1984

The Bender Gestalt and Symbol Digit Modalities Tests as Cerebral Dysfunction Screening Instruments: a Comparative Evaluation

Bruce R. Wright
Andrews University

Follow this and additional works at: https://digitalcommons.andrews.edu/dissertations
Part of the Analytical, Diagnostic and Therapeutic Techniques and Equipment Commons, and the Psychology Commons

Recommended Citation
https://digitalcommons.andrews.edu/dissertations/1562

This Dissertation is brought to you for free and open access by the Graduate Research at Digital Commons @ Andrews University. It has been accepted for inclusion in Dissertations by an authorized administrator of Digital Commons @ Andrews University. For more information, please contact repository@andrews.edu.
Thank you for your interest in the

Andrews University Digital Library

of Dissertations and Theses.

Please honor the copyright of this document by not duplicating or distributing additional copies in any form without the author’s express written permission. Thanks for your cooperation.
INFORMATION TO USERS

This reproduction was made from a copy of a document sent to us for microfilming. While the most advanced technology has been used to photograph and reproduce this document, the quality of the reproduction is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help clarify markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure complete continuity.

2. When an image on the film is obliterated with a round black mark, it is an indication of either blurred copy because of movement during exposure, duplicate copy, or copyrighted materials that should not have been filmed. For blurred pages, a good image of the page can be found in the adjacent frame. If copyrighted materials were deleted, a target note will appear listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed, a definite method of "sectioning" the material has been followed. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For illustrations that cannot be satisfactorily reproduced by xerographic means, photographic prints can be purchased at additional cost and inserted into your xerographic copy. These prints are available upon request from the Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases the best available copy has been filmed.

University Microfilms International
300 N. Zeib Road
Ann Arbor, MI 48106

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Wright, Bruce R.

THE BENDER GESTALT AND SYMBOL DIGIT MODALITIES TESTS AS CEREBRAL DYSFUNCTION SCREENING INSTRUMENTS: A COMPARATIVE EVALUATION

Andrews University

Copyright 1984 by Wright, Bruce R.

All Rights Reserved
PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark. 

1. Glossy photographs or pages  
2. Colored illustrations, paper or print  
3. Photographs with dark background  
4. Illustrations are poor copy  
5. Pages with black marks, not original copy  
6. Print shows through as there is text on both sides of page  
7. Indistinct, broken or small print on several pages  
8. Print exceeds margin requirements  
9. Tightly bound copy with print lost in spine  
10. Computer printout pages with indistinct print  
11. Page(s) _______ lacking when material received, and not available from school or author.  
12. Page(s) _______ seem to be missing in numbering only as text follows.  
13. Two pages numbered _________. Text follows.  
14. Curling and wrinkled pages  
15. Other ___________________________________________
THE BENDER GESTALT AND SYMBOL DIGIT MODALITIES
TESTS AS CEREBRAL DYSFUNCTION SCREENING
INSTRUMENTS: A COMPARATIVE EVALUATION

A Dissertation
Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Education

by
Bruce R. Wright
August 1984
THE BENDER GESTALT AND SYMBOL DIGIT MODALITIES
TESTS AS CEREBRAL DYSFUNCTION SCREENING
INSTRUMENTS: A COMPARATIVE EVALUATION

A dissertation
presented in partial fulfillment
of the requirements for the degree
Doctor of Education

by
Bruce R. Wright

APPROVAL BY THE COMMITTEE:

Chairman: Selma A. Chaib
Dean, School of Education

Committee Member: E. Stanley Chace

Committee Member: Robert C. Fadley
Date approved 6-26-84

Committee Member: W.G.A. Futcher

External Examiner: Conrad A. Reichert
ABSTRACT

THE BENDER GESTALT AND SYMBOL DIGIT MODALITIES TESTS AS CEREBRAL DYSFUNCTION SCREENING INSTRUMENTS: A COMPARATIVE EVALUATION

by

Bruce R. Wright

Chairman: Selma A. Chaij
Problem

One of the problems confronting the psychologist in an institutional setting is the selection of appropriate instruments for the screening and diagnosis of a variety of psychological problems. Among the more subtle and yet far-reaching problems of this sort are those associated with disease, trauma, or congenital defect of the central nervous system. It was the purpose of this study to evaluate the usefulness of two well-known psychological tests as screening instruments for neurological dysfunction in a large and varied institutional population.
Method

The Bender Visual-Motor Gestalt Test (BGT) and Symbol Digit Modalities Test (SDMT) were administered to one hundred incarcerated males from four institutions operated by the Michigan Department of Corrections. The attending neurologist at the prison hospital had diagnosed thirty-five subjects as having intracranial lesions, closed head trauma, or chronic cerebral disease process, and sixty-five as free from significant central nervous system disorder.

Scores obtained from the SDMT and from the BGT scored with the Hain criteria were compared to optimum cut-off scores provided by the test and scoring method authors to determine presence or absence of brain damage in the sample subjects. The frequencies of correct diagnoses for both tests, and for the combination of the two tests by adding the scores together, were compared by means of the Chi-square statistic for correlated samples.

Results

The two tests proved to be virtually opposite in their sensitivity. The SDMT was significantly more sensitive to the presence of brain damage in the sample subjects, while the BGT was significantly more sensitive to the absence of brain damage. Administration of the two tests together, and addition of their scores made no significant improvement in the frequency of correct diagnosis of the absence of brain damage in the study sample, and only a slight, statistically insignificant, improvement in the diagnosis of the presence of brain damage in the study sample.
Conclusions

The extent to which the two tests assess common neuro-behavioral functions appears to be significantly less than was expected. The BGT demonstrated a higher overall percentage of correct diagnoses, but by virtue of its greater sensitivity to the absence of brain damage in the sample, also produced a higher frequency of false negative diagnoses. The SDMT, in contrast, demonstrated a lower overall correct diagnostic rate and a higher frequency of false positive diagnoses.

From the data obtained, it appears that the greatest diagnostic accuracy can be obtained with the SDMT alone, or used in conjunction with other, more carefully constructed and standardized tests of visuo-grapho-motor neuropsychological functions than the Bender Visual-Motor Gestalt Test.
# TABLE OF CONTENTS

LIST OF TABLES ................................................................. v

ACKNOWLEDGEMENTS ......................................................... vi

Chapter

I. THE PROBLEM ................................................................. 1
   Introduction ................................................................. 1
   Statement of the Problem ............................................ 6
   Purpose of the Study .................................................. 7
   Research Hypotheses .................................................. 7
   Assumptions ............................................................... 7
   Definition of Terms .................................................... 9
   Significance of the Study ............................................. 10

II. REVIEW OF THE LITERATURE ........................................ 12
   Overview ......................................................................... 12
   Historical and Theoretical Perspectives ....................... 13
   Research in Neuropsychological Assessment ................. 20
   Research Relating to Development and Use of the Bender Gestalt Test ........................................ 23
   Research with the Symbol Digit Modalities Test .......... 29
   Summary ......................................................................... 32

III. THE RESEARCH METHODOLOGY .................................. 34
    Overview ......................................................................... 34
    Research Design ........................................................ 34
    Subject Selection ........................................................ 35
    Procedures ................................................................. 36
    Instrumentation .......................................................... 37
       The Bender Gestalt Test ............................................ 37
       The Hain Scoring System for the BGT ......................... 38
       The Symbol Digit Modalities Test .............................. 40
    The Null Hypotheses and Data Analysis ..................... 42
    The Statistical Analysis .............................................. 43
    Methodological Limitations ....................................... 44
    Summary ......................................................................... 46

IV. PRESENTATION OF THE FINDINGS ................................. 48
    Overview ......................................................................... 48
    Null Hypothesis 1 ....................................................... 48

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Review of the hypothesis</td>
<td>48</td>
</tr>
<tr>
<td>Results</td>
<td>48</td>
</tr>
<tr>
<td>Discussion</td>
<td>49</td>
</tr>
<tr>
<td>Null Hypothesis 2</td>
<td>50</td>
</tr>
<tr>
<td>Review of the Hypothesis</td>
<td>50</td>
</tr>
<tr>
<td>Results</td>
<td>50</td>
</tr>
<tr>
<td>Discussion</td>
<td>50</td>
</tr>
<tr>
<td>Null Hypothesis 3</td>
<td>51</td>
</tr>
<tr>
<td>Review of the Hypothesis</td>
<td>51</td>
</tr>
<tr>
<td>Results</td>
<td>52</td>
</tr>
<tr>
<td>Discussion</td>
<td>52</td>
</tr>
<tr>
<td>Null Hypothesis 4</td>
<td>53</td>
</tr>
<tr>
<td>Review of the Hypothesis</td>
<td>53</td>
</tr>
<tr>
<td>Results</td>
<td>53</td>
</tr>
<tr>
<td>Discussion</td>
<td>53</td>
</tr>
<tr>
<td>Null Hypothesis 5</td>
<td>54</td>
</tr>
<tr>
<td>Review of the Hypothesis</td>
<td>54</td>
</tr>
<tr>
<td>Results</td>
<td>54</td>
</tr>
<tr>
<td>Discussion</td>
<td>55</td>
</tr>
<tr>
<td>Null Hypothesis 6</td>
<td>55</td>
</tr>
<tr>
<td>Review of the Hypothesis</td>
<td>55</td>
</tr>
<tr>
<td>Results</td>
<td>56</td>
</tr>
<tr>
<td>Discussion</td>
<td>56</td>
</tr>
<tr>
<td>Summary</td>
<td>57</td>
</tr>
<tr>
<td>V. SUMMARY, CONCLUSIONS, AND IMPLICATIONS</td>
<td>58</td>
</tr>
<tr>
<td>Overview</td>
<td>58</td>
</tr>
<tr>
<td>Summary</td>
<td>58</td>
</tr>
<tr>
<td>Findings</td>
<td>60</td>
</tr>
<tr>
<td>Null Hypothesis 1</td>
<td>60</td>
</tr>
<tr>
<td>Null Hypothesis 2</td>
<td>62</td>
</tr>
<tr>
<td>Null Hypothesis 3</td>
<td>63</td>
</tr>
<tr>
<td>Null Hypothesis 4</td>
<td>64</td>
</tr>
<tr>
<td>Null Hypothesis 5</td>
<td>65</td>
</tr>
<tr>
<td>Null Hypothesis 6</td>
<td>67</td>
</tr>
<tr>
<td>Conclusions</td>
<td>68</td>
</tr>
<tr>
<td>Implications</td>
<td>72</td>
</tr>
<tr>
<td>Implications for Clinical Practice</td>
<td>72</td>
</tr>
<tr>
<td>Implications for Further Research</td>
<td>74</td>
</tr>
<tr>
<td>APPENDICES</td>
<td>76</td>
</tr>
<tr>
<td>1. Description of the Research Project</td>
<td>78</td>
</tr>
<tr>
<td>2. Certification of Informed Consent</td>
<td>80</td>
</tr>
<tr>
<td>3. Release of Medical Information</td>
<td>82</td>
</tr>
<tr>
<td>4. Test Figures for the Bender Gestalt Test</td>
<td>84</td>
</tr>
<tr>
<td>Table</td>
<td>Description</td>
</tr>
<tr>
<td>-------</td>
<td>------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>1</td>
<td>Hain Scoring System Signs and Weights for the BGT</td>
</tr>
<tr>
<td>2</td>
<td>Data for the Chi-square Analysis of the BGT and SDMT Scores of Brain-Damaged Subjects</td>
</tr>
<tr>
<td>3</td>
<td>Data for the Chi-square Analysis of the BGT and SDMT Scores of Non-Brain-Damaged Subjects</td>
</tr>
<tr>
<td>4</td>
<td>Data for the Chi-square Analysis for the Comparison of the BGT and BGT+SDMT for Brain-Damaged Subjects</td>
</tr>
<tr>
<td>5</td>
<td>Data for the Chi-square Analysis for the Comparison of the BGT and BGT+SDMT for Non-Brain-Damaged Subjects</td>
</tr>
<tr>
<td>6</td>
<td>Data for the Chi-square Analysis for the Comparison of the SDMT and SDMT+BGT for Brain-Damaged Subjects</td>
</tr>
<tr>
<td>7</td>
<td>Data for the Chi-square Analysis for the Comparison of the SDMT and SDMT+BGT for Non-Brain-Damaged Subjects</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

To the Michigan Department of Corrections for their interest in and support of this project; to Dr. Robert R. Walsh for the original suggestion of the research topic, and for his kind assistance in arranging for Department of Corrections approval of the project, as well as in arranging schedules for data collection; to my mother Mrs. Verna Wright, my wife Judy, and our children Jeffrey and Rebecca, whose continued financial and emotional support made the way so much smoother; to my committee members, E. Stanley Chace, Robert Fadeley, and W. G. A. Futcher, for their helpful criticism, guidance, and creative insights; and to my dissertation advisor Selma Chaïj, for her optimism, unfailing patience, and support during the preparation of this study, I express my heartfelt gratitude and appreciation.

vi

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
CHAPTER I

THE PROBLEM

Introduction

Development of neuropsychology as a discipline has essentially paralleled that of the various other branches of psychology. In characterizing neuropsychology as existing "at the intersection of the neurosciences . . . and the behavioral sciences" (1978, p. 1), Hecaen and Albert directly address this fact and simultaneously point out another source of neuropsychological development, namely medicine. Indeed, as one peruses the history of psychology and medicine, particularly the specialty area of neurology, one is struck by the large number of individuals whose names are common to both disciplines, and whose research has formed the theoretical basis for each.

As is the case in any true science, where systematic observation is the basis for inference, neuropsychology is based upon systematic behavioral observation and analysis of behavioral disturbances that are functionally related to alterations in normal brain function. Such a view is consistent with a definition of neuropsychology as "the study of neural mechanisms underlying human behavior" (Hecaen & Albert, 1978, p.1). Neuropsychological assessment has developed in response to the needs of various disciplines concerned with psychological functioning. Psychiatry, Clinical Psychology, Counseling Psychology, and Neurology have each contributed their
unique impetus to neuropsychological assessment practice. In the clinical setting, the practitioner is called upon to evaluate manifest behavior, assess symptom intensity, formulate diagnostic impressions, and develop remedial procedures. When this setting is such that the practitioner is confronted with large numbers of people displaying a wide variety of neurological or behavioral disorders, some means of discriminating strictly behavioral problems from those that are neurologically based is essential.

This differential diagnostic process may be viewed as occurring in two stages. The first stage, or screening phase, involves a more generalized, broad approach. At this level the diagnostician is concerned primarily with the necessity to determine (1) whether or not significant levels of dysfunction exist to warrant further evaluation, and (2) what broad categories of dysfunction are suspected.

The second phase of the process may be viewed as a compliment to or logical extension of the first. When it has been determined that a significant neurological problem exists or is indicated by the findings from the screening procedure, the actual diagnostic process begins. At this level, more specific information is sought. Among the issues to be settled are whether the problem is psychogenic, neurological, or hormonal in origin. If it is determined that the problem is essentially neurological, the location, extent, and specific nature of the lesion must be determined as accurately as possible. Lezak (1976) has pointed out that such data are essential not only to the diagnostic process, but to the formulation of rehabilitation plans as well.

In approaching this task, the clinical psychologist or
neuropsychologist shares much in common with the neurologist. Each is concerned with the manifest behavior of the patient albeit for different practical reasons, and each relies, wholly or in part, upon systematically obtained behavioral observations as their primary data source. Most often these systematic behavioral observations have taken the form of one or another test instrument or battery. One common concern shared by both disciplines is for the restoration of the patient to as near a normal functioning level as possible.

Yet, in spite of these commonalities, the differences between the disciplines often create a significant degree of confusion and ambiguity. This condition is nowhere more apparent than in the definitions given to key concepts which are, in turn, directly related to differences in the manner in which each discipline approaches their common goals. Smith (1962), in an extensive review of the then extant literature of neuropsychology, directly addressed a number of these ambiguities. While critically analyzing research designs, instrumentation, and proposed concept definitions, he observed:

The increasing application of psychological tests in routine neurological diagnostic studies reflects the universally acknowledged finding that, however defined, the mental functions these tests are assumed to measure are apparently impaired in the overwhelming majority of patients with brain lesions. (p. 312)

Yet, it is precisely the assessment aspect of neuropsychological diagnosis in which much of the ambiguity and conceptual confusion is perpetuated. A brief survey of the large and growing number of neuropsychological tests available leads to a recognition of the accuracy of Smith's observation even now, more than twenty years later. Two considerations seem critical in this regard. First is the fact that many of the instruments even now being used in neuropsychological
screening and diagnosis, particularly those that were developed and
standardized earlier than during the last ten or twenty years, were
probably designed for such purposes as intellectual or personality
assessment. Secondly, it must be kept in mind that no direct link
between specific neural cell assemblies and manifest behavior has
ever been conclusively established. Hence, the diagnostician is eval-
uating behavioral correlates of neuropsychological or electrochemical
events rather than the events themselves. Thus, to the already exist-
ing problems of conceptual ambiguity and confusion, the clinician
adds the risk of error related to indirect measurement.

