before we allow theses to be copied.

Please sign one of the following statements:

Booth Library of Eastern Illinois University has my permission to lend my thesis to a reputable college or university for the purpose of copying it for inclusion in that institution's library or research holdings.

July 30, 1982

Date

Author

I respectfully request Booth Library of Eastern Illinois University not allow my thesis be reproduced because _____

Date

Author

REYNOLDS BINDERY, INC. - 1703 LISTER - KANSAS CITY, MO. 64112

VALIDATION OF THE TACTUAL PERFORMANCE TEST AS AN

ORGANICITY SCREENING DEVICE

(TITLE)

BY

STEPHEN C. LIPPOLD

B. S. IN PSYCHOLOGY

PURDUE UNIVERSITY, 1972

THESIS

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF

MASTER OF ARTS

IN THE GRADUATE SCHOOL, EASTERN ILLINOIS UNIVERSITY
CHARLESTON, ILLINOIS

1982

YEAR

I HEREBY RECOMMEND THIS THESIS BE ACCEPTED AS FULFILLING
THIS PART OF THE GRADUATE DEGREE CITED ABOVE

7/21/82

DATE

ADVISER

7/21/82

DATE

COMMITTEE MEMBER

7/21/82

DATE

COMMITTEE MEMBER

7/21/82

DATE

DEPARTMENT CHAIRPERSON

Abstract

Organicity screening devices are those psychometric tests used to detect brain dysfunction during the initial assessment. The characteristics of such tests were listed and ones currently being used were reviewed. The review concentrated on the validity and clinical utility of these organicity tests. From the review it was concluded that the validity has not been thoroughly established, partially because weak criteria have been used. It was also concluded that current screening tests have questionable clinical utility, since an inordinant number of organically impaired people are not detected by the tests. The concept of organicity was discussed, and the Tactual Performance Test was proposed as a new organicity screening test. A validation and cross-validation experiment was done by comparing the Tactual Performance Test with the Halstead-Reitan Battery, a more appropriate criterion. This study investigated the validity and utility of the test.

Data from the Tactual Performance Test and the Halstead-Reitan Battery were collected for 200 male veterans; these subjects were randomly selected from the neuropsychology archives at a VA Medical Center. Two groups of 100 subjects each served as the Validation and Cross-validation samples. The groups did not differ significantly in age or education; the mean age was 44.24 years and the mean education level was 10.97 years. To compare the Tactual Performance Test with the Halstead-Reitan Battery, the three test scores, age and education were chosen as independent variables, and the Average Impairment Rating from the battery was chosen as the dependent variable. For the Validation group the

the variables were used to create a multiple regression equation which predicted the dependent variable. The independent variables for the Cross-validation group were entered into this equation to check its ability to predict the dependent variable. The validity of the Tactual Performance Test was judged by computing the multiple regression coefficient and the Pearson r for the two groups. The clinical utility of the test was judged by the overall accuracy and the number of Type I and Type II errors.

Of the five independent variables, age did not make a significant contribution to the predicted Average Impairment Rating; this variable was found to be a correlate of every other variable. The resultant equation was found very capable of predicting the dependent variable. The correlations for the Validation and Cross-validation groups were .85 and .86 respectively, each significant at the .001 level. The utility of the Tactual Performance Test and education was also impressive; there were 87% correct predictions for the Validation group and 86% correct for the Cross-validation group. In addition very few organically impaired individuals escaped detection by this method. Therefore the Tactual Performance Test is judged a valid and useful organicity screening device. Shortcomings of this study were discussed and a further study incorporating females was proposed.

Acknowledgements

Three groups of three individuals each have contributed amply to my well-being and knowledge during this research. I thank the members of my committee, Drs. Roberts, Hustmyer and Panek for their sincerity and guidance, as well as their friendship throughout this ordeal. They have created an interesting ambience within which to learn the byways of research. Drs. Roberts and Hustmyer read and critiqued the rough drafts and ensured the major errors were detected; those that remain I willingly own.

I thank also Drs. Claiborn, Rosenberg and Dwarshuis, colleagues at the VA Medical Center, Danville, Illinois, for the continued encouragement and empathic sympathy they have provided over the past three years.

Finally I thank Pat and our two sons who have stood by and tolerated my frequent absences into this mysterious land of "research".

Table of Contents

	Page
Abstract	i
Acknowledgements	iii
Table of Contents	iv
List of Tables	v
Introduction	1
Characteristics of Organicity Screening Devices	2
Current Status of Organicity Screening Devices	4
The Conceptualization of Organicity	16
Focus of This Research	19
Methods	23
Subjects	23
Procedure	24
Results	26
Multiple Regression Equation	26
Clinical Utility	26
Discussion	31
References	35

List of Tables

	Page
Table 1. Results of Some Validation Studies of Selected Screening Tests for Organicity	8
Table 2. Summary Table For Stepwise Multiple Regression Equation Created From Validation Group Data	27
Table 3. Means and Standard Deviations of the Experimental Variables for the Validation and Cross-validation Samples	27
Table 4. Results of Investigation of Tactual Performance Test as an Organicity Screening Device	29

Introduction

Within the realm of daily duties of the clinical psychologist, one of the most frequent is that of doing initial assessments on new clients. Many times the troubled client arrives at the mental health center or admission ward presenting a confusing array of complaints, seeking help. The clinician doing the assessment will note the complaints and pertinent behaviors in an attempt to arrive at a provisional diagnosis. This preliminary diagnosis will contain implications of the etiology or cause of the client's distress, and lead to recommendations for further assessment and treatment procedures.

Some of the confusion encountered in the initial interview arises from the often bewildering and etiologically indeterminate symptoms the client presents. That is, the complaints may represent a psychiatric disturbance or they may reflect an underlying abnormality of the central nervous system. In addition, the clinician's recommendations for further treatment can have drastic repercussions for the client, such as costly medical tests if an organic condition is suspected, or lengthy hospitalization and stigmatizing psychotherapy or chemotherapy if a functional disorder is suspected. Therefore the initial assessment should be as thorough as possible and the diagnostic impression should not be made lightly.

Most psychologists do not possess the requisite neurological acumen with which to make the sometimes subtle distinction between a functional versus an organic etiology based solely on the symptom complex presented by the client. Rather, the clinician will employ one of several ob-

jective psychometric tests available which are purportedly capable of differentiating between intact and impaired brain function. These tests, called organicity screening devices, are very much in demand and enjoy wide popularity among psychologists.

Recently Bigler and Ehrfurth (1981) reviewed one such screening test, the Bender-Gestalt, and concluded that the test was inappropriate for this purpose. They cited as evidence the alarming number of brain-damaged individuals who escaped detection by this particular test and also suggested that other psychometric instruments are equally poor as organicity screening devices. This suggestion will be investigated in the following paper.

