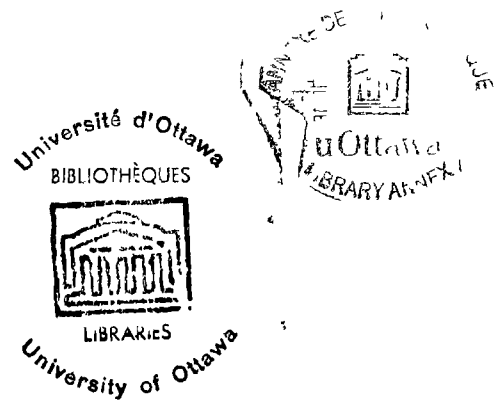


A STUDY OF THE RATIONALE
OF THE WAIS BLOCK DESIGN SUBTEST

by Jean Marie Labrecque

Thesis presented to the School of Psychology
and Education of the University of Ottawa as
partial fulfillment of the requirements for
the degree of Master of Arts in Psychology.



Ottawa, Canada, 1961

UMI Number: EC55355

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI[®]

UMI Microform EC55355
Copyright 2011 by ProQuest LLC
All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ACKNOWLEDGEMENTS

This thesis was prepared under the guidance of Doctor Maurice Chagnon whose assistance and encouragement have been greatly appreciated.

The writer wishes to express his gratitude to all who have assisted him in the realization of his undertaking. He especially acknowledges Mother St. Gabriel, C. N. D., Director of Studies, Notre Dame High School, and the psychologists who have helped in the collection of data.

CURRICULUM STUDIORUM

Jean Marie Labrecque was born on May 3rd, 1930, in Beaumont, Quebec. He went to primary school in Beaumont, studied at Levis College, and later attended the University of Ottawa where he obtained his Bachelor of Arts degree in the year 1954. He also obtained from the same University a Bachelor of Science degree in Physical Education and a Bachelor of Education, both in the year 1956.

TABLE OF CONTENTS

Chapter	page
INTRODUCTION	vi
I.-STATEMENT OF THE PROBLEM.	1
1. Importance of the study	2
2. The choice of Block Design	10
II.-REVIEW OF THE LITERATURE.	17
1. Hypothetical rationales	17
2. Experimental findings	27
3. List of interpretative hypotheses	38
4. Selection of the hypothesis to be investigated	42
III.-THE EXPERIMENTAL DESIGN	47
1. Test under investigation	47
2. Selection and description of tools	53
3. Selection and description of the population	71
4. Statistical methods of analysis	79
IV.-PRESENTATION AND ANALYSIS OF DATA	86
1. Presentation of data	86
2. Analysis of data	99
SUMMARY AND CONCLUSIONS	105
BIBLIOGRAPHY.	109
 Appendix	
1. THE RAW DATA	114
2. ABSTRACT OF <u>A Study of the Rationale of the WAIS</u> <u>Block Design Subtest</u>	120

LIST OF TABLES

Table	page
I. - Age Distribution of Sample in Intervals of Six Months	77
II. - Range of Scores, Means, and Standard Deviations for Each Test of the Battery with a Sample of One Hundred and Twenty- Two Subjects.	88
III. - Test of Linearity of Regression for the Simple Correlations Between Block Design and the Tests of Reasoning Ability. . . .	89
IV. - Test of Linearity of Regression for the Simple Correlations Between Block Design and the Tests of Reasoning Ability. . . .	90
V. - Correlations Between the Various Elements of the Test Battery and the Total Score of Block Design, with a Sample of One Hundred and Twenty-Two Subjects	92
VI. - Test of Significance for the Simple Correlations Between Block Design and Each of the Independent Variables	94
VII. - Test of Significance for the Obtained Multiple Correlation Coefficient with the Dependent Variable.	96
VIII. - Amount of Variance Loading of Each Test in Block Design.	97
IX. - Raw Data Obtained from the Test Battery . .	115
X. - Correlations Between and Among the Dependent and Independent Variables	119

INTRODUCTION

While the findings of many investigators, concerning the use of the Wechsler Batteries, have helped to bring about a recognition of the usefulness of these tools in diagnosis, numerous articles have been published which disagree strongly with such a use and emphasize the lack of validity of these instruments in differentiating various aspects of personality. This serious lack of agreement has pointed to the insufficiency of experimental studies concerning particular functions measured by the subtests and has indicated the need to investigate individually each one of them.

In the present study, the writer has undertaken to determine one of the functions presumably measured by Block Design, one of the subtests of the Wechsler Adult Intelligence Scale. He is concerned with the rationale or nature of this subtest. The survey of the literature has revealed numerous hypotheses which tend to overlap to a great extent and to create a state of confusion regarding the functions or processes it measures. Only one specific hypothesis will be investigated here.

The first chapter deals with the general problem and is mainly concerned with the importance of the investigation and the selection of the subtest Block Design as the object of the study. The second chapter presents the various

INTRODUCTION

hypotheses available through the literature. The main theoretical and experimental rationales are discussed, and a list of the interpretative hypotheses is suggested to the reader. This chapter also includes the selection of the specific research hypothesis and the reasons for its choice. In chapter three, Block Design, the test under investigation, is considered and the instruments used to examine the hypothesis are described in relation to the elements which they could purportedly measure in Block Design. A subsection deals with the manner of selection of the sample, while the statistical methods of analysis are presented subsequently. The fourth chapter is concerned with the presentation, analysis and discussion of the data obtained.

Finally, a summary of the findings is presented and suggestions for further research are offered.

CHAPTER I

STATEMENT OF THE PROBLEM

A review of the literature indicates that since the Wechsler-Bellevue was first introduced, its use has grown widely and extensively as an aid to psychiatric diagnosis. The desire of the psychologist to diagnose the whole personality and his recognition of the many potentialities of this tool soon led him to investigate its use in that particular field.

A survey on the research with the Wechsler-Bellevue for the years 1950-55, by Guertin, Frank, and Rabin¹ reported studies of Reliability and Validity of the test as a test of intelligence, as well as studies with a more restricted orientation such as short forms, applications with special population, sex differences, refinement and critiques concerning administration and scoring. Great interest was shown in scatter and pattern analysis. The great number of studies mentioned in that review concerning its applicability as a diagnostic aid reveals the constant interest of the clinical psychologists in attempting to evaluate its diagnostic validity.

¹ Wilson H. Guertin, George H. Frank, and Albert I. Rabin, "Research with the Wechsler-Bellevue Intelligence Scale: 1950-55," in the Psychological Bulletin, Vol. 53, No. 3, issue of May 1956, p. 235-257.

However, its usefulness in such situations has been widely criticized from many angles and considerable research designed to validate hypotheses concerning its applicability in differential diagnosis eventuated in inconclusive and inconsistent results.

The present paper represents an attempt to evaluate the rationale of one of the subtests of the Wechsler Adult Intelligence Scale², namely Block Design, with specific reference to one interpretative hypothesis. The writer, thus, intends to measure the element or elements which may be found in it and to determine more specifically its nature.

1. Importance of the Study

The importance of test rationale in clinical psychology has been stressed by several authors who have suggested that the lack of validity of our clinical instruments represents one of the greatest weaknesses of the clinician who is faced with evaluation of personality disorder, diagnosis or prediction of the outcome of treatment.

² David Wechsler, Manual for the Wechsler Adult Intelligence Scale, New York, Psychological Corporation, 1955, vi-110 p.

Wechsler³, while recognizing the necessity of defining the validity of the subtests used in his battery, nevertheless neglected to investigate what they measure and took for granted rationales already suggested in the literature by different writers.

However, more recently, Reichard and Schafer⁴ and especially Rapaport⁵ have focused sharply on the absence of adequate test rationale. Rapaport's work probably represents the most intensive and extensive attempt at formulation of test rationale and shows his concern about such a problem. His writings reveal a serious effort to arrive at further knowledge of the functions tapped by each of the subtests of the Wechsler-Bellevue. He stated:

The importance of test rationale cannot be overly stressed because the interpretation of test rationale depends upon the development of such a rationale which defines the psychological processes underlying the test performance.... Without such a

3 David Wechsler, The Measurement of Adult Intelligence, Third Edition, Baltimore, Williams and Wilkins Company, 1944, vii-258p.

4 S. Reichard and Roy Schafer, "The Clinical Significance of Scatter on the Wechsler-Bellevue Scale", in the Bulletin of Menninger Clinic, Vol. 7, No. 3, issue of May 1943, p. 93-98.

5 David Rapaport, Merton Gill, and Roy Schafer, "The Bellevue Scale", in Diagnostic Psychological Testing: The Theory, Statistical Evaluation, and Diagnostic Application of a Battery of Tests, Vol. 1, Chicago, Year Book Publishers, 1945, p. 5.

rationale, the tests and test results must remain meaningless to the psychiatrist, and a matter of automatic procedure and mechanical comparison to the clinical psychologist.⁶

Watson⁷, along with Rapaport and his co-workers, stressed the importance of research in defining the rationale of the subtests of the Wechsler-Bellevue.

Shapiro, who has also been much concerned about the adequate use of psychological tests for diagnostic purposes, discussed, in a paper presented in 1959, the needs which give rise to the use of the standardized validated tests. He wrote:

It is necessary to know as much as possible about the factors affecting the performance of the task in question, and to what extent it really does involve the function under investigation.⁸

He maintained that validation of a test before it can be used clinically is one of the three essential characteristics of a test. As he pointed out, the error of misclassification arising from unknown, imperfect or low validity are crucial. Every one would agree that a test of perfect reliability is of little use to the clinician if validity

6 Id., op. cit., p. 5.

7 Robert I. Watson, "The Use of the Wechsler-Bellevue Scale: A Supplement", in the Psychological Bulletin, Vol. 43, No. 1, issue of January 1946, p. 66-67.

8 M.B. Shapiro, "An Experimental Approach to Diagnostic Psychological Testing", in the Journal of Mental Science, Vol. 97, No. 409, issue of October 1951, p. 749-750.

is very low. However, many psychologists neglect to consider this factor in the clinical use of the Wechsler-Bellevue.

In the light of the many criticisms directed at the lack of adequate rationale and of poor validity of the subtests of the Wechsler-Bellevue, some writers have attempted to investigate the diagnostic validity of this test. The results are rather inconclusive.

Wittenborn and Holzberg reported their attempt to evaluate the diagnostic validity of the Wechsler-Bellevue test using two different types of criteria with a group of mixed patients from a mental institution. They concluded:

From the data provided by this sample, it appears that the possibilities of using the Wechsler-Bellevue subtests in differential diagnosis or in the prediction of the type of symptoms the patient will manifest are limited and may also be restricted as to be of no practical value. There is certainly little about the present data which would encourage the hope that these subtests would be used profitably in the descriptive diagnosis of an individual.⁹

More recently, Frank, Corris, and Fogel¹⁰ noted that most investigators who have been concerned with the appli-

⁹ J. Wittenborn, and J. Holzberg, "The Wechsler-Bellevue and Descriptive Diagnosis", in the Journal of Consulting Psychology, Vol. 15, No. 4, issue of August 1951, p. 329.

¹⁰ G.H. Frank, C.C. Corrie, and J. Fogel, "An Empirical Critique of Research with the Wechsler-Bellevue in Differential Psychodiagnostic", in the Journal of Clinical Psychology, Vol. 11, No. 3, issue of July 1955, p. 291.

cability of the Wechsler-Bellevue test in differential diagnoses, arrived at inconclusive and inconsistent results. They provide us with tentative explanations which emphasize the lack of validity of that tool, a test of intelligence, in differentiating modalities of personality. In order to attempt to explain further these erratic and confusing results found in the literature, Frank and his associates formulated an alternative hypothesis; they thought that the selection of subjects using their psychiatric diagnosis as a criterion, might have contributed to the conflicting nature of the results of the various investigations. They decided to investigate this aspect of past research data derived from the literature. They noted that "the investigators have been attempting to compare the characteristic performance of one group of unlike subjects with another."¹¹ One of their most obvious conclusions was that

The designs of the majority of the research conducted thus far are incapable of testing the question as to the efficacy of the Wechsler-Bellevue in differential diagnosis.¹²

The writers thus suggested that the research methodology utilized to essay the effectiveness of the Wechsler-Bellevue Scale might be responsible for the inconclusive results and not the instrument itself.

11 Id., op. cit., p. 292.

12 Id., op. cit., p. 293.

These conclusions seem to contradict previous statements found in the literature regarding the weaknesses of the instrument itself and especially with regard to the lack of validity of the subtests due to vague rationale.

It would seem reasonable, however, to agree with these writers that

Psychiatric diagnosis, per se, does not constitute an adequate criterion measure for the selection of subjects since this neglects to take into consideration many of the attenuating factors affecting performance on the Bellevue over and above psychopathology.¹³

Marks'¹⁴ criticisms of the Wechsler-Bellevue used as a diagnostic instrument appears to be one of the most severe to be found in the literature regarding this test. These criticisms stem from his understanding of the criteria required which are basically reliability and validity. He noted that

For justifiable inclusion in a battery a subtest should have a unique contribution, and where the battery is to be interpreted diagnostically, this requirement is quite compelling.¹⁵

Thus each test should as much as possible be saturated with

13 Id., op. cit., p. 293.

14 Melvin R. Marks, "A Criticism of the Use of Wechsler-Bellevue Scale as a Diagnostic Instrument", in the Journal of General Psychology, Vol. 49, First Half, issue of July 1953, p. 143-144.

15 Id., op. cit., p. 144.

a single factor in order to measure different factors or functions.

He reaches the conclusion that the Wechsler-Bellevue Scale is inefficient, lacks validity and univocality because of high subtest intercorrelations and inconsistent rationale of inclusion of subtests which are not separated entities.

Other writers have also pointed out similarly that studies of the mental functioning of various groups demand more extensive and more reliable measures than are provided by various subtests of the Wechsler.

The clinician who comes across such definite statements is compelled to question further his use of this instrument in his routine diagnostic investigations and to investigate more thoroughly the nature of these criticisms. It appears evident to him that the great number of studies reporting on the applicability of the Wechsler-Bellevue Scale in differential diagnosis emphasize the lack of validity of the tool in differentiating different aspects of personality. Its usefulness in such situations has been widely criticized from many angles and this would seem to be one of its most commonly accepted limitations.

On the other hand, it has been suggested by many psychologists who used the Wechsler-Bellevue as a diagnostic instrument that marked variability in subtest scores can be considered as an evidence of mental pathology; they suggested

that certain types of pathology are commonly characterized by conspicuous failures in one or two kinds of mental functions. The profiles, thus, would be profitable aids to diagnosis because they would reveal disorder in certain modalities of intellectual functioning as characterized by the specific failures on the subtests.

However, the tenants of the later hypotheses have difficulty in demonstrating the objectivity of their approach due to the lack of validity of the subtests in measuring specific modalities of intellectual functioning. It was impressively evident, in August 1949, at the Boulder Conference on Training in Clinical Psychology, that the validation of its own techniques and theories was considered to be the most urgent problem confronting clinical psychology. Myers, one of the members of the conference, summed the matter up:

What emerged from the discussion was a sober recognition of the fact that, though society may need the services of clinical psychology, what clinical psychology needs most urgently is more dependable knowledge and more valid methods with which to do its job.¹⁶

The numerous references in the literature to the lack of adequacy of most of our tools should bring about a recognition of the necessity of valid diagnosis and

¹⁶ C.R. Myers, "Prediction on Clinical Psychology", in the Canadian Journal of Psychology, Vol. 4, No. 3, issue of September 1950, p. 102.

prediction. Thorne,¹⁷ Payne,¹⁸ and Meehl¹⁹ have recently expressed their awareness of the weakness of our psychodiagnostic methods and have urged again the research oriented psychologist to assess the validity of our diagnostic instruments. It appears very important to determine the functions or processes underlying the various subtests performances on the Wechsler Adult Intelligence Scale. The present writer is thus justified in investigating the rationale of one of the subtests of the Wechsler Adult Intelligence Scale if it furthers our understanding of the intellectual function or functions involved in its solution.

2. The Choice of Block Design

Although it has been mentioned earlier that the purpose of this investigation was to study the rationale of Block Design, one of the subtests of the Wechsler Adult Intelligence Scale, no reason was given for its choice.

¹⁷ Frederick C. Thorne, "Professional Insecurity: An Occupational Neurosis of Clinical Psychologists", in the Journal of Clinical Psychology, Vol. 12, No. 4, issue of October 1956, p. 404.

¹⁸ Robert W. Payne, "Diagnostic and Personality Testing in Clinical Psychology", in the American Journal of Psychiatry, Vol. 115, No. 1, issue of July 1958, p. 25-26.

¹⁹ Paul E. Meehl, "The cognitive Activity of the Clinician", in American Psychologist, Vol. 15, No. 1, issue of January 1960, p. 19.