Associated with the trade-off between breadth of coverage and
precision already noted, the degree of overlap between two or more
instruments used in conjunction with one another also confronts the
clinician. That is, the tests may be viewed as overlapping to the
extent that they are measuring the same neuropsychological events.
While different tests are unlikely to overlap completely, if they are
measuring shared behaviors some degree of overlap can be anticipated.
Obviously then, there is an inverse relationship between the extent
of overlap (the extent to which they all measure the same things)
between two or more tests and their combined discriminating power.

Although development of some single "best" test or battery for
screening or diagnosing brain damage has long been an intriguing pros-
pect, to date no such battery or test has been produced. In spite of
continuing research and development that has produced numerous tests
and batteries designed for assessing various types of neurological
dysfunction, none has achieved universal acceptance. Moreover, in
spite of increasing sophistication in both instruments and evaluation
procedures, the neuropsychologist still must rely upon measurement of the behavioral correlates of neural events rather than the neural events themselves for data, and he or she still often relies upon tests that were designed and standardized for other uses. Some of these instruments have become so widely used in neuropsychological assessment, that they have acquired "standard" status. In many cases these tests have repeatedly demonstrated acceptable levels of validity and reliability in neuropsychological diagnostic applications.

In the present study, aside from general theoretical considerations essential to understanding historical perspectives and evaluating the current status of the neuropsychological assessment field, the focus is essentially empirical. The two tests to be compared may be viewed as representative, in a sense, of the two classes of diagnostic tests just noted. The Bender Visual-Motor Gestalt (hereafter referred to as the Bender Gestalt Test, the Bender Gestalt, or simply the BGT) has continued to enjoy wide popularity among clinicians even in the face of continuing equivocal evidence regarding its usefulness as a neurodiagnostic instrument and in spite of originally being designed for other uses. On the other hand, the Symbol Digit Modalities Test (SDMT), a carefully designed and standardized test specifically intended for neurodiagnostic use, is supported by a growing body of research attesting to its validity and reliability.

Additive or subtractive selection is one practical consequence of evaluation, and selection of the most parsimonious, economical, and effective tests possible is a problem faced by every clinician. In a clinical setting, such concerns are of somewhat greater importance than is the specific theoretical heritage of the test or battery.
It is within this general context that the present study addressed the question of the validity and reliability of these two widely used psychological tests when used as neuropsychological screening instruments either individually or in combination.

**Statement of the Problem**

Examination of the Bender Gestalt and Symbol Digit Modalities tests suggests that both instruments tap common visual-motor, visual-perceptual, integrative, and coordination skills. At the same time, whereas the Bender Gestalt is perhaps more involved in eliciting information about perceptual qualities, the Symbol Digit Modalities Test appears particularly sensitive to deficits in language-related symbolic functions and non-language abstract symbolic cognitive functions. To the extent that the two tests do not overlap then, a higher degree of discriminating power may be expected from the combined use of the two tests.

In the face of findings to be presented in chapter II, it is somewhat surprising that, although it has been suggested that use of the two tests together would provide a useful, quick screening battery for use in a neurology practice (Canter, 1978), apparently no attempt has been made to directly compare the two tests or to assess the degree of overlap between them, if any. Neither has there apparently been any attempt to directly compare the performance of the two tests using the same sample of subjects. In attempting to address both of the above issues, the present study yielded data relevant to the utility of both instruments used individually or together to evaluate neurological dysfunction. Moreover, the study provided a means of further testing of Smith's assertion (cited in Walker, 1983) that
"The SDMT is probably the most economic, sensitive, valid and useful single test for differential diagnosis and for general clinical assessment" (p. 485).

**Purpose of the Study**

The purpose of this study was to evaluate the ability of the Bender Gestalt and Symbol Digit Modalities tests to discriminate brain damaged subjects representing a broad spectrum of etiology, type, extent, and locus of lesion from non-brain-damaged subjects.

**Research Hypotheses**

In order to provide focus and direction to the current study, two research hypotheses were proposed:

1. The SDMT will discriminate the presence or absence of brain damage in the sample subjects significantly better than the Bender Gestalt Test when the two are used individually.

2. Combining the obtained scores for the two tests will produce a significantly higher frequency of correct diagnoses of the presence or absence of brain damage than will be obtained by either test used alone.

**Assumptions**

Ideally, in research endeavors, all variables except the experimental variable or variables are controlled by some means so that a cause-effect relationship can be posited between the treatment procedure and observed change in the experimental variable. In practice however, such a straightforward procedure is seldom possible. This is particularly true in social or educational research, where the use of human subjects precludes absolute control of all potentially
relevant variables. When such conditions prevail, the researcher is
called upon to design controls into his procedures by means of special
subject selection techniques, statistical treatment of the data, and
by the acceptance of certain a priori assumptions regarding the nature,
relationship, and relative influence of any known uncontrolled vari-
ables upon the outcomes of the study.

The present study unavoidably incorporated a somewhat circular
relationship between the criterion and experimental variables. That
is, the neurologist, in establishing a diagnosis of the presence or
absence of brain damage, frequently relies upon neuropsychological
test results to guide in the final diagnosis, and it may be that in
some cases one or the other of the two tests being evaluated in the
present study provided part of the diagnostic data. Again, variances
in clinician competence and thoroughness in performing evaluations
for such diagnostic data were beyond the control of this researcher.

Because the present study was designed to approximate a
screening procedure, wherein such information as previous contact
with a neurologist or evaluation by a neuropsychologist was unavail-
able or at best incomplete, control of these variables was impossible.
The researcher was thus forced to delineate assumptions regarding
them.

For the purposes of the present study, four a priori assump-
tions were made:

1. Diagnosis of cerebral dysfunction or brain damage by a
qualified neurologist is tacit evidence of brain damage in the sub-
jects so diagnosed.

2. Diagnosis of no significant disorder or trauma by a
qualified neurologist is tacit evidence that the subject so diagnosed is free of significant brain damage.

3. Prior knowledge of the nature of the research as a function of the informed consent provisions of the study will have no appreciable effect, positive or negative, upon test performance.

4. The specific nature of any particular neuropsychological tests or test batteries used to assist the neurologist in arriving at the final diagnosis will have no appreciable effect upon the outcomes of the current testing procedure.

**Definition of Terms**

Recognizing that the present study intrudes, to a degree, upon several disciplines and that confusion relating to the meaning or use of shared terms or concepts thus becomes likely, the following operational definitions are offered for certain of these terms used in the present study.

**Brain Damaged** - shall refer to a functional diminution or loss related to intracranial lesion or trauma regardless of etiology.

**Cerebral Impairment** - shall be considered an equivalent of brain damage and shall be so defined.

**Demographic Factors** - shall refer to a history of head injury, brain disease, congenital anomaly, chronic alcoholism, or long-term substance abuse in the absence of an accompanying diagnosis of brain damage.

**Diagnosis** - shall be understood to mean an exhaustive and thorough neuropsychological evaluation designed to localize and evaluate a suspected lesion and the associated behavioral deficits.

**Screening** - shall be understood to refer either to the tests...
or to the process involved in ascertaining whether or not significant levels of neurological deficit exist in the subject to warrant further diagnostic referral.

**Standard Scoring** - shall refer to the use of the scoring procedures for a given test, developed by the test author or publisher, or to standardized scoring methods developed for a given test by persons other than the original test author or publisher.

**Significance of the Study**

The literature review to be presented in chapter II reveals that the field of neuropsychological assessment is characterized at the present time by conditions of flux. It has been asserted that the discipline is in its infancy. Moreover, it has been suggested that, although one can look back some seventy years evaluating the field's development, progress has been slow. Corollary to this assertion is the notion that available measurement techniques are therefore somewhat crude by comparison to those available to the more "mature" disciplines.

While varying degrees of refinement and sophistication may be found in the various tests available to the neuropsychological diagnostician or researcher, the fact still remains that in neuropsychological assessment, behavioral correlates of neural processes and not the processes themselves continue to be measured. The two tests used in the present study in a sense represent an "early", more intuitive, perhaps less well-standardized instrument (the Bender Gestalt Test) which continues to be used, and a more recently developed instrument (the Symbol Digit Modalities Test) that has been specifically designed
to discriminate neurological deficits.

Even as the two tests to some extent represent differing foci, they are also among the most widely used tests in neuropsychological screening, and are often included in diagnostic batteries. Therefore, in light of the often conflicting findings reported in the various comparative studies, and particularly the frequent criticism of the continued use of the Bender Gestalt Test as a screening or diagnostic instrument, the findings from the study provide additional data regarding the relative efficacy of both tests for such applications. Beyond this, the study further illuminates issues of neuropsychological assessment inherent in the somewhat different natures of the two tests, and thus the potentially different neuropsychological processes evaluated. Finally, the study tends to support rather than call into further question the assertion of Smith (cited in Walker, 1983) that "The SDMT is probably the most economic, sensitive, valid and useful single test for differential diagnosis and for general clinical assessment" (p. 485), and the supporting opinion of Bigler and Ehrfurth (1980) that "regardless of how the BG is scored or analyzed, as a single measure or as an adjunct screening measure . . . its clinical use as a neurodiagnostic technique is specious at best" (p. 58).
CHAPTER II

REVIEW OF THE LITERATURE

Overview

Reviewing existing literature relating to a given research problem presents the researcher with a number of dilemmas. On the one hand is the obligation to review and present research which bears directly upon the question under consideration. On the other hand, redundancy or repetition serves no useful purpose. The present review should be viewed as representative rather than exhaustive in the sense that the literature cited provides sufficient breadth to establish the necessary historical and theoretical milieu for the study, and also a reasonably comprehensive review of current research in neuropsychological assessment. Although emphasis has been placed upon studies involving one or the other of the tests being evaluated in the present study, further research utilizing other instruments, but bearing generally upon issues of neuropsychological assessment or neuro-diagnostic problems, have been included.

In the present review a discussion of sources relating to historical and theoretical development and studies generally related to neuropsychological assessment are followed by sections presenting studies related to the development and use of the Bender Gestalt Test, various objective scoring systems for the Bender Gestalt, and finally,
studies dealing with the development and use of the Symbol Digit Modalities Test.

**Historical and Theoretical Perspectives**

Although it is possible to trace the development of neuropsychological assessment from antiquity, to do so in the context of the present study would add little to understanding the current problem. Hence, only cursory attention is given to historical issues—prior to the beginning of the nineteenth century—and to philosophical trends—only to the extent of establishing the primary schools of neurological or neuropsychological thought in the twentieth century. Wertheimer (1979) speaks to this issue in summarizing the historical roots of experimental psychology, and thus indirectly of neuropsychology as well, when he says:

... although the origins of many of the trends culminating in the experimental psychology of 1860 and beyond, can be seen in Greek and pre-renaissance thought, it is primarily the history of ideas since the renaissance which casts light upon the academic ortegeist and zeitgeist of mid-nineteenth century Europe, where experimental psychology was born. (p. 49)

In this regard, one can consider five lines of development in scientific thought that bear directly upon the problem at hand. From physiology came the demonstration of anatomic differences between sensory and motor neurons, the study of the electrical nature of nerve impulses, development of the neuron theory, and the growing controversy between globalist and localizationist explanations for observed neurological phenomena. From biology and natural history came Darwin's concepts of natural selection which led directly to the theory of evolution, and revolutionized scientific thought. From physics came the atomic theory of matter, and from mathematics came more
sophisticated quantification based on knowledge produced in the developing field of statistics. Finally, there was in Europe at that time, a strong tendency toward academic dominance in most scientific investigation. That is, most of the research of the time tended to be done in academic settings or under the auspices of universities, an observation to which establishment of a laboratory by Wilhelm Wundt at the University of Leipzig in 1879 bears witness.

The influence of all these various factors may be traced in the fields of neurology and neuropsychology. A dominant controversy has been between globalist and localizationist views of neurological phenomena. In general terms, the localizationist position is that any behavioral manifestation that may be observed involves a more or less specific neural locus, whereas the globalist position interprets these same phenomena from an all-encompassing, essentially unitary conceptualization of the nervous system. While the modern observer with the benefit of history and of the post-renaissance scientific revolution to aid his perception, may conclude that each position contributes some significant concepts to current understandings of nervous system function, such eclecticism has not always prevailed. Indeed, the historical development of neuropsychology has witnessed an ebb and flow in the influence of each position.

Reise (1959) speaks of "Rufus of Ephesus . . . the first medical writer that adopted a global view of the nervous system" (p. 27). This view of the nervous system appears to have dominated neurological thought more or less consistently until the middle of the nineteenth century, when the mounting body of experimental research began to more clearly support the concept of localization of function. In a
brief review of the history of neuropsychology, Hecaen and Albert (1978) assert that:

The localizationists can date the beginning of their movement as April 18, 1861, when Broca described a patient who had lost the faculty of articulated language and who had a left hemisphere lesion. Broca concluded that the posterior-inferior portion of the frontal lobe controlled this behavioral faculty. (p. 14)

Franz Gall's doctrine of organology, later refined and rechristened "phrenology" by his pupil Spurzheim, may be viewed as a precursor of the localizationist approach although his work later came to be regarded as pseudoscientific. However, it is in the work of Broca, noted above, and the later work of Wernicke (cited in Hecaen & Albert, 1978, p. 14) that localizationist theory found early and perhaps its strongest support.

Yet, even as the localizationist movement was gaining strength and beginning to dominate neuropsychological thought during the latter part of the nineteenth century, other investigators were reporting findings which challenged the dominant viewpoint, and which set into motion intellectual and theoretical forces that would eventually force a return to more globalistic views of neural function. Wertheimer (1979) notes the work of Pierre Flourens who about 1820 moved away from the extreme localizationist views exemplified by the phrenologists. Flourens "... used the method of extirpation (surgical removal), studying the relation of loss of brain tissue to changes in behavior, and concluded that the brain acts essentially as a whole, although certain lobes do have fairly specific functions" (p. 27).

The localizationist position continued to dominate neuropsychology, but a growing undercurrent of globalist research characterized the field until after World War I had ended. Perhaps as a
function of the simple accumulation of on-going research begun prior to the war, but more likely influenced by observations of brain damage among returning servicemen, an abandonment of extreme localizationist positions came about. The trend toward a more holistic view of nervous system function begun by Flourens thus resulted, a century later, in a reassertion of globalist theory. Many factors doubtless contributed to this change, but Heilman and Valenstein (1979) single out the discrediting of phrenology as a core factor, suggesting that by it "... other localizationist theories became suspect" (p. 27).

The more holistic trend in neuropsychological research begun by Flourens and continued after World War I, is perhaps nowhere better exemplified than in the work of Karl Lashley, who in 1938 proposed his theory of mass action. This formulation states that "The behavioral result of a lesion depends on the amount of brain removed more than on the location of the lesion" (cited in Heilman & Valenstein, 1979, p. 7). In their own discussion of the effect of brain lesions upon language, Hecaen and Albert (1978, p. 35) echo this position.

During the years since the end of World War I, the two principal schools of neuropsychology have continued to develop. In addition to Lashley's work noted above, the emerging school of Gestalt psychology seems to have had a pronounced effect upon mid-twentieth century neuropsychological theory. Certainly, the mass-action theory and the notion of total organization derived from Gestalt theory had the effect of further diluting the potency of the localizationist position. However, in the period between the two world wars "... the two main schools of thought ... came to lose all heuristic value by their
detachment from concrete facts and their reliance on a priori conceptual notions" (Hecaen & Albert, 1978, p. 5).

The theoretical warfare characteristic of the field during the latter part of the nineteenth century and continuing until the beginning of World War II, appears to have spent itself, and a more moderate holistic position which appears to incorporate certain aspects of both theoretical positions has taken its place. Continuing research has confirmed many of the earlier localizationist positions, while Lashley's mass-action theory has been somewhat modified, to account for the behavioral effects associated with lesions in specific cortical regions.