This paper will review several currently popular organicity screening devices. It will begin by enumerating the desirable properties of such instruments, and examine articles that have investigated the validity and clinical utility of these tests. We will then examine the underlying conceptualization of organicity which existed when these measures were proposed as screening tests and note changes in this construct over the past few decades. Finally we will propose a new organicity screening device and investigate its validity and clinical utility.

Characteristics of Organicity Screening Devices

To be of service to clinicians in detecting brain damage an organicity screening device must meet several requirements. First, any apparatus used in the test should be relatively simple and inexpensive to ensure the availability of the test to a wide variety of settings. Second, the test should be objective, yielding quantitative scores and having standardized administration and scoring procedures. In this way the re-

sults obtained on a given client can be directly compared to empirically derived norms, which also should be available. Third, the test should use up little time so that it may be of use to clinicians performing interviews in time-pressured settings. Finally, the instrument must be valid for the purpose of detecting or predicting brain damage in individuals.

This last point needs some clarification. The validity of a test as an adequate measure of brain dysfunction must be empirically demonstrated by comparing the performance of the experimental measure with that of a criterion known to be capable of detecting brain damage. The selection of an appropriate criterion is an important part of any validation study. There are several criteria which can be used in the study of brain damage, and these can be classified into two categories, internal and external (Yates, 1954). Since the tests being considered here are psychometric measures, an internal criterion would be another psychometric instrument which has been substantially validated for the purpose of detecting brain damage. An example of this is the Halstead-Reitan Battery. An external criterion would be other than a psychometric test such as psychiatric diagnosis based upon anamnestic data, or medical diagnosis based upon data from neurological procedures such as the Brain Scan or Angiogram. For purposes of test validation the use of a psychometric criterion would provide a more stringent proof of the validity of the experimental psychometric test. However, the most accurate criteria, external or internal, should be employed in the investigation of a test's validity. In the case of organicity there has been evidence that the Halstead-Reitan Battery is more accurate in detecting brain

damage than any of the currently available neurological procedures (Filskov & Goldstein, 1974).

Current Status of Organicity Screening Devices

At the present time several psychological tests are being used to screen for the presence of brain damage as part of the initial assessment process. Of this number five tests were selected for review; these five are considered representative and were chosen either because they are widely used or were designed specifically to predict organicity. The five tests are the Bender-Gestalt Test (BG), the Graham-Kendall Memory-For-Designs Test (MFD), the Minnesota Percepto-Diagnostic Test (MPD), the Hooper Visual Organization Test (VOT), and the Benton Visual Retention Test (BVRT).

These tests share several characteristics. They are all brief and consume little of the client's or clinician's time. They have been standardized and have objective scoring systems, although in some cases the clinician uses subjective evaluation in deciding whether the client's performance is indicative of brain damage or not. The scoring systems employ a cut-off score as the decision maker; that is, scores above or below a certain level are indicative of brain dysfunction, thus utilizing an objective interpretation of an individual's performance. Finally, these tests are routinely used by psychologists both as criterion measures in experiments investigating organicity and as screening tests for organicity.

However, these psychodiagnostic procedures share two characteristics which militate against their use when determinations of brain dysfunction are to be made. First, the tests have been proven to be rela-

tively ineffective in the few studies which have dealt specifically with their validity as measures sensitive to the presence of brain dysfunction. It would seem that psychologists presume the validity of the devices as measures of organicity (Tolor & Schulberg, 1963), and use them armed with this presumption. Second, validation studies report a substantial number of false positive and false negative errors, calling into question their clinical utility. The number of errors in classification is an important consideration, since the treatment and further diagnostic procedures that follow often hinge upon the initial assessment of the organicity screening device. Many misclassifications can result in much lost time and unnecessary hospital bills for the clients being assessed.

Before examining the validation studies reported on these tests a word must be said about construct validity. To accurately assess the validity of a test as a predictor of some construct, it is necessary to secure an external measure which is itself known to be a valid predictor of that construct, so that the experimental and control subjects can be differentiated. The construct dealt with here is brain damage, or organicity. As yet there are no assessment procedures in medicine or psychology which are perfectly capable of detecting brain damage in the individual, and which could serve as the criterion measure. The neurodiagnostic tests currently available are exceptionally capable of detecting specific types of lesions (e.g., the Brain Scan and Angiogram can identify most vascular diseases and lesions; the Pneumoencephalogram and electroencephalogram can detect most space-occupying lesions) but none of these tests is able to detect all lesion types. In a landmark study

Filskov and Goldstein (1974) evaluated several neurodiagnostics to determine how effective they were individually in detecting brain damage. The criterion for each subject was the final diagnosis which was made based on the results of all medical and neuropsychological tests any particular individual was given. The accuracy of the neurodiagnostics ranged from 16% for the Skull X-Ray to 85% for Angiograms and 80% for Pneumoencephalograms. To be kept in mind is the fact that the latter two procedures carry morbidity/mortality rates of 4% and .25% respectively (Tavaras & Wood, 1964). The recent introduction of the CT Scan has helped to replace some of the riskier procedures with a safe test of comparable efficacy (Tsushima & Popper, 1980). The relative effectiveness of these tests should be borne in mind when viewing organicity studies which employ them as the external criterion.

Another method of classifying subjects as brain damaged or not is by diagnostic history. In this process a medical or psychiatric diagnosis is given to the patient based on information provided by the patient or significant others. Gross lesions of the brain may be easily detected, but discrete lesions or congenital anomalies may escape detection, even by a neurologist. In addition the validity of psychiatric diagnoses hinges on the demonstration of the reliability of the diagnoses. This means that agreement of diagnosis by several physicians or by the same physician at different times should be obtained experimentally. To date the reliability of psychiatric diagnostic systems has not been proven to be high (Zubin, 1967; Spitzer & Fleiss, 1974). Therefore differentiation of organicity by diagnosis alone is not considered to be an adequate criterion.

Presented in Table 1 are results of validation studies that investigated the ability of the BG, MFD, MPD, VOT and BVRT to predict the presence of brain damage in various populations. For each study the size of the organic and control groups are given; in some cases two control groups (psychiatric and "normals") were collapsed into one, since the purpose of these studies was to differentiate brain damage from non-brain damage. The external criteria employed in these studies were either psychiatric or medical diagnoses. For each experiment the reported discriminative efficiency of the test is given. This number is the overall rate of correct predictions, and is interpreted as a rough measure of the test's ability to discriminate between brain damaged and non-brain damaged subjects. The discriminative efficiency is not a good measure of the utility of a test, however, because it can be influenced by the base rate of brain damage in the sample. And, as will be seen, the overall prediction rate has a tendency to mask the number of false negative classifications.

