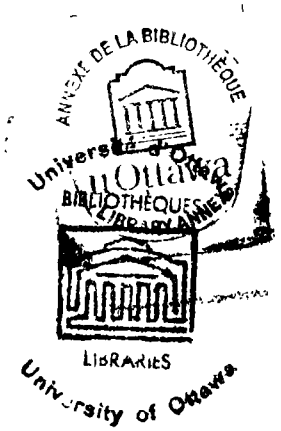
C3P-2
594

THE EFFECTIVENESS OF THE VERDUN PROJECTIVE BATTERY AS A PERSONALITY SCREENING DEVICE AT THE HIGH SCHOOL LEVEL

by

Arthur James Melbourne

Thesis presented to the Faculty of Arts of the University of Ottawa through the Institute of Psychology as partial fulfilment of the requirements for the degree of Doctor of Philosophy.



Verdun, Canada, 1955

UMI Number: DC53310

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 DC53310
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

ACKNOWLEDGMENTS

This thesis was prepared under the direction and guidance of Dr. Lawrence T. Dayhaw of the Institute of Psychology, University of Ottawa.

Permission to use the Verdun Projective Battery was generously granted by the authors, Dr. Heinz Lehmann and Dr. Herbert Dorken jr. of the Verdun Protestant Hospital staff; the latter and his assistants also provided material help in the administration of the tests to the students.

The authorities in the Department of Studies of the Montreal Catholic School Commission kindly gave permission for the testing to take place in The Thomas D'Arcy McGee Boys' High School and every cooperation was likewise received from those in charge of the direction of the school.

Finally, the writer's teacher-colleagues at this school assisted greatly by their contribution to the delicate task of rating the pupils on a personality basis and their patience and forebearance during the various interruptions in their teaching schedules necessitated by the project.

Deep gratitude is here expressed to all for their interest and cooperation.

CURRICULUM STUDIORUM

Arthur James Welbourne, born in Verdun, Que., on October 29, 1919, was awarded the degree of Bachelor of Arts by Loyola College, Montreal, on June 2, 1941, and the degree of Master of Arts by McGill University, Montreal, on October 5, 1946. The title of the M.A. thesis was: "A Study of Educational Practices in the Schools on the Island of Montreal

Other degrees include: Bachelor of Education on March 18, 1944 (Institute of Pedagogy, University of Montreal) Baccalauréat en Pédagogie on October 15, 1952 and Licence en Pédagogie on June 3, 1953 (Institut pédagogique Saint-Georges, Université de Montréal).

TABLE OF CONTENTS

Chapter		page
	INTRODUCTION	vi
I. -	DESCRIPTION OF THE TEST MATERIAL	1
	1. General aims of the problem	2
	2. Verdun Projective Battery	8
	3. Cornell Index - Form N2	16
	4. Rotter Incomplete Sentences Blank - High School Form	20
II. -	DESIGN OF THE EXPERIMENT	26
	1. Selection of the tests	26
	2. Selection of the subjects	31
	3. Administration procedure	40
	4. Method of analysis	49
III. -	ANALYSIS OF THE DATA	53
	1. Validity studies	53
	2. Academic standing comparisons	64
	3. Group versus individual administration	68
	4. Item analysis for reliability	71
IV. -	DISCUSSION OF THE DATA	75
	1. Validity appraisal	76
	2. Estimate of scholastic standing relationship	81
	3. Influence of type of administration	84
	4. Reliability evaluation	89
	SUMMARY AND CONCLUSIONS	93
	ANNOTATED BIBLIOGRAPHY	97

Appendix

1.	RATING INSTRUCTIONS TO TEACHERS	103
2.	VERDUN PROJECTIVE BATTERY	105
3.	CORNELL INDEX - FORM N2	107
4.	ROTTER INCOMPLETE SENTENCES BLANK - HIGH SCHOOL FORM	109
5.	ABSTRACT OF The Effectiveness of the Verdun Battery as a Personality Screening Device at the High School Level	111

LIST OF TABLES

Table		page
I.	- Classes Screened by Each Teacher-Rater.....	34
II.	- Breakdown by Grade of the "Normal" and "Abnormal" Student Groups.....	37
III.	- The Class, Age and Teacher-Selection Ratio of the "Normal" and "Abnormal" Student Groups...	39
IV.	- Scores Made on All Tests and Subtests by the "Normal" and "Abnormal" Subjects.....	55
V.	- Frequency Distribution of Verdun Projective Battery Subtest Scores for the "Normal" and "Abnormal" Groups.....	57
VI.	- Frequency Distribution of V.P.B., Cornell and Rotter Total Scores for the "Normal" and "Abnormal" Groups.....	58
VII.	- Means and Other Statistics Leading to the "t" Ratios for All Tests and Subtests.....	60
VIII.	- Obtained Frequencies Leading to Chi Squares for All Tests and Subtests.....	61
IX.	- Test and Subtest Intercorrelations for the "Normal" and "Abnormal" Groups.....	63
X.	- Academic Results in June 1954 of the "Normal" and "Abnormal" Subjects.....	66
XI.	- Obtained Frequencies Leading to Chi Squares for the Sex Difference and Campers vs. Students Comparisons.....	69
XII.	- Percentages of "Failure" on Each Item of the V.P.B. Subtests Leading to the Establishment of the Reliability Split Halves.....	73