Many reasons have led the writer to select Block Design as the object of his investigation. First of all, the numerous references to this test in the literature reveal both its importance in clinical diagnosis and a serious lack of agreement concerning its rationale. Secondly, its consistent reoccurrence in the short forms as one of the leading subtests of the Wechsler, as well as its use for diagnosis of organic brain pathology, call for a more adequate and more precise knowledge of the intellectual functions or modalities tapped by it. Thirdly, in order to respond to some urgent needs in clinical psychology, which call for more accurate and more objective instruments validated for the purpose of diagnosing specific forms of human behavior, it appeared important to the present writer to investigate further a test which is employed in diagnosing unrelated pathological behavior.

a) The importance of Block Design and the lack of agreement concerning its rationale. -- Since Kohs²⁰ introduced the Block Design test in 1927, and presented it as a measure of intelligence, many references have been made to this test and its use has spread widely attesting its popularity. It has been included in many test batteries and offered as measures of intelligence. As a single test, it

²⁰ Kohs, Intelligence Measurement, New York, The MacMillan Company, 1927, xii-312p.

has also become popular in the investigation of brain pathology.

Numerous investigators have offered various interpretative hypotheses in order to explain what it measures. While these studies succeeded in isolating certain factors for the whole Wechsler battery, difficulties in naming these factors have resulted in a complexity of hypotheses which seem to overlap to a great extent. Moreover, so far it has been impossible to find any study whose purpose was to investigate the rationale of Block Design alone.

b) Its importance in short forms and its use for diagnosis of organic brain pathology. -- The interest in and need to develop short form scales had arisen very soon after the publication of the Wechsler-Bellevue. The hope was to devise an abbreviated scale which would meet the requirements of time economy, accuracy of intelligence rating and diagnostic screening capacity. Thus number of studies were made with various samples in order to determine which combination of subtests had highest correlation with the Full Scale.

McNemar²¹, reporting on the best combinations found in the standardization group of Wechsler, mentioned that Block Design appears as the most stable subtest, being present in the best three groups of correlations for the

21 Quinn McNemar, "On Abbreviated Wechsler-Bellevue Scales", in the Journal of Consulting Psychology, Vol. 14, No. 2, issue of April 1950, p. 80.

four different teams. A study on the Wechsler Adult Intelligence Scale by Maxwell²² reported the same trend as for the Wechsler-Bellevue. Maxwell, using a variation of McNemar's formula, determined validities of all possible abbreviated WAIS scales of 2, 3, 4, and 5 subtests and observed that the subtest Block Design appears in the four combinations of subtests having the highest correlation with the Full Scale. On the basis of these similarities, it is reasonable to conclude that Block Design has a major contribution to make in the constitution of both scales and that further knowledge of the intellectual modalities it measures would improve the efficacy and accuracy of our predictions.

Numerous investigators, who were interested in diagnosing organic brain pathology with the Wechsler-Bellevue, have indicated repeatedly that the performance on Block Design is impaired in organic brain cases. Aita and his co-workers²³ have reported that Block Design was one of the tests markedly affected by brain damage. Two studies by

22 Eileen Maxwell, "Validities of Abbreviated WAIS Scales", in the Journal of Consulting Psychology, Vol. 21, No. 2, issue of April 1957, p. 125.

23 J.A. Aita, S.G. Armitage, R.M. Reitan, and A. Rabinowitz, "The Use of Certain Psychological Tests in the Evaluation of Brain Injury", in the Journal of General Psychology, Vol. 37, First Half, issue of 1947, p. 27-35.

Allen²⁴²⁵ supported such findings. Similarly, severe impairment of the performance of brain-injured patients on Block Design was demonstrated by Klebanoff, Singer, and Wilensky²⁶ who analyzed psychological consequences of brain lesions and ablations. Goldstein and Scheerer²⁷ have also made an extensive use of Block Design in the investigation of organic deterioration.

c) The need for more accurate and more objective instruments. -- The studies previously reported, as well as a great number of references concerning the Wechsler-Bellevue and the WAIS, demonstrate the persistent preoccupation with scatter and pattern analysis and indicate a constant effort at better understanding of psychiatric syndromes through study of variations in functioning efficiency.

24 R.M. Allen, "The Test Performance of Brain Injured", in the Journal of Clinical Psychology, Vol. 3, No. 3, issue of July 1947, p. 225-230.

25 Idem, "The Test Performance of the Brain Diseased", in the Journal of Clinical Psychology, Vol. 4, No. 3, issue of July 1948, p. 281-284.

26 S.G. Klebanoff, J.L. Singer, and H. Wilensky, "Psychological Consequences of Brain Lesions and Ablations", in the Psychological Bulletin, Vol. 51, No. 1, issue of January 1954, p. 1-41.

27 Kurt Goldstein and Martin Scheerer, "Abstract and Concrete Behavior: an Experimental Study with Special Tests", in the Psychological Monographs, Vol. 53, No. 2, Whole No. 239, 1941, 151 p.

The reports in the literature are far from agreeing completely as to the diagnostic significance of an impairment of the performance on Block Design. As noted previously, various investigators have reported that it is a good instrument for the diagnosis of organic brain damage. On the other hand, Rapaport, in his analysis of the data collected from his numerous groups of subjects, concluded that "Great impairment on Block Design is a generally reliable indicator of depressive trends, the severity of which may be estimated by the extent of impairment."²⁸ In contrast to depressed subjects, schizophrenics did not show any significant impairment. These findings most obviously contradict Goldstein and Scheerer's²⁹ reports. They mentioned that the Cube test is actually a test of abstract and concrete behavior, particularly applicable to organic deterioration and schizophrenia which shows deficiencies in abstract thinking.

All these studies demonstrate the importance of Block Design in the investigation of pathological intellectual functioning and indicate the necessity of studying its rationale more adequately.

²⁸ David Rapaport, Merton Gill, and Roy Schafer, op. cit., p. 280.

²⁹ Kurt Goldstein and Martin Scheerer, op. cit., p. 31, passim.

In this first chapter the writer discussed the importance of test rationale in clinical psychology and more specifically the need to investigate the intellectual functions underlying the subtest of the Wechsler Battery. The reasons for the choice of Block Design as the object of the study were presented to the reader.

CHAPTER II

REVIEW OF THE LITERATURE

The writer attempted to review all the literature available which might be directly related to the investigation of the rationale of Block Design, a subtest of the Wechsler Adult Intelligence Scale. It is expected that the sample of literature reviewed here incorporates most, if not all, of the ideas and hypotheses which may be pertinent to this study. The discussion will present first the purely hypothetical rationales. Then the experimental findings will be discussed. These will be followed by a presentation of the list of interpretative hypotheses suggested in the literature. Finally, there will be a discussion of the hypothesis to be investigated. This approach necessitates some concession to the chronological sequence of the development of the subject and does not require a rigid listing of opinions and experiments in order to establish possible relationships between the various ideas.

1. Hypothetical Rationales

When Kohs¹ first introduced the Block Design Test, he offered it as a measure of non-verbal intelligence.

¹ S.G. Kohs, Intelligence Measurement, New York, The MacMillan Company, 1927, xii-312p.

Indeed, its high correlation of $.84 \pm .01$ with the well known Stanford-Binet, of $.57 \pm .03$ with the Trabue Language Completion Test, and of $.71 \pm .02$ with the Military Test, sufficiently revealed the new test's potentiality as a measure of intellectual functioning.

Objecting to different theories of intelligence, then existent, Kohs thought that

Intelligent acts of all sorts require both an analysis of the situation which confronts one, a critical inquiry into methods of solving the problem, and a final synthesizing of details into a consistent whole.²

He pointed out that a careful analysis of the demands made upon the mental operations of the individual being tested by the Block Designs

Will clearly reveal that attention, adaptation and auto-criticism are all involved in the successful accomplishment of each task. That point in the grade series of designs at which a child will begin failing to achieve further success, will be a rough measure of the development of his ability to attend, to adapt, and to critically survey his general plan of performance and his ultimate accomplishment.³

In further explanation of the functions enclosed in his concept of intelligence, he suggested that

If "intelligence" involves the following mental operations: analyzing, combining, comparing, deliberating, completing, discriminating, judging,

2 Id., op. cit., p. 169.

3 Id., op. cit., p. 172.

criticizing, and deciding, then the block design tests may, with justice, be said to call upon the functioning of intelligence and to that extent they are a measure of that mental capacity.⁴

Thus, from Kohs' concept of intelligence, it would seem that the process of solving the Block Design tasks involves most, if not all, mental functions. However, since he considered the processes of analyzing and synthesizing of the greatest importance in mental functioning, it seems logical to hypothesize that he recognized these two processes as forming the cornerstone of the Block Design rationale.

However, Kohs' attempt to measure intelligence with the Block Design Test was criticized by Brown⁵ who found a low correlation of $.46 \pm .04$ between the Kohs Test and the National Intelligence Test. He concluded that the test had doubtful clinical value. Brown's main criticism was directed at the standardization procedures.

Wechsler⁶ would seem to disagree also with Kohs as to the ability of Block Design to measure intelligence because he believes that no single test can measure it. However, he, himself, admits that "Block Design is not only an

⁴ Id., op. cit., p. 195.

⁵ A.J. Brown, "An Inquiry into the Standardization of the Kohs' Block Design Test", in the Journal of Applied Psychology, Vol. 14, No. 2, issue of April 1930, p. 180-181.

⁶ David Wechsler, The Measurement of Adult Intelligence, Third Edition, Baltimore, Williams and Wilkins Company, 1944, p. 91.

excellent test of general intelligence, but one that lends itself admirably to qualitative analysis."⁷ He explained that

The test turns out to be our best single performance item. It conforms to all statistical criteria for a good test. It not only correlates well with total score (.73) and individual test items, but is one of the few performance tests that seemingly does measure very much the same sort of thing that verbal tests measure.... It is hard to explain the basis for its merits other than that the reproduction of the type of design devised by Kohs in some way involves both synthetic and analytic ability.⁸

However, although he made no serious attempt to investigate the validity of these assertions, he later on suggested that

The role of the abstract approach has been greatly overestimated. It is, of course, reflected in certain types of cases, but in most, low scores on Block Design are due to difficulty in visual-motor coordination.⁹

As it has been shown by Kohs and Wechsler, there is evidence that Block Design correlates highly with varied measures of intelligence. There is also evidence that both Block Design and the Progressive Matrices measure, to a certain extent, similar functions. The latter test is

7 Id., op. cit., p. 92.

8 Id., op. cit., p. 92.

9 David Wechsler, The Measurement and Appraisal of Adult Intelligence, Fourth Edition, Baltimore, Williams and Wilkins Company, 1958, p. 80.

regarded by its author as measuring

A person's capacity at the time of the test to apprehend meaningless figures presented for his observation, see the relations between them, conceive the nature of the figure completing each system of relations presented, and, by so doing, develop a systematic method of reasoning.¹⁰

This test, which is not considered as a measure of intelligence by Raven, however, "correlates 0.86 with the Terman-Merrill Scale and has been found to have a "g" saturation of 0.82." ¹¹ Stacey and Gill¹² reported a correlation of .60 for the Progressive Matrices and Block Design with a group of 172 subnormal adult subjects. Another investigation by Levine and Iscoe¹³ indicates a similar trend for a group of 60 undergraduates whose average age was 19.1. The correlation between these two tests was .63.

A substantial correlation between Block Design and the full score on the Wechsler Adult Intelligence Scale was

¹⁰ J.C. Raven, Guide to the Standard Progressive Matrices, Sets A, B, C, D, and E, London, H.W. Lewis and Sons, 1958, p. 1.

¹¹ Id., op. cit., p. 2.

¹² C.L. Stacey and Marie R. Gill, "The Relationship Between Raven's Colored Progressive Matrices and Two Tests of General Intelligence for 172 Subnormal Adult Subjects", in the Journal of Clinical Psychology, Vol. 11, No. 1, issue of January 1955, p. 87.

¹³ B. Levine and I. Iscoe, "A Comparison of Raven's Progressive Matrices (1938) with a Short Form of the Wechsler-Bellevue", in the Journal of Consulting Psychology, Vol. 18, No. 1, issue of February 1954, p. 10.

also reported by Wechsler.¹⁴ The correlation was found to be .72 for two hundred subjects in the 18-19 age group.

Goldstein and Scheerer followed a similar trend of thought to Kohs. In a monograph¹⁵, they described Block Design as a measure of abstract ability or attitude. They claimed that the normal individual may assume two distinct types of behavior, the concrete and the abstract, and that Block Design can measure a qualitative difference between abstract and concrete levels of mental behavior. They suggested that the abstract attitude

Implies conscious activity in the sense of reasoning, awareness and self-account of one's doing. We transcend the immediately given situation, the specific aspect or sense impression: we abstract common from particular properties; we are oriented in our action by a rather conceptual viewpoint, be it a category, a class, or a general meaning under which the particular object before us falls. We detach ourselves from the given impression, and the individual thing represents to us an accidental example or representation of a category.¹⁶

The concrete attitude, on the other hand, does not imply conscious activities in the sense of reasoning, awareness

14 David Wechsler, WAIS Manual: Wechsler Adult Intelligence Scale, First Edition, New York, the Psychological Corporation, 1955, p. 15.

15 Kurt Goldstein and Martin Scheerer, "Abstract and Concrete Behavior: an Experimental Study with Special Tests", in the Psychological Monographs, Vol. 53, No. 2, Whole No. 239, 1941, 151 p.

16 Id., op. cit., p. 3-4.

or self-control on one's doing. It has none of the above mentioned characteristics of the abstract attitude but it is characterized by rigidity, lack of shifting and quite often by fluidity.

There is an obvious lack of demonstrable evidence for many of the statements made by Goldstein and Scheerer. They provided the reader neither with norms, quantitative data nor standardization data; moreover, there was no description of the population from which the hypotheses were drawn; one could only find qualitative descriptions of the behavior of individual cases based exclusively on the qualitative findings. Under these circumstances, it is evident that the validity of the test becomes a subjective proposition and is liable to change with each interpreter. Thus, as a result, their rationale of Block Design can only be accepted as an hypothesis which needs to be investigated scientifically.

Different attempts to quantify the Block Design Test as a measure of abstract and concrete behavior have been reported in the literature. Probably one of the most significant contributions was made by Grassi's work¹⁷. Essentially, the main innovation was that the patient copy not

17 Joseph R. Grassi, "The Fairfield Block Substitution Test for Measuring Intellectual Impairment", in the Psychiatric Quarterly, Vol. 21, No. 3, issue of July 1947, p. 474-489.

a drawing but a set of blocks taking into account the characteristic of the three faces. An attempt was made to measure both concrete and abstract reproduction at two levels. Grassi maintained that the Fairfield Block Substitution Test depicts intellectual impairment due to functional or organic defects. He thus fully agreed with Goldstein and Scheerer as to the rationale of Block Design. He, however, introduced some important innovations and used sound standardization procedures which may bring new light in the interpretation of results.

Rapaport's clinical findings with Block Design seem to have led him to believe that this is a test of visual-motor co-ordination. He indicated that visual-motor co-ordination in the Block Design subtest can be characterized as occupying a middle position between its role in Digit Symbol and its role in Object Assembly.¹⁸ He believed that visual-motor co-ordination and particularly visual organization play a 'productive' role in Object Assembly while, in Digit Symbol the role of visual-motor co-ordination is an 'imitative' one. On the other hand, in Block Design, the

¹⁸ David Rapaport, Merton Gill, and Roy Schafer, "The Bellevue Scale", in Diagnostic Psychological Testing: The Theory, Statistical Evaluation, and Diagnostic Application of a Battery of Tests, Vol. 1, Chicago, Year Book Publishers, 1945, p. 271.

role of visual-motor co-ordination is 'reproductive'. It would appear from these considerations that Rapaport offers a completely new rationale for the test. However, it may be misleading to come to this conclusion in the light of some further comments he made:

Our emphasis on visual-motor co-ordination should not veil the fact that the work of visual organization in this test implies concept formation as evidence in the analytic and synthetic phases of performance.¹⁹

Indeed, he appears to believe that Block Design is a test of concept formation as he writes:

These tests which Goldstein considers tests of "abstract attitude" are considered by the present authors as tests of concept formation. The analytic and synthetic task implied by Block Design is related to the conceptual analysis underlying successful performance of concept formation tests.²⁰

There appears to be no strict evidence, from his writing, of what he considers the main rationale. It may be that he consider visual-motor co-ordination as the necessary and most important step in performing the task, and as leading gradually to the formation of a concept about a specific design; thus, what Block Design would measure would be that first step. On the other hand, Block Design might measure the analytic and synthetic phases of

19 Id., op. cit., p. 275.

20 Id., op. cit., foot note 101, p. 273.

performance; visual-motor co-ordination would appear, then as a result of concept formation. However, it is usually recognized that he considers Block Design as a measure of visual-motor co-ordination, and as such, he adds a new element to the rationale already hypothesized.

Other writers have expressed opinions similar to the ones already discussed and have contributed to a much smaller degree to the formation of a Block Design rationale.