For purposes of comparing the different validation studies more efficiently, three further statistics were calculated from the preceding information. The Chi Square with Yates' correction (McNemar, 1949) was computed to determine whether the dichotomous predictions were significantly different from that expected by base rate classification. In the second study by Hain (1964) and three other experiments (Holland & Wadsworth, 1974; Holland, Wadsworth & Royer, 1975; McManis, 1974) the sample sizes were too small to approximate the statistical significance by Chi Square; therefore Fisher's exact test (Hays, 1963) was computed on these data. To examine the strength of association between predicted

Table 1

Results of Some Validation Studies of Selected Screening Tests for Organicity

Study	Organics _a	Controls _a	Overall Accuracy	False Neg _a	False Pos _a	Chi Square	Phi	Incremental Validity
Bender-Gestalt Test								
Adams, 1971	22	22	52	15	6	0	0	.02
	22	22	75	7	4	9.26*	.46	.25
	22	22	84	3	4	17.86*	.64	.34
Brilliant & Gynther, 1963	34	76	82	11	6	40.18*	.60	.26
Hain, 1964	20	63	89	7	2	35.13*	.65	.23
	21	21	62	10	6	1.58	.19	.12
Johnson, et al., 1971	120	120	64	52	35	17.42*	.27	.14
Korman & Blumberg, 1963	40	40	74	9	12	16.29*	.45	.24
Kramer & Fenwick, 1966	18	24	76	10	0	11.71*	.53	.22
Mosher & Smith, 1965	140	119	59	99	7	20.79*	.28	.12
	140	119	43	108	8	11.62*	.21	.08
Pardue, 1975	20	20	63	7	15	3.02	.22	.12
	20	20	88	5	2	28.82*	.69	.31

Table 1 (cont)

Study	Organics _a	Controls _a	Overall Accuracy	False Neg _a	False Pos _a	Chi Square	Phi	Incremental Validity
Graham-Kendall Memory For Designs								
Brilliant & Gynther, 1963	34	76	78	13	9	27.05*	.50	.21
Korman & Blumberg, 1963	40	40	82	9	5	31.56*	.63	.33
Grundvig, et al., 1973	50	50	66	18	16	9.01*	.30	.16
	50	50	74	17	9	21.72*	.47	.24
McManis, 1974	20	20	68	8	5	3.68	.30	.18
Minnesota Percepto-Diagnostic Test								
Crookes & Coleman, 1973	44	137	80	26	11	24.37*	.37	.12
Fuller & Laird, 1963	52	234	85	9	33	95.03*	.58	.20
	100	440	85	18	62	180.49*	.58	.21
Holland & Wadsworth, 1974	20	20	72	6	5	6.42 _b	.40	.22
Holland, et al., 1975	20	20	75	6	4	6.46 _b	.40	.25
Uyeno, 1963	52	52	83	12	6	42.45*	.64	.33

Table 1 (cont)

Study	Organics _a	Controls _a	Overall Accuracy	False Neg _a	False Pos _a	Chi Square	Phi	Incremental Validity
Benton Visual Retention Test								
Brilliant & Gynther, 1963	34	76	81	13	2	46.16*	.65	.25
L'Abate, et al., 1963	30	36	72	9	9	11.61*	.42	.22
Watson, 1973	23	394	87	14	39	15.47*	.19	.03
Hooper Visual Organization Test								
Boyd, 1981	40	40	74	13	8	16.46*	.45	.24
Love, 1970	29	86	74	8	21	19.53*	.41	.18

a Number of subjects

b $p < .01$ by Fisher's exact test

* $p < .01$

outcomes and criterion outcomes the phi coefficient (Dinham, 1976) was computed. This coefficient is a measure of the relative strength of the test as a predictor of brain damage. Finally the incremental validity (Wiggins, 1973) of the test was computed. The incremental validity in each study is a measure of the benefit to be realized, in terms of increased predictive efficiency, by use of the test as compared to simple assignation of subjects to groups based on the known base rate and selection rate. The higher this incremental validity, the more worthwhile a test is as a predictor of brain damage, both for the clinician interested in a valid test and for the client interested in obtaining an accurate diagnosis.

Table 1 shows that for the Bender-Gestalt validation studies the authors found overall prediction rates ranging from 43% (Mosher & Smith, 1965) to 89% (Hain, 1964). Admittedly different scoring systems were used in the various experiments, but no one system has been proven to be more accurate than the others. For example Hain (1964) devised his own scoring procedure of the BG and, using a cut-off score which minimizes classificatory errors, found that he had correctly identified 74 of 83 subjects. In a cross-validation study, however, this system was found to be less accurate as only 26 of 42 subjects were correctly identified. Mosher and Smith (1965) used Hain's scoring method in their second analysis and reported that this system misclassified 57% of all subjects.

All but three of the experiments found that the BG was able to significantly improve classification of organic and control subjects over that expected from random assignment to groups based on base rates alone. As mentioned above, however, the overall accuracy of prediction and the

significant improvement over chance distribution of subjects tend to hide one of the shortcomings of this test. In most of the studies the rate of false negatives, those brain-damaged subjects misclassified as non-organic by the test, far exceeds the rate of false positives. In other words the BG is failing to do what it purportedly is intended to do, identify brain-damaged individuals.

The strength of association between test decisions and the external criteria, as measured by the phi coefficient, varies markedly over the experiments. Interestingly the strength of this correlation tends to decrease as the sample size increases, suggesting that generalization to a wider population should be done with caution. The last column in Table 1 shows the benefit to be realized by utilizing the test instead of random assignment to groups according to base rate. It can be seen that this incremental validity of the BG ranges considerably over studies. Also, the size of this increase in validity varies inversely with the size of the sample.

In summary the gain to be realized in terms of correct predictions or identification of brain dysfunction in varying populations appears to outweigh the amount of time required to administer and score the Bender-Gestalt Test. No one scoring system has been found superior in this endeavor. However the number of brain-damaged subjects who escaped detection by the test leads to the conclusion that the BG is not extremely sensitive to organicity. It would be hoped that very few individuals with brain damage would be miss with a test used to screen for organicity. Ideally the rate of false negatives would be less than that of false positives, as it is to be expected that some organically impaired per-

sons would not be identified by neurodiagnostic means, but could be found out by sensitive psychometric tests.