INTRODUCTION

The departments of psychological services of modern school systems are seldom able to meet the many demands and calls made upon their facilities and techniques by harassed parents and educators, who are deeply concerned with the personal and scholastic problems of the eager and vibrant but often troubled and bewildered young adolescents entrusted to their charge. The discovery of a simple, rapid and accurate personality device that even in non-clinical hands would dependably screen those with serious psychopathic deviation, potential or actual, from the ones with more minor and less acute disorder would prove of great value in the present educational setting. The mental health and hygiene personnel of the school organizations could then concentrate all their attention on the former cases, being freed from the tedium of the less serious, but time-consuming, minor maladjustments, which could probably be very well handled under supervision by the guidance and counselling services with the assistance of the school nurse or social worker.

It was the intent of this research project to attempt to ascertain if the new and as yet commercially unavailable Verdun Projective Battery is a personality screening instrument sensitive to and capable of discriminating between these case differences at the high school level. The general technique

was to administer the Battery and two other personality screening tests to two teacher-selected "extreme" groups of male adolescent pupils of a large urban high school and subsequently make a statistical analysis and interpretation of the results obtained by means of these measuring instruments. The description of the procedure followed and the evaluation of the data secured constitute the essence of this report of the investigation.

The material obtained has been crystallized into four distinct chapters. The first of these provides a description of the test material utilized in the project. This comprised the three subtests (Word Association, Multiple Choice Rorschach and the Conformity Index) making up the Verdun Projective Battery, Form N2 of the Cornell Index, and the Rotter Incomplete Sentences Blank-High School Form. The general aims of the problem are also outlined and discussed. In addition to ascertaining the degree of validity and reliability of the Battery, three other associated aims form part of the project. These are the investigation of the influence of type of administration (group versus individual), the estimate of scholastic standing relationship to test and teacher rating, and finally, some brief study and examination of the degree of similarity in the results obtained by the projective and non-projective testing approaches.

The design of the experiment is dealt with in Chapter Two. Here are described how both the tests and the subjects were chosen and selected for the experiment. The administration procedure is presented in exact detail and the chapter concludes with the theory underlying the method of analysis proposed: "t" ratio, chi square analysis and correlation procedure, primarily.

The third and fourth chapters are very closely related, the former being devoted to the straightforward analysis of the data obtained, while the latter embraces the discussion and interpretation of their significance. Thus, Chapter Three opens with a treatment of the validity studies made, which is matched by the introductory appraisal of validity in the fourth chapter. This process is repeated in the three other matching sections of these chapters, namely, the academic standing relationship and estimate, the results and influence of administration type and procedure, and the evaluation of the data in the important area of reliability.

The main body of the dissertation closes with a short summary of the conclusions reached as a result of the whole investigation. Following the annotated bibliography, pertinent illustrative material is appended in the form of specimen copies of the rating instructions sheet issued to the teachers to assist them in their preliminary screening of the students and the three testing devices used in the project, together with, finally, the essential abstract of the thesis.

CHAPTER ONE

DESCRIPTION OF THE TEST MATERIAL

In this opening chapter, the matter to be treated has been subdivided for convenience into four distinct, though interrelated, sections. The first of these concerns itself with a short evaluation and highlighting of the principal aims and intentions, both general and specific, of the entire research project. Next in line will be found a brief description of the experimental test, the Verdun Projective Battery, together with some analysis of its three constituent subtests (Word Association, Conformity Index, and Multiple Choice Rorschach). The third section is devoted to an outline of the first of the companion testing devices, Form N2 of the Cornell Index. Finally, in the fourth and last section, similar treatment has been accorded the other measuring instrument made use of in the investigation, namely, the High School Form of the Rotter Incomplete Sentences Blank. Statistical data concerning validity and reliability are reported in each instance.