The existence of a new middle-of-the-road focus for neuropsychology may pose some theoretical or logical conflicts; but, on a practical level, these conflicts, where they have existed, do not seem to have significantly hampered continuing eclectic approaches. Indeed, in their summary statement regarding the globalist-localizationist controversy, Heilman and Valenstein (1979) comment:

Almost every serious student of aphasia made observations supportive of both viewpoints: only in the overview was there a clash. Indeed, any number of influential workers are linked theoretically with one approach but produced meaningful work in the other. . . . Thus, while it is accurate to divide prevailing philosophies . . . almost all investigators utilize both approaches to a greater or lesser extent. (p. 25)

Establishment of the "Gestalt Movement" just prior to World War I, provided yet another dimension to neuropsychological theory and, by extension, to neuropsychological assessment. In addition to the move away from a strict localizationist position toward the more holistic position noted earlier, incorporation of Gestalt theories and concepts into neuropsychological theory forced addition of another
dimension, namely, perception, to be considered as part of the total picture of neurological function and dysfunction.

Perusal of current neuropsychological literature leads to the conclusion that the primary focus of neuropsychological research has been away from consideration of perceptual factors as useful indicators of neurological function, if indeed it has not actively minimized their importance. This appears to be the position taken by Smith (1982) and by Bigler and Ehrfurth (1980, 1981). Certainly, most test-development effort has been focused upon other, more readily quantified correlates of cerebral dysfunction. Nonetheless, some investigators have continued to see utility for such tests as the Bender Gestalt and have continued to publish research supporting the utility of that test as a psychodiagnostic instrument (Brilliant & Gynther, 1963; Evans & Marmorston, 1963; Hain, 1964; Keogh & Smith, 1961; Kramer & Fenwick, 1966; Lacks, 1979; Landis, Baxter, Patterson & Tauber, 1974; Mosher & Smith, 1965; Schulberg & Tolor, 1961; Snortum, 1965).

While the research just cited lends credence to the idea that perceptual deficits may be as legitimately correlated with neural events as other activities such as repeating digits forward and backward, or symbol-digit substitution tasks, the fact remains that the latter are more easily objectified and thus more readily quantified. This fact alone can account for some of the noted tendency away from the former in test development. Indeed, it may have been this difficulty with quantification and resulting problems with precision that contributed to the observation that the introduction of Gestalt theory into neuropsychology was "a necessary movement that was nonetheless
retrograde and dangerous' (Hecaen & Albert, 1978, p. 4).

That introduction of the Gestalt concept of the whole as greater than the sum of its parts is viewed as "retrograde and dangerous" can best be understood through examination of the context of the assertion. The authors, as physicians, have strong ties to the physical sciences, and a heritage of empirical research. The biological sciences have long recognized a biological equivalent to the principle in the concept of "emergents", qualities of the organism, which can be observed in some form at all levels of development from single cells to complex systems. These emergent qualities of organisms are described from the perspective of systems analysis by Miller (1978), who says "a measure of the sum of a system's units is larger than the sum of that measure of its units . . . the more complex systems at higher levels manifest characteristics, more than the sum of the characteristics of the units, not observed at lower levels" (p. 28).

While recognizing "... the fundamental similarity of the living systems studied by biologists and those studied by social scientists," Miller goes on to say, "I oppose any conceptualization of emergents that involves inscrutable characteristics of the whole, greater than the sum of the parts, which are not susceptible to the ordinary methods of scientific analysis" (p. 28).

Thus, Hecaen and Albert's (1978, p. 4) assertion must be viewed from their perspective, and that of many other natural scientists as well, that Gestalt psychological theory was essentially non-scientific. As a matter of fact, after asserting that Gestalt psychological theory substituted "psychological concepts for explanations," it is interesting that they go on to acknowledge that the introduction
of Gestalt theory did produce "new methods of observation and demanded consideration of total behavior, not simply an isolated performance in a single domain" (p. 4).

Both the Bender Gestalt and the Symbol Digit Modalities tests should be considered products of globalistic thinking. However, in the case of the SDMT, one deals with an instrument which more frankly incorporates features of both the globalist and localizationist positions. In one sense, the two tests represent different ways of assessing neurological dysfunction. Yet, although manifestly different in certain respects, the two instruments seem to tap enough common functional skills to suggest that both may have utility as neurodiagnostic screening devices.

Research in Neuropsychological Assessment

As noted earlier, neuropsychological assessment developed in response to the needs of various behavioral and medical disciplines for standardized data relating to brain-behavior interactions posited from the accumulating findings of neurological and psychological experimentation. Early work reported by Flourens (cited in Riese, 1959; Wertheimer, 1979), Broca and Wernicke (cited in Hecaen & Albert, 1979) and Lashley (cited in Hecaen & Albert, 1979), has laid the groundwork for the major theoretical systems found in modern scientific neuropsychology.

More recent research has followed the lead of these pioneers in generally emphasizing deficits in language and language-related symbolic skills in most investigations. In one study of the effect upon both language and cognitive functions, on non-dominant
hemispherectomy (surgical removal of cortical tissue) for tumor, Smith (1969) reported that "... all three hemispherectomy patients with no signs of tumor recurrence one or more years post-operatively, showed consistent, specific non-language defects" (p. 44). A later study, taking a somewhat different approach to assessing the neuropsychological effects of cerebral damage, reported that a sample of eighty-seven incarcerated male, long-term drug abusers, showed no significant deficit in either intellectual functioning as measured with the WAIS or in neuropsychological function as measured with a neuropsychological test battery made up of selected tests from the Halstead-Reitan Battery (Bruhn & Maage, 1975).

Much neuropsychological research has sought to find or develop a single "best" test or battery for differential diagnosis of neurological defects. This emphasis has resulted in a large number of studies which compare the performance of various available tests in the differential diagnosis of various combinations of neurological and psychiatric disorders, and in different patient populations, as well as detailing the development of new tests. Among the more prolific investigators during recent years in the area of psychometric assessment of brain damage, Charles Watson and various associates at the Veterans Hospital, St. Cloud, Minnesota, and elsewhere, have reported several studies detailing attempts to select individual tests or batteries that would prove effective in discriminating between psychotic and organically impaired patients (Watson, 1973; Watson, Plemmel & Jacobs, 1973; Watson, Thomas, Anderson & Felling, 1968). These investigations studied the usefulness of personality measures such as the Minnesota Multiphasic Personality Inventory (MMPI) and various
intelligence and neuropsychological tests as discriminators between psychosis and organicity. The reported results are equivocal.

Along similar lines, Spellacy (1978) reported using a "thirty-one variable neuropsychological test battery in conjunction with the MMPI, to discriminate violent from non-violent incarcerated males." He concluded that while it was possible to discriminate between the two groups with a high degree of accuracy using both behavioral tests and the MMPI, "the behavioral tests comprised in the neuropsychological battery, taken as a group, give better discrimination than does the MMPI" (pp. 51-52).

Using the Wechsler Adult Intelligence Scale (WAIS) and other ability measures, along with the MMPI to discriminate organic from depressed patients, Watson, Davis, and Gasser (1978) reported somewhat contrary results. They indicated that "the capabilities of ability-oriented brain damage tests to separate psychiatric hospital organics from depressives are limited" (p. 396). They further concluded that "the results for the . . .WAIS digit span were particularly disappointing" (p. 397). Further clarification of these findings is afforded when it is noted, on closer examination of the sample of depressives, that they were selected in such a way as to eliminate any cases in which brain damage was suspected to be coexistent with psychosis. Thus, inclusion of data from an instrument that is sensitive to depression of a functional origin would be expected to improve the rate of correct diagnoses.

The several studies cited above bear witness to the equivocal, if not contradictory, findings characterizing the broad spectrum of current research in neuropsychological assessment. Aside from
theoretical issues already raised, the state of flux characterizing the current research may be accounted for, in part, in terms of wide variation in the quality and sophistication of research methodologies used in the various studies noted. This view is congruent with that of Parsons and Prigatano (1978) who state that "some of the methodological problems . . . if carefully considered and acted on could substantially improve the quality of the research" (p. 608).

**Research Relating to the Development and Use of the Bender Gestalt Test**

Yet another line of inquiry particularly important to the present study is that which has followed introduction of the Bender Gestalt Test (Bender, 1938). From the outset, research investigating the test's validity and reliability as a diagnostic instrument for assessing neuropsychological dysfunction has produced equivocal results. Nevertheless, since its initial publication, the Bender Gestalt Test (BGT) has enjoyed, and continues to enjoy, wide popularity among practitioners. Noting that the BGT "has been found to be the third or fourth most frequently used psychological instrument in the whole armamentarium," Schulberg and Tolor (1961) go on to point out that "more than a third" of their sample of 176 members of American Psychological Association Division 12 "felt very confident" using the test for "differential diagnosis involving organicity" (p. 351).

Similarly, reporting a survey of psychologists' opinions regarding test use, Wade and Baker (1977) found that the BGT ranked fourth in a group of twenty common psychological tests that clinicians recommended should be mastered by clinical psychology students (p. 876).

Reflecting the variety of modes by which the posited
relationship between brain function and behavior has been theorized, research in neuropsychology has proceeded on a number of fronts. Studies have sought to elucidate relationships between such variables as sex, age, type, and location of lesion, various scoring methods, and the like. In one study of twenty-seven Veterans' Hospital patients with a variety of organic brain-disease diagnoses, it was claimed that "The moderately high commonality found among the subjects of the study suggests that the Bender Gestalt is successful in pointing out common features in the performance of these organic patients" (Guertin, 1954, p. 367). Another investigation using the BGT drawings of children was reported by Quast (1961), who found that on ten of seventeen attributes of the children's drawings, the test successfully discriminated between brain-damaged and emotionally disturbed samples of subjects.

Comparing the BGT to other well-known neuropsychological tests (the Benton Visual Retention Test and the Graham-Kendall Memory for Designs Test), Brilliant and Gynther (1963) demonstrated that all three "tests for organicity . . . predict clinical diagnosis of organicity at a very high level of significance" (p. 477).

Approaching the issue from the standpoint of variation in the standard administration of the BGT designs, Armstrong (1965) used copied and recalled figures; Snortum (1965) presented the figures with a tachistoscope; and Canter (1966, 1971) developed and used a background interference procedure. All reported that the BGT was generally successful in discriminating brain-damaged from normal and certain psychiatric populations. Again, in a study relating clinical judgment to recall scores, a "statistically significant difference
was seen in mean BG recall scores between normal and Huntington's disease groups" (Lyle & Quast, 1976, p. 230).

Looking at the problem of cerebral dysfunction as a consequence of other disease or trauma processes that are not themselves neurological, two studies are of interest. In the first, a study of 204 subjects with histories of cerebral thrombosis or myocardial infarction showed that all of the psychological tests used, one of which was the Bender Gestalt, discriminated between the brain-damaged (cerebral thrombosis) and assumed non-damaged (myocardial infarction) patients at "significantly better than chance levels" (Evans & Marmourston, 1963, p. 915). In the second investigation, evaluation of two different models of filters for the heart-lung machine employed during open-heart surgery (the filters remove microemboli from the patient's circulating blood) by evaluating post-operative neurological impairment by means of the BGT figures scored with Pascal and Suttell's (1951) scoring system, "striking results" were reported (Landis, Baxter, Patterson & Tauber, 1974, p. 560).

Braun and Reed (1975), reporting an interesting attempt to assess the sensitivity of the BGT to brain damage, indicated that their sample of "non-brain-damaged college subjects could not simulate organicity on the Bender-Gestalt test when instructed to do so" (p. 244).

Reflecting much more equivocal conclusions, Weinstein and Johnson (1964) reported that a panel of twenty psychologists of varying levels of training and experience could not "significantly distinguish between drawings of epileptic patients before or after unilateral temporal lobectomy", nor could the panel distinguish these
drawings from drawings produced by a sample of psychotic patients (p. 819). These findings corroborate those reported earlier by Goldberg (1959), who found that "Staff psychologists, psychology trainees, and nonprofessional persons did not differ from each other in their ability to differentiate organic from nonorganic patients by means of their Bender protocols" (p. 28).

Reporting a 1963 study that investigated the comparative efficiency of a battery of tests, including the Bender Gestalt, for discriminating cerebral damage, Korman and Blumberg (1963) indicated the most and least effective instruments in their study, but only mention the results with the BGT in passing, suggesting that they were not significant in comparison to those obtained from the other instruments, even though they reported an overall correct diagnostic rate of 73.75 percent for the BGT (pp. 305-6). These findings were echoed by Watson (1968), who found the Bender Gestalt's ability to discriminate schizophrenic from organic patients "quite unimpressive" (p. 413).

During the decades since the introduction of the Bender Gestalt Test, there have been numerous attempts to standardize or quantify scoring of the figures for purposes of differential diagnosis. Among the more notable of these are those published by Billingslea (1948), Gobetz (1953), Hain (1964), Hutt and Briskin (1960), Pascal and Suttell (1959), and Peek and Quast (1951). Several studies have also been reported comparing these methods in various combinations and across similar samples. Butler, Coursey, and Gatz (1976) reported an interesting study in which they compared two different methods of scoring the BGT across cultural (black-white) lines. Their findings
indicated that, (1) the Bender did not discriminate against blacks in their study, and (2) "both scoring methods failed to differentiate brain-damaged from psychiatric patients" (p. 283).

Investigating the effect of differing scoring procedures and modes of administration of the BGT to children, Keogh and Smith (1961) reported that "no significant differences between group administration methods were demonstrated" and that "the scoring system proposed by the authors is a reliable instrument for evaluating performance of kindergarten children on the Bender Gestalt Test" (pp. 174-5).

Kramer and Fenwick (1966), reporting an investigation in which eighteen organic and twenty-four functional psychiatric patients were administered the BGT, found that both of the two scoring systems which were compared demonstrated the ability to discriminate between the organic and psychiatric samples "beyond the .001 level of significance" (p. 59). Adding the factor of examiner experience to the above formulation, Lacks and Newport (1980) reported that "psychologists regardless of their level of experience can achieve high levels of diagnostic accuracy with the Bender Gestalt through use of a brief . . . objective scoring system" (p. 356). These findings contrast with those reported by Mosher and Smith (1965) that, while the two methods evaluated "differentiated between the groups at a better than chance level of significance, diagnostic errors were so frequent as to preclude the BGT's being useful for individual diagnosis" (p. 530).

Finally, Lacks (1979) has published a brief review of the literature relating to the current status of the Bender Gestalt Test, especially with respect to diagnostic accuracy. She concludes that "the Bender Gestalt test appears to be most useful in detecting
diffuse cortical damage rather than localized lesions ... the kinds of neuropathology most frequently seen in psychiatric settings" (p. 33).

Smith (1975, 1983) has published two discussions of neuropsychological testing practice that are highly critical of the continued use of the Bender Gestalt or similar tests, either as single screening instruments or as part of more extensive batteries. Perhaps his attitude in this regard is most precisely summed up by Bigler and Ehfurth (1980), who state that "regardless of how the BG is scored or analyzed, as a single measure or as an adjunct screening measure ... its clinical use as a neurodiagnostic technique is specious at best" (p. 88).

In an effort to add another perspective to Bigler and Ehfurth's assertion, it may be helpful to examine their article more closely. It should be noted that the authors were responding to Lack's (1979) review. In support of their assertion, they present five case histories of patients whose computer assisted tomography (CAT) scans clearly showed the presence of serious brain pathology, while their reproductions of the Bender Gestalt figures showed "few errors and no indication for the severity of the damage" (Bigler & Ehfurth, 1980, p. 88). While the case histories presented are certainly evidence that differential diagnosis of brain pathology with the Bender Gestalt Test is subject to error, it must be pointed out that such a charge is true of any test or battery currently available. In any case, a sample of five case histories cannot be seriously regarded as adequate support for a conclusion of this sort.
Research with the Symbol Digit Modalities Test

Since the middle-to-late 1960s, one of the more promising areas of neuropsychological assessment research has centered around use of symbol-digit substitution tasks to provide the behavioral data for assessing cerebral dysfunction. An obvious advantage inherent in the use of such tasks for testing intellectual or neuropsychological function is the ease with which they can be quantified, and thus also standardized and normed. That such tasks might also lend themselves to greater accuracy in differential diagnosis, either through more precise measurement or by tapping neurological functions that are more sensitive to the presence of brain pathology, is also a distinct possibility. Reviewing 101 studies dating from 1910 to 1968, dealing with organic impairment and the role of learning and motor factors in test performance, Kaufman (1968) has this to say:

One test which is used in the diagnosis of organic intellectual deficit is the well known substitution test. . . . The work of several investigators indicates that subjects with cerebral pathology tend to do badly on substitution tests. (p. 47)

The time span (1910—1968) suggested in Kaufman's review indicates that tests involving symbol substitution tasks are not new. In one form or another, they have been used for many years in psychological tests of various kinds. "The digit symbol or substitution test is one of the oldest and best established of all psychological tests . . . and the speed with which he (the subject) does it serves as a measure of his intellectual ability" (Wechsler, 1958, cited in Smith, 1968, p. 84).