Looking next to the Graham-Kendall Memory-For-Designs Test, a similar picture develops. The reported overall accuracy rates in the five experiments shown in Table 1 range from 66% to 82%. In four of the studies the MFD was found to give a significant improvement over chance in the differentiation of brain-damaged and non-brain-damaged subjects. However, as seen with the BG, the overall correct prediction rate tends to hide the fact that there is a much higher rate of false negative than false positive error. As suggested by this accuracy rate, the strength of association between the MFD and the criteria is moderate. There is also a considerable range of increases in validities reported for the test but the range of sample sizes is too small to determine if this variable is inversely related to the number of subjects used. It is concluded that the MFD is as effective as the BG in detecting brain dysfunction in mixed populations; however it appears to be somewhat insensitive in this endeavor.

Results of six validation studies of the Minnesota Percepto-Diagnostic Test as an organicity screening device are reported in Table 1. This test uses two designs from the BG, each presented in three orientations. In the original validation study Fuller and Laird (1963) used three groups of subjects differentiated on the basis of diagnosis. The groups of personality disorders and normals were collapsed to form the control group shown in Table 1. The authors report that the MPD was found to be remarkable efficient in discriminating between subjects diagnosed as either brain-damaged or non-brain-damaged. Unfortunately the

mean age of the organic group was 10 years greater than that of the personality disorder group and 26 years greater than the normals. Similar disparities are found for mean education level for the three groups, with the organics by far the less educated group. Consequently it is impossible to tell whether the MPD is sensitive to brain damage, age, education, or any combination of these. The same methodological problem occurred in their cross-validation of the test (Fuller & Laird, 1963). In other experiments investigating the validity of the MPD as an organicity screening test some familiar results are seen. In these studies the overall accuracy rate clusters around 77% and for all of them the test is found to assign subjects to groups significantly better than that expected by chance. As seen above there is a marked discrepancy between the false negative and false positive rates, the former consistently higher. The phi coefficients and the values of incremental validity are not appreciably different than those found on the other two tests.

In summary the MPD appears to share the fate of the BG and MFD in that it can predict brain damage at a better than chance rate, but fails to correctly identify a substantial number of organically impaired subjects. It must be concluded that the MPD is not sensitive enough to the variety of deficits found in brain-damaged individuals.

Table 1 also contains the results of validation studies done on the Benton Visual Retention Test and the Hooper Visual Organization Test. There are only a few relatively accessible articles for the BVRT and the VOT, and therefore they will not be discussed. Mention should be made of a gross error in one study, that done by Brilliant and Gynther (1963),

that investigated the validity of the BG, MFD and the BVRT as organicity screening devices. These authors relied solely upon diagnosis to separate their experimental and control groups, and committed the error of including 26 persons diagnosed as chronic alcoholics in the non-organic group. The authors noted that post-hoc the scores obtained by the alcoholics agreed well with the other non-organic subgroups, reinforcing the alignment of the alcoholics with non-brain-damaged subjects. However the alcoholic subgroup had a mean IQ score at least 10 points higher than any other subgroup. Therefore the results of this study are uninterpretable.

Several conclusions may be drawn concerning the appropriateness of the BG, MFD, MPD, BVRT, and VOT as organicity screening devices. On a practical level it may be concluded that these tests are not efficient as screening tests. They make too many classificatory errors of both types, calling brain damaged those subjects who are found to be non-organic by neurological or neurodiagnostic investigation, and classifying as non-brain-damaged those subjects known to be organically impaired. To be considered as an effective screening device a test should make minimal classification errors in both organic and non-organic individuals. Whereas the overall accuracy reported by these validation studies tends to fall into a range considered acceptable, the disparity between the frequency of false negatives and false positives points out the common failure of all of these tests. In most cases there were many more false negatives. False positives are to be expected since some individuals with mild dysfunction may escape detection by the traditional medical and neurological methods with early manifestations of organicity detec-

table only by behavioral measures. Given that these instruments are used extensively for the detection of mild and moderate degrees of brain dysfunction, it must be concluded that they fail in this endeavor.

The strength of association between the tests' ability to predict brain damage and the external criteria should not be taken lightly. In some of the studies the criterion used was diagnosis only. The phi coefficients tended to be higher in these studies and lower in studies where more extensive and more valid criteria were used. Looking only at the latter, one can conclude that the organicity screening devices may be sensitive to brain damage, but that they may also be detecting some other factor, such as perceptual abnormalities, motor deficiencies, or the normal deterioration of the brain with age. For all five tests the correlations between test and criteria are in the moderate range, not sufficiently robust enough to argue for their validity in detecting brain damage. And it should be recalled that all of these studies employed criteria that are themselves less than perfectly accurate.

Finally, looking at the incremental validities reported above, it can be concluded that considerable gain in precision or detection of brain damage can be realized, when compared to the accuracy expected from the assignment to groups made by consideration of the base rates alone. Overall however, this increase may be marginal if many organically impaired individuals remain undetected by these measures.

The Conceptualization of Organicity

The rationale for the use of the Bender-Gestalt Test and the other tests lies in the conception of brain damage as a single entity, a concept that was prevalent in the early decades of this century. This con-

ceptualization of organicity has its roots in some of the then-current psychological theories. Most notable among these are Kurt Goldstein's concept of the abstract attitude, and Lashley's theories of mass action and equipotentiality. During the era when these concepts enjoyed popularity the brain was conceptualized as a unitary organ which performed effectively when it was healthy. If the brain was damaged by any appreciable lesion, however, the efficiency of the brain would be reduced. It followed logically that brain-impaired individuals were different somehow from their brethren, that this difference was detectable and measureable, and that a single test could be designed which would be sensitive to this difference. The most popular of the tests which were felt to be sensitive to this notion of organicity are, as we have seen, not sufficiently accurate or valid for this purpose.

With the emergence of neuropsychology as a scientific discipline, a new conceptualization of the brain as a functional organ has emerged. Not an undifferentiated organ, the brain is now perceived as a complex organ with a myriad of functions, some of which are rather specifically dependent upon certain areas of the brain being intact, others of which are not readily localizable. Given this current conceptualization of the functionally multifaceted brain, it is probable that the failure of the screening tests reviewed above is due to the fact that these tests tap or assess one or only a few of the many behaviors of the brain. indeed, in commenting on the size of his false negative errors, Hain (1964) observed that this error is understandable with tests that measure one or only a few types of impairment associated with damage of the brain. Any test measuring only selected behavioral manifestations of brain fun-

ction can be expected to yield a sizeable number of false negatives (Golden, 1977; Lezak, 1976; Spreen & Benton, 1965).