The rationales that have been discussed so far are all hypothetical ones; they are based on clinical observation or empirical findings although some attempt was made to validate them through qualitative analysis of performance in individual cases. However, there appears to be some similarity of opinion in regards to the functions measured by Block Design. Indeed, Kohs, Wechsler, and Rapaport, all suggested that the analytic and synthetic functions are somehow being measured by the Block Design Test. Goldstein and Scheerer, in somewhat different terms, also suggested that these functions are revealed by the subject's performance since the abstract ability is fundamentally dependent upon them. Rapaport, however, brought a new element when he suggested that visual-motor co-ordination is the main function measured in Block Design performance.

2. Experimental Findings

Balinsky²¹ reported a study of the Wechsler-Bellevue consisting of an analysis of the various factors appearing in different age groups and their importance for such groups. Using 600 cases from the standardization group of the Wechsler, he found that Block Design measures different functions or modalities at different ages. He suggested that Block Design measures in a certain way visualization involving perception of form and space relations especially at the age of 12 where the factor has its highest loading in Block Design. He called this factor the non-verbal or performance factor. It appears consistently in every group except at the age of 9, and it has, in each of these groups, a high loading in Block Design. The highest loadings appear in the age groups 12, 25 to 29, 35 to 44, and 50 to 59; they are respectively .602, .743, .707, and .428. A 'G' factor, which seems to correspond to the Spearman 'G', is present in age group 9 with its highest loading of .801 in Block Design. A factor called reasoning with high loading in Block Design is also found in age groups 35 to 44 and 50 to 59. For the age group 15 he also found the factor 'awareness of

21 Benjamin Balinsky, "An Analysis of the Mental Factors of Various Age Groups from Nine to Sixty", in the Genetic Psychology Monographs, Vol. 23, First Half, issue of February 1941, p. 191-234.

environment' with significant loadings in Information, Object Assembly, and Block Design.

It is evident from Balinsky's findings that different factors seem to appear at different ages. This is explained by the author as a change and reorganization of mental traits and functions over a span of years. This has a direct bearing on the rationale of Block Design since it might be measuring different functions at different ages.

Findings by Smith²² indicate the importance of the reasoning element in the Block Design subtest. In a study of grade 8 and 9 students, he found a correlation of .50 between Block Design and Thurstone's test of reasoning ability. The correlation was significant at the .01 level of confidence.

Divergent findings were reported by Becker²³ who found a very low and non-significant relationship, for a group of 50 high school juniors, between the Thurstone's space factor and Block Design. The Block Design correlation with reasoning factor was still lower. These results

22 Arthur Smith, A Comparison of the SRA Primary Mental Abilities Test with the Wechsler-Bellevue Intelligence Scale, unpublished Master's thesis presented to Illinois State Normal University, 1947, p. 17.

23 George Becker, The Relationship Between the Thurstone, SRA Primary Mental Abilities and the Wechsler-Bellevue Intelligence Test, unpublished Master's thesis presented to the Fordham University, New York, 1948, p. 27.

contradict the findings reported by Estes²⁴ who claimed that Block Design is the most valid test, among a battery of five tests of spatial ability, in predicting achievement in descriptive geometry. An experiment was conducted in which the Wechsler-Bellevue Block Design Test and the Minnesota Paper Form Board were administered. The correlation found between these two tests was .40. DeWit²⁵ appears to agree with both Estes and Balinsky. In a study of the results of 63 college students between the ages of 18 and 24, recruited from three art education courses, he revealed a correlation of .624 between Block Design and Revised Minnesota Paper Form Board, a measure of spatial perception ability. He also found a significant relationship between Block Design and drawing ability and concluded that this test seems to be the best predictor of drawing ability if used with college students.

This spatial factor was also identified by Corter²⁶

24 Stanley G. Estes, "A Study of Five Tests of Spatial Ability", in the Journal of Psychology, Vol. 13, Second Half, issue of April 1942, p. 270.

25 Frederick DeWit, The Block Design Test Drawing Ability, unpublished Master's thesis presented to the Pennsylvania State College, Pennsylvania, 1951, p. 4-5.

26 Harold M. Corter, "A Factor Analysis of Some Reasoning Tests", in the Psychological Monographs, Vol. 66, No. 8, issue of March 1952, p. 1-31.

who thinks that the ability to recognize size, shape, and congruence plays a predominant role in Block Design performance. He also believed that Block Design measures 'productiveness' as well as 'concept' ability and 'flexibility'. Concept ability apparently involves the ability to recognize essential similarities, to abstract and generalize and to think inductively. This factor is very similar to the 'G' factor suggested by Kohs, Wechsler, and Balinsky. The factor flexibility was also believed to be found in Block Design both by Goldstein and Meili.

Meili²⁷ published a report on studies involving Block Design using two groups of children between the ages of 9 and 13 and one group of 30 adults. The reporter identified the factor 'plasticité' and the factor 'complexité' in Block Design. The factor plasticity, which is the most important, is defined as the ability to break a structure and reorganize the material into a new organization or structure. Meili explained: "Nous pensons que cette facilité de briser une structure ou tout simplement de la transformer, dépend de la plasticité de l'organisation mentale".²⁸

²⁷ Richard Meili, "L'analyse de l'intelligence", in Les Archives de Psychologie, Vol. 31, No. 121, issue of March 1946, p. 1-64.

²⁸ Id., op. cit., p. 20.

This factor corresponds to the ability to shift reported by Goldstein. Meili indicated that this factor is more important for children than for adults because children have more difficulty in breaking down a structure in order to analyze its components.

The factor complexity is defined in the following terms:

Il se rapporte à la capacité de réaliser clairement et avec précision des structures complexes.²⁹ ...Il détermine la facilité avec laquelle une personne peut réaliser des structures complexes ou le degré de complexité que les ensembles intellectuels d'un sujet peuvent atteindre.³⁰

This factor seems to be associated with the ability to analyse and to synthesize suggested as a rationale of Block Design by Kohs and Wechsler, and implied by Rapaport.

Birren's³¹ findings in his investigation of an elderly sample revealed that four main factors have significant loading in Block Design. These included a verbal factor indicating ability to deal with stored information thought to be similar to the 'G' factor found by Balinsky, as reported

29 Id., op. cit., p. 21.

30 Id., op. cit., p. 47.

31 J.E. Birren, "A Factorial Analysis of the Wechsler-Bellevue Scale Given to an Elderly Population", in the Journal of Consulting Psychology, Vol. 16, No. 5, issue of October 1950, p. 399-405.

earlier. Another factor called closure was believed to be based mainly on visual processes and to include perceptual speed and spatial relations. It involved the ability to fuse the perceptual field into a single percept. This factor has been identified by Balinsky as a non-verbal factor. Birren's analysis also reported a factor called memory which he associated with Balinsky's reasoning. A final factor of induction would appear to support Goldstein's rationale of Block Design.

Davis³² with a population of 356 junior high school students isolated a factor of visualization which has been interpreted by Birren, Cohen, and Balinsky respectively as closure, non-verbal, and performance. This factor was also reported by Simkin, but his findings were not available for analysis. According to Davis, Block Design would also measure a mechanical element as well as perceptual speed and information. The mechanical and perceptual functions have been reported by many authors under different names; the factor mechanical might be what Rapaport called reproductive visual-motor co-ordination.

32 Paul C. Davis, "A Factor-Analysis of the Wechsler-Bellevue Scale", in the Educational and Psychological Measurement, Vol. 16, No. 1, issue of Spring, 1956, p. 127-146.

In his study, Bolle³³ found that the ability to evaluate and organize visual material had a high loading in Block Design. He suggested that evaluation, pictorial, planning, visual-motor co-ordination and ability to pick up information were all involved in Block Design performance.

Holzberg and Belmont compared Block Design to specific factors found in the Rorschach. The results did not support the hypothesis that Block Design is significantly related to W % and M. These writers believed that all measures were related to conceptualized thinking and to the abstract attitude. They also hypothesized that Block Design and M were a good measure of creative thinking. However, their results were not conclusive as they explain:

The three Rorschach factors which were predicted to be related to the Block Design test do not show a significant relationship.... Several Rorschach factors for which no relationship with Block Design was predicted yield significant correlations. One of the refined measures of F % shows a relationship at the .05 level of confidence. In addition FC-, CF-, is negatively related at the .01 level and FC approaches significance at the .10 level of confidence.³⁴

33 Arthur Bolle, A Factor Analysis of the Echelle d'Intelligence Ottawa-Wechsler, unpublished Master's thesis presented to the Institute of Psychology and Education of the University of Ottawa, 1954, p. 59.

34 J. Holzberg and L. Belmont, "The Relationship Between Factors of the Wechsler-Bellevue and Rorschach Having a Common Psychological Rationale", in the Journal of Consulting Psychology, Vol. 16, No. 1, issue of February 1952, p. 28.

Karson³⁵ and his co-workers reported a somewhat similar investigation with a group of 50 male patients whose age mean was 31.1 years. Their purpose was to search for common factors in a wide variety of clinical psychological instruments which are assumed to measure the abstract thinking abilities of psychiatric patients. Their results were somewhat inconclusive. However, they found three factors of which only one, the factor of abstract ability or general intelligence, had loading in Block Design. This finding supports the rationale formulated by Kohs, Wechsler, Goldstein, and to a lesser degree, by Rapaport and Birren.

A more serious attempt was made by Cohen³⁶ to investigate the functions measured by the Wechsler-Bellevue with psychiatric patients between the age of 20 and 40. Among the factors isolated by Cohen, two have direct bearing on the rationale of Block Design. A non-verbal organization factor, with loading in Block Design in all three groups,

35 Samuel Karson, Pool and Kenneth Bryner, "The Abstract Thinking Abilities of Mental Patients", in the Journal of Clinical Psychology, Vol. 13, No. 2, issue of April 1957, p. 126-132.

36 Jacob Cohen, A Comparative Analysis of the Factors Underlying Intelligence Test Performance of Different Neuropsychiatric Groups. Multiple Factor Analyses of the Wechsler-Bellevue Intelligence Scale Performance of Schizophrenic, Psychoneurotic, and Brain-Damaged Groups, microfilmed Doctor's thesis presented to the University of New York, New York, March 1950, vii-162 p.

was interpreted as "involving the ability to organize non-verbal material, visually perceived into meaningful wholes against a time limit."³⁷ More specifically, Block Design was thought to involve abstract spatial visualization in the organizational task and some kind of Gestalt. This non-verbal organization factor had been defined as performance by Balinsky.

Another factor, factor C, or freedom from distractability, had loading in Block Design in the psychoneurotic group only.³⁸ Cohen³⁹ thought that the complexity of Block Design as seen in the many relationships among the block faces, and the many possibilities of pattern organization, as well as the element of speed, were factors that permit the test to measure the sensitivity of the effect of distractability. This factor had not been isolated previously as such. However, Balinsky, as was reported earlier in this study, found, in comparable age groups, a factor which he identified as a memory factor. Cohen suggested that the discrepancy may well be solely on the interpretive level, since freedom from distractability is a necessary prerequisite to memorization. Balinsky's factor had no loading in

37 Id., op. cit., p. 69.

38 Id., op. cit., p. 73.

39 Id., op. cit., p. 75.

Block Design.

In a later report of his findings, Cohen, realizing the difficulty involved in interpreting the factors found, concluded that Block Design

Offers some ambiguity in its interpretation. It measures Non-Verbal Organization well in the Schizophrenic and brain-damaged, and no other factor. In the psychoneurotic, however, it measures both this factor and Freedom from Distractability, both relatively poorly.... Despite the ambiguity of common factor interpretation, this test is useful because of its correlation with "G". In all the groups, it is at least the most closely related to "G" among the performance tests, and for the schizophrenics, it is best in this regard to all the test. The Block Design Test is therefore an excellent single measure of present general functioning.⁴⁰

Some of the ideas in the interpretation presented above were suggested by Kohs, Wechsler, and Rapaport in their discussion of the rationale of Block Design.

An attempt, by Cohen⁴¹, to study the factorial structure of the Wechsler Adult Intelligence Scale between early adulthood and old age resulted in the identification of five factors in all but the 60 to 75 years old group, which proved deficient in factor E. The perceptual

⁴⁰ Jacob Cohen, "A Factor Analytically Based Rationale for the Wechsler-Bellevue", in the Journal of Consulting Psychology, Vol. 16, No. 4, issue of August 1952, p. 276.

⁴¹ Jacob Cohen, "The Factorial Structures of the WAIS Between Early Adulthood and Old Age", in the Journal of Consulting Psychology, Vol. 21, No. 4, issue of August 1957, p. 283-290.

organization factor that he identified previously in a study of the Wechsler-Bellevue was present again and Block Design with Object Assembly were the only two subtests loading it consistently in all four age groups. Many factor analyses of the Wechsler-Bellevue have isolated this factor although it has been identified as performance, spatial-perceptual, closure, non-verbal organization or ability to evaluate and organize visual material, and visualization. Cohen explained that

The common content of these interpretations is the non-verbal, perceptual, organizational characteristics. Because of the loadings of Picture Arrangement, and Digit Symbol this factor resists interpretation as a spatial factor. Clearly, however, both speed and quality of perceptual performance and organization are involved; hence it is here named Perceptual Organization. It very likely represents a few highly correlated factors such as Thurstone's Perceptual Speed, Closure, and Spatial Relations, which do not emerge separately because of the paucity of tests in this subdomain of the intellectual sphere.⁴²

These consistent findings seem to disagree with the main hypotheses put forward by such authors as Kohs, Goldstein, and Wechsler who based their rationale on observation and empirical clinical findings. They are not consistent either with other experimental findings reported by Balinsky, Corter, Meili, Birren, and Karson who suggested that Block

⁴² Jacob Cohen, "The Factorial Structures of the WAIS", p. 285.

Design measures some form of reasoning, abstract ability, plasticity, and induction.

Analysis of Cohen's findings suggests that the organization of intellectual functions is essentially invariant between the age of 18 and 54. As he pointed out, "The present findings although incomplete, cast serious doubt on the notion that the general factor "differentiates" into quasi-independent factors by the late teens".⁴³ These findings appear to cast a shadow on Balinsky's claims that mental traits change and undergo reorganization over a span of years and that the same test may measure different abilities or modalities at different ages.

Thus far, it seems very difficult to isolate factors that will be generally agreed upon. The studies reported have attempted to identify factors pertaining to the whole test and there has been very few attempts to study separately each individual subtest. Evidently, there is a need for studies oriented toward the evaluation of the specific functions measured by each subtest.

3. List of Interpretative Hypotheses

The present survey of the literature has revealed a great number of hypotheses based on either empirical findings

⁴³ Jacob Cohen, "The Factorial Structures of the WAIS", p. 281.

or on experimental evidence. Very often, however, different suggested rationales appear to be representing a similar function. Difficulties in naming derived factors have resulted in a complexity of hypotheses which seem to overlap to a great extent. An attempt will be made here to put under the same category suggested rationales which appear to be closely related and to represent similar functions or mental processes. These hypotheses could be grouped in the following manner:

1. Block Design measures "general intelligence", (Kohs, Wechsler, Balinsky); "ability to analyse and synthesize", (Kohs, Wechsler, Schafer); "ability to abstract, generalize and think inductively", (Corter); "concept formation" or "analytic and synthetic behavior", (Rapaport); "concept ability" or "ability to recognize essential similarities, to abstract and generalize, think inductively", (Corter); "complexity" or "ability to realize clearly and concisely. complex structures", (Meili).

2. Block Design measures "abstract and concrete approach or attitude", (Goldstein and Scheerer); "abstract approach" (Wechsler, Karson, Grassi); "abstract conceptual approach", (Schafer).

3. Block Design measures "reasoning", (Balinsky, Goldstein and Scheerer); "evaluation and planning", (Bolle).

4. Block Design measures "flexibility", or "ability to change one's mental set", (Corter); "plasticity" or "ability to break a structure, to reorganize the material into a new structure", (Meili); "ability to shift", (Goldstein and Scheerer).

5. Block Design measures "spatial ability", (Estes); "spatial-perception ability", (DeWit); "spatial", or "ability to recognize size, shape, and congruence", (Corter); "abstract spatial visualization", (Cohen); "pictorial", (Bolle).

6. Block Design measures "reproductive visual-motor co-ordination", (Rapaport, Schafer); "mechanical", (Davis); "visual-motor co-ordination, (Bolle).

7. Block Design measures "perceptual organization" or the "organization of non-verbal, visually perceived material against a time limit", (Cohen); "non-verbal organization, (Cohen); "visualization involving perception of form and space relations", "non-verbal", or "performance", (Balinsky); "closure or perceptual speed and spatial relations", (Birren); "visualization", (Davis); "perceptual speed", (Davis); "ability to evaluate and organize visual material", (Bolle); "productiveness", (Corter).

8. Block Design measures "memory", (Birren); "information", (Davis).

9. Block Design measures "gestalt", (Cohen).

10. Block Design measures "freedom from distractibility", (Cohen).

11. Block Design measures "induction", (Birren).

12. Block Design measures "awareness of environment", (Balinsky).

13. Block Design measures "verbal ability" or "ability to store information", (Birren); "ability to pick up information", (Bolle).