Royer (1971) and Royer, Gilmore, and Bruhn (1981) report two studies using the symbol-digit substitution task to investigate
information processing capabilities in groups of brain damaged and schizophrenic subjects.

The Royer studies used the substitution tasks taken from the WAIS and from the Army Beta Test. Smith (1968) reported development of a "Symbol Digit Modalities Test" (SDMT) consisting of a series of nine symbols paired with numbers from one to nine. The subject is asked to match the numbers and symbols in timed written and oral tests.

The SDMT was developed . . . for two specific purposes: (1) to provide an effective and economic instrument for screening apparently normal children and adults for learning and other cerebral disorders, and (2) to contribute to definitions of the nature of defects resulting from brain lesions and reflected in other impaired neuropsychological test performances. (Smith, 1968, p. 85)

Since its first separate publication in 1973, the SDMT has undergone two subsequent revisions, first in 1976 and again in 1982. The SDMT has been widely used both as a diagnostic and as a research instrument. Numerous studies have been undertaken in which the SDMT has been used alone or as part of more extensive neuropsychological test batteries, for discriminating organic impairment related to aging in the presence of educational level as a modifying variable (Cento-fanti, 1975), investigation of neuropsychological deficits associated with an endocrine disorder (Whelan, Schteingart, Starkman, & Smith, 1980), discriminating minimal brain dysfunction from emotional disturbance in children (Owens, 1970), screening senile dementia patients (Pfeffer, Kurosaki, Harrah, Chance, Bates, Detels, Filos & Butzke, 1981), and documenting "initial and long term neuropsychological sequelae of commissurotomy" (Campbell, Bogen & Smith, 1981). The findings reported in these studies are not uniformly supportive of
the SDMT, and thus only partially corroborate the assertion that "The SDMT is probably the most economic, sensitive, valid and useful single test for differential diagnosis and for general clinical assessment" (Smith, 1983, p. 485).

Such a claim is indeed strong, but perhaps not entirely unexpected, inasmuch as the author and developer of the test has much at stake in its acceptance. In any case, the comment provides an interesting format for further consideration of the Owen (1970) and Whelan et al. (1980) studies.

In the first instance, the SDMT and various other selected tests from the Michigan Neuropsychological Test Battery were used in an attempt to discriminate between samples of brain-damaged and emotionally disturbed pre-adolescent boys. It was reported that "none of the combinations of tests was able to discriminate between the two groups" (Owens, 1970, p. 383). In the second study, the authors, reporting results of an investigation of neurological deficits associated with Cushing's syndrome, found that "comparison of mean scores of the four groups reveals a strikingly consistent pattern of progressive declines in almost all tests of cognitive and non-cognitive function" (Whelan et al., 1980, p. 755). However, of greater relevance to the present discussion, closer examination of the scores for the various tests in the battery including the complete WAIS as well as the SDMT, reveals that even at the extreme, that is, marked deficit level, the SDMT proved no more sensitive to the deficits associated with Cushing's syndrome than were ten of the fourteen WAIS subtests.

Finally, two articles reporting studies by Charles Watson and his associates, in which the SDMT and other ability and personality
measures were used in attempts to discriminate organic from psychiatric patients, are of interest. In the first, portions of the Michigan Neuropsychological Test Battery and the MMPI depression scale were used in an attempt to discriminate organics from depressives. At one level, it was attempted to discriminate between the two groups alone, and it was reported that "... the most powerful ability-based discriminatory [sic] was the . . . SDMT" (Watson, Davis & Gasser, 1978, p. 394). The second study investigated a similar question, but included a much broader spectrum of syndromes in the psychiatric sample. It was reported that:

The Smith Symbol Digit Modalities Test results were particularly encouraging. The SDMT proved capable of separating the organics from each of the functional groups both before and after matching for demographic variables. (Watson, Gasser, Schefer, Buranen, & Wold, 1981, p. 352)

From the foregoing, it is clear that while the digit-symbol substitution task generally, and the Symbol Digit Modalities Test in particular have received considerable support as a means of screening cerebral dysfunction, the research has by no means been unanimous in that support. Doubtless, continuing research will serve to further clarify the issue.

Summary

Chapter II has presented a review of the literature relevant to the current study. The sources cited in the review illustrate the general trend of research and development as they relate to neuropsychological assessment. In particular, research using or relating to the use of the Bender Visual-Motor Gestalt and Symbol Digit Modalities tests has been emphasized. While on the basis of the sources surveyed
it may be acknowledged that support for the SDMT is perhaps stronger, more consistent and more recent than much that is reported for the Bender Gestalt, it cannot be argued at this juncture that the findings unequivocally support or oppose the use of either instrument in differential diagnosis.

Chapter III presents the research methodology used in the present study.
CHAPTER III

THE RESEARCH METHODOLOGY

Overview

Chapter III discusses the methodology developed for the present study. The discussion includes a description of the research design, identification of the dependent and independent variables, subject selection procedures, including selection criteria and where subjects were obtained. A description of the field procedures used (including the interview format), securing informed consent, and testing instructions for each subject is followed by a discussion of the instrumentation employed in the present study. Finally, the null hypotheses are stated and the data analysis procedures are described in detail.

Research Design

The present study employed an ex-post-facto design for examining posited relationships between neuropsychological functioning and test performance. An a priori assumption was made that the examining neurologist's diagnosis established the presence or absence of significant brain damage. Thus, for the purposes of the present research, the criterion condition was assumed to be an established fact for each subject participating in the study. Since the criterion condition was established by prior neurological examination and diagnosis, it was
possible to compare the decisions obtained from the two tests used in the study with the criterion. The degree of agreement between the two tests provided a measure of the usefulness of the two tests, both individually and collectively, for screening brain damage in the target population.

The Independent Variable

The presence or absence of brain damage (the criterion), was established by neurological examination.

The Dependent Variables

Two dependent variables were studied:

1. The obtained scores for the Bender Gestalt Test, using the criteria established by the Hain scoring system.

2. The obtained scores on the Symbol Digit Modalities Test.

In each case, the decision regarding presence or absence of neurological dysfunction was made in accordance with cut-off scores and norms established for each test.

Subject Selection

Subjects for the present study were selected from patients referred to the neurology clinic of the prison hospital at the State Prison of Southern Michigan during 1983 and who were still incarcerated in the Michigan corrections system at the time of the data collection. First, hospital records were reviewed to verify the names and numbers of residents appearing on the referral sheets, regardless of the basis for referral. Second, the current location of these individuals was obtained from the computerized data processing system.
of the Corrections Department. Each individual thus identified was then placed on a master list noting, in addition to name, identification number, and location in the system. From this list, names were taken in the order listed and the relevant medical records reviewed to ascertain whether or not the resident would be suitable for the study based upon the type of diagnosis recorded in his file by the neurologist. In general, any subject thus identified could be used in the study so long as the presence or absence of central nervous system disorder had been established by the examining neurologist. Residents with diagnoses of neurological disorders involving peripheral structures were included in the study sample as non-brain-damaged subjects, as were those with diagnoses of narcolepsy or headache without further diagnosis stated. Subjects with diagnoses of brain tumor, abscess, cerebral atrophy, cerebellar atrophy, intrusive or concussive trauma, viral or bacterial infection, thrombosis, aneurism, and cerebral palsy were classified as brain damaged for the purposes of the present study.

Procedures

After the population of potential test subjects had been identified, the following steps were taken in obtaining the test data.

1. Each prospective subject was interviewed by the researcher, given a written description of the proposed study (appendix 1), and provided with an oral explanation of any aspects of the study that remained unclear to him, except that specific details regarding the nature of the test instruments were not disclosed.

2. Each subject who gave verbal consent to participate in the study was asked to sign a standard form (appendix 2) attesting to
his informed consent to participate in the study.

3. In those cases where it was potentially necessary to obtain health or hospital records from private practitioners or hospitals, the subject was asked to sign a standard information release form (appendix 3) authorizing the researcher to obtain those records.

4. Testing of the consenting subjects generally took place immediately after signing of the consent forms, unless institutional security concerns dictated otherwise.

5. Each subject was assigned a code number that was written on the face of the test protocol and placed on the master list maintained in the Office of Healthcare, Department of Corrections, and the subject's name and identification number were then removed from the test protocol.

6. All test responses and scoring were recorded on standard test forms where applicable.

7. All test protocols were scored blind. That is, the identity of the subject was unknown to the scorer at the time the tests were scored.

Instrumentation

For the present study, two widely used, brief psychological tests, the Bender Gestalt Test and the Symbol Digit Modalities Test, were administered to all study subjects.

The Bender Gestalt Test

The Bender Gestalt Test (Bender, 1938) consists of a series of nine different designs selected by Bender from those developed by Wertheimer (cited in Bender, 1938) in his studies of visual perception.
In its present form, each of the designs (appendix 4) is printed on a 4 X 6 inch card. The design cards are presented to the subject one at a time in a specified order with instructions to "draw what you see on the card." Bender (1938) has identified certain drawing characteristics which she found to be associated with aphasia and certain other psychological conditions.

Although Bender did not attempt to score drawings of her test figures quantitatively, others, as noted in chapter II (p. 26) have done so. Virtually all of the scoring systems developed involve assignment of numerical weights to various drawing characteristics such as position, rotations, perseveration, loss of closure, distortion of angles, and the like. Examination of the more prominent and more frequently used of these methods reveals that the primary differences between most of them lies in the number of signs used and the weights assigned to them. In most cases, an optimum or "best" cut-off score, above or below which some degree of neurological disorder is indicated, is provided by the author of the scoring system.

The Hain Scoring System for the BGT

Hain (1964) reported validation studies of a scoring system utilizing fifteen signs which were assigned numerical weights from one to four depending upon the degree to which the sign discriminated between brain damaged and normal subjects. Table 1 summarizes the fifteen signs used and their numerical weights.

When the Bender Gestalt protocol had been obtained, the figures were evaluated according to Hain's (1964) criteria (appendix 5) for the presence of any of the fifteen signs indicating the presence
of possible neurological dysfunction in the subject. The protocol score in each case was the sum of the weighted scores for all signs appearing in the drawings of the test figures.

TABLE 1

<table>
<thead>
<tr>
<th>Sign</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Perseveration</td>
<td>4</td>
</tr>
<tr>
<td>2. Rotation and reversal</td>
<td>4</td>
</tr>
<tr>
<td>3. Concretism</td>
<td>4</td>
</tr>
<tr>
<td>4. Added angles</td>
<td>4</td>
</tr>
<tr>
<td>5. Unit line separation</td>
<td>3</td>
</tr>
<tr>
<td>6. Design overlap</td>
<td>3</td>
</tr>
<tr>
<td>7. Distortion</td>
<td>3</td>
</tr>
<tr>
<td>8. Embellishments</td>
<td>2</td>
</tr>
<tr>
<td>9. Partial rotations</td>
<td>2</td>
</tr>
<tr>
<td>10. Omissions</td>
<td>1</td>
</tr>
<tr>
<td>11. Abbreviation of designs 1 and 2</td>
<td>1</td>
</tr>
<tr>
<td>12. Separation of elements</td>
<td>1</td>
</tr>
<tr>
<td>13. Absence of erasures</td>
<td>1</td>
</tr>
<tr>
<td>14. Closure loss</td>
<td>1</td>
</tr>
<tr>
<td>15. Figure A point of contact</td>
<td>1</td>
</tr>
</tbody>
</table>

An optimum cut-off score of eight, established by Hain (1964), was used in arriving at the decision regarding presence or absence of possible brain damage in the study subjects. Attainment of a score of eight or less did not rule out the possibility of brain damage, but such scores were not viewed as indicative of significant brain pathology. Conversely, attainment of a score of nine or more was taken as indicative only of the possibility that significant neurological dysfunction was present.

Hain (1964) reported an interscorer reliability coefficient of .91 between himself and an independent scorer for a sample of forty-five complete protocols for both the original validation and the cross
validation phases of his study. Other, more recent investigations reporting correct diagnosis percentages as opposed to reliability coefficients have reported successful diagnosis rates ranging from 55-88 percent (Lacks, 1979, p. 30).

The Symbol Digit Modalities Test

The SDMT (Smith, 1973) was specifically designed as a brief neuropsychological test to be used in screening and evaluating both children and adults suspected of being brain-damaged or learning-disabled. The test (appendix 7) consists of a series of nine abstract symbols, each of which has been paired with a number from one to nine. At the top of the test sheet, a key is presented in which the symbols and appropriate numbers are shown. For the test, the symbols are presented in a random arrangement without the numbers. The subject is asked to match each symbol with the appropriate number in two separate (one written and one oral) test runs, each of which is timed for 90 seconds. The subject's score is the total number of correct responses for each test, and neurological dysfunction is indicated if the score is 1.5 standard deviation units or more below the norms for either of the test runs.

Interpretation of raw scores is based on a comparison of the obtained scores to norms for both children (age 8-17) and adults (age 18+), whichever is appropriate. The norms are stated in standard deviation units adjusted for age, educational level, and in the case of children, sex. The cut-off score for children is 1.0 or more standard-deviation units below the norm for the subject's age and sex. The scores obtained were recorded as ratios of the number of correct
responses achieved to the total number attempted. Thus, the scores provided an assessment of the subject's speed and accuracy in completing the test.

The SDMT manual delineates three scoring patterns to be noted in the interpretation of children's protocols that are suggestive of modality-specific neurological deficits: (1) subnormal written and normal oral scores, (2) subnormal oral and normal written scores, and (3) subnormal scores on both the oral and written test. In patterns one and two, while some deficit may be indicated in the communications mode where the subnormal score was obtained, neurological dysfunction per se may not be indicated. On the other hand, when both the written and oral scores are significantly depressed, the possibility of brain damage and the advisability of a more detailed neurological examination are indicated.

When adult protocols are being evaluated, three levels of subnormal test performance are designated by the manual: (1) low scores—those falling at or near 1.0 standard deviation units below the norm, (2) moderately low scores—those falling between 1.5 and 2.0 standard deviation units below the norm, and (3) very low scores—those falling more than 2.0 standard-deviation units below the norm. In general, scores falling more than 1.5 deviation units below the norm (appendix 8) suggest the possibility of significant cerebral damage and indicate the advisability of further, more detailed neurological examination. Thus, while specific localization or assessment of the magnitude of the lesion must be left to other medical or neuropsychological procedures, the SDMT appears to perform an appropriate screening function.
The Null Hypotheses and Data Analysis

In chapter I of this study, the stated purpose was to evaluate the ability of the Bender Gestalt and Symbol Digit Modalities tests to identify brain-damaged subjects or to discriminate between brain-damaged and non-brain-damaged subjects representing a broad range of conditions and levels of severity. In order to further focus and facilitate the study, two "research hypotheses" were proposed. The conceptualization of the problem embodied in those hypotheses provided the context for formulation of the following null hypotheses:

1. There is no significant difference between the Bender Gestalt and Symbol Digit Modalities tests, with respect to the frequency of correct diagnosis of the presence of brain damage.

Hypothesis 1 was tested by means of the Chi-square analysis for correlated samples (Ferguson, 1976, p. 201) since this statistic is appropriate for data relating to categorical or dichotomous variables. In this analysis the two instruments were compared in terms of obtained frequencies of correct versus incorrect diagnoses.

2. There is no significant difference between the Bender Gestalt and Symbol Digit Modalities tests, with respect to the frequency of correct diagnosis of the absence of brain damage.

Hypothesis 2 was tested in the same manner as hypothesis 1.

3. There is no significant difference between the Bender Gestalt Test and the combination of the Bender Gestalt and Symbol Digit Modalities tests, with respect to the frequency of correct diagnosis of the presence of brain damage.

Hypothesis 3 was tested in the same manner as was hypothesis 2 except that the presence of brain damage was assumed when one or both
instruments yielded scores below the established cut-off points.

4. There is no significant difference between the Bender Gestalt Test and the combination of the Bender Gestalt and Symbol Digit Modalities tests, with respect to the frequency of correct diagnosis of the absence of brain damage.

Hypothesis 4 was tested in the same manner as hypothesis 3.

5. There is no significant difference between the Symbol Digit Modalities Test and the combination of the Symbol Digit Modalities and Bender Gestalt tests, with respect to the frequency of correct diagnosis of the presence of brain damage.

Hypothesis 5 was tested in the same manner as hypothesis 3.