This revised conceptualization of brain damage asserts that behavioral deficits manifested by the impaired brain are numerous. The exact deficit to be observed in any brain-injured person will depend in part on many aspects of the lesion: the type and size of the lesion; its location within the cortex or subcortical structures; the age of the individual when stricken and the age of the lesion; the premorbid condition of the person; and the particular neuropsychological organization of functions the individual possesses (Reitan, 1966). This revised concept of organicity also suggests the type of test which would be appropriate as an initial assessment device: a test that is complex, tapping into many behaviors known to be dependent upon an intact cortex. A test which taps only one function as a screening device will yield few false positive but many false negative errors; many individuals with discrete brain lesions will slip through the screen. Conversely the use of a generalized test will yield few if any false negative but many false positive errors; almost everyone will be included.

Focus of This Research

An adequate screening test for detecting or predicting the presence or absence of brain damage in individuals can be seen from the discussion above to have several requisites. The screening device must be a complex task for which an adequate performance depends upon the intactness of many behaviors subtended by the cerebrum. It must be a standardized test that is scored objectively, removing the subjective interpretation from the tasks of the clinician and improving its accuracy. It should require relatively inexpensive and simple apparatus. In the interest of psychologists who work in time-pressured settings, the screening test should take up a minimum of time for administration and scoring. And it must be valid for the purpose of predicting brain impairment in mixed populations of organically and functionally impaired individuals. The Tactual Performance Test (TPT) meets the first four requirements; however, a study of its validity for predicting brain damage has not been done. We propose to study the validity of the TPT and investigate the clinical utility of the test.

The Tactual Performance Test is one of the most complex tasks of the Halstead-Reitan Neuropsychological Battery. The apparatus for the test consists of a form board adapted by Halstead from the Seguin-Goddard form board, ten wooden blocks of various shapes which fit into the spaces of the board, and a stand upon which the board rests. The Halstead-Reitan Battery examines skills known to be dependent upon intact brain function by utilizing many tests, each of which evaluates one or a few of these skills. There is a modicum of redundancy built into the

battery because most of the brain-related skills are tapped by more than one test. The TPT, however, taps into several of these skills. An adequate performance on this test requires tactile form discrimination, motor coordination and manual dexterity, and visualization of the spatial configuration of the shapes in terms of their spatial interrelationships on the board (Reitan & Davison, 1974). Since the skills tapped by the TPT are scattered over a large part of both cerebral hemispheres, lesions in many loci can produce or result in an impaired performance. Consequently this test may be a good predictor of brain dysfunction due to any of a number of lesions, and resembles a global assessment technique of the type sought after in the early decades of this century. The short amount of time needed to administer the test (about 40 minutes) and the minimal space and apparatus necessary suggest that the TPT would be a reasonable organicity screening device for the busy clinician.

The validity of the TPT as a screening test has not been firmly established. It has been found capable of discriminating between brain-damaged and non-brain-damaged groups (Bigler & Tucker, 1981; Reed & Reitan, 1962; Reitan, 1959), but it has not been subjected to a rigorous investigation against a valid criterion. Other studies have shown that the level of education of the subject apparently has no effect upon the performance on the TPT (Finlayson, Johnson & Reitan, 1977), whereas age has been found to have a significant influence upon performance (Cauthen, 1978). Thus, poorer performance on the TPT would appear to be associated with loss of cerebral efficiency or an increase in brain dysfunction whether as a result of increasing age or of acquired brain damage. In addition, the three scores from the TPT have not been found to load on

any measures of hemispheric function in factor analytic studies of the Halstead-Reitan Battery (Golden, 1977; Goldstein & Shelly, 1973).

The validity of the TPT as a screening test for organicity will be done by comparing its performance against the Halstead-Reitan Battery. This battery was chosen as the criterion because it has established itself as a very valid assessment tool in the detection of brain dysfunction, and it is a psychometric procedure. In their classic study of the ability of various neurodiagnostic procedures to detect brain damage, Filskov and Goldstein (1974) also investigated the Halstead-Reitan Battery and found it far superior. This battery correctly identified all 89 subjects whose ultimate diagnosis was made on the basis of the complete set of neuropsychological and neurodiagnostic tests. The Halstead-Reitan Battery has been the subject of several independent researches which have compared brain-damaged groups to a wide variety of non-brain-damaged control groups (Chapman & Wolff, 1959; Chelune, Heaton, Lehman, & Robinson, 1979; Claiborn & Greene, 1981; Golden, 1977; Matthews, Shaw, & Klove, 1966; Reitan, 1955; Vega & Parsons, 1967). These studies have established the battery as a valid instrument in the prediction of brain damage. In selecting the Halstead-Reitan Battery it was presumed that since the tasks are similar in nature to the TPT (ie., all are psychometric measures sensitive to the functioning of the brain), this criterion would constitute a more rigid measure of validity than a criterion based upon the structural or physiological properties of the brain, such as the neurodiagnostics (Yates, 1954). To investigate the clinical utility of the TPT as an organicity screening device, we will look at the classificatory errors made by the TPT. The errors of interest are both

false negatives, those brain-damaged persons whom the TPT classifies as organically intact, and the false positives, those intact individuals whom the TPT classifies as brain-damaged. The judgment of clinical utility will be based upon the magnitude of both types of errors; higher clinical utility can be claimed only if these errors are kept to a minimum. The appropriateness of the TPT as a screening test for organicity will be assessed by comparing its performance to the performances reported above for the Bender-Gestalt, the Memory-For-Designs, and the Minnesota Percepto-Diagnostic tests.

Methods

Subjects

For this retrospective study subjects were selected from the archives of the neuropsychological laboratory at the V.A. Medical Center, Danville, Illinois. Files in the archives contain test data from veterans who had been examined because the presence of brain damage was known or suspected by the referring physician or psychologist. The veterans included both inpatients and outpatients, all of whom were receiving treatment for medical and/or psychiatric problems at the time they were examined.

Files were drawn for inclusion in this study if they contained complete data for the neuropsychological battery used at the Medical Center. This battery is composed of the Halstead-Reitan Battery (Reitan, 1955), the Trail Making Test (Reitan, 1958), the Aphasia Screening Test (Halstead & Wepman, 1949) and one of the Wechsler intelligence scales (Wechsler, 1944, 1955). No other selection criteria were used. The selection process was stopped after 200 complete files were drawn. These files were then divided into two groups, the first 100 arbitrarily called the Validation group and the second 100 the Cross-validation group. The means and standard deviations for age and level of formal education for the Validation group are 44.80 years (13.66) and 10.99 years (2.72) respectively. For the Cross-validation group the means and standard deviations for age and level of formal education are 43.68 years (13.39) and 10.95 years (2.98) respectively. The differences between the groups on these two parameters are not statistically significant.