1. General Aims of the Problem

The field of personality testing seems to have been largely dichotomized into two major schools of thought, each with its own particular viewpoint, special emphasis, and able practitioners. There are those who favor the use of the personality inventory or questionnaire with its facility of group administration, ease of comprehension, and simplicity and objectivity of scoring. Included among its limitations are the conscious or unconscious modifications by the subject of his answers. Then, of recent years, has appeared and grown in very great proportions the myriad devices that group themselves under the heading of projective techniques. Frank N. Freeman refers to projection as being an unconscious process and defines the method thus:

> A projective test, then, is one that provides the subject with a stimulus situation giving him an opportunity to impose upon it his own private needs and his particular perceptions and interpretations. The several forms of the projective method (pictures, ink-blots, word associations, one's own writings and drawings) are intended to elicit responses that will reveal the individual's feelings, values, motives, characteristic modes of adjustment, or "complexes". He is said to project the inner aspects of his personality through his interpretations and creations, thereby involuntarily revealing traits that are below the surface and incapable of exposure by means of questionnaire type of personality test.[1]

---

1 Freeman, Frank S., Theory and Practice of Psychological Testing, New York, Holt, 1950, p. 400.

Lawrence Abt, in turn, speaks of projective psychology as being interested "in an inquiry with respect to the role of all the psychological functions and processes that operate within the context of the total personality".[2] John E. Bell further elucidates the operational technique in these words:

> Generally the examiner gives some inaccurate purpose as part of the structuring of the stimulus situation for the individual, or the subject is left to guess the examiner's intent. This reduces conscious control by the subject over the behavior which is to be analyzed and produces responses reflecting his own individuality. The interpretation the subject makes of the test situation provides the first reflection of the personality. It is assumed in these techniques that the individual organizes events in terms of his own motivations, perceptions, attitudes, ideas, emotions, and all other aspects of his personality..... Projective techniques emphasize primarily the uniqueness of the responses-those qualities that discriminate between individuals. Thus the best technique is that which will command the greatest range of responses in the shortest possible time.[3]

This development, in turn, has led to speculation on the part of psychometricians as to the feasibility of constructing an instrument which would combine the ease and objectivity and quantification of the questionnaire method, on the one hand, and the accuracy and qualification (because of the "unaware truths" revealed) and presumably greater

---

2 Abt, Lawrence Edwin and Leopold Bellak, Projective Psychology, New York, Knopf, 1950, p. 37.

3 Bell, John Elderkin, Projective Techniques, New York, Longmans, 1948 p. 4-5.

clinical validity, on the other, of the later projective approach. Is such a device possible? The authors of the Verdun Projective Battery have taken a preliminary step in that direction with their new instrument, which is as yet commercially unavailable for distribution.

It is the major aim of this project to attempt to ascertain the degree of validity and reliability this Battery -which represents a combination of the two approaches- has as a personality screening device at the high school level. This will bear some amplification. The test, which is described more fully in the section immediately following this one, purports to be a rapid but accurate means of screening for potential neuropsychiatric disturbance and maladjustment and its ten-year background has been concerned with the testing of psychotic patients in a large mental hospital. But the possibility of its practical application (the result is reported by means of a quantified score) to the real problem of solving decisions about referrals to the school psychological services seemed worth investigating. If the test in question is sensitive to real or incipient psychopathy or psychosis in a school population, where the degree of incidence and the level of severity is expectedly much less than that in an institution dedicated to the care of the mentally-ill, then one of the important concerns of present-day school life, namely, the prevention or early detection and treatment of such disturbance, will have been relatively alleviated.

It is pertinent to refer to the facts of the incidence of mental illness in our times. Coleman in discussing the problem stresses its importance as follows:

> It is startling to note that mental patients occupy almost as many hospital beds as all other patients combined; that approximately one out of twenty persons now living in the United States will at some time be hospitalized for mental illness alone. And for each of the 700,000 or more persons hospitalized for mental illness in the United States, there are at least twenty more outside who are not so severely maladjusted but are nevertheless in need of psychiatric aid. And these figures say nothing of the many kinds of organic illness brought on by emotional conflict or of other serious types of maladjustments, such as we see in juvenile delinquency and crime.[4]

He refers to these statistics as conservative estimates, incidentally. That these conditions also pertain, unfortunately, to Canada as well may be gathered from a recent leaflet of the Canadian Mental Health Association which frankly states:

> Out of every 100 Canadian school children, it is estimated, that at some time in their life, 5 will be committed to a mental hospital, 2 will go to prison, and at least 5 more will be an emotional problem to themselves, their families, or their teachers.[5]

School psychologists are few in number at the present time. Obviously, it is physically impossible for them to examine individually the many thousands of pupils crowding the schools to-day; even the ones at the critical adolescent

---

4 Coleman, James C., Abnormal Psychology and Modern Life, Chicago, Scott, Foresman, 1950, p.3.

5 Canadian Mental Health Association, "Let's Stop This Human Waste Now!", (no publisher),(no place),(no date), p.2.

levels cannot all be reached. But an accurate test, simple enough to be safely entrusted to the school guidance counselor or even the school nurse or social worker, would go far towards solving the problem of ascertaining which pupils stand in need of specialized treatment and should therefore be referred to the school psychologist or the mental health services and which ones have difficulties that are more amenable to environmental solution in school or home. The statistical evaluation of the test from the point of view ~~of view~~ of its degree of validity and reliability with boys in a large urban high school is, then, the basis of the research.