6. There is no significant difference between the Symbol Digit Modalities Test and the combination of the Symbol Digit Modalities and Bender Gestalt tests, with respect to the frequency of correct diagnosis of the absence of brain damage.

Hypothesis 6 was tested in the same manner as hypothesis 3.

The Statistical Analysis

Analysis of the raw data obtained from the scores for the two test instruments was undertaken by means of the Chi-square analysis for correlated data. In situations where it is desirable to compare the frequencies of some observed phenomenon in which alternative or dichotomous conditions exist, and where the samples are not independent, as when two different tests are administered to the same subjects, the Chi-square statistic is useful. In this context, Chi-square is a descriptive measure of the magnitude of the discrepancies between the two sets of observations. Chi-square is, in this case,
the ratio of the squared difference, to the sum of cell frequencies from the contingency table, and is represented by the formula:

\[ \chi^2 = \frac{(A-D)^2}{A+D} \]

where A and D represent the frequencies for the cells in the Chi-square matrix representing interaction or disagreement between the two conditions of interest. In general, as the magnitude of the discrepancies increases, so also does the magnitude of the Chi-square value. Evaluation of the significance of the calculated Chi-square value is dependent, among other things, upon sample size and the number of degrees of freedom, which are in turn related to the number of cells in the contingency table or matrix and the a priori significance criterion (α value). The significance criterion was α = .05 with one degree of freedom for each statistical test. In the present analysis significance of the calculated Chi-square value was determined by comparing it with the critical value for the statistic at the established significance criterion. This critical value was obtained from tables published in Ferguson (1976, p. 488).

In the present study, the criterion variable (neurologists' diagnosis) was dichotomous. Moreover, the numerical data for the dependent variables, were obtained by administering two measures to both the brain-damaged and non-brain-damaged groups. Hence, the present study involved comparison of dependent or correlated samples.

**Methodological Limitations**

Certain of the a priori assumptions presented in chapter I of this study focused upon methodological problems inherent to the design employed and represent an attempt to establish parameters for logical
interpretation of the findings obtained. Among the specific problems addressed by those assumptions was that of designation of the criterion measure. The criterion adopted was the diagnosis recorded by the examining physician. As is the case in any endeavor involving human observations, these diagnoses by their nature, are subject to some degree of error. As often as not, in the present study, these diagnoses were based upon the physician’s neurological examination, physical findings, and history, with no corroborating neuropsychological test data recorded, and often no more than an EEG report indicating questionable findings.

Given the inherent imperfection of such observations, they nevertheless may be viewed as an acceptable standard against which the performance of the two tests could be evaluated. To whatever extent the physician’s diagnosis erred, that error was transmitted to the data obtained in the present study.

Yet another limiting aspect to the present methodology relates to the issue of sampling procedure. In general, unless certain conditions relating to sample size and subject selection have been met, the obtained data may not be readily generalized beyond the sample itself with any degree of confidence, and in some cases may be uninterpretable. In the present study, random sampling was not possible. Two observations regarding the nature of the sample and population seem appropriate at this point. First, the population of inmates referred by medical staff to the neurology clinic for evaluation and diagnosis was relatively small (approximately 300), and a large proportion of these men were either unavailable to the researcher or were otherwise inappropriate for inclusion in the study sample because the
referring problem did not meet the selection criteria of intracranial lesion or trauma and absence of recent psychiatric problems. Thus, those individuals selected for the study were those inmates whose referring problems were, or appeared to be, central nervous system disorders. More importantly, the only residents who could be used were those who were willing to volunteer and were able to give informed consent.

Secondly, the characteristics of the study sample were not the focus of primary interest in the present research. Rather, the evaluation of the ability of the Bender Gestalt and Symbol Digit Modalities tests to discriminate between brain-damaged and normal or non-brain-damaged subjects was the primary thrust of the study. Thus, while the subjects selected were required to possess the common characteristic of a neurology clinic referral and diagnosis, generalization of the results was aimed toward an applied situation and only secondarily toward a particular population. That is to say, if either or both tests successfully discriminated between normal and brain-damaged subjects in the present study, they could reasonably be expected to perform in a similar manner in any other screening application regardless of the specific location or characteristics of the population.

Summary

Chapter III has been a presentation of the research methodology employed in the present study. A description of the research design was followed by a discussion of the subject selection process, instrumentation, data collection procedures, hypotheses and
statistical analysis. Finally, a discussion of two methodological limitations inherent to the present study, was presented.

Chapter IV presents in detail, the data analysis and findings of the present study.
CHAPTER IV

PRESENTATION OF THE FINDINGS

Overview

This chapter presents the findings relating to the hypotheses stated in chapter III. The data are presented in six sections, each dealing with one of the null hypotheses tested. Each section includes a review of the hypothesis—in which the null hypothesis and its associated research hypothesis or question are restated, a presentation of the results obtained from the statistical analysis, and a brief discussion of the significance of the results.

Null Hypothesis 1

Review of the Hypothesis

Null hypothesis 1 stated that there is no significant difference between the Bender Gestalt and Symbol Digit Modalities tests, in terms of the frequency of correct diagnosis of the presence of brain damage. The associated research hypothesis (chapter I, p. 7) proposed that the Symbol Digit Modalities Test would discriminate significantly better than the Bender Gestalt Test between brain-damaged and normal subjects.

Results

Data pertaining to null hypothesis 1 are presented in table 2. The data are the appropriate frequencies arranged in a 2 X 2
contingency table similar to that used to calculate the Chi-square value for the test of null hypothesis 1.

**TABLE 2**

DATA FOR THE CHI SQUARE ANALYSIS OF THE BGT AND SDMT SCORES OF BRAIN-DAMAGED SUBJECTS

<table>
<thead>
<tr>
<th></th>
<th>Incorrect</th>
<th>Correct</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDMT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inc</td>
<td>3</td>
<td>12</td>
<td>15</td>
</tr>
<tr>
<td>Totj</td>
<td>7</td>
<td>28</td>
<td>35</td>
</tr>
<tr>
<td>BGT</td>
<td>4</td>
<td>16</td>
<td>20</td>
</tr>
</tbody>
</table>

The obtained Chi-square for the data is 8.89, with p = .003. Null hypothesis 1 was rejected at the required level of significance.

**Discussion**

The data presented in table 2 clearly show that the frequency of correct diagnoses of the presence of brain damage in the sample subjects achieved by the Symbol Digit Modalities Test, and that achieved by the Bender Gestalt Test differ significantly. This finding partially supports research hypothesis 1 (chapter I, p. 7) which proposed that the SDMT would discriminate the presence or absence of brain damage in the sample subjects significantly better than the BGT. Indeed, the SDMT discriminated the presence of brain damage in the sample subjects more frequently than the BGT by a factor of nearly 2 to 1.
Null Hypothesis 2

Review of the Hypothesis

Null hypothesis 2 stated that there is no significant difference between the Bender Gestalt and Symbol Digit Modalities tests, in terms of the frequency of correct diagnosis of the absence of brain damage. This hypothesis is essentially the reverse of null hypothesis 1 and represents the second test related to research hypothesis 1.

Results

The data pertaining to the test of null hypothesis 2 are presented in table 3. Again, these data are cell frequencies obtained from the contingency table analysis used to compute the Chi-square value for the comparison of the two tests under the null hypothesis.

TABLE 3

DATA FOR THE CHI SQUARE ANALYSIS OF THE
BGT AND SDMT SCORES OF NON-BRAIN-DAMAGED SUBJECTS

<table>
<thead>
<tr>
<th></th>
<th>Incorrect</th>
<th>Correct</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDMT</td>
<td>23</td>
<td>34</td>
<td>57</td>
</tr>
<tr>
<td>BGT</td>
<td>8</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Tot</td>
<td>31</td>
<td>34</td>
<td>65</td>
</tr>
</tbody>
</table>

The obtained Chi-square is 23.0 with $p < .0005$. This value was found to be significant at the required level, and null hypothesis 2 was therefore rejected.

Discussion

In general, when the null hypothesis is rejected, the corresponding research hypothesis must be supported. However, inspection
of the data in table 3 reveals that this did not in fact occur. The
data show that the BGT agreed with the criterion, thus correctly iden-
tifying subjects without brain damage in 57 of 65 cases for a true
negative hit rate of 87.7 percent. Against the same criterion, the
SDMT correctly identified only 34 of 65 cases for a true negative hit
rate of 46.2 percent. Moreover, in every case where the SDMT cor-
rectly diagnosed the absence of brain damage, it was in agreement with
the BGT. Thus, it is clear that the BGT was the more sensitive of
the two tests to the absence of brain damage in the sample subjects,
a condition which contradicts research hypothesis 1 in spite of the
statistical significance obtained.

These findings are essentially the converse of those reported
for null hypothesis 1, and suggest that the degree of overlap between
the two tests may not be as great as was originally anticipated.

Null Hypothesis 3

Review of the Hypothesis

Null hypothesis 3 stated that there is no significant differ-
quence between the Bender Gestalt Test and the combination of the Bender
Gestalt and Symbol Digit Modalities Test, with respect to the fre-
cquency of correct diagnosis of the presence of brain damage. This
hypothesis relates to research hypothesis 2 (chapter I, p. 7) and
addresses the question of whether or not combining the two tests as a
screening battery will produce a higher proportion of true positive
or negative diagnoses than would be realized by the use of the Bender
Gestalt Test alone.
Results

The data pertaining to the test of null hypothesis 3 are presented in table 4.

TABLE 4

DATA FOR THE CHI SQUARE ANALYSIS FOR THE COMPARISON OF THE BGT AND BGT+SDMT FOR BRAIN-DAMAGED SUBJECTS

<table>
<thead>
<tr>
<th></th>
<th>Incorrect</th>
<th>Correct</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>BGT+SDMT</td>
<td>C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inc</td>
<td>0</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Tot</td>
<td>4</td>
<td>16</td>
<td>20</td>
</tr>
</tbody>
</table>

From the data in table 4, the calculated Chi-square is 16.0 with p < .0005. This value was significant at the required level, and null hypothesis 3 was therefore rejected.

Discussion

The test of null hypothesis 3 partially supported research hypothesis 2. That is, the null hypothesis addressed the issue of the presence of brain damage, whereas the research hypothesis addressed either presence or absence.

From the data in table 4, it is quite clear that in testing for the presence of brain damage in the sample subjects, combination of the scores from the SDMT and BGT more than doubled the frequency of correct diagnoses. Indeed, table 4 shows that the SDMT alone accounted for 16 of 31 correct diagnoses, or for 51.6 percent under the combined condition. Interestingly, both tests agreed on all false negative decisions.
Null Hypothesis 4

Review of the Hypothesis

Null hypothesis 4 stated that there is no significant difference between the Bender Gestalt Test and the combination of the Bender Gestalt and Symbol Digit Modalities tests, in terms of the frequency of correct diagnosis of the absence of brain damage. Null hypothesis 4 is the conceptual counterpart of null hypothesis 3, and addresses the question of potential improvement of diagnostic accuracy by combining the results from both tests.

Results

The data pertaining to the test of null hypothesis 4 are presented in table 5.

TABLE 5

DATA FOR THE CHI SQUARE ANALYSIS FOR THE COMPARISON OF THE BGT AND BGT+SDMT FOR NON-BRAIN-DAMAGED SUBJECTS

<table>
<thead>
<tr>
<th></th>
<th>Incorrect</th>
<th>Correct</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>BGT+SDMT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>23</td>
<td>34</td>
<td>57</td>
</tr>
<tr>
<td>Inc</td>
<td>8</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Tot</td>
<td>31</td>
<td>34</td>
<td>65</td>
</tr>
</tbody>
</table>

The computed Chi-square is 23.0 with an associated probability of p < .0005. This value was significant at the required level, and null hypothesis 4 was rejected.

Discussion

Again, as was the case with the test of null hypothesis 3, the test of the hypothesis produced statistically significant results.
that did not support the corresponding research hypothesis. The data presented in table 5 show that the BGT correctly diagnosed the absence of brain damage in the sample subjects 87.7 percent (57 of 65 cases) of the time whereas the combined BGT and SDMT did so only in 52.3 percent of the cases (34 of 65). Of somewhat greater import to the present discussion is the fact that in no case did the SDMT provide a correct diagnosis of the absence of brain damage, or a true negative hit, when the BGT failed to do so. Thus, it is clear that in the present study, rather than improving the rate of correct diagnoses, addition of the SDMT to the BGT actually reduced the accuracy of the diagnostic procedure for establishing the absence of brain damage.

Null Hypothesis 5

Review of the Hypothesis

Null hypothesis 5 stated that there is no significant difference between the Symbol Digit Modalities Test and the combination of the Symbol Digit Modalities and Bender Gestalt tests, with respect to the frequency of correct diagnosis of the presence of brain damage. It is readily apparent that null hypothesis 5 is simply the reverse of null hypothesis 3. Research hypothesis 2, related to both, stated that combining the obtained scores from both tests would produce a significantly higher frequency of correct diagnoses than would the use of either test alone.

Results

The data pertaining to null hypothesis 5 are presented in table 6. The obtained Chi-square was 3.0 with $p = .083$. This value was not significant at the required level, and null hypothesis 5 was
therefore not supported in the present analysis.

TABLE 6
DATA FOR THE CHI SQUARE ANALYSIS FOR THE COMPARISON OF THE SDMT AND SDMT+BGT FOR BRAIN-DAMAGED SUBJECTS

<table>
<thead>
<tr>
<th></th>
<th>Incorrect</th>
<th>Correct</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDMT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>28</td>
<td>28</td>
</tr>
<tr>
<td>Inc</td>
<td>4</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Tot</td>
<td>4</td>
<td>31</td>
<td>35</td>
</tr>
</tbody>
</table>

Discussion
The data presented in table 6 and the related calculations clearly show that when the SDMT was used to diagnose the presence of brain damage in the sample subjects, addition of the Bender scores made no significant improvement in the frequency of correct diagnoses. Indeed, the BGT added only 3 of 31 correct diagnoses of the presence of brain damage, or only about 9.7 percent, whereas the SDMT correctly diagnosed 28 out of 31, or 90.3 percent of the brain-damaged cases in the sample.

Null Hypothesis 6
Review of the Hypothesis

Null hypothesis 6 stated that there is no significant difference between the Symbol Digit Modalities Test and the combination of the Symbol Digit Modalities and Bender Gestalt tests, with respect to the frequency of correct diagnosis of the absence of brain damage. This hypothesis is essentially the reverse of null hypothesis 4 and is thus conceptually associated with research hypothesis 2. That is,
it was hypothesized that combining the scores from the two tests would produce a higher frequency of correct diagnoses of both the presence and absence of brain damage in the sample subjects. The test for the presence of brain damage was carried out under null hypothesis 5.

Results

The data pertaining to the test of null hypothesis 6 are presented in table 7.

TABLE 7

<table>
<thead>
<tr>
<th>SDMT+BGT</th>
<th>Incorrect</th>
<th>Correct</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>0</td>
<td>34</td>
<td>34</td>
</tr>
<tr>
<td>Inc</td>
<td>31</td>
<td>0</td>
<td>31</td>
</tr>
<tr>
<td>Tot</td>
<td>31</td>
<td>34</td>
<td>65</td>
</tr>
</tbody>
</table>

The computed chi-square is zero, with \( p > .9999 \). This value was not significant at the required level, and null hypothesis 6 was supported. Under the test of the null hypothesis, the SDMT did not contradict the diagnosis of the combined SDMT and BGT in any case.

Discussion

The data in table 7 clearly show that with respect to the diagnosis of the absence of brain damage in the study sample, the two tests were virtually in total agreement. Although the BGT did in fact correctly diagnose the absence of brain damage in the sample subjects significantly more frequently than the SDMT when the two tests were
used independently, when, under the present hypothesis, the two tests were combined, assumption of the absence of brain damage required that both tests so indicate. In the present case, all of the 34 cases correctly diagnosed by the SDMT were also correctly diagnosed by the BGT. Because of the requirement that both tests must indicate the absence of brain damage in order to establish the diagnosis, the 23 cases where the BGT correctly diagnosed the absence of brain damage when the SDMT did not, did not appear in the analysis and thus did not contribute statistically to the improvement of the accuracy of the diagnosis.

**Summary**

Chapter IV has presented the data pertaining to the tests of the six null hypotheses proposed in chapter I (p. 7). Four of the six null hypotheses were rejected, while two were retained. Chapter V presents a detailed summary of the study findings, a discussion of their significance in the light of the research reviewed in chapter II, and a discussion of implications for professional practice and for further research.
CHAPTER V

SUMMARY, CONCLUSIONS, AND IMPLICATIONS

Overview

This chapter is a presentation of the conclusions and implications to be drawn from the findings of the study. A brief summary of the rationale and research design are followed by a recapitulation of the major findings from the tests of each of the six null hypotheses, discussions of the conclusions, and implications for clinical practice and further research.