Procedure

From each subject's file five scores were taken which served as the independent variables. These scores are three measures from the Tactual Performance Test (Time, Memory and Location), the subject's age at the time of testing, and his level of formal education. The dependent variable was the Average Impairment Rating, an omnibus index from the Halstead-Reitan Battery. In this battery there are seven tests which yield 12 indices of brain damage. These 12 indices are ratings of severity of impairment on each of the 12 scores from the battery, the ratings being based upon normative data. The 12 indices or ratings are then averaged to create the Average Impairment Rating (AIR). This AIR is customarily viewed as a measure of the presence or absence of brain damage. Scores on the AIR can range from 0.0 (above average) and 1.0 (intact) to 5.0 (profoundly impaired), and an AIR of 1.55 serves as the cut-off for determining brain damage.

The five independent variables and the AIR for the subjects in the Validation group were used to create a multiple regression equation to predict the AIR. A stepwise regression method was employed to determine which of the independent variables contributed significantly to the predictive ability of the resultant equation, and which did not. The final equation was composed of those independent variables which made significant contributions to the prediction of the AIR. The pertinent independent variables from the Cross-validation group were entered into this equation.

Analysis of the validity of the TPT as a predictor of brain damage was accomplished by computation of the multiple correlation coefficient

for the Validation group, and the correlation coefficient between the predicted and actual AIR for the Cross-validation group. Analysis of the clinical utility of the TPT was done by computing the Chi Square and the phi coefficient of the contingency tables (distribution tables) for each group.

Results

Multiple Regression Equation

The summary table of the stepwise multiple regression equation is presented in Table 2. For each variable entered into the equation the incremental increase in the overall predictive ability is shown under the column, R^2 change, and the significance of the variables are shown under the column, F to enter. With these as guides it was decided that all variables made significant contributions to the prediction of the AIR except Age, so Age was dropped as an independent variable. The variable Age does correlate moderately with the AIR ($r = .535$) but inspection of the correlation matrix revealed that Age is also a correlate of the three TPT variables. The resultant multiple regression equation for predicting the AIR is:

$$\text{AIR}_{\text{est}} = 2.16 + .0457(\text{Time}) - .096(\text{Memory}) - .045(\text{Education}) \\ - .0546(\text{Location})$$

The means, standard deviations and t-values of the variables are presented in Table 3. There are no significant differences between the two groups on any of the variables.

Clinical Utility

The TPT and education level are found to predict the actual AIR for the Validation sample with a remarkable degree of precision. The multiple regression coefficient between the predicted AIR and the actual AIR is .85 ($F = 62.66$, $p < .001$), the predicted AIR being computed by the regression equation above. With this method the TPT and education account for 72% of the total variance in the AIR, itself a derivative of the

Table 2

Summary Table For Stepwise Multiple Regression Equation

Created From Validation Group Data

Step	Variable Entered	<u>F</u> to enter	<u>R</u>	<u>R</u> ²	<u>R</u> ² change	Pearson <u>r</u>	Overall <u>F</u>
1	Time	126.793*	.751	.564	.564	.751	126.793*
2	Memory	30.366*	.817	.668	.103	-.681	97.577*
3	Education	13.610*	.842	.709	.041	-.334	78.045*
4	Location	5.513**	.852	.725	.016	-.650	62.664*
5	Age	2.096	.855	.731	.006	.535	51.128*

* $p < .001$ ** $p < .05$

Table 3

Means and Standard Deviations of the Experimental Variables
for the Validation and Cross-validation Samples

Variable	<u>Validation</u>		<u>Cross-validation</u>		<u>t</u>
	Mean	<u>SD</u>	Mean	<u>SD</u>	
Time	19.35	6.51	19.62	7.40	-.27 (ns)
Memory	6.04	2.05	6.56	1.88	-1.86 (ns)
Location	2.00	2.02	2.48	2.37	-1.53 (ns)
Education	10.99	2.72	10.95	2.98	.10 (ns)
AIR	1.85	.65	1.86	.71	-.10 (ns)

complete Halstead-Reitan Battery. The ability of the TPT and education to account for this much of the variance in the AIR might possibly be due to the fact that the TPT is a multifaceted test and that there is a moderate amount of redundancy in the abilities tapped by the various parts of the Halstead-Reitan Battery.

Using the regression equation, for the Cross-validation sample the correlation coefficient between the predicted AIR and the actual AIR is .86 ($z = 8.55, p < .001$). Here the TPT and education level account for 74% of the total variance in the AIR, reinforcing the suggestion that the TPT taps several abilities known to be subtended by the brain, and that there is a substantial amount of inter-correlation among the tests comprising the Halstead-Reitan Battery.

The results of the current investigation are reported in Table 4, along with measures of the significance of these results. In both samples the base rate of brain damage as determined by the neuropsychological data of the subjects is 65 percent. For the Validation group there were 87 correct predictions, a significant improvement over what could be expected by random assignation of subjects by base rate alone ($\chi^2 = 47.10, p < .001$). The phi coefficient is .71, a strong indication of the degree of association between TPT-based predictions and actual neuropsychological decisions, and of sufficient magnitude to attribute confidence to this test as a predictor of brain damage. The incremental validity, or increase in accuracy of prediction over that to be expected by consideration of the base rate, is .21, again substantial.

For the Cross-validation group there were 86 correct predictions, again a significant improvement ($\chi^2 = 44.43, p < .001$) over the number of

Table 4

Results of Investigation of Tactual Performance Test as an Organicity Screening Device

Sample	Organics _a	Controls _a	Overall Accuracy	False Neg _a	False Pos _a	Chi Square	Phi	Incremental Validity
Validation	65	35	87	4	9	47.10*	.71	.21
Cross-validation	65	35	86	6	8	44.43*	.69	.31

a Number of subjects

* $p < .001$

correct choices by base-rate consideration alone. Here the phi coefficient of .69, and the incremental validity of .31 offer strong evidence in support of the contention that the TPT is remarkably accurate in predicting or identifying brain damage.

Most importantly, inspection of the percentages of false negatives and false positives reveals a reversal of trends found on the other organicity screening tests. For the Validation sample there were 4 false negatives and 9 false positives, and for the Cross-validation sample these figures were 6 and 8, respectively. Very few people who were found to be neuropsychologically impaired actually escaped detection by the TPT. Therefore the TPT is considered to be very sensitive to the myriad of behavioral deficits found in brain-damaged individuals.

Discussion

This study evaluated the Tactual Performance Test as a potential screening device for brain dysfunction. Evaluation included the determination of the validity of the test by comparing its predictive accuracy against that obtained by using the complete Halstead-Reitan Neuropsychological Battery. The comparison showed that the TPT was valid for the purpose of detecting brain damage.