The presence of so many factors which have been suggested as rationales for the Block Design Test has created a state of confusion regarding the functions or processes it measures and points to the need for further investigations concentrated on the study of the test alone. It has been suggested by different writers that the subtests of the Wechsler do not measure distinct psychological functions. Cohen concluded in one of his reports that

The relatively small degree of subtest specificity makes suspect any rationale which attributes quasi-unique measurement functions to the subtests, and helps account for the disappointing outcome of ten years of research in pattern analysis with the Wechsler-Bellevue.⁴⁴

Although one might be tempted to agree with this statement, it seems premature to conclude that the investigations pursued with the intent of determining the functions measured by the Wechsler subtests have given the clinician a satisfactory answer. First of all, most of the studies have been done with psychiatric patients. Frank, Corrie, and Fogel have demonstrated the ambiguity of such studies. They noted that

Investigators have been attempting to compare the characteristic performance of one group of unlike subjects with another. The effect of the attempt to differentiate two such heterogeneous groups of subtest scores is to structure the data for failure by attempting to compare incomparables. There is no reason to assume that psychiatric diagnosis, as a criterion measure for the selection of groups, will automatically eventuate in a random sampling of homogeneous subject's subtests when we know variables as age, sex, and even length of hospitalization affect subtest patterns in their own right.⁴⁵

Secondly, it is evident from this review that there has been very few investigations directly concerned with the

⁴⁴ Jacob Cohen, "The Factorial Structures of the WAIS", p. 290.

⁴⁵ G.H. Frank, C.C. Corrie, and J. Fogel, "An Empirical Critique of Research with the Wechsler-Bellevue in Differential Psychodiagnosis", in the Journal of Clinical Psychology, Vol. 11, No. 3, issue of July 1955, p. 292.

rationale of any one single subtest of the Wechsler. Such studies with a normal population would possibly be a great asset. There is a great need for well defined rationales. It is necessary to know as much as possible about the function or functions measured by the tools the psychologist uses and about the factors affecting the performance on these tools. Knowing exactly the particular functions, processes, or modalities underlying a particular subtest in regard to normal subjects would open the door to better understanding of the functional or organic disturbances of the patients. Such knowledge would presumably permit the clinician to arrive at a more valid diagnosis and help him to predict more accurately the outcome of treatment.

The present research intends to investigate some aspect of the rationales suggested for Block Design, and to gain further knowledge of the mental process involved in the accomplishment of the task.

4. Selection of the Hypothesis to be Investigated

Among the numerous assumptions which have been made regarding the processes involved in the performance on Block Design some can more easily be investigated than others. Yet, in most cases, very few instruments are available which satisfy the requirements of construct validity and

which are adequate for the measurement of the suggested hypotheses. The availability of such instruments is then, an important factor in the selection of the rationale to be investigated.

The choice of the specific hypothesis for the study has been suggested mainly by the nature of the assumptions reported in the literature. The multiplicity of the hypotheses and the obvious overlapping of functions have demonstrated that Block Design does not measure a single ability or factor. However, it is important to determine as clearly as possible which function or functions are mainly involved in solving the task and to what extent they are measured by the test.

As demonstrated by Kohs and Wechsler and consistently reported by others, Block Design has a high correlation with varied measures of intelligence. On the other hand, factor analyses of the main tests of intelligence with which Block Design correlates well have revealed that many functions or abilities are measured by these tests.

As a result, since Block Design correlates well with these tests, it is reasonable to assume that it measures to a great extent some of the factors found in them. Moreover, it has been suggested in a number of investigations reported in the review of the literature, that

reasoning ability was one of the functions being tapped by these tests. It is thus hypothesized that reasoning is one function being measured by Block Design.

Balinsky's findings would tend to support such a hypothesis, except that the present study deals with a sample of subjects of a different age. It would appear that this assumption does not hold here, because, as he suggested himself, mental traits change and undergo reorganization over a span of years and the same test measures different abilities at different ages. However, as indicated before, Cohen disagreed with Balinsky and claimed that the organization of intellectual functions is essentially invariant for the ages 18 to 54. Thus, the reasoning factor which has been isolated for two age groups by Balinsky could appear for other age groups under different conditions.

Considering these factors, the writer arrived at the specific hypothesis under investigation, namely, whether or not Block Design measures the reasoning ability, and, if so, to what extent. The null hypothesis would be that Block Design does not correlate with any of the measures of reasoning ability employed nor does it measure to any extent any of the independent variables. No specific definition of reasoning ability is given here; however, it should be understood that the operational definitions accepted for the present study will be the ones provided by the respective

authors of the tests used to verify the hypothesis.

The selection of reasoning as an assumption to be verified does not necessarily exclude some other functions which might be very closely related to it. Indeed, due to the nature of the tools used in this study, other rationales might also be indirectly involved. For, example, planning and anticipation, and the ability to analyze may be measured to some extent since Thurstone's Primary Mental Abilities, Reasoning subtest⁴⁶ supposedly contains these factors. This test will be discussed later. Similarly, the judgment factor, which has been found in the reasoning factor⁴⁷, and "...has weight in the Reasoning test of Thurstone"⁴⁸, might be revealed indirectly. Other suggested rationales such as: ability to abstract and generalize and think inductively, as well as some perceptual factor may also be indirectly under investigation. The tests of reasoning ability, which are used in this study and found in the Differential Aptitude

⁴⁶ L.L. Thurstone and T.G. Thurstone, Examiner Manual for the SRA Primary Mental Abilities, Preliminary Edition, Chicago, Science Research Associates, 1947, p. 1.

⁴⁷ J.P. Guilford, Editor, and J.I. Lacey, Research Associate, Army Air Forces Aviation Psychology Program Research Reports, Printed Classification Tests, Report No. 5, Washington, Superintendent of Documents, U.S. Government Printing Office, 1947, p. 129.

⁴⁸ Harold M. Corter, "A Factor Analysis of Some Reasoning Tests", in the Psychological Monographs, Vol. 66, No. 8, issue of March 1952, p. 28.

Tests battery⁴⁹, which are used in this study, measure different modalities of the reasoning ability.

The material which has been discussed in this chapter consisted of a review of the different hypotheses suggested in the literature in order to explain the function or functions measured by Block Design. This was followed by a list of the interpretative hypotheses and an attempt was made to group together the ones which appeared to represent similar functions or mental processes. Finally, the selection of the specific hypothesis under investigation was presented and reasons for its choice were given.

The elements which contributed to the general hypothesis have been considered in this chapter. In the next portion of this report, the experimental design will be discussed and special attention will be paid to the test under investigation, to the tools used, to the selection of the population, and to the methods of analysis employed.

⁴⁹ George K. Bennett, Harold G. Seashore, and Alexander G. Wesman, A Manual for the Differential Aptitude Tests, New York, The Psychological Corporation, 1952, iv-78 p.

CHAPTER III

THE EXPERIMENTAL DESIGN

The discussion, thus far, has been centered around two main topics, namely, the statement of the problem and the review of the literature directly concerned with the problem under investigation. Block Design, the test to be investigated, and the instruments used for the verification of the assumption will now be considered. This will be followed by a presentation and a justification of the chosen population. Finally the procedure which was followed in investigating the assumption will be considered and special attention will be paid to the methods of statistical analysis employed.

1. Test under Investigation

a) Description - The Block Design subtest consists of ten items. There are two sample items which are demonstrated to the subject for which he receives credit upon completion within the time limit. These two sample items and the next four test items consist of 4 block, each to be juxtaposed so as to form a given pattern; the last four items consist of 9 blocks each. Some changes have been introduced in the Block Design subtest of the WAIS. The blocks are somewhat different than the ones used in the

Koh's Block Test and in the Wechsler-Bellevue. The colors blue and yellow have been eliminated and only red and white remain. Each block has two red, two white and two red/white sides; any side can be used to reproduce the design according to the requirements of the pattern. Each block takes the form of a cube and the blocks in each completed design form a square. All the designs show some symmetry.

As to the items themselves, seven patterns have been retained from the Wechsler-Bellevue and three new ones have been introduced; the new items being the first sample and the last two designs which appear to be more difficult. The items are timed, as previously, but the time allowance has been reduced to 60 seconds for the first six designs and to 120 seconds for the last four tasks. By introducing these changes Wechsler has increased the discriminating quality of the test.

Successful completion of each design within the time limit is credited with four points, and additional time-credits from one to two points are awarded for speedy performance on the last four test items. No credit is given for partially correct or incomplete patterns. Successful performance on all items without time credit - forty points - yields a weighted score of twelve, while a maximum score of forty-eight points gives a weighted score of seventeen.

c) Reliability - The reliability of Block Design presents another important problem. Such subtest reliability information is important since it determines in great part, the confidence which may be attached to statements about the abilities possessed by the subject and about patterns of subtest variation. As Lindquist pointed out: "If a test is unreliable, that is, if it is not measuring anything consistently, it, of course, cannot be valid, that is, it cannot be measuring accurately what it is intended to measure...."¹ This problem was overlooked by Wechsler who quoted reliability figures for the full scale of the Wechsler-Bellevue but not for the separate subtests.

Investigators, using psychiatric samples, have reported some measure of reliability. Two studies by Rabin on test-retest with the Wechsler-Bellevue are available. In the first,² the reliability for a group of miscellaneous patients was .74. A second study³ compared test-retest results of 30 schizophrenic with 30 non-schizophrenic state hospitalized subjects. The Block Design reliability

1 E.F. Lindquist, A First Course in Statistics, Boston: Houghton Mifflin Co., 1942, p. 207.

2 A.I. Rabin, "Test Constancy and Variation in the Mentally Ill", in the Journal of General Psychology, Vol. 31, Second Half, issue of October 1944, p. 233.

3 A.I. Rabin, "Fluctuations in the Mental Level of Schizophrenic Patients", in the Psychiatric Quarterly, Vol. 18, No. 1, issue of January 1944, p. 78-81.

coefficient was .71 for the schizophrenics and .70 for the non-schizophrenics. Hamister⁴, using a neuropsychiatric sample, found a test-retest reliability of .71 for the group of miscellaneous patients and a reliability of .67 for the schizophrenic group.

The few reliability studies reported here have employed abnormal subjects with little attention paid to reliability obtained with normal populations. Derner, Aborn, and Canter⁵ investigated that problem with a group of 158 normal subjects and found a coefficient of reliability of .84 which is much higher than the coefficients reported above. This test placed third after Vocabulary and Information, with reliabilities of .88 and .86, respectively. A split-half method of estimating reliability was used by Webb and DeHaan⁶ with two groups of female subjects. The coefficient of reliability for the normal subjects was .76 and for the psychotic group, .86. Block Design again appeared as

⁴ Richard C. Hamister, "The Test-Retest Reliability of the Wechsler-Bellevue Intelligence Test (Form I) for a Neuropsychiatric Population", in the Journal of Consulting Psychology, Vol. 13, No. 1, issue of February 1949, p. 43.

⁵ Gordon F. Derner, Murray Aborn, and Aaron H. Canter, "The Reliability of the Wechsler-Bellevue Subtests and Scales", in the Journal of Consulting Psychology, Vol. 14, No. 3, issue of June 1950, p. 176.

⁶ Wilse B. Webb and Henry DeHaan, "Wechsler-Bellevue Split-Half Reliabilities in Normals and Schizophrenics", in the Journal of Consulting Psychology, Vol. 15, No. 1, issue of February 1951, p. 70.

one of the most reliable subtests, placing second for the psychotic group and fourth for the normal subjects. These writers pointed out:

These results, showing the higher reliabilities of psychotics in performance tests are quite surprising in view of the fact that one would logically expect paranoid schizophrenics to obtain test reliabilities either similar to or lower than normals⁷.

However, in an examination of these data, Helmick⁸ suggested that the failure to consider the differences in variability between the two groups was primarily responsible for the obtained differences in reliability figures. By computing the standard errors of measurement and by correcting the coefficients of correlation for differences in range, he found that the psychotics were more variable in their test performance than the normals. He obtained a reliability of .85 with the normals and of .77 with the psychotics.

In 1953, Botwinick⁹ investigated the retest reliability of three groups of subjects: a young normal group,

7 Id., op. cit., p. 70.

8 John S. Helmick, "Reliability or Validity?", in the Journal of Consulting Psychology, Vol. 16, No. 2, issue April 1952, p. 154.

9 J. Botwinick, "Wechsler-Bellevue Split-Half Subtest Reliabilities: Differences in Age and Mental States," in the Journal of Consulting Psychology, Vol. 17, No. 3, issue of June 1953, p. 226.

an old normal group and a group of senile patients. He found a reliability of .76 for the first group, of .84 and .89 respectively for the other two groups. Armitage and Pearl¹⁰ reported study results on groups of schizophrenics, non-schizophrenics, and normals. Although the purpose of the study was not to determine the reliability of the subtests, the test-retest correlations can, nevertheless, be accepted as a measure of reliability. They indicated that Block Design has a reliability of .84 with the normals. The reliability for the group of 53 schizophrenics was .71.

These references have dealt with Block Design as a subtest of the Wechsler-Bellevue Intelligence Scale. No report could be found in the literature, except for the results published by Wechsler, regarding the reliability of Block Design as a subtest of the Wechsler Adult Intelligence Scale. The reliability coefficients presented by Wechsler¹¹ are based on his re-standardization of the Wechsler-Bellevue Intelligence Scale. The data are available for three age groups only. For the age group 18-19, the reliability coefficient of Block Design is .86. The reliability for the age

10 Stewart G. Armitage and David Pearl, "Wechsler-Bellevue Changes Over Time", in the Journal of Clinical Psychology, Vol. 14, No. 1, issue of January 1958, p. 23.

11 David Wechsler, WAIS Manual: Wechsler Adult Intelligence Scale, First Edition, New York, The Psychological Corporation, 1955, p. 13.

groups 25-34 and 45-54 is .83 and .82 respectively.

In general, it appears that the reliability of the Block Design subtest merely satisfies the sophisticated statistician who recommends reliability coefficients no lower than .80 and preferably above the .90 level. Since all the reported coefficients are below .90, judgments with respect to the validity of the test must be made with caution.

2. Selection and Description of Tools

An important problem to be considered is the selection of the tools. The choice of the tests will be explained with special reference to the rationale under investigation. The definition of the function or modality measured by these tests is also important because it determines the rationale of the tools themselves and the final interpretation of the results. Since for each test a definition has usually been given by its author, it appears quite essential to accept this meaning for the present study.

Closely connected with the question of definition is the problem of finding a measure of a unique factor. The purpose of this investigation being to determine the importance of the reasoning function in Block Design performance, special attention has been given to this problem.

Although it is questionable whether or not such tests exist, some do correspond better to this requirement than others and an effort was made to select the most adequate ones. In spite of the severe criticisms formulated against the method of factor analysis one test which has been standardized through such a method has been included here because it appeared adequate for the purpose of this study. Three other tests, for which numerous data are available, have been found suitable because of the nature of their rationale.

The tools selected to investigate the hypothesis are the Reasoning subtest of the Primary Mental Abilities Test¹² and the Abstract Reasoning, Verbal Reasoning, and Mechanical Reasoning subtest of the Differential Aptitude Tests¹³.

a) Reasoning subtest of the Primary Mental Abilities Test - The early literature concerning the Thurstone Test was quite controversial. Thurstone intended to isolate the primary abilities by objective experimental procedures. His test is a result of his efforts to isolate

12 L.L. Thurstone and T.G. Thurstone, Examiner Manual for the SRA Primary Mental Abilities, Preliminary Edition, Chicago, Science Research Associates, 1947, p. 5-6

13 George K. Bennett, Harold G. Seashore, and Alexander G. Wesman, A Manual for the Differential Aptitude Tests, New York, The Psychological Corporation, 1952, p. 5.

the fundamental abilities, to describe their nature, to ascertain the loading of each fundamental ability in each test, and to describe each subject as regards each of the fundamental abilities.¹⁴ While recognizing Thurstone's outstanding experimental contribution, McNemar expressed some criticism. He indicated that

Aside from the subjectivity involved in giving meaning to the factor, one also notes a disturbing arbitrariness in the selection of tests utilized in rationalizing some of the factors.... So far as the factorial methods are concerned it would seem that the weakest links in an otherwise strong chain are those which bridge the gap between mathematical factors and psychological meaning and the chasm between sample and universe values. In fact, the latter link is practically missing, and from the reviewer's viewpoint this constitutes a serious limitation.¹⁵

Spearman's analysis revealed a somewhat similar trend of thought. He deplored such conditions as small correlations, large increase of errors of sampling, and the exceedingly large number of factors introduced. He noted that

14 L.L. Thurstone, "Primary Mental Abilities", in the Psychometric Monographs, No. 1, Chicago, The University of Chicago Press, 1938, p. 4.

15 L.L. Thurstone, "Primary Mental Abilities", in the Psychometric Monographs, No. 1, as reviewed by Quinn McNemar, in the Journal of Educational Psychology, Vol. 30, No. 1, issue of January 1939, p. 76.