Summary

Screening for neurological dysfunction is difficult at best for the practitioner because of the vast complexity of the brain-behavior interactions which constitute the primary data source at his or her disposal. Although there are numerous tests currently being marketed that are alleged to be useful for such screening applications, many of them were originally designed for other uses and their utility as neuropsychological screening instruments is thus open to question. The present study was an attempt to evaluate the usefulness of two well-known psychological tests, the Bender Visual Motor Gestalt and Symbol Digit Modalities tests, as screening instruments for neuropsychological dysfunction. The two tests represent different approaches to the neurodiagnostic process. The BGT is one of that group of tests
designed for other uses that has come to be regarded as useful neuro-
diagnostically, and it retains much of its popularity in spite of
equivocal research support. On the other hand, the SDMT was origi-
nally designed as a neuropsychological screening test and has shown
quite consistent patterns of diagnostic accuracy in the research lit-
erature.

The present study was essentially an ex-post-facto design.
That is, the criterion, the neurologist's diagnosis was already
established and known to the researcher before the subjects were
selected. The dependent variables were the scores obtained by the
study subjects on the two tests. These variables were treated as
dichotomous in that, on the basis of established cut-off scores,
decisions as to the presence or absence of brain damage in the study
subjects were made. The null hypotheses and data analyses were placed
in a counter-balanced pairs arrangement such that each comparison
looked first at the evidence of the presence, and then the absence of
cerebral dysfunction in the study subjects.

The subjects for the study were drawn from the list of resi-
dents of the Michigan Corrections Department's facilities, that had
been referred to the neurology clinic of the prison hospital at the
State Prison of Southern Michigan during 1983. From the approximately
300 residents so identified, one hundred were selected whose diagnosis
indicated whether or not their referring complaints originated in the
central nervous system. Those whose diagnoses indicated the presence
of central nervous system disorder, intracranial lesion, or closed-
head trauma were classified as brain-damaged, while those whose diag-
nosis indicated either peripheral neurological disorder or absence of
any neurological problem were classified as non-brain-damaged. Potential participants in the study were interviewed by the researcher and asked to volunteer for the study. Each resident so interviewed was provided with both a written and oral description of the study, and then asked to sign an informed consent form.

The data were collected with the Bender Gestalt Test (Bender, 1938) and the Symbol Digit Modalities Test (Smith, 1973, 1982). The BGT was scored by means of the criteria and sign weights developed by Hain (1964).

Because the data were dichotomous and were not independent, the Chi-square analysis for dependent samples (Ferguson, 1976) was used. In the data analysis, obtained Chi-square values were compared to a critical value established for one degree of freedom and a significance criterion of $\alpha = .05$. Actual significance levels presented as probabilities were obtained for each computed Chi-square and reported along with each significance decision.

Findings

Null Hypothesis 1

Research hypothesis 1 (chapter I, p. 7) stated that the SDMT would discriminate between neurologically impaired and normal subjects significantly better than the BGT. In order to address the basic question implied in this hypothesis, null hypothesis 1 was formulated, stating that there is no significant difference between the BGT and the SDMT, in terms of the frequency of correct diagnosis of the presence of brain damage.

The computed Chi-square for the test of null hypothesis 1 was
8.89 with $p = .003$. This value was found to be significant at the required level, and null hypothesis 1 was rejected.

**Discussion**

Rejection of the null hypothesis implied support of the underlying research hypothesis. In the present case, the research hypothesis was directional in terms of the test identified, whereas the null hypothesis was not. From the data obtained and the calculated Chi-square value, it was clear that the frequencies of correct diagnoses reported for the two tests differed significantly. Moreover, the instrument designated by the research hypothesis (the SDMT) proved to be the more sensitive of the two tests to the presence of brain damage in the sample subjects, as had been anticipated.

The data reveal that the SDMT agreed with the criterion diagnosis in 28 of 35 cases. The seven cases missed by the SDMT represent a false negative hit rate of 20 percent. Conversely, the SDMT was in agreement with the criterion for presence of brain damage 80 percent of the time. In contrast, the BGT agreed with the criterion in 15 of 35 cases for a true positive hit rate of 43 percent and a corresponding false negative rate of 57 percent. Clearly, significantly higher accuracy in discriminating brain-damaged subjects was achieved by the Symbol Digit Modalities Test.

These findings with respect to the SDMT are comparable to those reported in the recent research literature (Watson, Davis & Gasser, 1978; Watson, Schaefer, Buranen, & Wold, 1981). On the other hand, the present findings regarding the BGT are somewhat less definitive than expected in the light of findings reported by Brilliant &
Gynther (1963), Evans and Marmorston (1965), Hain (1964), Korman and Blumberg (1973), Lacks (1979), and Landis, Baxter, Patterson, and Tauber (1974).

Null Hypothesis 2

Null hypothesis 2 stated that there is no significant difference between the Bender Gestalt and Symbol Digit Modalities tests, in terms of the frequency of correct diagnosis of the absence of brain damage. This hypothesis is the conceptual counterpart of null hypothesis 1, and is also derived from research hypothesis 1. The concern here, is whether or not either of the two tests will in fact identify the normal or non-brain-damaged subject when such is encountered.

The computed Chi-square was 23.0, and \( p < .0005 \). This value was found to be highly significant and null hypothesis 2 was therefore rejected.

Discussion

The rejection of null hypotheses 1 and 2 implies support for the underlying research hypothesis. However, the results are not straightforward in this regard. While research hypothesis 1 was sustained by the rejection of null hypothesis 1, rejection of null hypothesis 2 did not imply support for the underlying research hypothesis. This was so because in the present analysis, it was the BGT that proved to be the more sensitive of the two tests to the absence of brain damage, the opposite of the expected outcome.

It is somewhat cumbersome logically, to conceptualize a test as sensitive to the extremes of a presence-absence continuum. A more parsimonious approach seems to be to consider a given test as
sensitive to both the presence and absence of brain damage simultaneously and that these qualities are present to some proportional degree. This would seem to be the objective of the established optimum cut-off scores published for each test. That is, such cut-off scores are those that have been shown to produce the highest proportion of true positive diagnoses. The present results, therefore, indicate that the two instruments demonstrated significantly different levels of sensitivity to the criterion measure.

The consensus at the present time regarding the BGT is that it is particularly sensitive to a relatively specific class of visuo-grapho-motor deficits related to diffuse cerebral dysfunction that has been associated with chronic, stabilized lesions. The SDMT on the other hand has been shown to be quite sensitive to a broad range of neuropsychological deficits including oral and written language, symbolic reasoning and higher cognitive processes related to more highly localized cerebral problems. Thus, some degree of overlap particularly in terms of the visual-motor functions measured, was expected. The data, however, suggest that such overlap may be minimal.

Null Hypothesis 3

Null hypothesis 3 stated that there is no significant difference between the Bender Gestalt Test and the combination of the Bender Gestalt and Symbol Digit Modalities tests, in terms of the frequency of correct diagnosis of the presence of brain damage. Null hypothesis 3 was derived from research hypothesis 2 (chapter I, p. 7) which states that combining the scores from the two tests would produce a significantly higher frequency of correct diagnoses of the presence
of brain damage than would be realized using either test alone.

The computed Chi-square of 16.0, where $p < .0005$, was found to be highly significant, and null hypothesis 3 was rejected.

**Discussion**

Under this hypothesis, it was expected that combining the two tests would produce a more sensitive screening procedure than would be realized through individual use. Rejection of the null hypothesis implied support of the research hypothesis that the differences between the two were significant. The data show that when the SDMT scores were added to those obtained from the BGT, the frequency of correct diagnoses of the presence of brain damage more than doubled. This finding also supports the idea that the two tests may be measuring different neural-behavioral events, a conclusion that agrees with the findings reported for null hypotheses 1 and 2.

**Null Hypothesis 4**

Null hypothesis 4 stated that there is no significant difference between the Bender Gestalt Test and the combination of the Bender Gestalt and Symbol Digit Modalities tests, in terms of the frequency of correct diagnosis of the absence of brain damage. As was the case with null hypothesis 3, research hypothesis 2 was its source.

The data from the test of null hypothesis 4 yielded a computed Chi-square of 23.0, with $p < .0005$. This value was found to be significant beyond the required level, and null hypothesis 4 was rejected.

**Discussion**

Under null hypothesis 4, it was proposed that addition of the
SDMT scores to those obtained with the BGT would increase the frequency of correct diagnosis. The data, however, show that in this analysis addition of the two sets of scores did not in fact enhance the sensitivity of the diagnostic procedure. Rather, the SDMT was in complete agreement with the BGT in 34 of the 57 cases correctly diagnosed. Moreover, in the present analysis the SDMT did not disagree with the BGT in any case. This observation may be understood in the light of the criterion for determining absence of brain damage.

Since the criterion for presence of brain damage was that the subject must be so diagnosed by either or both tests, establishment of the absence of brain damage required that both tests so indicate. Under this criterion, combining the scores of the two tests must lead to an increase in the probability of false positive diagnoses and to a decrease in the frequency of diagnoses of the absence of brain damage.

Once again, the null hypothesis of no significant difference between the observed frequencies of correct diagnoses was rejected, and the implied support of the underlying research hypothesis was misleading. The fact that the two tests were in total agreement as to the diagnosis, whether they together agreed with the criterion or not, shows that the SDMT made no contribution to enhancing the frequency of correct diagnosis of the absence of brain damage in the study sample.

Null Hypothesis 5

Null hypothesis 5 stated that there is no significant difference between the Symbol Digit Modalities Test and the combination of
the Symbol Digit Modalities and Bender Gestalt tests, in terms of the frequency of correct diagnosis of the presence of brain damage. As was the case with null hypothesis 3 and 4, null hypothesis 5 is a derivative of research hypothesis 2 which stated that the combined effect of the two tests would be to increase the frequency of true positive diagnoses.

The obtained Chi-square from the test of the null hypothesis was 3.0, with $p = .083$. This value was not significant at the .05 level, and null hypothesis 5 was therefore rejected.

**Discussion**

In the present analysis, research hypothesis 2 was not supported. The present finding is, in a sense, a mirror image of that for null hypothesis 3. Indeed, the two differ only in terms of the instrument designated for single use. The fact that addition of the BGT scores to those from the SDMT had no appreciable effect upon the frequency of correct diagnosis, of course, supports the null hypothesis. However, in the light of findings reported earlier, the present finding is not entirely unexpected. Review of the frequencies analyzed under this hypothesis may be helpful at this point. The SDMT correctly diagnosed 28 of 35 cases for a true positive hit rate of 80 percent, while the BGT correctly diagnosed 15 of 35 cases for a correct diagnostic rate of 42.8 percent. However, the data show that both tests together only identified 31 of 35 cases correctly, and in only 3 cases did the BGT correctly diagnose presence of brain damage when the SDMT did not. Thus, the BGT contributed only a little more than 9 percent of the total number of correctly diagnosed cases.
Thus, although as anticipated, the addition of the scores for the two tests did, in fact, increase the frequency of correct diagnoses under the hypothesis, the increase was found to be statistically not significant. However, lack of statistical significance in no way detracts from the clinical significance of such an increase. Where the continuing health and welfare of the patient may be at stake, such increases should not be overlooked.

Null Hypothesis 6

Null hypothesis 6 stated that there is no significant difference between the Symbol Digit Modalities Test and the combination of the Symbol Digit Modalities and Bender Gestalt tests, in terms of the frequency of correct diagnosis of the absence of brain damage. Null hypothesis 6 was the last of four formulated to test the proposition that combining the scores from the two tests would result in significantly higher frequencies of correct diagnoses of both presence and absence of brain damage in the sample subjects. Null hypothesis 6 is conceptually related to null hypothesis 4 which compared the BGT to the combination of the BGT and SDMT with respect to the diagnosis of absence of brain damage. In the present analysis, both of the contingency table cells occupied by the frequencies for interaction of the conditions contained no cases. This not only resulted in a Chi-square of zero, but also showed that the effects of the two tests essentially segregated themselves. This situation appeared to be a function of the decision criterion which required that in order for a diagnosis of absence of brain damage to be made, both tests had to so indicate. The present results can be further understood if those obtained for
the tests of null hypotheses 1 and 2 are recalled. In those analyses the two instruments also segregated themselves, at least in terms of the pattern of sensitivity shown, the SDMT showing sensitivity to the presence and the BGT showing sensitivity to the absence of brain pathology.

Conclusions

The research findings discussed in chapter IV and summarized in the present chapter lead to several conclusions, both of a general and specific nature, regarding the two instruments evaluated. First, an underlying assumption of the current study was that there was some degree of overlap between the BGT and SDMT in terms of the visuomotor-skills assessed, but that other essential differences between them would suggest their usefulness as a neuropsychological screening battery, as has been suggested by Canter (1966), when slight modifications are made in the Bender testing procedure. The current findings suggest that the degree of overlap between the two tests may be somewhat less than anticipated in the design of the present study. The pattern of discrimination demonstrated by the two tests suggests that quite different neuropsychological events are being assessed by each test.

Of perhaps greater importance is the issue of accuracy and consistency of diagnosis. If the correct diagnostic percentages are compared, it appears that the BGT is the more accurate of the two, having agreed with the criterion 72 percent of the time, while the SDMT did so only 62 percent of the time. The finding with respect to the BGT is generally consistent with results reported in chapter II,
whereas the findings with respect to the SDMT are somewhat less defin-
itive than anticipated on the basis of the literature review.

Two questions may be raised at this point. The first relates
to the issue of what is being measured. In this regard, it seems log-
ically cumbersome to conceptualize two tests as possessing differing
levels of sensitivity to the extremes of a presence-absence continuum.
Rather, it seems more appropriate to focus upon the general property
of sensitivity since brain damage, if it is present, tends to be mani-
fest to varying degrees. Thus, both presence and absence of pathology
are measured simultaneously, and a given instrument may be evaluated
as more or less sensitive to the behavioral signs of neuropsycholog-
ical function.

The second question relates to the statistical and practical
problem of true versus false diagnostic decisions, both positive and
negative. Whereas, on a statistical level, validity of research find-
ings is a primary concern, on a practical level, particularly in
neuropsychology, the stakes associated with incorrect diagnosis may
be somewhat higher. In this context, a false positive decision will
result in a superfluous neurology clinic referral. On the other hand,
a false negative decision may result in a missed diagnosis, and at
worst, debilitating or lethal consequences for the patient. Obviously,
some cost is associated with either alternative. Avoidance of the
ethical costs of the latter seems to justify the increased economic
costs of the former.

A test that produces a high frequency of accurate diagnoses
of the absence of brain damage (essentially a negative decision) is
more likely to also produce a higher frequency of false negative
decisions, that is, decisions that brain damage is absent when there is in fact, a lesion present. Conversely, a test which produces a high frequency of positive diagnoses, that is, diagnoses of the presence of brain damage, is also more likely to produce a higher rate of false positive diagnoses.

Under the current line of reasoning, the Symbol Digit Modalities Test appears to be the more usefully sensitive of the two tests since the errors it is more likely to make are ethically more acceptable than those that the Bender Gestalt seems most likely to make.

A second conclusion to be drawn from the current study relates to the issue of combined use of the two tests as a screening battery. It was hypothesized that use of the two tests together would make it possible to obtain a higher frequency of correct diagnoses of either the presence or absence of brain damage than could be obtained using either of the two tests alone. The data relating to this question were mixed. With regard to diagnosis of the presence of brain damage, it was shown that combined use of the two tests produced statistically significant increases in correct diagnosis only when the scores for the SDMT were added to those obtained from the BGT. When the scores for the BGT were added to those obtained from the SDMT, no statistically significant increase over the individual use of the SDMT was noted, although addition of the BGT scores did in fact increase the frequency of correct diagnosis by 9.7 percent (3 cases). Clearly, in the present study, the SDMT contributed more toward increasing the accuracy of the diagnostic procedure than did the BGT although the contribution of the latter should not be ignored in the clinical setting.
The third conclusion taken from the present study relates to the use of the various objective scoring systems available for the BGT. As has been noted, these scoring methods vary in terms of the number of signs included and the manner in which weighted values are assigned to each of the signs. The literature review in chapter II shows considerable variance in the results obtained with any of these methods under experimental conditions. While these variances may be partially explained in terms of the unique characteristics of the study samples involved, the findings from the present study suggest that a positive relationship exists between the number of signs included in a given system and its overall diagnostic accuracy. Thus, in retrospect, use of the Hain system rather than one of the more extensive methods available may constitute a weakness in the present research design. However, the original purpose of the study was to evaluate the two tests as screening devices rather than strictly diagnostic instruments. If the screening process is understood to be somewhat broader and more general, as well as briefer in duration than the diagnostic process, the logic of using the shorter scoring method is apparent. This observation corroborates the observations of Lacks (1979) regarding the Hain scoring system for the BGT.