Yates (1954), in a review of some tests of brain damage, had proposed making a distinction between external and internal criteria, the former being of a different nature than the experimental measure and the latter being of the same nature. Yates suggests that internal criterion measures are more appropriate for studies concerned with test validity, and that they present a more stringent demand upon the performance of the experimental measure. In the present study an internal criterion was utilized, that the TPT may be subjected to the more appropriate validation. Against this criterion the TPT was found to be highly valid and, owing to the nature of the data, the TPT was also found to be very reliable. The latter is important since the outcome of the regression equation (the estimated AIR) yields a reasonably accurate measure of the degree of brain dysfunction present in the client.

The correlations between test and criterion for the TPT are higher than those found with the Bender-Gestalt Test and the other measures reviewed above and, as mentioned, the criterion employed for the validation of the TPT was more stringent. In comparison to tests currently employed as organicity screening devices, the TPT may be considered as

probably the more valid instrument.

This study also examined the clinical utility of the TPT by looking at the number and type of prediction errors found, and the overall accuracy of prediction. The overall accuracy surpassed that reported for other instruments currently used to predict brain damage, suggesting that the TPT is more efficient in this endeavor.

In comparing the number or frequency of false positive and false negative prediction from all of these measures, the TPT appears to perform the best. The disparity between false negative and false positive rates is found on the TPT, as it was found on the others, but the discrepancy is not as large. More importantly, the number of false negatives is lower than the number of false positives, in contradistinction to the trend noticed for the other screening devices. Therefore the TPT is more sensitive to the variety of behavioral deficits found in brain-damaged persons.

Looking at the incremental validity, the TPT yields a drastic improvement in correct identification over that to be expected by assignment to groups based upon consideration of base rates alone. Admittedly the base rate of brain damage in the present samples (65%) is higher than the base rates used in the other studies.

In general this study affirms the current hypothesis of "organicity" as a multifaceted construct, rather than a unitary entity. In as much as older screening tests of organicity were based upon the unitary entity model, they were found to be reasonably effective. With the shift in the conceptualization of the nature of organicity based upon neuropsychological research, these tests are reinterpreted as relatively in-

effective since they fail to identify many brain-damaged people. The TPT, on the other hand, is a broad-spectrum test which is sensitive to many different aspects of brain function and better capable of detecting the vast majority of organically impaired individuals. It would appear that the TPT is the more appropriate screening device and should merit precedence over the Bender-Gestalt or other older tests.

There are some drawbacks or shortcomings to this study which may limit the extent to which the results are generalizable to other populations. All of the subjects in this study are males, and this may restrict the use of the regression equation to males. There have been some studies suggesting or reporting that females as a whole perform the TPT differently than males (Cauthen, 1978), whereas to date there is no firm evidence of differences in performance between the sexes on the Halstead-Reitan Battery. Therefore a further study incorporating women only, or both sexes as subjects, would be necessary to determine the applicability of this method to female clients.

Second, the subjects used in this study had received an initial screening for brain dysfunction. They were referred for neuropsychological testing by physicians or psychologists, and had evidenced either known brain damage by history or aroused the suspicion of organicity among the referees. The latter group of subjects are essentially the same as those clients whom the clinician would meet in an initial assessment setting, and for whom the determination of the presence or absence of brain damage would be called for. Since patients with both known and suspected brain damage were included in this study, a partial contamination is present. A further study is necessary in an outpatient or triage setting, in which all clients could be given the entire Halstead-

Reitan Battery and the exact applicability of the TPT as a screening test could be better examined.

The age and education range of the subjects used in this study need to be borne in mind when looking at different populations. These subjects represent a cross-section of veterans receiving treatment at a large VA Hospital, whose mean age and education levels are probably typical of non-VA settings. It is possible that younger or better educated samples might be sufficiently different than this sample to render the regression equation weights less accurate. Selective data analysis, looking at subgroups of patients with differing ages and educations, could better determine whether this method is generalizable across all age and education ranges.

In conclusion, it was found that the TPT and education level could be used as a valid screening test for organicity. This procedure is more efficacious than the Bender-Gestalt Test and others on three counts. First, it is more valid, having been compared against classification based upon a neuropsychological battery rather than diagnosis or neurodiagnostic data. Second, this method yields a higher rate of accurate classification than other tests currently used to detect or predict brain damage. Third, there were far fewer organic subjects missed by this procedure than were undected by the conventional tests. Presumably this would lead to fewer incidences of inappropriate therapies and neurological procedures being employed, and a savings to the clients served in terms of cost and time. Given these points, the TPT should be the test of choice in initial assessment settings where the determination of brain damage is required of the clinician.

References

- Adams, J. Comparison of task-central and task-peripheral forms of the Canter-HP in diagnosing brain damage in adults. Perceptual and Motor Skills, 1971, 33, 1259-1267.
- American Psychiatric Association. Diagnostic and statistical manual of mental disorders. (3rd ed.) Washington, D.C.: American Psychiatric Association, 1980.
- Bigler, E. D., & Ehrfurth, J. W. The continued inappropriate singular use of the Bender visual motor gestalt test. Professional Psychology, 1981, 12, 562-569.
- Bigler, E. D., & Tucker, D. M. Comparison of verbal IQ, tactual performance, Seashore rhythm and finger oscillation tests in the blind and brain-damaged. Journal of Clinical Psychology, 1981, 37, 849-851.
- Boyd, J. L. A validity study of the Hooper visual organization test. Journal of Consulting and Clinical Psychology, 1981, 49, 15-19.
- Brilliant, P. J., & Gynther, M. D. Relationships between performance on three tests for organicity and selected patient variables. Journal of Consulting Psychology, 1963, 27, 474-479.
- Cauthen, N. Normative data for the tactual performance test. Journal of Clinical Psychology, 1978, 34, 456-460.
- Chapman, L. F., & Wolff, H. G. The cerebral hemispheres and the highest integrative functions in man. Archives of Neurology, 1959, 1, 19-35.
- Chelune, G. J., Heaton, R. K., Lehman, R. A., & Robinson, A. Level versus pattern of neuropsychological performance among schizophrenic and diffusely brain-damaged patients. Journal of Consulting and Clinical Psychology, 1979, 47, 155-163.