On the whole, then, the present revolutionary results would appear to derive not from any cogent facts or logic, but only from a statistical accident.... The new method generates many further factors which can claim little or no support anywhere else.... Thurstone not only analyses the data in his own preferred manner: he makes little suggestion that other modes of analyses are also possible, even at the same time. Again, he not only adopts his principle of maximizing the number of zero entries, but does so, apparently without concern for any ensuing violation of the standard principle of minimizing the number of factors.... The new method ... is liable to go completely astray in a certain application of questionable principles to inadequate data.¹⁶

These comments represent early adverse opinions which have now subsided as the application of factor analysis has gained popularity and the theory of multiple abilities has gained acceptance. Many studies have reported findings that support Thurstone's results. Michael and his associates¹⁷ also identified factors which were similar to those reported by Thurstone except for S, (spatial relations), and Vs, (visualization). More primary abilities were found through intensive work by Guilford and

¹⁶ C. Spearman, "Thurstone's Work Reworked", in the Journal of Educational Psychology, Vol. 30, No. 1, issue of January 1939, p. 10-14 and 16.

¹⁷ William B. Michael, Wayne S. Zimmerman, and J. P. Guilford, "An Investigation of the Spatial Relations and Visualization Factors in Two High School Samples," in the Educational and Psychological Measurement, Vol. 11, No. 4, issue of Winter 1952, p. 572.

his co-workers¹⁸. However, the results were not conclusive and identification of abilities are still in process.

The choice of this test as a measure of reasoning ability rests mainly on its definition. Thurstone states that the Reasoning test measures

The ability to solve logical problems,... to foresee and plan... The person with good reasoning ability can solve problems, foresee consequences, analyse a situation on the basis of past experiences, and make and carry out plans according to recognizable facts¹⁹.

It has been mentioned earlier that foresight and planning, the ability to analyse, and judgment would not be specifically investigated. However, these different abilities have all been found to exist to a certain degree in the factor reasoning and might be essential to the process of reasoning. A number of investigators, such as Green²⁰,

18 J.P. Guilford, Editor, and J.I. Lacey, Research Associate, Army Air Forces Aviation Psychology Program Research Reports, Printed Classification Tests, Report No. 5, Washington, Superintendent of Documents, U.S. Government Printing Office, 1947, p. 823-838, passim.

19 L.L. Thurstone and T.G. Thurstone, "Examiner Manual", op. cit., p. 1.

20 Russel Frank Green, A Factor-Analytic Study of Reasoning Abilities, Unpublished Doctoral Dissertation, presented to the University of Southern California, 1951, p. 191, passim.

Adkins and Lyerly²¹, and Guilford and his associates²², have indicated the difficulty of defining the reasoning ability and have suggested that it involves a complexity of modes. Thurstone recognized such a possibility when he stated that

Recent research has shown that R is really two separate abilities: inductive reasoning, the ability to reason from specific cases to a general rule; and deductive reasoning, the ability to reason from stated premises to a logical conclusion. The present test is a composite measure of both abilities.²³

These investigations have revealed that the nature of the reasoning process varies with quite a number of variables and is almost too complex to permit a highly positive and adequate evaluation with any single test.

The reliability of the test is quoted at .93 by Thurstone²⁴, using a sample of 500 tenth grade students. This was presumably based on a test-retest situation

21 Dorothy C. Adkins and Samuel B. Lyerly, Factor Analysis of Reasoning Tests, Chapel Hill, University of North Carolina Press, 1952, pp. IV, 122, as reviewed by Edward E. Cureton, in the Journal of Educational Psychology, Vol. 44, No. 3, issue of March 1953, p. 189-191.

22 J.P. Guilford, Norman W. Kittner, and Paul R. Christensen, "The Nature of the General Reasoning Factor", in the Psychological Review, Vol. 63, No. 3, issue of May 1956, p. 169-172.

23 L.L. Thurstone and T.G. Thurstone, "Examiner Manual", op. cit., p. 2.

24 Id., op. cit., p. 13.

although it is not clearly stated in the Manual whether this was the actual procedure. This reliability coefficient is high. However, an investigation by Traxler²⁵, who used a sample of 104 high school girls, revealed lower reliability. He found a test-retest reliability coefficient of .58 for grade ten, of .73 for grade eleven, and of .76 for grade twelve. In a similar study by Jacobs²⁶ with 97 tenth grade students the reliability was found to be .87. This finding supports Thurstone's claims of the test's dependability and is also in agreement with other investigation reported in the literature.

Thurstone's Reasoning test consists of thirty items to be solved by the subject within a six minutes period. Each item is made up of a series of letters which follow each other according to a certain rule. The subject must discover the rule and indicate his answer by crossing on the answer pad the letter that would logically follow if the line were to be continued. For example, if a line read "cdcdcd", it would be correct to cross the letter "c"

25 Arthur Traxler, "Stability of Scores on the Primary Mental Abilities Tests", in School and Society, Vol. 53, No. 1365, issue of February 1941, p. 256.

26 Robert Jacobs, "The Reliability and Intercorrelation of the Scores on the SRA Primary Mental Abilities Test", in the Educational Record Bulletin, Vol. 49, No. 3, issue of May 1949, p. 54.

on the answer pad; or if it read "axbyasbyaxb" the letter that would logically follow would be "y". The score is the total number of correct answers with no correction formula for guessing. A maximum score of 30 points is possible.

b) Reasoning subtests of the Differential Aptitude Tests - Although the Differential Aptitude Tests battery is still at an early stage of development, it has been widely used and promising results have been reported in the literature. Bennett and his co-workers, in constructing the battery, were primarily concerned with the evaluation of the individual's power in the particular ability or skill. They stated that

The principle underlying the Differential Aptitude Tests is that the level of operation is usually the most important aspect of abilities which the counselor and his client need to consider. For this reason the Differential Aptitude Tests are power tests, except for the Clerical Speed and Accuracy test which is intended as a direct measurement of speed.²⁷

As they pointed out, their intention was not so much to construct pure tests of the various dimensions as to develop highly reliable, valid, and useful tests. Numerous studies have been reported which suggest that such a goal has been

²⁷ George K. Bennett, Harold G. Seashore, and Alexander G. Wesman, Differential Aptitude Tests, Manual, New York, Psychological Corporation, 1952, p. 4.

achieved. Super²⁸ considers that the Differential Aptitude Tests battery is one of the only two batteries which is ready for use in counseling. He believes that, with one or two possible exceptions, the tests are excellent in format, item construction, standardization, and validation. However, he noted that the overlap of abilities apparently measured by the subtests of the battery might at times be disturbing. In some cases, therefore, the separate tests appear to measure a combination of factors or abilities. Carroll stated:

At the present time, it can be said that considering the tests themselves and all the supporting data, the DAT constitutes the best available foundation battery for measuring the chief intellectual abilities and learned skills which one needs to take account of in high school counseling.²⁹

Most of the available studies are concerned with the problem of reliability and the prediction of achievement with certain academic courses. No investigation could be found whose specific purpose was the validation of the rationale of

28 Donald E. Super, "The Differential Aptitude Tests: An Overview", in the Personnel Guidance Journal, Vol. 35, issue of October 1956, p. 81-93.

29 G.K. Bennett, H.G. Seashore, and A.G. Wesman, Differential Aptitude Tests, Manual, New York, Psychological Corporation, 1952, pp.IV, 78 p., as reviewed by John B. Carroll, and quoted by O.K. Buros, Ed., The Fifth Mental Measurements Yearbook, Highland Park, N.J., The Gryphon Press, 1959, p. 673.

the tests. Nevertheless, as the authors suggested

There is no 'validity of a test', there are only empirical 'validities' describing the usefulness of the test in numerous situations. Out of the mass of data, generalizations emerge regarding the test's effectiveness. Validation is a never ending process.³⁰

Three subtests of the Differential Aptitude Tests battery which measure different modalities of the reasoning ability were selected; Abstract Reasoning, Verbal Reasoning, and Mechanical Reasoning. Each of these will now be considered.

Bennett and his co-workers believe that the Abstract Reasoning test measures certain functions which are associated with general intelligence. They indicate that

The Abstract Reasoning test is intended as a non-verbal measure of the student's reasoning ability. The series presented in each problem requires the perception of an operating principle in the changing diagrams. In each instance, the student must discover the principle or principles governing the change of the figures and give evidence of his understanding by designating the diagram which should logically follow.... It involves the ability to perceive relationships in abstract figure patterns - generalization and education of principles from nonlanguage designs.³¹

The authors have compiled numerous data which indicate that this test predicts best the future performance of students

³⁰ George H. Bennett, Harold G. Seashore and Alexander G. Wesman, Differential Aptitude Tests, Manual, New York, Psychological Corporation, 1952, p. 39.

³¹ Id., op. cit., p. 6-7.

in science courses. High correlations have also been found with industrial arts and mechanical drawings. Comparison between the Reasoning subtest of the Primary Mental Abilities and the Abstract Reasoning test has been reported in Bennett's Manual³² and a correlation of .60 has been quoted for a group of male subjects. This test also appears to measure certain functions which are closely associated with intelligence as measured by the OTIS and the Henmon-Nelson test. Comparing the Abstract Reasoning test with the Henmon-Nelson IQ, a correlation of .71 has been quoted for a group of female subjects, while one of .69 was found between the former and the OTIS S-A: Higher Examination with a group of 208 females.³³

The Abstract Reasoning portion of the Differential Aptitude Tests battery is made up of fifty items to be solved by the subject within a period of twenty-five minutes. It is a non-verbal type of test which consists of rows of designs. Each row is a problem which is made of four figures to form a series. The subject must discover the operating principle or principles which govern the evaluation of the diagram in each series. He also is required to indicate

32 Id., op. cit., p. 72.

33 Id., op. cit., p. 72.

which of a series of five choices properly carries out the logical development shown by the sequence of figures.

The test is well constructed and the authors have succeeded in making all diagrams clear and the differences between diagrams self-evident. The progressive changes within a series from one design to another are clues to the operating principles which will influence the choice of the last diagram.

The scoring method is quite simple. The maximum score is 50. The formula is the sum of right answers minus one quarter the sum of errors.

Bennett and his associates reported reliabilities of the Abstract Reasoning test for groups of boys and girls. A reliability of .64 was obtained³⁴ on test-retest data gathered from a group of 90 grade nine girls whose scores were compared with those received when they reached grade twelve. A reliability of .62 was indicated³⁵ for a group of 71 grade nine boys under similar conditions. The three years period that elapsed between test and retest sessions might account for these low reliabilities. Substantial test-retest coefficients of .91 for grade ten and eleven, and of .85 for grade twelve were quoted³⁶ by the same authors. They also

34 Id., op. cit., p. 64.

35 Id., op. cit., p. 64.

36 Id., op. cit., p. 66.

mentioned an average coefficient of reliability of .89 for a group of 1,064 high school girls.³⁷

The Verbal Reasoning test has something in common with Abstract Reasoning in that it is supposed to measure intellectual functions which are associated with general intelligence. Bennett and his associates indicate that it is

A measure of ability to understand concepts framed in words. It is aimed at the evaluation of the student's ability to abstract or generalize and to think constructively, rather than at simple fluency or vocabulary recognition. The analogies form of test item is peculiarly appropriate for the measurement of reasoning ability.... The item thus samples the student's knowledge and his ability to abstract and generalize relationships inherent in that knowledge.³⁸

This test appears to predict with reasonable accuracy successes in many academic fields. Its correlation with successful achievement in college courses such as English, Science and Social Studies probably confirms the authors' hypothesis that it measures verbal intelligence. Comparing the Reasoning subtest of the Primary Mental Abilities with the Verbal Reasoning test, a correlation of .49 has been quoted for female subjects.³⁹ There is statistical evidence that these tests both measure to a certain extent a similar

37 Id., op. cit., p. 66.

38 Id., op. cit., p. 5-6.

39 Id., op. cit., p. 72.

function. Nevertheless, the coefficients are sufficiently low to suggest that they also measure something different; therefore, their inclusion is worthwhile in order to tap different aspects of the reasoning ability. Very substantial correlations with measures of intelligence have been reported by Bennett. When a group of 204 female subjects was tested with both the Verbal Reasoning test and the ACE Psychological Examination test a coefficient of correlation of .86 was found.⁴⁰ Comparisons of the Hemmon-Nelson I.Q. and the OTIS S-A: Higher Examination with the Verbal Reasoning test were also reported⁴¹ and the correlations were .77 and .84 respectively for two groups of girls.

The Verbal Reasoning test consists of an analogies form of test item. There are fifty sentences whose first word and last word are left out. The subject must select, from two sets of four words each, the appropriate words for the analogy. The authors claim that the items are highly reliable "in that the chances are only one in sixteen that any correct answer will be guessed; in most multiple-choice test items, the chances are one in four or five".⁴² Special attention was devoted, in the construction of the test, to

40 Id., op. cit., p. 72.

41 Id., op. cit., p. 72.

42 Id., op. cit., p. 6.

the selection of versatile, varied, and relatively complex items without loading it with rare items of knowledge or very unusual vocabulary terms.

The scoring system of Verbal Reasoning is simple. The maximum score that a person may get is fifty points, one mark being given for each correct answer. There is no retaliation for wrong answers.

The coefficients of reliability reported in Bennett's Manual indicate that Verbal Reasoning is one of the most reliable tests of the battery and certainly the most stable of the three reasoning tests. A reliability of .82 for girls and .87 for boys was obtained on test-retest data gathered from grade nine boys and girls whose scores were compared with those they received in grade twelve.⁴³ An average test-retest coefficient of reliability of .90 was quoted by Bennett and his associates⁴⁴ for a population of female subjects from grade eight to twelve inclusively. Reliabilities as high as .93 and .92 have been reported by the same authors⁴⁵ for two groups of grade eleven and grade twelve female subjects respectively.

The Mechanical Reasoning test used in this

43 Id., op. cit., p. 68.

44 Id., op. cit., p. 66.

45 Id., op. cit., p. 66.

investigation is essentially a new form of a test prepared earlier by Bennett. The authors express considerable confidence in it and suggest that it fulfils a need in the investigation and understanding of mechanical and physical principles in familiar situations. Their claim is supported by reports in the literature which have shown the usefulness of this approach to reasoning in the mechanical field. Although this test is supposed to measure the capacity for understanding mechanical relationships, the rationale of the test, nonetheless, appears quite unspecific, as the authors write that

The ability measured by the Mechanical Reasoning test may be regarded as one aspect of intelligence, if intelligence is broadly defined. The person who stands high in this characteristic finds it easy to learn the principles of operation and repair of complex devices. The score is affected by the previous experience of the subject but not to a degree that introduces serious difficulties in interpretation. Formal training in physics produces an increase in score of only a few points.⁴⁶

This test, used mainly in school situations, has been useful in predicting achievement in physical science, technical, and manual training courses. Comparisons with some tests of intelligence indicate that the ability measured by the Mechanical Reasoning test may be closely associated to some aspect of intelligence as suggested by Bennett and his

⁴⁶ Id., op. cit., p. 7.

associates. A sample of 204 female subjects has been tested both with the Mechanical Reasoning test and the ACE Psychological Examination test and a correlation of .60 has been found.⁴⁷ Similar comparisons with the Henmon-Nelson I.Q. test and the OTIS S-A: Higher Examination test with two groups of female subjects resulted in correlations of .56 and .61 respectively.⁴⁸ On the other hand, a correlation of .40 was found for a group of 87 females in a comparison of their results on Mechanical Reasoning and the Reasoning subtest of the Primary Mental Abilities.⁴⁹

The Mechanical Reasoning test consists of a number of pictures presenting a mechanical situation together with simple questions about those pictures. The subject has 30 minutes to complete the test of sixty-eight problems which requires understanding of mechanical and physical principles. Considerable care was taken to prevent complicated and rare mechanisms that require special knowledge. Two forms are available for this test.

The scoring system takes into account right and wrong answers. The formula is the sum of correct answers minus one half the sum of errors. The highest possible score is 68.

⁴⁷ Id., op. cit., p. 72.

⁴⁸ Id., op. cit., p. 72.

⁴⁹ Id., op. cit., p. 72.

Bennett⁵⁰ reported a reliability of .63 which was obtained in test-retest data gathered from a group of 90 grade nine girls whose scores were compared to those they received in grade twelve. Reliabilities of .70, .72, and .69 were also found for three samples from grade ten, eleven, and twelve respectively.⁵¹ The average reliability coefficient for a population of 1064 girls was quoted⁵² as .71, compared to .85 for a sample of 960 boys. From Manual data, this difference between boys and girls can be found from grade eight to twelve inclusively. This test's reliability is low.

At this point, the writer wants to re-emphasize that he had no desire to question the rationale of the tests selected for the investigation of the hypothesis under study, nor did he intend to debate their relative validities and reliabilities. He, thus, accepted the definitions suggested by the authors of the tests as possible measures of diverse modalities of the reasoning ability as measured by Block Design.