Finally, the present study corroborates those findings noted earlier by Goldberg (1959), Korman and Blumberg (1963), Watson (1968), and the opinions expressed by Bigler and Ehrfurth (1980, 1981) and Lezak (1976), which suggest that the BGT is of limited use as a neuro-diagnostic instrument. On the other hand, the present findings relating to the overall diagnostic accuracy of the SDMT are somewhat less conclusive than expected in light of the findings reported by

The somewhat lower frequency of correct diagnoses attained by the SDMT in the present study may be accounted for, in part, in terms of the unique characteristics of the study sample. That is, many of the studies cited in chapter II used subjects drawn from populations of hospital neurology service patients with clearly established diagnoses or known post-operative conditions. Thus, a more plausible connection could be made between lesion type or location and test response patterns. In the present study, no attempt was made to so refine the sample inasmuch as one objective of the study was to evaluate the test's performance in a simulated screening process. Moreover, the selection criterion made no distinction between subjects with regard to the rigor of the diagnostic process to which they had been exposed. Thus, potential for a higher frequency of false diagnostic decisions either positive or negative, was increased. Nonetheless, the correct results suggest that the SDMT is indeed useful as a neuropsychological screening test whether it is used alone or in conjunction with a more extensive battery of tests.

Implications

Implications for Practice

Because the current research took place in an actual applied setting, and because a deliberate attempt was made to approximate the intended application of the tests evaluated, certain implications for clinical practice are suggested. In the applied setting, several considerations interface in the process of selecting and using neuropsychological tests. Over and above the issue of validity and reliability
addressed in the present study, the unique characteristics of the setting and target population, and ease of administration and scoring are considerations that the potential user will face. The findings of the present study suggest that, in any setting where group administration and rapid scoring and interpretation must be balanced with a need for accuracy, the SDMT is clearly the better choice. Along these same lines, if only one screening instrument can be used, or at best a very limited battery, the SDMT is again the test of choice. In this latter regard, the findings of the present study suggest that, if broader coverage than is possible with the SDMT alone is desired, the practitioner or clinic administrator may be well advised to seek other instruments that appear to tap the same neuropsychological functions as does the BGT, but which may do so with greater precision. Two tests which come to mind are the Benton Visual Retention Test (BVRT) and the Minnesota Percepto-Diagnostic Test (MPD). Both of these instruments are ability-oriented, as is the BGT. Their major advantage is the fact that they have been carefully standardized and designed to provide quantified scoring. The MPD is perhaps the more useful of the two. Whereas the BVRT is primarily a test of visual memory, the MPD is designed for four different styles of administration, each of which taps different aspects of neuropsychological function. Thus, in addition to the factor of more careful standardization and quantified scoring, it provides somewhat broader coverage than may be possible with the Bender Gestalt. The trade-off however, is that they may require a slightly greater sacrifice of time.

Whereas the extensive body of research relating to the BGT has demonstrated rather equivocal conclusions regarding its usefulness...
as a neurodiagnostic test, the weight of the research has demonstrated its usefulness as an adjunct projective technique. It may be that the BGT can best be included in diagnostic batteries for purposes of screening for psychiatric rather than neurological problems.

Implications for Research

The practical implications just discussed suggest certain directions for future research. Whereas the usefulness of the BGT as a neurodiagnostic or screening instrument has been further challenged by the present findings, avenues of further research relating to or utilizing the SDMT remain quite open.

1. In recent years, neuropsychological research has tended more and more to look toward the possibility that neurochemical or hormonal factors may be common to many types of psychotic or violent behavior. Whelan, Schteingart, Starkman, and Smith (1980) have shown that the SDMT is sensitive to neuro-behavioral correlates of Cushing's syndrome, a hormonal disorder associated with malfunction of the adrenal cortex. Further research relating to the neuro-behavioral correlates of other biochemical or hormonal disorders including some forms of mental deficiency may show the SDMT to be particularly sensitive to these disorders as well. Certainly, the possibility that the SDMT can provide a useful means of comparing the differing effects of various syndromes to which it may be sensitive should not be overlooked.

2. Another direction for research in the neurosciences, in which the SDMT could prove useful, is in the prediction of violent and assaultive behavior. During recent years, as crime rates have soared, court dockets become jammed, and correctional facilities overcrowded,
the possibility of being able to predict violent or assaultive behavior has become especially intriguing to law enforcement, correctional, and parole/probation personnel. Recent neuropsychological research has sought to determine whether or not some neurochemical factors may be present in certain individuals and tend to be activated by environmental stressors. Since the SDMT has already demonstrated sensitivity to some syndromes having neurochemical bases, it may be that it would prove useful in identifying persons with violent or assaultive tendencies. Certainly, clear demonstration of such neurochemical or hormonal factors in the causation of violence, assaultiveness, or other criminal behaviors would have a profound effect upon current educational, rehabilitation, and psychological practice.
APPENDIX 1

DESCRIPTION OF RESEARCH PROJECT
DESCRIPTION OF RESEARCH PROJECT

The purpose of the present study is to compare two psychological tests that are presently used at the Reception and Guidance Center, State Prison of Southern Michigan, to obtain information about their ability to discriminate between persons who have a history of brain injury, and those who do not.

Each participant in the project will be asked to complete two short psychological tests. These tests, the Bender Gestalt Test and the Symbol Digit Modalities Test, are paper-and-pencil tests like those you may have taken while in school, or at the time you entered the corrections system. Although one of the tests is timed, there is no set time limit for the second, and most people are able to complete both tests in about thirty minutes.

Scores from the two tests will be analyzed by statistical procedures designed to indicate which one most accurately discriminates participants with histories of brain damage from those without such histories.

Once the participant has completed the two tests, all identifying information will be deleted from the test protocol so that when the tests are scored, the identity of the respondent will be unknown to the scorer.
APPENDIX 2

CERTIFICATION OF INFORMED CONSENT TO
PARTicipate in NEUropsychological
Assessment Research
APPENDIX 2

CERTIFICATION OF INFORMED CONSENT
TO PARTICIPATE IN NEUROPSYCHOLOGICAL ASSESSMENT RESEARCH

I certify that I have received both a written description and an oral explanation of the above research project, that said description and explanation were full and complete, and that I understand the essential nature of the research project.

I further certify that at no time have I been in any manner subject to coercion or pressure to participate in the above research project.

I hereby consent and agree to participate in the above research. I understand that I have the right to withdraw from, and to refuse further participation in the project at my discretion.

Signed:__________________________
Date:__________________________
Witnessed:__________________________
(principal researcher)
APPENDIX 3

RELEASE OF MEDICAL INFORMATION
To Whom it May Concern:

Please provide Mr. Bruce R. Wright with the health/medical/surgical records specified below.

The above named person has my permission to seek access to my health/medical/surgical records for the purpose of providing data pertinent to his research in neuropsychological assessment.

Signed:________________________

Res. No:_______________________

Date:__________________________
APPENDIX 4

TEST FIGURES FOR
THE BENDER GESTALT TEST
APPENDIX 4

TEST FIGURES FOR
THE BENDER GESTALT TEST

Reproduced from the Bender Visual Motor Gestalt Test, copyright 1938 by the American Orthopsychiatric Association. Reproduced by permission.
APPENDIX 5

HAIN SCORING SYSTEM MANUAL FOR
THE BENDER GESTALT TEST
APPENDIX 5

SCORING SYSTEM FOR THE BENDER GESTALT TEST

University of Virginia School of Medicine

Jack D. Hain, Ph.D.

Instructions

The instructions used here in presenting the Bender Gestalt test are somewhat different from Bender's original instructions. Unless the instructions given here are used in administering this test, this scoring system should not be utilized.

The cards are presented one at a time and are aligned at the top of the sheet of paper. The paper is placed in front of the subject in a vertical position. The subject is told, "Here are some figures (designs) for you to copy. Make your copy as much like the drawing on the card as you can." If the subject asks questions such as, "Should I count the dots," or "Does it have to be exactly like the card," answer, "Just make your drawing as much like this one (indicating the design on the card) as you can."

The subject is allowed to pick up the card, but he is not allowed to rotate the card or paper. If he does, the examiner should put the card or paper or both back into the correct orientation and say, "It must be copied this way." The paper may be temporarily turned by the patient to complete one part of a design; but if it seems as if the subject intends to execute the entire design with the paper turned, the examiner should reorient or direct the subject to reorient the paper into its original position.
Scoring Directions

Each sign is scored once per record. The Score Sheet (see attached form) is designed to facilitate scoring. The easiest and most accurate procedure for scoring is to follow the order of the signs on the record sheet. Consider one sign at a time and review the entire protocol one design at a time by looking for the sign under consideration. When a sign is noted, its weight as given on the Score Sheet is recorded on the Score Sheet in the appropriate place. The procedure is repeated for the next sign and so on until the protocol has been reviewed for all fifteen signs. The total score is the sum of the weights (points) recorded on the Score Sheet.

Scoring Criteria

Signs score only on designs indicated in parentheses.

1. Perserveration

Two types of perseveration are scored. If both occur, the sign is scored only once.

(a) Type 1 - Intra-design perseveration (designs 1, 2, 3, 5, and 6)

Design 1: More than 14 dots

Design 2: More than 12 (13) rows where the stimulus figure has 10 (11) rows

Design 3: More than 1 dot in 1st element

4 dots in 2nd

7 dots in 3rd

9 dots in 4th

More than 5 elements
Design 5: More than 27 dots in the arc or more than 11 dots in the diagonal

Design 6: More than 6 crests of sinusoidal waves in either horizontal or vertical waves

(b) Type 2 - Repeated design perseveration (all designs)

Scored when the subject makes and fails to erase or cross out more than two reproductions of any design. If the subject asks if he may repeat the design, the sign is not scored.

2. Rotations and Reversals (all designs)

This item is scored if a design is rotated more than 45 degrees from its axis. Reversals (90 degree - 270 degree rotations) are scored as rotations. The axes are illustrated below by dotted lines.

Do not score if only one subpart of a design is rotated. For example, the following do not score for this sign. They score for sign 89 (partial rotation) instead.
3. Concretism (all designs)

This design is rarely encountered. To score, a design must be reproduced as a concrete object other than the stimulus object. Examples are one long unbroken line to represent the line of dots in Figure 1; diagonal lines to represent the rows of circles in Figure 2; a formation of bird-like elements for Figure 3 or Figure 3 drawn to resemble a tree; Figure 6 drawn to resemble snakes. Substitution of a written word such as "dot," "circle," or "angle" for the drawn representation of such elements also scores. Where objects such as birds, trees, or snakes are drawn, the subject will usually verbalize the object he is attempting to draw or often will label the drawing as such. The sign is not scored if a subject only comments that a design resembles some object and does not distort the reproduction in an attempt to make it look like the object.

4. Added Angles (designs A, 4, 7, and 8)

This item scores when an extra angle is added. The lines forming the extra angle must be approximately straight and must form a definite angle. Earns score. Curves or arcs resulting in change of direction of a line do not score. Gross tremor resulting in a jagged line does not score.

Examples which score:

Examples which do not score:
5. Separation of Lines (designs A, 4, 6, 7, and 8)

This item scores when any one-unit line of the stimulus is reproduced as two separate lines which do not touch. If the lines touch, do not score.

Examples which score:

6. Overlap (all designs)

This item scores when the elements of one design overlap or run into the "space" of another design. Examples which score:

7. Distortion (all designs)

This item is scored when there is a destruction of the gestalt of any design so that there is an extreme departure from the stimulus. If the shape of the original stimulus is not recognizable, this item is scored.

If the figure has lost its form because of concretism (item 3), this item does not score. Examples which score:
8. Embellishments (designs A, 4, 6, 7, and 8)

This deviation occurs rarely; but when it does, it usually is found on curved designs. To score, an extra meaningless line must be included in the design. Extra lines that are not integrated into a design are scored. Such lines are usually in an entirely opposite direction from the line which they are near or attached. They are easily overlooked in scoring, for they are often small and drawn lightly. Examples which score:

9. Partial Rotations

Two types of partial rotation score.

Type 1 (all designs) - When a design is rotated more than 20 degrees but less than 45 degrees from its axis (see axis illustrated in Sign 2 above).

Type 2 (designs A, 4, 5, 6, 7, and 8) - When one, and only one, subpart is rotated more than 20 degrees, this item scores.

Examples that score:

10. Omission of a Subpart (designs A, 4, 5, 6, 7, and 8)

This item scores when either of the two subparts of a design is omitted. The two subparts of each design which score when omitted are:

Design A  -  circle, square
4  -  open square, curve

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
5 - arc, diagonal extension
6 - horizontal wavy line, vertical wavy line
7 - vertical hexagon, leaning hexagon
8 - hexagon, center diamond

11. Abbreviation (designs 1 and 2)
   
   Design 1: Scores when the design is reproduced with less than 11 dots
   Design 2: Scores where the design is reproduced with less than 9 (10) rows of circles where the stimulus figure has 10 (11) rows

12. Separation (designs A, 4, 5, 6, 7, and 8)
   
   This sign scores when there is a separation between subparts of a design. If the subparts touch at all, the sign does not score. The designs and subparts involved in this score are the same as those outlined in Omission of Subpart item above.

13. Absence of Erasures
   
   This item scores if there is a failure to attempt to erase blatant errors. If there is any erasure on any design, this item does not score.

14. Closure (designs A, 4, 7, and 8)
   
   This item scores when on any one design more than one angle is not closed.

15. Design A Point of Contact (design A)
   
   This item scores when the square and circle touch so that either (a) the point of the square penetrates the circle with space seen within the overlap or (b) the square is imperfectly formed so that two of its sides do not join each other but touch the circle.

Example: (a)  
(b)
APPENDIX 6

MAIN SCORE SHEET FOR THE BENDER GESTALT TEST
<table>
<thead>
<tr>
<th>Points</th>
<th>Description</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Perseveration</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rotation or Reversal</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Concretism</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Added Angles</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Separation of Lines</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Overlap</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Distortion</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Embellishments</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Partial Rotation</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Omission</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Abbreviation #1 or #2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Separation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Absence of Erasure</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Closure</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Point of Contact on Figure A</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>TOTAL</strong></td>
<td></td>
</tr>
</tbody>
</table>
SYMBOL DIGIT MODALITIES TEST
Symbol Digit Modalities Test
Aaron Smith, Ph.D.
University of Michigan

APPENDIX 7

Published by

WARNING: Reproduction of this instrument in any form may result in inconsistent ink densities, distortion of the stimulus materials, or the introduction of extraneous marks thereby confusing the stimulus fields. Any such variation could affect the performance of the subject and invalidate the normative data. Additionally, any unauthorized reproduction of this instrument is a violation of copyright law.

<table>
<thead>
<tr>
<th>Name</th>
<th>Sex</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Birthdate</td>
</tr>
<tr>
<td>School</td>
<td></td>
</tr>
<tr>
<td>Grade</td>
<td></td>
</tr>
<tr>
<td>Hand Used</td>
<td>(circle one)</td>
</tr>
<tr>
<td></td>
<td>Left</td>
</tr>
<tr>
<td></td>
<td>Right</td>
</tr>
<tr>
<td>Date</td>
<td></td>
</tr>
</tbody>
</table>

Written Score

Oral Score

Copyright © 1973 by WESTERN PSYCHOLOGICAL SERVICES
Not to be reproduced in whole or in part without written permission of Western Psychological Services. All rights reserved.
APPENDIX 8

ADMINISTRATION AND SCORING INSTRUCTIONS
FOR THE SYMBOL DIGIT MODALITIES TEST
ADMISSION

The SDMT requires the examinee to substitute a number, either orally or written, for randomized presentations of geometric figures. The appropriate number is shown in a key containing the Arabic numbers 1 through 9, each of which is paired with a different geometric symbol.