- Claiborn, J. M., & Greene, R. L. Neuropsychological changes in recovering men alcoholics. Quarterly Journal of Studies on Alcohol, 1981, 42, 757-765.
- Crookes, T. G., & Coleman, J. A. The Minnesota percepto-diagnostic test (MPD) in adult psychiatric practice. Journal of Clinical Psychology, 1973, 29, 204-206.
- Dinham, S. M. Exploring statistice. An introduction for psychology and education. Monterey, Ca.: Brooks/Cole, 1976.
- Filskov, S. B., & Goldstein, S. G. Diagnostic validity of the Halstead-Reitan neuropsychological battery. Journal of Consulting and Clinical Psychology, 1974, 42, 382-388.
- Finlayson, M. A. J., Johnson, K. A., & Reitan, R. M. Relationship of levels of education to neuropsychological measures in brain-damaged and non-brain-damaged adults. Journal of Consulting and Clinical Psychology, 1977, 45, 536-542.
- Fuller, G. B., & Laird, J. T. The Minnesota percepto-diagnostic test. Journal of Clinical Psychology, 1963, 16, 1-33 (Monogr. Suppl.).
- Golden, C. J. Validity of the Halstead-Reitan neuropsychological battery in a mixed psychiatric and brain-injured population. Journal of Consulting and Clinical Psychology, 1977, 45, 1043-1051.
- Goldstein, G., & Shelly, C. H. Univariate versus multivariate analysis in neuropsychological test assessment of lateralized brain damage. Cortex, 1973, 9, 204-216.
- Grundvig, J. L., Ajax, E. T., & Needham, W. E. Screening organic brain impairment with the memory-for-designs test: Validation of comparison of different scoring systems and exposure times. Journal of Clinical Psychology, 1973, 29, 350-354.

- Hain, J. D. The Bender-gestalt test: A scoring method for identifying brain damage. Journal of Consulting Psychology, 1964, 28, 34-40.
- Halstead, W. C., & Wepman, J. M. The Halstead-Wepman aphasia screening test. Journal of Speech and Hearing Disorders, 1949, 14, 9-15.
- Hays, W. L. Statistics for psychologists. New York: Holt, Rinehart, & Winston, 1963.
- Holland, T. R., & Wadsworth, H. M. Incidence versus degrees of rotation on the Minnesota percepto-diagnostic test in brain-damaged and schizophrenic patients. Perceptual and Motor Skills, 1974, 38, 131-134.
- Holland, T. R., Wadsworth, H. M., & Royer, F. L. The performance of brain-damaged and schizophrenic patients on the Minnesota percepto-diagnostic test under standard and BIP conditions of administration. Journal of Clinical Psychology, 1975, 31, 21-25.
- Johnson, J. E., Helkamp, D. T., & Lottman, T. J. The relationship between intelligence, brain damage, and the Hutt-Buskin errors on the Bender-gestalt. Journal of Clinical Psychology, 1971, 27, 84-85.
- Korman, M., & Blumberg, S. Comparative efficiency of some tests of cerebral damage. Journal of Consulting Psychology, 1963, 27, 303-309.
- Kramer, E., & Fenwick, J. Differential diagnosis with the Bender-gestalt test. Journal of Projective Techniques, 1966, 30, 59-61.
- L'Abate, L., Friedman, W. H., Vogler, R. E., & Chused, T. M. The diagnostic usefulness of two tests of brain damage. Journal of Clinical Psychology, 1963, 19, 87-91.
- Lezak, M. D. Neuropsychological assessment. New York: Oxford University Press, 1976.

- Love, H. G. I. Validation of the Hooper visual organization test on a New Zealand psychiatric hospital population. Psychological Reports, 1970, 27, 915-917.
- McManis, D. L. Memory-for-designs performance of brain-damaged and non-damaged psychiatric patients. Perceptual and Motor Skills, 1974, 38, 847-852.
- McNemar, Q. Psychological statistics. New York: Wiley, 1949.
- Matthews, C. G., Shaw, D. J., & Klove, H. Psychological test performances in neurologic and "pseudo-neurologic" subjects. Cortex, 1966, 2, 244-253.
- Mosher, D. L., & Smith, J. P. The usefulness of two scoring systems for the Bender-gestalt test for identifying brain damage. Journal of Consulting Psychology, 1965, 29, 530-536.
- Pardue, A. M. Bender-gestalt test and background interference procedure in discernment of organic brain damage. Perceptual and Motor Skills, 1975, 40, 103-109.
- Reed, H. B. C., Jr., & Reitan, R. M. The significance of age in the performance of a complex psychomotor task by brain-damaged and non-brain-damaged subjects. Journal of Gerontology, 1962, 17, 193-196.
- Reitan, R. M. An investigation of the validity of Halstead's measures of biological intelligence. Archives of Neurology and Psychiatry, 1955, 53, 28-35.
- Reitan, R. M. Validity of the trail making test as an indicator of organic brain damage. Perceptual and Motor Skills, 1958, 8, 271-276.
- Reitan, R. M. Effects of brain damage on a psychomotor problem-solving task. Perceptual and Motor Skills, 1959, 9, 211-215.

- Reitan, R. M. Problems and prospects in studying the psychological correlates of brain lesions. Cortex, 1966, 2, 127-154.
- Reitan, R. M., & Davison, L. A. (Eds.) Clinical neuropsychology: Current status and applications. Washington, D.C.: Winston, 1974.
- Spitzer, R. L., & Fleiss, J. L. A re-analysis of the reliability of psychiatric diagnosis. British Journal of Psychiatry, 1974, 125, 341-347.
- Spreeen, O., & Benton, A. L. Comparative studies of some psychological tests for cerebral damage. Journal of Nervous and Mental Diseases, 1965, 140, 323-333.
- Tavaras, J., & Wood, E. Diagnostic radiology. Baltimore: Williams and Wilkins, 1964.
- Tolor, A., & Schulberg, H. An evaluation of the Bender-gestalt test. Springfield: C. C. Thomas, 1963.
- Tsushima, W. T., & Popper, J. S. Computerized tomography: A report of false negative errors. Clinical Neuropsychology, 1980, 3, 130-133.
- Uyeno, E. Differentiating psychotics from organics on the Minnesota percepto-diagnostic test. Journal of Consulting Psychology, 1963, 27, 462.
- Vega, A., & Parsons, O. Cross-validation of the Halstead-Reitan tests for brain damage. Journal of Consulting Psychology, 1967, 31, 619-625.
- Watson, C. G. A simple bivariate screening technique to separate NP hospital organics from other psychiatric groups. Journal of Clinical Psychology, 1973, 29, 448-450.
- Wechsler, D. The measure of adult intelligence. Baltimore: Williams and Wilkins, 1944.

Wechsler, D. Manual for the Wechsler adult intelligence scale. New York: Psychological Corporation.

Wiggins, J. S. Personality and prediction: Principles of personality assessment. Reading, Ma.: Addison-Wesley, 1973.

Yates, A. J. Validity of psychological tests of brain damage. Psychological Bulletin, 1954, 51, 359-377.

Zubin, J. Classification of the behavior disorders. Annual Review of Psychology, 1967, 18, 373-406.