50 Id., op. cit., p. 68.

51 Id., op. cit., p. 66.

52 Id., op. cit., p. 66.

3. Selection and Description of the Population

The method of selecting the subjects as well as explanations of the reasons for choosing a selected rather than random sample are matters that will be considered along with the description of the population. The study involved a certain number of subjects to whom Block Design and the experimental test battery were administered. A description of this sample will be included. For reasons that need to be discussed it was thought necessary to control variables such as age, education, sex, and language.

Numerous criticisms of studies dealing with the Bellevue Tests have stressed vigorously the need to control certain variables. Johnson noted that

Most authors have contributed to the confusion by not carefully defining their groups and in general, samples are limited. Little effort has been made to equate groups for such factors as age, sex, education, or intelligence.⁵³

Schofield was not less severe and suggested that

There have been sufficient demonstrations of the need to control the age, general ability, and educational achievement variables as to discourage acceptance of findings when these factors have been permitted to introduce distorting variance.⁵⁴

⁵³ Thomas F. Johnson, "Some Needs in Research with the Wechsler-Bellevue Scale", in the Journal of General Psychology, Vol. 41, First Half, issue of July 1949, p. 34.

⁵⁴ William Schofield, "Critique of Scatter and Profile Analysis of Psychometric Data", in the Journal of Clinical Psychology, Vol. 8, No. 1, issue of January 1952, p. 19.

a) Age - The need to control the factor age has been repeatedly mentioned in the literature. Benton⁵⁵ reported that the ability to deal with Block Design declines significantly with age. Balinsky⁵⁶ gathered data that led him to believe that mental traits change and undergo reorganization over a span of years. On the other hand, Cohen suggested that "organization of intellectual functioning is essentially invariant between the ages of 18 and 54."⁵⁷ Such difference of opinion stresses the importance of controlling the age variable and indicates the necessity of determining more specifically the traits or factors measured by Block Design at different ages.

Data published by Wechsler⁵⁸ indicated that the highest mean for this test was found with the 17 to 19 years

55 A.L. Benton, "A Study of the Performance of Young Adults on the Kohs Block Designs Tests", in the Journal of Applied Psychology, Vol. 25, No. 4, issue of August 1941, p. 425.

56 Benjamin Balinsky, "An Analysis of the Mental Factors of Various Age Groups from Nine to Sixty", in the Genetic Psychology Monographs, Vol. 23, First Half, issue of February 1941, p. 231.

57 Jacob Cohen, "The Factorial Structure of the WAIS between Early Adulthood and Old Age," in the Journal of Consulting Psychology, Vol. 21, No. 4, issue of August 1957, p. 289.

58 David Wechsler, The Measurement of Adult Intelligence, Third Edition, Baltimore, Williams and Wilkins, 1944, p. 222.

age group. Of importance, also is the fact that the highest Block Design reliability coefficient was reported for the age group 18 to 19, in the standardization of the WAIS.⁵⁹ It may be, then, that the different traits of factors measured by Block Design are more accurately tapped at this age. Wechsler commented that

Beyond the age of 15, mean scores on practically all intelligence scales cease to increase significantly with age. Psychologists have generally interpreted this fact to mean that intellectual ability stops growing at that age.⁶⁰

Considering all these facts, it appeared important to select a population whose age mean would be in the late teens. The sample in the present investigation was thus limited in age range and consisted of a group of fifteen to nineteen year old subjects.

b) Education - In a study of the performance of young adults on the Kohs Block Designs Tests, Benton⁶¹ found that educational level is a significant cause of score variance. This finding was re-emphasized later on by other writers. Referring to the use of the Wechsler-Bellevue,

⁵⁹ David Wechsler, The Measurement and Appraisal of Adult Intelligence, Fourth Edition, Baltimore, Williams and Wilkins, 1959, p. 103.

⁶⁰ David Wechsler, The Measurement of Adult Intelligence, op. cit., p. 21.

⁶¹ A.L. Benton, op. cit., p. 425.

both Garfield⁶² and Schofield⁶³ also stressed the factor education. Although little is actually known of the effects of schooling on Block Design, there has been sufficient demonstration of the high correlation between Block Design and the intellectual quotient, which is highly correlated with educational achievement, to render its control necessary, especially in investigations of the rationale of this test. In the present study, the educational level was therefore controlled by limiting the choice of the subjects to a High School population. The sample included all the students of a High School from grade ten to thirteen inclusively.

c) Sex - Wechsler⁶⁴ reported no significant differences between male and female subjects in the Block Design subtest of the Wechsler-Bellevue Scale. Jastak found that "the percentage of successful performance in Block Design test is somewhat higher for men than for women in five out of seven times. Yet women tend to get higher total scores

62 Sol L. Garfield, "An Evaluation of Wechsler-Bellevue Patterns in Schizophrenia," in the Journal of Consulting Psychology, Vol. 13, No. 4, issue of August 1949, p. 286.

63 William Scholfield, op. cit., p. 19.

64 David Wechsler, The Measurement of Adult Intelligence, op. cit., 258 p.

on this test than men."⁶⁵ Norman's study⁶⁶ on sex differences also supports Jastak's findings. A similar investigation by Goolishian and Foster⁶⁷ revealed that sex differences exist strengthening the latter conclusions. Moreover, Wechsler⁶⁸ himself found a significant difference between male and female scores on this test in the standardization of the WAIS. Because of the nature of these comments, it seemed important to control the influence of the sex factor. The sample was composed exclusively of female subjects in order to obtain greater homogeneity.

d) The Language - The language is a factor that needed to be controlled since one of the instruments of the study, the Verbal Reasoning subtest of the D.A.T. battery, is a test which depends significantly on the mastering of the English language and on the ability to deal with verbal

⁶⁵ Joseph Jastak, "An Item Analysis of the Wechsler-Bellevue Tests", in the Journal of Consulting Psychology, Vol. 14, No. 2, issue of April 1950, p. 93.

⁶⁶ R.D. Norman, "Sex Differences and Other Aspects of Young Superior Adult Performance on the Wechsler-Bellevue", in the Journal of Consulting Psychology, Vol. 17, No. 6, issue of December 1953, p. 412.

⁶⁷ H.A. Goolishian and Austin Foster, "A Note on Sex Differences on the Wechsler-Bellevue Test", in the Journal of Clinical Psychology, Vol. 10, No. 3, issue of July 1954, p. 299.

⁶⁸ David Wechsler, The Measurement and Appraisal of Adult Intelligence, op. cit., p. 147.

concepts. This study excluded all students who could not speak English fluently. The judgment of the director of studies regarding the subject's ability to understand and to speak adequately the English language, was used as the definitive criterion. A certain number of subjects were, however, bilingual, the French or Spanish languages being the mother tongue in such cases.

The choice of subjects was thus made according to the principles of selection discussed in the foregoing paragraphs. The sample consisted of one hundred and twenty-two High School English speaking female students whose age ranged from fifteen years and two months to nineteen years and eight months. They were all students from grade ten to thirteen inclusively and were unevenly distributed among the six classes composing the High School population. The age distribution is presented in Table I. The subjects have been divided into ten classes using an interval of six months per class. The mean age of the sample is 17.24 years, with a standard deviation of 11.32 months.

Analysis of the table indicates a distribution which tends toward the younger side, the bulk of subjects belonging to the sixteen and seventeen year age groups. A tendency toward bimodality is suspected when a curve is drawn, although the most apparent characteristic is small negative skewness.

TABLE I.-
Age Distribution of Sample in Intervals of Six
Months.

Age		Frequency
19 yrs. 6 mos.	to 19 yrs. 11 mos.	1
19 yrs. 0 mos.	to 19 yrs. 5 mos.	4
18 yrs. 6 mos.	to 18 yrs. 11 mos.	9
18 yrs. 0 mos.	to 18 yrs. 5 mos.	14
17 yrs. 6 mos.	to 17 yrs. 11 mos.	20
17 yrs. 0 mos.	to 17 yrs. 5 mos.	25
16 yrs. 6 mos.	to 16 yrs. 11 mos.	22
16 yrs. 0 mos.	to 16 yrs. 5 mos.	18
15 yrs. 6 mos.	to 15 yrs. 11 mos.	7
15 yrs. 0 mos.	to 15 yrs. 5 mos.	2
Total Number		122
Arithmetic Mean		17 . 24 yrs.
Standard Deviation		11 . 32 mos.

In order to get a fairly large sample representative of the Kingston High School female population, the investigator had planned to use all the students of a High School whose population consisted of one hundred and sixty-five females. However, this number was gradually reduced to a total of one hundred and twenty-two subjects which formed the experimental sample. The testing sessions coincided with the regular school hours, but were carried out during a period of time devoted to special projects, except for the last testing session which was held on a Saturday morning.

Different factors influenced the reduction of the chosen population to a sample of 122 subjects. Since the testing sessions were spread over a period of approximately seven weeks, thirty-five students were eliminated missing one or more sessions through absenteeism. From this group twenty were absent from testing sessions during school hours and fifteen missed the last testing session which coincided with a week-end holiday. The latter were boarders whose families either lived outside the city or the province of Ontario, or resided in another country. Eight students were dropped for other reasons. Among these were one subject who was eliminated due to an arm injury which impaired her ability to deal with the manipulation of the blocks, and seven students who were rejected because of their inability to

speak English and to understand adequately the meaning of some words.

The method of selection described above indicates that the aim of the writer was to obtain an homogeneous group with respect to certain characteristics such as age, sex, education, and language. Obviously, the essential purpose in controlling these variables was to reduce or to eliminate the variations or effects due to their influence. This control of variables would thus make it possible to infer that, when differences are found with respect to the homogeneous group, greater differences would be expected in an heterogeneous group where such variables have not been controlled.

4. Statistical Methods of Analysis

The main purpose of the study was to determine the degree of relationship existing between Block Design and the reasoning factor as indicated by four different measures of reasoning ability. In the treatment of the statistical methods, the choice of formulae and procedures was made according to the requirements of the very nature of the investigation. The investigator was thus primarily concerned with the discovery and analysis of:

1. The degree of relationship existing between each of the measures of reasoning ability used for the study and

the total score on the Block Design test;

2. The degree of relationship existing between all the tests of reasoning ability taken together and the total score on the Block Design test;

3. The contribution of each variable to the total predicted variance in the total score on the Block Design test.

The statistical approach to the analysis of data will now be considered in relation to each of these specific aims and explanations will be provided where necessary for clarification.

a) The degree of relationship existing between each of the measures of reasoning ability and the full score on the Block Design test was one of the important aspects of the investigation. The nature of this problem seemed to favour a correlational approach. It was thus decided that the Pearson product-moment coefficient of correlation would be employed if the two most important requirements for its use were met, that is, if the conditions of linearity and homoscedasticity were present. The analysis was also to be carried on with the use of the raw-score formula:

$$r = \frac{N\sum XY - (\sum X)(\sum Y)}{\sqrt{[N\sum X^2 - (\sum X)^2][N\sum Y^2 - (\sum Y)^2]}}$$

In order to determine whether the obtained correlations were genuine ones, the standard error of an assumed simple

correlation of zero was estimated by the formula:

$$\sigma_{r_0} = \frac{1}{\sqrt{N - 1}} ,$$

while the reliability of the obtained significant relationships and those approaching significance was estimated by the general formula:

$$\sigma_r = \frac{1 - r^2}{\sqrt{N - 1}} .$$

The test for significance of the simple correlation was performed and the following formula was employed:

$$F_r = \frac{r_{12}^2}{(1 - r_{12}^2)/(N - 2)} .$$

b) Since the investigator was mainly interested in the amount or strength of relationship existing between Block Design and the numerous independent variables taken together, a multiple coefficient of correlation needed to be employed. The Doolittle method seemed to fulfil the purpose well and was used. In the verification of the reliability of the multiple correlation, the standard error was measured by the formula:

$$\sigma_r = \frac{1 - R^2}{\sqrt{N - m}},$$

which is suggested for most samples and more specifically for small samples.

The test for significance of the multiple correlation was performed by the use of the following formula:

$$F_R = \frac{R^2/m}{(1-R^2)/(N-m-1)}.$$

Since it is generally accepted that the multiple correlation is an inflated value, thus, a biased estimate of the multiple correlation in the population, a correction was applied. A common procedure of reducing it to a more probable population value is by the formula:

$${}_cR^2 = 1 - (1 - R^2) \left(\frac{N - 1}{N - m} \right),$$

which was used in the present study.

c) The last step in the statistical approach to the investigation had to do with the discovery of the contribution of each variable to the total predicted variance in the total score on the Block Design test. This was accomplished through the analysis of the beta weights in the multiple correlation formula. The formula reads as follows:

$$R^2 = \beta_{12}r_{12} - \beta_{13}r_{13} - \beta_{14}r_{14} - \beta_{15}r_{15} .$$

The analysis of the significance of the results was based on the accepted confidence levels which are described below. The relationship was interpreted as:

i) approaching significance if the simple or multiple correlation is 1.96 times as great as the standard error of the correlation, placing it at the .05 level of confidence,

ii) significant if the simple or multiple correlation is 2.58 times as great as the standard error placing it at the .01 level of confidence, and

iii) very significant if the simple or multiple coefficient is 2.81 times as great as the standard error of the correlation, placing it at the .005 level of probability.

In order to obtain a relationship which would approach significance at the .05 level of confidence, a simple coefficient of .1781 would be required, while one of .2345 would be necessary to achieve significance at the .01 level of confidence. The relationship at the .005 level of confidence would be considered very significant if the simple coefficient reaches .2554. As for the multiple correlation, the importance of the relationship would vary according to the values obtained due to the use of a slightly different formula. A multiple coefficient of .1031 would be necessary to be considered as approaching significance at the .05 level

of confidence, while one of .1357 would be required to be statistically significant at the .01 level of confidence. Finally, in order to obtain a relationship which would be very significant at the .005 level of confidence a multiple coefficient of correlation of .1478 and above is essential.

This method of evaluating the significance of the relationship between variables which has just been presented dealt with the number of standard deviations away from the mean, that is, 2.81, 2.58, or 1.96 times the deviations obtained. This is one of the methods used for the purpose of testing the significance of the obtained correlations; it was included in the analysis mainly because it has been extensively used. However, another method, the analysis of variance or what is commonly called the "F" test was employed in order to obtain a more adequate measurement of the significance of the relationships. In this case, the significance of the obtained relationships will be determined by the use of a table of "F" according to the levels of confidence already suggested.

This third chapter was divided into four main sections. The first two sections included material related to the description of the Block Design subtest, the Primary Mental Abilities, Reasoning test, and the three subtests of the Differential Aptitude Tests, Abstract Reasoning, Verbal Reasoning, and Mechanical Reasoning, together with some

aspects of the literature concerning them. There was mention of the reasons for the choice of each test and a brief review of the reliabilities and validities available through the literature. In the next two sections, the discussion was centered around two main topics: the population used for the investigation and the statistical methods. With reference to the population, the problem of selection of variables such as education, age, sex, and language was discussed and a description of the sample was presented. The last section of the chapter dealt with the statistical methods. This involved explanation of the choice of formulae and the procedures of analysis, as well as indication of the methods of evaluating relationships between variables. The criteria for evaluating the significance of relationships were outlined taking into account the bias estimate of the multiple correlation coefficient.

The foregoing discussions were introduced in order to put the investigation in its perspective. The next chapter will be dealing with the presentation, analysis, and discussion of the results obtained.

CHAPTER IV

PRESENTATION AND ANALYSIS OF DATA

The primary aim of the preceeding chapters was to indicate to the reader the nature of the problem, to acquaint him with the literature relevant to the investigation and more specifically to reveal to him the method of study employed for the investigation of the hypothesis. A discussion of the tools, population, and statistical methods used in the treatment of the data has finally led the investigator to the presentation and analysis of the results obtained. One of the purposes of the first section is to discover the relationship of each of the tests of reasoning ability to the total Block Design score. Then, the extent of the relationship of all the tests taken together to the total Block Design score will be studied. Lastly, a discussion of the contribution of the independent variables to the total variance in Block Design will follow. The second section, dealing with the analysis of the data, will give meaning to the results obtained with the experimental sample.

1. Presentation of Data

Although the raw scores obtained by the subjects are necessary to measure the relationships existing between and among the tests, they have not been included in this

discussion, except as a résumé which will be found in Table II. This table contains information such as the range of scores, the means and standard deviations for each test administered to the sample. However, since the raw data might be of interest to some readers, they will be available for reference in the appendix.

a) Degree of relationship existing between each of the measures of reasoning ability and the total score on Block Design. - This aspect of the investigation deals with the simple correlations computed from the raw scores, using the formula described previously in the section of statistical methods.

Since the Pearson coefficient of correlation cannot legitimately be used unless the trend of relationship between the two variables being compared is rectilinear and homoscedasticity is present, it was necessary to investigate each of the coefficients in order to find out if these conditions were fulfilled. The Dayhaw Correlation Chart was employed, and the F test of linearity of regression was used, basing the regression function for the dependent variable (the total Block Design score) on each of the independent variables (the tests of reasoning ability). The same test was repeated, basing the regression function for each of the independent variables on the dependent variable. In either case, as it is seen in Table III and Table IV, no evidence

TABLE II.-

Range of Scores, Means, and Standard Deviations for Each Test of the Battery with a Sample of One Hundred and Twenty-Two Subjects.