For large-scale screening applications, the suggested standard procedure is initial group administration with the written version of the SDMT followed by re-examination with the oral version of the test when necessary. Group written tests are easily administered and scored in a relatively short time. Thus, the SDMT offers an economical method for early screening of apparently normal children and adults for possible covert manual motor, visual, learning, and/or other cerebral defects. Subnormal scores on the written version of the SDMT may be obtained for a variety of reasons including "crabbed" writing by left-handed examinees or awkward writing with the right hand. Examinees, especially children, with written scores more than 1 standard deviation below their appropriate age-sex norm should be re-examined with the oral SDMT in order to avoid unnecessary referrals. This procedure should especially be followed when the SDMT is used alone as a screening procedure.

ADMINISTRATION

This version of the test may be either group or individually administered. The test form is handed to the examinee(s) and the following instructions are read aloud by the examiner (examiner reads only the boldface words):

Please look at these boxes at the top of the page. You can see that each box in the upper row has a little mark in it. Now look at the boxes in the row just underneath the marks. Each of the boxes under the marks has a number. Each of the marks in the top row is different, and under each mark in the bottom row is a different number.

Now look at the next line of boxes (examiner points to line of boxes) just under the top two rows. Notice that the boxes on the top have marks, but the boxes underneath are empty. You are to fill each empty box with the number that should go there according to the way they are paired in the key at the top of the page. For example, if you look at the first mark, and then look up at the key, you will see that the number 1 goes in the first empty box. So write the number 1 in the first box. Now, what number should you put in the second box? (Number 5) That's right. So write the number 5 in the second box. What number goes in the third box? (Number 2) Two, right. That is the idea. You are to fill each of the empty boxes with the numbers that should go in them according to the key. Now for practice, fill in the rest of the boxes until you come to the double line. When you come to the double line, stop.

The examiner should check to see that each examinee understands the task. Any errors made in the first 10 practice responses should be immediately pointed out by the examiner and corrected by the examinee. If an examinee has not understood the nature of the task, the instructions are repeated with further examples until the nature of the test is clearly understood. The examiner then continues with the following instructions:

Now when I say "Go!" write in the numbers just like you have been doing as fast as you can until I say "Stop!" When you come to the end of the first line, go quickly to the next line without stopping, and so on. If you make a mistake, do not erase, just write the correct answer over your mistake. I repeat, DO NOT ERASE as you will waste time. Just write the correct answer over your mistake. Do not skip any boxes and work as quickly as you can. Ready? Go!
Exactly 90 seconds from starting, the examiner says:
Stop!

In group administrations, care should be taken to ensure that each examinee understands the task, and that all start promptly and stop immediately at the “Go!” and “Stop!” commands.

Administering the Oral Version of the SDMT

The oral version of the SDMT can only be administered individually since it requires that the examiner record the examinee’s responses. If the examinee has not been given the written version of the test and the oral version is the only form administered, as in the case of the examinee who is unable to write, the examinee completes the first 10 boxes for practice just as in the written version. The instructions are essentially the same as for the written version except that the examinee is told to speak the responses instead of writing them down. The test form is placed before the examinee and the following instructions are read aloud by the examiner (examiner reads only the boldface words):

Please look at these boxes at the top of the page. You can see that each box in the upper row has a little mark in it. Now look at the boxes in the row just underneath the marks. Each of the boxes under the marks has a number. Each of the marks in the top row is different, and under each mark in the bottom row is a different number.

Now look at the next line of boxes (examiner points to the line of boxes) just under the top two rows. Notice that the boxes on the top have marks, but the boxes underneath are empty. You are to fill each empty box with the number that should go there according to the way they are paired in the key at the top of the page and tell me what the number is. For example, if you look at the first mark, and then look at the key, you will see that the number 1 goes in the first box. So you call out the number 1 for the first box. Now, what number should you put in the second box? Just call it out to me. (Number 5) That’s right. So you would say “5” to me. What number goes in the third box? (Number 2) Two, right. That is the idea. You are to fill each of the empty boxes with the numbers that should go in them according to the key and call the numbers out to me. Now for practice, tell me the numbers that fill in the rest of the boxes until you come to the double line. When you come to the double line, stop.

If the examinee has not understood the nature of the task, the instructions are repeated with further examples.

Now when I say “Go!” call out the numbers just like you have been doing until I say “Stop!” I will write the numbers down for you. When you come to the end of the first line, go quickly to the next line without stopping, and so on. If you make a mistake, tell me what you think the correct answer is. Do not skip any boxes and work as quickly as you can. Ready? Go!

Exactly 90 seconds from starting, the examiner says:
Stop!

If the oral version of the SDMT is a retest given shortly after the administration of the written version, the examinee is not allowed to practice again. A blank SDMT form is handed to the examinee and the examiner says the following:

Now we’ll do the same thing again, but this time, instead of you writing the numbers, you are to call them out to me and I’ll write them down. When I say “Go!” begin with the first box on the line and go just as quickly as you can, saying the numbers aloud so that I can write them down. Ready? Go!

Exactly 90 seconds from starting, the examiner says:
Stop!

SCORING

The score in both written and oral administrations of the test is the number of correct substitutions in each 90-second interval. This does not include those substitutions made during the practice period (i.e., the first 10 boxes). The total number of correct responses can be easily found by placing the scoring key over the test form and counting the number of correct responses. This score is recorded as a proportion of the total number of responses. For example, a score of 36/39 indicates that the examinee made a total of 39 responses, 36 correct and 3 incorrect. The total score provides a measure of the speed and accuracy of symbol-digit substitutions.
APPENDIX 9

ADULT NORMS FOR MEAN WRITTEN AND ORAL
(RETEST) SDMT SCORES BY AGE GROUP
AND EDUCATIONAL LEVEL
## Table 3

### Adult Norms for Mean Written and Oral (Retest) SDMT Scores by Age Group and Level of Education

<table>
<thead>
<tr>
<th>Deviation</th>
<th>18-24</th>
<th>25-34</th>
<th>35-44</th>
<th>45-54</th>
<th>55-64</th>
<th>65 or More</th>
</tr>
</thead>
<tbody>
<tr>
<td>Units*</td>
<td>Written</td>
<td>Oral</td>
<td>Written</td>
<td>Oral</td>
<td>Written</td>
<td>Oral</td>
</tr>
<tr>
<td>-3.0 SD</td>
<td>29</td>
<td>33</td>
<td>27</td>
<td>28</td>
<td>19</td>
<td>22</td>
</tr>
<tr>
<td>-2.5 SD</td>
<td>34</td>
<td>38</td>
<td>31</td>
<td>34</td>
<td>23</td>
<td>28</td>
</tr>
<tr>
<td>-2.0 SD</td>
<td>38</td>
<td>42</td>
<td>35</td>
<td>39</td>
<td>28</td>
<td>33</td>
</tr>
<tr>
<td>-1.5 SD</td>
<td>42</td>
<td>47</td>
<td>39</td>
<td>44</td>
<td>33</td>
<td>38</td>
</tr>
<tr>
<td>-1.0 SD</td>
<td>46</td>
<td>51</td>
<td>43</td>
<td>49</td>
<td>38</td>
<td>44</td>
</tr>
<tr>
<td>-0.5 SD</td>
<td>50</td>
<td>56</td>
<td>47</td>
<td>55</td>
<td>42</td>
<td>49</td>
</tr>
<tr>
<td>0.0 SD</td>
<td>54</td>
<td>61</td>
<td>52</td>
<td>60</td>
<td>47</td>
<td>54</td>
</tr>
<tr>
<td>+0.5 SD</td>
<td>59</td>
<td>67</td>
<td>56</td>
<td>65</td>
<td>52</td>
<td>59</td>
</tr>
<tr>
<td>+1.0 SD</td>
<td>63</td>
<td>73</td>
<td>60</td>
<td>70</td>
<td>57</td>
<td>64</td>
</tr>
<tr>
<td>+1.5 SD</td>
<td>67</td>
<td>78</td>
<td>64</td>
<td>76</td>
<td>62</td>
<td>70</td>
</tr>
<tr>
<td>+2.0 SD</td>
<td>71</td>
<td>84</td>
<td>68</td>
<td>81</td>
<td>66</td>
<td>75</td>
</tr>
<tr>
<td>+2.5 SD</td>
<td>75</td>
<td>90</td>
<td>72</td>
<td>86</td>
<td>71</td>
<td>80</td>
</tr>
<tr>
<td>+3.0 SD</td>
<td>79</td>
<td>95</td>
<td>76</td>
<td>91</td>
<td>76</td>
<td>85</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Deviation</th>
<th>13 Years or More of Education</th>
</tr>
</thead>
<tbody>
<tr>
<td>Units*</td>
<td>Written</td>
</tr>
<tr>
<td>-3.0 SD</td>
<td>31</td>
</tr>
<tr>
<td>-2.5 SD</td>
<td>37</td>
</tr>
<tr>
<td>-2.0 SD</td>
<td>42</td>
</tr>
<tr>
<td>-1.5 SD</td>
<td>47</td>
</tr>
<tr>
<td>-1.0 SD</td>
<td>52</td>
</tr>
<tr>
<td>+0.5 SD</td>
<td>57</td>
</tr>
<tr>
<td>+1.0 SD</td>
<td>62</td>
</tr>
<tr>
<td>+1.5 SD</td>
<td>67</td>
</tr>
<tr>
<td>+2.0 SD</td>
<td>72</td>
</tr>
<tr>
<td>+2.5 SD</td>
<td>77</td>
</tr>
<tr>
<td>+3.0 SD</td>
<td>82</td>
</tr>
</tbody>
</table>

*SD = standard deviation from the mean.

Copyright 1973, 1976, 1982 by Western Psychological services. Reprinted by permission of Western Psychological Services, 12031 Wilshire Blvd., Los Angeles, CA 90025.
APPENDIX 10

MEMORANDA AND CORRESPONDENCE RELATING TO
DEPARTMENT OF CORRECTIONS APPROVAL
OF THE PROJECT
TO: Robert R. Walsh, Ph.D., Administrator
    Psychological Services
    Reception and Guidance Center

DATE: 11/9/83

FROM: Robert H. Richards, Ph.D., Chief of Research
      Program Bureau

SUBJECT: Research Proposal by Drue Wright

I have reviewed Mr. Wright's research proposal entitled, "The Bender Gestalt
and Symbol Digit Modalities Tests as Cerebral Dysfunction Screening
Instruments: A Comparative Evaluation" dated October, 1983 and find that it
meets all of the criteria for approval by the Program Bureau. I would
therefore urge Mr. Wright to proceed with his study subject to the
following: (1) That he advise us of the institutions in which he will be
working so that we can confirm that the Warden/Superintendent of each insti-
tution has given his/her approval; (2) That Mr. Wright submit a copy of the
final draft of the study to the Program Bureau to confirm the accuracy of
statements made about the Department of Corrections; and (3) That Mr. Wright
submit a file copy of the final draft of the study to the Program Bureau for
inclusion in the Departmental library.

Please tell Mr. Wright that I thoroughly enjoyed reading his excellently
prepared proposal and look forward to reading the final product.

RHR:ls1

cc: Regional Administrator Foltz
    Superintendent Prelesnik
    J. Ritenour, Ph.D.
The attached dissertation research proposal was submitted to this office to generate approval for access to prisoners' health records for identification of subjects to be included in the study.

Since no authorization to release confidential information will be possible until the researcher has reviewed records to identify the patients to be included, policy dictates that approval be sought from the Corrections Commission (PD-OHC-42.01, Prisoner Health Record, page 3, #8).

Betty Horton of our office has completed the requisite inquiries and advises me that Bill Kime feels Program Bureau approval does not constitute approval by the Commission. Since there will be no experimentation with patients, I do not feel it is necessary to approach the Protocol Protection Review Committee. It appears that the wisest course is to go directly to the Corrections Commission.

Please note that on page 12 of the research proposal, Mr. Wright has suggested that a master list of subjects be maintained in the office of the Chief Psychologist, R&GC, in case he needs to return to the records. It is customary for the health record department to maintain such a list in accord with ACA and DPH standards, to ensure that the patients' privacy and confidentiality rights are observed. Mr. Wright has been advised that he would have access to that list at any time, and has agreed to this amendment.

I'm sure Mr. Wright would appreciate receiving approval from the Commission as soon as possible.

Thank you for your assistance.

cc: Bruce Wright; R&GC
    Robert Ort, MD, Ph.D.; COHC
    R. Richardson, Ph.D.; Program Bureau
    E. Horton, MHPE; COHC
Cerebral Dysfunction Screening Instrument Research

Bruce Wright, clinical psychologist at SPSM, has requested access to medical records to try to analyze which of two screening devices is most effective in identifying persons who have cerebral dysfunction which has been confirmed by neurological study. This request is for a research of records and involves no experimental protocol for prisoners themselves. The project was approved by the Commission on a motion from Commissioner Waters, supported by Commissioner Eardley.
January 24, 1984

Mr. Bruce Wright
Clinical Psychologist
Reception and Guidance Center
4000 Cooper Street
Jackson, Michigan 49201

Dear Mr. Wright:

This is to inform you that the Corrections Commission, on January 6, 1984, approved your request for access to medical records as described in your proposal to do research on cerebral dysfunction screening instruments.

Best wishes for a successful project.

Sincerely,

DEPARTMENT OF CORRECTIONS

William L. Kime
Deputy Director
Program Bureau

WLK:1s1

cc: J. K. Harness, M.D.
John Prelesnik, Ph.D., R&GC
APPENDIX 11

CORRESPONDENCE GRANTING PERMISSION TO REPRODUCE AND USE COPYRIGHTED MATERIALS
March 1, 1984

Bruce R. Wright
119 South Grennell
Jackson, MI 49203

Dear Mr. Wright:

In response to your letter inquiring about revisions of my preliminary manual for the Bender Gestalt test, it has not been revised since the original publication in 1965. Since then there have been numerous studies using my system and reported in the literature but I do not have a bibliography of these on hand.

You have my permission to reproduce any portion of the scoring manual and tables that you wish in your dissertation, with appropriate credit given. Best of luck with your research.

Sincerely,

Jack D. Hain

Jack D. Hain, Ph.D.
Professor

JDH/da
April 10, 1984

Bruce R. Wright
119 South Grinnell
Jackson, MI 49203

Dear Mr. Wright:

Thank you for your letter of April 3, 1984 requesting permission to reproduce portions of the SYMBOL DIGIT MODALITIES TEST in the appendix of your dissertation.

Permission for the reproduction of the following material is granted provided each reproduction includes the entire copyright notice, at it appears on the published material, along with the following information: "Reprinted by permission of Western Psychological Services, 12031 Wilshire Blvd., Los Angeles, CA 90025."

1. Sections of the manual addressing the administration and scoring of the test on pages 2 and 3.

2. Table 3 entitled "Adult Norms for Mean Written and Oral (retest) Scores by Age Group and Level of Education."

Thank you, again, for your interest in the SYMBOL DIGIT MODALITIES TEST and, best of luck in the successful completion of your studies.

Sincerely,

Kathleen McWilliams
Rights and Permissions

KM:se
April 26, 1984

Bruce R. Wright, M.Ed.
119 South Grinnell
Jackson, MI 49203

Dear Mr. Wright:

Permission has been granted to reproduce the test figures if taken from Plate I, page 4 of Dr. Lauretta Bender's monograph, A Visual Motor Gestalt Test and Its Clinical Use.

Permission is contingent on your listing this as taken from the Bender® Visual Motor Gestalt Test published by the American Orthopsychiatric Association, 1938.

Thank you for your interest and cooperation.

Sincerely yours,

Marion F. Langer, Ph.D.
Executive Director
LIST OF REFERENCES
LIST OF REFERENCES


Peek, R. M., and Quast, W. A. "A Scoring System for the Bender Gestalt Test." Minneapolis, Minnesota: the authors, 1951.


VITAE

NAME: Bruce R. Wright

DATE OF BIRTH: April 9, 1939

PLACE OF BIRTH: Kalispell, Montana

EDUCATION:

1964 Bachelor of Arts - Biology
Union College
Lincoln, Nebraska

1969 Master of Education - Educational Psychology
Wayne State University
Detroit, Michigan

1984 Doctor of Education Candidate
Educational Psychology and Counseling
Andrews University
Berrien Springs, Michigan

PROFESSIONAL EXPERIENCE:

1965-70 Classroom Teacher
Michigan Conference of Seventh-day Adventists

1970-75 Staff Psychologist
Plymouth Center for Human Development
Northville, Michigan

1975-76 Unit Supervisor and Behavioral Clinician
Northern Indiana State Hospital and
Developmental Disabilities Center
South Bend, Indiana

1977 Psychology Assistant
Apple Creek State Institute
Apple Creek, Ohio

1977-84 Clinical Psychologist
State Prison of Southern Michigan
Jackson, Michigan