Test	Range	Mean	Sigma
B. D. ^a Total Score	20 to 47	33.41	7.33
D. A. T. ^b Abstract Reasoning	0 to 46	33.80	8.53
Verbal Reasoning	4 to 46	29.44	9.05
Mechanical Reasoning	0.5 to 54.5	27.46	10.69
P. M. A. ^c Reasoning	8 to 30	19.68	4.51

a B.D. stands for Block Design

b D.A.T. stands for Differential Aptitude Tests

c P.M.A. stands for Primary Mental Abilities

TABLE III.-

Test of Linearity of Regression for the Simple Correlations Between Block Design and the Tests of Reasoning Ability.

Variables	F. Test Ratio ^a
B.D. and P.M.A., Reasoning	.2763 ^b
B.D. and D.A.T., Verbal Reasoning	.4126 ^b
B.D. and D.A.T., Abstract Reasoning	1.1628 ^b
B.D. and D.A.T., Mechanical Reasoning	.8490 ^b

^a Test ratio when the regression function for the dependent variable is based on each of the independent variables.

^b Not significant.

TABLE IV.-

Test of Linearity of Regression for the Simple Correlations Between Block Design and the Tests of Reasoning Ability.

Variables	F. Test Ratio ^a
P.M.A., Reasoning and B.D.	1.0424 ^b
D.A.T., Verbal Reasoning and B.D.	.6721 ^b
D.A.T., Abstract Reasoning and B.D.	1.4031 ^b
D.A.T., Mechanical Reasoning and B.D.	1.4148 ^b

^a Test ratio when the regression function for each of the independent variables is based on the dependent variable.

^b Not significant.

was found to reject the hypothesis of a straight line of regression in any of the correlations.

As for the condition of homoscedasticity, the scattergrams indicated clearly that for each coefficient of correlation computed the dispersions of the raw scores were approximately equal. The distributions were thus fairly symmetrical and homoscedasticity prevailed.

After verification of the assumptions underlying the product-moment correlation, the intercorrelations were obtained and tabulated. Since only the relationships between the dependent and independent variables are of immediate interest in this investigation no reference has been made to intercorrelations among the independent variables. However, these are available for reference since a complete tabulation has been included in Appendix I. The degree of relationship between Block Design and each of the independent variables has nevertheless been reported and will be found in Table V. The standard error of an assumed coefficient of zero for the sample used in this study is .0909.

As it has been mentioned earlier, two methods were used in the evaluation of the significance of the relationship between the variables. Where the standard error of a coefficient of zero was employed, all the correlations were found to be statistically significant at the .005 level of confidence and above. The relationship between Block Design

TABLE V.-

Correlations Between the Various Elements of the Test Battery and the Total Score of Block Design, with a Sample of One Hundred and Twenty-Two Subjects.

Independent Variables	Block Design
Differential Aptitude Test	
Abstract Reasoning	.5329 ^a
Verbal Reasoning	.5292 ^a
Mechanical Reasoning	.6174 ^a
Primary Mental Abilities	
Reasoning	.3436 ^a

The Standard Error for a correlation coefficient of zero for a sample of 122 subjects is .0909.

^a Significant at the .005 level of confidence.

and the P.M.A. Reasoning was the lowest of the group although it was significant at the .005 level of probability.

The use of the F test to verify the significance of the relationship between each of the independent variables and Block Design revealed a similar trend. Table VI shows that the F ratio for all the correlations was significant beyond the .005 level of probability. A definite relationship was found to exist between each of the measures of reasoning ability and the dependent variable (Block Design). The relationship between the D.A.T., Mechanical Reasoning and Block Design was found to be the most significant.

b) Relationship of all the tests of reasoning ability taken together to the total Block Design score. - The main purpose here was to evaluate the strength of relationship found between all the independent variables and the Block Design score. The Doolittle method for computation of the multiple coefficient was employed.

A multiple coefficient of correlation (R) of .6567 was obtained when all the tests of reasoning ability were compared to the total Block Design score.

As previously, two methods were employed to evaluate the significance of the relationship among the variables. First of all, the multiple coefficient of correlation was checked for significance by means of the F test.

TABLE VI.-

Test of Significance for the Simple Correlations Between Block Design and Each of the Independent Variables.

Independent Variables	Simple Correlations	Standard Error of r	Variance Ratio ^a
P. M. A. Reasoning	.3436	.0802	16.18 ^b
D. A. T. Verbal Reasoning	.5292	.0654	46.68 ^b
D. A. T. Abstract Reasoning	.5329	.0651	47.33 ^b
D. A. T. Mechanical Reasoning	.6174	.0562	74.74 ^b

a For formula used refer to page 81.

b Significant at the .005 level of confidence.

Table VII indicates a significant relationship beyond the .005 level of confidence. The same trend could be found when the standard error of a multiple coefficient was used as the criterion of significance. The relationship is also significant beyond the .005 level of probability. After the formula for the correction of the inflated value had been applied to the multiple coefficient obtained, the F test revealed once more that the relationship between the dependent variable (Block Design) and the independent variables (tests of reasoning ability) is significant at the .005 level of confidence and above.

c) Contribution of each variable to the total predicted variance in the total score on Block Design. - One of the main purposes of the investigation was to discover the amount of variance loading of each test in Block Design. This was accomplished through the analysis of the beta weights in the multiple correlation formula, which was discussed earlier in the section concerning the statistical methods. The results presented here in the form of a table, will be found in Table VIII.

Considering that all the relationships were found to be significant beyond the .005 level of probability it can be seen that the contribution of each variable to the total predicted variance in Block Design varies greatly. As revealed by Table VIII, the D.A.T., Abstract Reasoning

TABLE VII.-

Test of Significance for the Obtained Multiple Correlation Coefficient with the Dependent Variable.

Dependent Variable	Multiple Correlation	Standard Error of R	Variance Ratio ^a
Block Design Total Score	.6567	.0526	17.59 ^b

a For formula used refer to page 82.

b Significant at the .005 level of confidence.

TABLE VIII.-
Amount of Variance Loading of Each Test in Block
Design.

Independent Variable	Variance Loading
Differential Aptitude Test	
Abstract Reasoning	.0303
Verbal Reasoning	.1212
Mechanical Reasoning	.2667
Primary Mental Abilities	
Reasoning	.0131
Total Variance Loading	.4313

controlled 3% of the variance while the D.A.T., Verbal Reasoning accounted for 12.1% of the variance loading. The D.A.T., Mechanical Reasoning test explained 26.7% of the variance as compared to the P.M.A., Reasoning which controlled 1.3% of the variance in the Block Design total score. The sum total of the variance accounted for by all four tests was 43.1%. Since only 43.1% of the variance loading in Block Design was accounted for, 56.9% of the total predicted variance in the total Block Design score still remained to be controlled.

The first section of this chapter has been dealing with the results obtained from the sample studied and various modes of presentation have been used. The raw data was summarized in a table which contained information such as the range of scores, the means, and standard deviations for each test administered. The various correlations between each of the independent variables and the Block Design total score were also reported in a tabular form and the size of relationship needed for significance was noted. The multiple correlation coefficient accompanied with its standard error and variance ratio was also introduced in a similar manner. The contribution of each independent variable to the total predicted variance as well as the sum total of the variance accounted for by all four tests of reasoning ability were also considered and presented in a table.

Having presented all the data for consideration, there still is an important question to be discussed; that is, the analysis of results, which forms the core of the following section.

2. Analysis of Data

The discussion at this point will be centered around the significance of the relationships obtained in terms of the criteria used. More precise explanations will thus be given as to the meaning of the relationships between the Block Design score and the different modalities of the reasoning ability as measured by the different tests of the battery employed.

a) Block Design and each independent variable. - As previously noted all relationships were significant at the .005 level of confidence. The highest correlation occurred for the D.A.T., Mechanical Reasoning and was found to be $.6174 \pm .0562$. The D.A.T., Abstract Reasoning and Verbal Reasoning had a correlation of $.5329 \pm .0651$ and $.5292 \pm .0654$, respectively. The lowest relationship was reported for the P.M.A., Reasoning test with a correlation of $.3436 \pm .0802$.

As suggested by the size of the correlations obtained, a substantial relationship exists between the three D.A.T. measures of reasoning and the dependent variable.

The relationship between P.M.A. Reasoning and Block Design is, however, rather small. The results, thus, suggest that the factors which are supposedly measured by these four tests are also measured to a certain degree by Block Design. Block Design, then, to some extent, measures "the ability to solve logical problems -- to foresee and plan"¹, "the ability to perceive relationships in abstract figure patterns - generalization and eduction of principles from non-language designs"², as well as, "the ability to understand concepts framed in words -- to abstract or generalize and to think constructively".³ It also measures the ability "to learn the principles of operation and repair of complex devices".⁴ The reasoning ability, as described by the authors of the measures employed, appear to be involved in the solution of the task.

b) Multiple Correlation and the relative contribution of each independent variable. - Although each component of the multiple correlation has been analyzed in terms

1 L.L. Thurstone, and T.G. Thurstone, Examiner Manual for the SRA Primary Mental Abilities, Preliminary Edition, Chicago, Science Research Associates, 1947, p. 1.

2 George K. Bennett, Harold G. Seashore, and Alexander G. Wesman, A Manual for the Differential Aptitude Tests, New York, Psychological Corporation, 1952, p. 7.

3 Id., op. cit., p. 5.

4 Id., op. cit., p. 7.

of its contribution to the total predicted variable, nothing has been said of the nature of this contribution.

Table VIII, as presented in the preceding section, enabled the reader to obtain a more definite idea of the relative importance of each variable. The table revealed that the P.M.A., Reasoning test contributed .0131 or 1% to the total predicted variance, thus indicating that "the ability to solve logical problems -- to foresee and plan"⁵ contributes a little to solving the Block Design task. A somewhat higher contribution was reported for the D.A.T., Abstract Reasoning. In this case, "the ability to perceive relationships in abstract figure patterns -- generalization and eduction of principles from non-language designs"⁶ contributed .0303 or 3% to the total predicted variable. However, the D.A.T., Verbal and Mechanical Reasoning taken together appear to be the main components; they represent 39% of the total explained variance. The D.A.T., Verbal Reasoning, or the "ability to understand concepts framed in words -- to abstract or generalize and to think constructively"⁷ contributes 12% or .1212 to the Block Design score, while the D.A.T., Mechanical Reasoning or the ability "to

5 L.L. Thurstone and T.G. Thurstone, op. cit., p. 1.

6 George K. Bennett, Harold G. Seashore, and Alexander G. Wesman, op. cit., p. 7.

7 Id., op. cit., p. 5.

learn the principles of operation and repair of complex devices"⁸ account for .2667 or 27% of it.

The analysis of the results has indicated so far that .4313 or 43% of the variance is accounted for by whatever is measured by the four tests of the battery which has been used in this investigation. It appears quite clearly also that two tests or whatever is measured by them, eliminating from double consideration things that they have in common, explain over one third of the variance. It is interesting to note that the verbal and mechanical factors are prominent.

These factors have been reported previously although they might, at times, have been interpreted differently. As it has been noted in the review of the literature, a verbal element has been suggested by Wechsler⁹ and Birren¹⁰ in order to explain the performance on Block Design. Wechsler¹¹ has also suggested, as did Goldstein and

⁸ Id., op. cit., p. 7.

⁹ David Wechsler, The Measurement of Adult Intelligence, Third Edition, Baltimore, Williams and Wilkins Company, 1944, p. 92.

¹⁰ J.E. Birren, "A Factorial Analysis of the Wechsler-Bellevue Scale Given to an Elderly Population", in the Journal of Consulting Psychology, Vol. 16, No. 5, issue of October 1950, p. 402.

¹¹ David Wechsler, op. cit., p. 93.

Scheerer¹², that it measures in a certain way the abstract attitude. A similar factor has been isolated by Corter¹³ who named it "concept ability". Meili¹⁴ discovered a factor which could be associated with the verbal reasoning element identified by the test used in this investigation. The reasoning element was also reported by Balinsky¹⁵ for the age groups 35 to 44 and 50 to 59.

While the present findings indicate that Block Design measures to a certain extent the Verbal Reasoning ability or the ability to abstract or generalize and to think constructively, they also strongly suggest that the Mechanical Reasoning ability is being tapped to a greater degree by this test. Cohen's interpretation of the non-verbal organization factor that he isolated might contain, to a certain extent, the mechanical reasoning ability or

12 Kurt Goldstein and Martin Scheerer, "Abstract and Concrete Behavior: an Experimental Study with Special Tests", in the Psychological Monographs, Vol. 53, No. 2, Whole No. 239, 1941, p. 32-33.

13 Harold M. Corter, "A Factor Analysis of Some Reasoning Tests", in Psychological Monographs, Vol. 66, No. 8, issue of March 1952, p. 23.

14 Richard Meili, "L'analyse de l'intelligence", in Les Archives de Psychologie, Vol. 31, No. 121, issue of March 1946, p. 20.

15 Benjamin Balinsky, "An Analysis of the Mental Factors of Various Age Groups from Nine to Sixty", in the Genetic Monographs, Vol. 23, First Half, issue of February 1941, p. 229.

the ability to learn the principles of operation and repair of complex devices. His later suggestion, however, that Block Design is useful for the purpose of measuring perceptual organization seems to agree better with the present findings. Davis¹⁶ also reported that Block Design measures a mechanical element.

The foregoing discussion has been mainly concerned with the presentation and analysis of the results obtained. A first section studied the significance of the relationship of each independent variable to Block Design, as well as of all the independent variables taken together with the dependent variable. This was followed by a discussion of the contribution of the independent variables to the total variance in Block Design. The second section dealt with the significance of the relationships obtained in terms of the criteria used. More precise explanations were furnished as to the meaning of the relationships discovered and the findings were interpreted in the light of some studies reported in the literature.

16 Paul C. Davis, "A Factor-Analysis of the Wechsler-Bellevue Scale", in the Educational and Psychological Measurement, Vol. 16, No. 1, issue of Spring, 1956, p. 141.

SUMMARY AND CONCLUSIONS

This report deals with an investigation of some aspect of the rationale of Block Design, a subtest of the Wechsler Adult Intelligence Scale. The difficulty of explaining adequately the nature of the function or functions involved in Block Design performance has pointed to the necessity of studying further the rationale of this test which appears as an important tool in clinical practice. Both the importance of such a study and the choice of the problem to be investigated were considered in the first chapter.

A review of the literature has brought to the fore numerous interpretative hypotheses which were discussed in the second chapter. An extensive survey included the most significant assumptions made on empirical grounds and findings based on experimental data. The suggested rationales were presented and the choice of the specific research hypothesis was proposed.

The procedure necessary for the investigation of the specific hypothesis was analyzed in the following chapter. In a first section, Block Design, the test under investigation, was described and its reliability was considered. Then, the instruments related to the hypothesis to be studied were introduced and special attention was given to the problem of the reliability of the tools. The population was described subsequently and certain factors such

as age, education, language, and sex were taken into consideration. A last section dealt with the choice of formulae and statistical procedures according to the requirements of the nature of the investigation.

The last chapter was concerned with the presentation and analysis of the obtained data. The writer discussed the degree of relationship existing between each of the measures of reasoning ability and the total score on Block Design. This discussion led to the consideration of the two conditions necessary for the use of the Pearson coefficient of correlation, namely, linearity of regression and homoscedasticity. Both conditions were met. Two methods for evaluating the significance of the relationship between the dependent and independent variables were used and, in either cases, all relationships were found significant. The strength of relationship found between all the independent variables taken together and the Block Design score was then presented and checked for significance. All relationships were found significant. The contribution made by each of the independent variables to the total predicted variance was also examined.

The discussion, in the last section, was centered around the significance of the relationships obtained in terms of the criteria used. Although a substantial relationship exists between the three D.A.T. measures of

reasoning and Block Design, the most important correlation occurred with the Mechanical Reasoning test. A rather small correlation between the P.M.A., Reasoning test and the dependent variable was reported. The results, thus, indicated that the functions which are supposedly measured by these four tests are also measured to a certain degree by Block Design. The relative contribution of each independent variable was also reported and the Mechanical Reasoning factor accounted for 27 percent of the total explained variance. The analysis of the results indicated that 43 percent of the total variance is accounted for by whatever is measured by the four tests employed.

From the nature of these findings emerges a general conclusion, namely, Block Design measures a reasoning element or factor. This factor, moreover, can be differentiated into various modalities. It is possible to say that the mechanical reasoning ability is the reasoning modality which stands out in this study. There is also evidence that Block Design measures a verbal modality and to a lesser extent it includes an abstract reasoning element. These conclusions, however, could be applicable only to a female population whose age range is between fifteen and nineteen.

Many assumptions remain to be investigated if Block Design is to be used with maximum efficacy in clinical diagnosis. Some suggestions for further research can be

made in relation to the findings reported here. Since there is evidence that three reasoning modalities are found in Block Design performance when the results of fifteen to nineteen years old high school girls are analyzed, it is suggested that similar findings could be duplicated with experimental groups of a different age. Investigations whose purpose would be to verify this hypothesis are in order. Assuming that the same modalities are found for different age groups it would be of diagnostic interest to investigate the relative importance of these modalities with various clinical groups. Investigations with a male sample which would compare in terms of criteria to the one used in this study would presumably be of value. It might be hypothesized that the relative contribution of each independent variable would be different due to the sex differences. Studies which would take into account the qualitative approach to the solution of the task might also be revealing.

BIBLIOGRAPHY

Balinsky, Benjamin, "An Analysis of the Mental Factors of Various Age Groups from Nine to Sixty", in the Genetic Psychology Monographs, Vol. 23, First Half, issue of February 1941, p. 191-234.

This study of the Wechsler-Bellevue consists of an analysis of the various mental factors present in different age groups and their reorganization over a period of time. It is thought that Block Design measures a verbal factor and a performance factor at the age of fifteen and some form of reasoning in the age group thirty-five to forty-four.

Benton, A.L., "A Study of the Performance of Young Adults on the Kohs Block Designs Tests", in the Journal of Applied Psychology, Vol. 25, No. 4, issue of August 1941, p. 420-427.

Using a sample of men and women within the age range of twenty to twenty-nine, it was found that sex and the educational level were two factors which influenced the performance on the Kohs Block Designs Tests.

Birren, J.E., "A Factorial Analysis of the Wechsler-Bellevue Scale Given to an Elderly Population", in the Journal of Consulting Psychology, Vol. 16, No. 5, issue of October 1952, p. 399-405.

A factor analysis of the mental abilities measured by the Wechsler-Bellevue in an elderly group of subjects was made with the specific purpose of aligning these factors with the Thurstone factors. It is suggested that Block Design measures closure, memory, and to some extent, induction.

Bolle, Arthur, A Factor Analysis of the Echelle D'Intelligence Ottawa-Wechsler, unpublished Master's thesis presented to the Institute of Psychology and Education of the University of Ottawa, 1954, xi-60 p.

A study of the Ottawa-Wechsler consisting of an analysis of the factors present in the subtests revealed that different components were present at various age levels for the same tests. It is indicated that Block Design measures mainly the ability to evaluate and organize visual material, planning, and visual-motor co-ordination.

Cohen, Jacob, A Comparative Analysis of the Factors Underlying Intelligence Test Performance of Different Neuro-Psychiatric Groups. Multiple Factor Analyses of the Wechsler-Bellevue Intelligence Scale Performance of Schizophrenia, Psychoneurotic, and Brain-Damaged Groups, microfilmed Doctor's thesis presented to the University of New York, New York, March 1950, vii-162 p.

The factorial structures underlying the performance of a neuropsychiatric population on the various subtests of the Wechsler-Bellevue were investigated and various factors were isolated. Block Design was found to measure non-verbal perceptual organization and distractability.

Cohen, Jacob, "The Factorial Structures of the WAIS Between Early Adulthood and Old Age", in the Journal of Consulting Psychology, Vol. 21, No. 4, issue of August 1957, p. 283-290.

A factor analysis of the WAIS with four age groups of normals revealed that the same factors were present over the entire age range and that these factors were essentially the same factors as had been identified for a neuropsychiatric population on the Wechsler-Bellevue. It is indicated that Block Design is useful for the purpose of measuring perceptual organization over the entire age range.

Corter, Harold M., "A Factor Analysis of Some Reasoning Tests", in the Psychological Monographs, Vol. 66, No. 8, issue of March 1952, p. 1-31.

A group of reasoning tests were investigated and common factors were identified. The author reported that Block Design measures four different modalities: a spatial factor or the ability to recognize size, shape, and congruence; productiveness; concept ability or the ability to recognize essential similarities, to abstract and generalize, to think inductively; and flexibility or plasticity.

Davis, Paul C., "A Factor-Analysis of the Wechsler-Bellevue Scale, in the Educational and Psychology Measurement, Vol. 16, No. 1, issue of Spring 1956, p. 127-146.

An attempt was made to relate Wechsler's findings to these factors well established by other research. It is reported that Block Design had the highest communality of all the performance tests. It contributed significantly to the measurement of its three leading factors: visualization, mechanical knowledge, and perceptual speed.

Frank, G.H., C.C. Corrie, and J. Fogel, "An Empirical Critique of Research with the Wechsler-Bellevue in Differential Psychodiagnosis", in the Journal of Clinical Psychology, Vol. 11, No. 3, issue of July 1955, p. 291-293.

The authors attempted to investigate the influence of the selection of subjects using the psychiatric diagnosis as a criterion. It was concluded that the inconsistent, inconclusive, and sometimes contradictory results obtained in testing the efficacy of the Wechsler-Bellevue in differential diagnosis, were due to research methodology rather than to faults inherent to the tool itself.

Garfield, Sol L., "An Evaluation of Wechsler-Bellevue Patterns in Schizophrenia", in the Journal of Consulting Psychology, Vol. 13, No. 4, issue of August 1949, p. 279-287.

Studies concerned with diagnostic patterns were criticized because of the neglect of sampling problems, the inadequacy of criteria for diagnostic classification of subjects, and the limitation of various methods of study employed. The importance of controlling variables such as age, education, I.Q., and proper classification were emphasized.

Goldstein, Kurt and Martin Scheerer, "Abstract and Concrete Behavior: an Experimental Study with Special Tests", in the Psychological Monographs, Vol. 53, No. 2, Whole No. 239, 1941, 151; p.

The rationale of a battery of tests is explained in terms of the concrete and abstract intellectual attitude involved in the solution of a task. Observation and empirical findings are reported. Block Design is offered as one of the instruments which measures the concrete and abstract intellectual behavior.

Kohs, S.G., Intelligence Measurement, New York, the MacMillan Company, 1927, xii-312 p.

A manual which includes an analysis of the current definitions of intelligence, comparison of Block Design results with other measures of intelligence, and discussion of the value of this test as a tool to evaluate the mental capacities. The Block Design Test is offered as a non-verbal measure of intelligence. More specifically, it is considered as a measure of the ability to analyse and to synthesize.

Marks, Melvin R., "A Criticism of the Use of the Wechsler-Bellevue Scale as a Diagnostic Instrument", in the Journal of General Psychology, Vol. 49, First Half, issue of July 1953, p. 143-152.

The use of the Wechsler-Bellevue as diagnostic instrument was analyzed for adequacy with respect to suggested criteria of efficiency, validity, and univocality. The author reached the conclusion that the test is not efficient, lacks validity, and is not univocal as a diagnostic tool.

Meili, Richard, "L'analyse de l'intelligence", in Les Archives de Psychologie, Vol. 31, No. 121, issue of March 1946, p. 1-64.

In an attempt to explain the fundamental intellectual organization, thirteen tests were used with children and adult subjects. The factors plasticity and complexity were found with a very high loading in Block Design both with children and adults.

Rapaport, David, Merton Gill, and Roy Schafer, "The Bellevue Scale", in Diagnostic Psychological Testing: The Theory, Statistical Evaluation, and Diagnostic Application of a Battery of Tests, Vol. 1, Chicago, Year Book Publishers, 1945, p. 44-318.

A thorough analysis of the Wechsler-Bellevue consisting mainly in the description of each subtest, a statistical demonstration of the significance of scatter, and a statistical analysis of data obtained from control and clinical samples. Block Design is seen as a measure of visual-motor co-ordination which implies concept formation.

Schofield, William, "Critique of Scatter and Profile Analysis of Psychometric Data", in the Journal of Clinical Psychology, Vol. 8, No. 1, issue of January 1952, p. 16-22.

A critique of pattern studies based mainly on the discussion of the lack of well defined rationales, the inadequate control of pertinent variables and the deficiencies of research designs. The patterning effect of age, educational level, I.Q. level, actual symptomatic status, and duration of illness is considered very important.

Shapiro, M.B., "An Experimental Approach to Diagnostic Psychological Testing", in the Journal of Mental Sciences, Vol. 97, No. 409, issue of October 1951, p. 748-764.

This article was concerned with the utilization of psychological tests for diagnostic purposes. The author discussed the needs which give rise to the use of the standardized validated test, and the difficulties that arise when the results are being evaluated. The main thesis put forward was that the difficulties can, to a large degree, be met by making explicit the principles of experimental method.

Watson, Robert I., "The Use of the Wechsler-Bellevue Scales: A Supplement", in the Psychological Bulletin, Vol. 43, No. 1, issue of January 1946, p. 61-68.

It is suggested that the psychological functions underlying the subject's performance should be investigated more thoroughly either by means of a correlation approach or by factorial analysis. Educational and cultural background are believed to influence the research findings even if the subtests were pure functional entities.

Wechsler, David, The Measurement of Adult Intelligence, Third Edition, Baltimore, Williams and Wilkins Company, 1944, vii-258 p.

A manual which contains administration and scoring methods, interpretative data, standardization procedures, as well as practical applications of the Wechsler in the clinical field. The author suggests that Block Design is an excellent test of general intelligence; it involves both synthetic and analytic ability.

Wechsler, David, The Measurement and Appraisal of Adult Intelligence, Fourth Edition, Baltimore, Williams and Wilkins Company, 1958, x-297 p.

This manual includes a discussion of the nature of intelligence and of its measurement, a presentation of the methods of administration, scoring, and interpretation of the WAIS, as well as suggestions and advices regarding the practical utilization of the Bellevue Scales. Block Design is regarded as an excellent test of general intelligence. Difficulty in visual-motor organization, however, affects the scores.

APPENDIX I

THE RAW DATA

The raw data is presented in the following pages. The scoring methods used by the writer in this study are the ones devised by the respective authors. In order to obtain further information about the scoring methods, the reader could refer to chapter three which deals with the experimental design and more specifically with the description of the tools.

The Block Design subtest scores tabulated in Table IX are composed of both the accuracy and time bonus scores.

The test used in the battery can be identified in Table IX by symbols which are explained in the following manner:

Code: Test's identification number;
T1 : Block Design;
T2 : P. M. A. Reasoning;
T3 : D. A. T. Verbal Reasoning;
T4 : D. A. T. Mechanical Reasoning;
T5 : D. A. T. Abstract Reasoning.

Table X includes the correlations which were computed between the dependent and independent variables as well as the intercorrelations between the independent variables.

TABLE IX.-

Raw Data Obtained from the Test Battery

Code	T-1	T-2	T-3	T-4	T-5
1	40	20	16	32.5	31
2	29	15	21	17	24
3	34	24	21	17	25
4	32	16	18	37	16
5	24	21	16	12.5	24.5
6	24	17	22	12.5	24
7	24	17	17	20	17.5
8	24	19	14	7	21
9	20	8	15	7	7.5
10	37	13	26	12	30
11	30	19	28	38	32
12	39	24	41	30.5	36.5
13	26	9	4	2	0
14	40	22	25	29	35.5
15	39	20	41	20	33.5
16	32	20	41	33.5	39
17	22	21	16	8	28
18	30	13	31	27.5	32.5
19	40	14	36	29	36
20	33	20	43	36.5	40
21	37	19	32	34.5	34.5
22	36	19	20	17	21
23	28	19	28	12.5	22.5
24	35	24	33	38	39
25	28	19	28	23	44
26	47	26	39	39.5	37.5
27	47	25	41	41.5	38
28	28	26	30	21.5	33
29	42	19	33	30.5	33.5
30	33	18	24	20.5	29
31	24	18	23	27.5	30
32	28	12	21	.5	16.5
33	28	15	24	24.5	23.5

TABLE IX.-

Raw Data Obtained from the Test Battery (Continued)

Code	T-1	T-2	T-3	T-4	T-5
34	38	19	21	35.5	28
35	30	20	28	40	39.5
36	38	17	17	32	20
37	24	24	20	5.5	29
38	28	11	36	20	24
39	28	21	24	24	31
40	24	15	17	8	29
41	47	21	43	53	39.5
42	24	22	25	23	37.5
43	22	10	12	16	25
44	24	19	34	24.5	29
45	24	21	16	24	28
46	28	27	28	20.5	36
47	33	28	37	35	38
48	22	13	24	11	16.5
49	43	14	30	34.5	29
50	26	22	33	33.5	35
51	46	22	38	42.5	45
52	45	20	32	48.5	39
53	24	12	27	17	32
54	29	22	21	17	30
55	41	16	24	26	31.5
56	28	16	18	12.5	19.5
57	26	11	11	10	3
58	35	17	41	23	43
59	37	13	16	44	35
60	36	18	26	31.5	37
61	28	26	37	23	35
62	28	20	36	29.5	36.5
63	32	14	29	28	35
64	32	22	29	24.5	37.5
65	40	19	30	17.5	40.5
66	24	17	34	36.5	35

TABLE IX.-

Raw Data Obtained from the Test Battery (Continued)

Code	T-1	T-2	T-3	T-4	T-5
67	20	21	26	18.5	20
68	25	16	18	25.5	30
69	36	20	42	33	42
70	26	18	12	19	22.5
71	32	19	26	29.5	33
72	44	20	27	42	28
73	25	14	38	29.5	33
74	28	25	19	32.5	35
75	32	26	21	41	41
76	41	22	30	42.5	35
77	38	21	34	26	34
78	35	20	21	36.5	45
79	39	20	26	27.5	40
80	24	13	25	14.5	27.5
81	41	23	28	21.5	31
82	34	14	28	16	31
83	37	15	21	14	21
84	30	12	26	22	35
85	45	24	36	30	43
86	45	18	46	54.5	34.5
87	46	22	31	39.5	42
88	35	30	43	48.5	45
89	39	23	38	40.5	39
90	34	20	40	27	32
91	39	16	34	23.5	32.5
92	37	24	38	35	31.5
93	30	17	33	29.5	39
94	39	22	44	34.5	41
95	44	18	40	44	45
96	35	24	33	26.5	35
97	20	24	26	20	39
98	33	14	32	35	31
99	33	21	34	23	32
100	32	20	26	21	34

TABLE IX.-

Raw Data Obtained from the Test Battery (Continued)

Code	T-1	T-2	T-3	T-4	T-5
101	39	28	35	29.5	44
102	38	21	31	33.5	34
103	39	14	40	26.5	35
104	43	19	33	43.5	29
105	25	21	25	27	29
106	32	28	38	28	34
107	38	25	45	33	45
108	38	26	29	39	39
109	30	15	30	20	36
110	29	24	46	30	34
111	42	26	38	39.5	40
112	39	22	44	40	43
113	33	24	31	25.5	37.5
114	46	30	45	25	39
115	45	26	42	40.5	45
116	38	20	28	18.5	41
117	34	23	34	31.5	41
118	46	29	38	36.5	46
119	38	22	45	37	44
120	36	24	25	21.5	44
121	38	22	44	36	35.5
122	34	16	22	36.5	30.5

TABLE X.-

Correlations Between and Among the Dependent and Independent Variables.

	F. M. A. Reasoning	D. A. T. Verbal Reasoning	D. A. T. Abstract Reasoning	D. A. T. Mechanical Reasoning
Block Design	.3436	.5292	.5329	.6174
P. M. A. Reasoning		.4496	.4364	.4117
D. A. T. Verbal Reasoning			.6779	.5662
D. A. T. Abstract Reasoning				.7044

APPENDIX 2

ABSTRACT OF

A Study of the Rationale of the WAIS Block
Design Subtest. Form I¹

In the present study, the writer undertook to determine one of the functions presumably measured by Block Design a subtest of the Wechsler Adult Intelligence Scale, and to what extent it is measured by this subtest. A survey of the literature has revealed numerous hypotheses which could have been investigated. The choice of the specific hypothesis for this investigation has been suggested mainly by the availability of adequate instruments and the nature of the assumptions reported in the literature.

Thus, considering these factors, the writer arrived at the specific hypothesis under investigation namely, whether or not Block Design measures the reasoning ability, and, if so, to what extent. This hypothesis was studied through the use of the P. M. A. Reasoning subtest and the D. A. T. Abstract Reasoning, Verbal Reasoning, and Mechanical Reasoning subtests.

The sample selected consisted of one hundred and twenty-two High School English speaking female students whose

¹ M.A. Thesis presented by Jean Marie Labrecque, in 1961, to the School of Psychology and Education of the University of Ottawa, 121 pages.

age ranged from fifteen years and two months to nineteen years and eight months. They were all students from grade ten to thirteen inclusively and were unevenly distributed among the six classes composing the High School population.

The results obtained from the sample studied were tabulated and intercorrelations were computed in order to determine the extent of relationship among the dependent and independent variables. The multiple correlation coefficient was also reported. The contribution of each independent variable to the total predicted variance as well as the sum total of the variance accounted for by all four tests of reasoning ability were also considered.

From the analysis of the data emerges a general conclusion, namely, Block Design apparently measures a reasoning element or factor. This factor, moreover, can be differentiated into various modalities. The mechanical reasoning ability stands out and is being tapped to greater degree by this test.