While such a problem as the one highlighted above is assuredly most important and timely under modern educational conditions, it is not the sole aim of the project. Three other related matters formed part of the investigation: 1) the respective roles played by group versus individual administration in the resulting scores on the Battery (the former with the exception of the timed subtest was the practice followed in the school testing situation, while the latter procedure was essential for the hospitalized psychotics); 2) the relationship of academic or scholastic standing to test and rater evaluation of the respondents; and 3) the examination of the projective (Verdun Projective Battery and Rotter Incomplete Sentences Blank) versus the non-projective (Cornell Index) test material utilized, from the standpoint of similarity of results obtained from the testing of the same subjects. The other tests mentioned

are described in the survey of the literature in the remaining sections of this first chapter.

This raises the question of why these particular instruments were chosen for comparison with the Verdun Projective Battery. However, as these matters will require considerable elaboration in their own right, and the description of the test material will take up the rest of this chapter, to avoid unwieldiness of presentation it has been decided to devote the entire second chapter to a consideration in detail of the experimental design, subsuming such matters as the reasons for and techniques of test and case selection, administration procedure, general method of analysis, etc. For the moment, then, let it suffice to say that the tests were chosen on their merits as screening and not diagnostic devices. This ruled out, for example, such excellent items as the Minnesota Multiphasic Personality Inventory, the Cattell Sixteen Personality Factors Questionnaire, etc. The aspects measured by these and similar instruments are not directly comparable to the more generalized screening technique of the Verdun Projective Battery.

2. Verdun Projective Battery

The Verdun Projective Battery (V.P.B.)[6] was designed by its authors, Heinz Lehmann, M.D. (Clinical Director) and Herbert Dorken jr., Ph.D. (Psychologist in Charge) for use with psychotic patients in a mental institution.[7] It is primarily a rapid (approximately fifteen minutes for individual administration) screening device to detect potential neuro-psychiatric casualties and psychotic disorders and the material has been gathered over a ten-year period (1944-1954). Brevity is essential in testing those with psychotic disorders, for their attention span is often insufficient for psychological tests generally. It is to be emphasized and stressed that this test is designed for screening and not diagnostic purposes; this is a fundamental point that must be borne in mind throughout this report.

---

6 The data of this section were taken from:
a) Lehmann, Heinz and Herbert Dorken, jr., "The Administration, Scoring, and Percentile Standardization of the Verdun Projective Battery", in the Canadian Journal of Psychology, Vol. 7, No. 2, issue of June 1953, p. 69-80.
b) Lehmann, Heinz and Herbert Dorken, jr., "The Clinical Application of the Verdun Projective Battery", in the Canadian Journal of Psychology, Vol. 6, No. 4, issue of December 1952, p. 164-172.

7 The Verdun Protestant Hospital, a large institution (1700 beds) devoted to the care of all the certified non-Catholic mental patients of the Province and located in the City of Verdun (hence the title), a suburb of Montreal, Quebec, the metropolis of Canada.

The subtests comprising the Battery in its present form are: 1) a ten-word Word Association Test, which is timed with a stop-watch; 2) a ten-word untimed word-association test (from the Kent-Rosanoff tables) described as a Conformity Index; and 3) series A of Harrower's Multiple Choice Rorschach. Though the Battery is at present not commercially distributed, permission was readily granted for its use in the present situation.

Possibly it should be pointed out that this research is only one among several studies currently being conducted in an effort to ascertain the full assets and applicability of the Verdun Projective Battery. Though the subjects tested, both in the standardization period and in actual clinical practice, were for the most part normal or psychotic young adults, normative data have been prepared and extended from the age of majority to beyond the age of retirement, i.e., the chronological levels from 21 to 79 years inclusive. Other specialized investigations are currently being conducted on such related segments as the geriatric group (even up to 94 years of age), the years of childhood and youth, and, of course, in the present case, the adolescent level. Other analyses are being made of possible sex differences in the scores obtained. These, incidentally, seem to be negligible; however, the results with the youngsters present some deviation and have not as yet been stabilized; furthermore, the aged,

as might have been foreseen, tend to a quite marked prolongation of reaction-time on the timed Word Association subtest.

Word Association (W.A.).- This subtest purports to measure mental control and refers to the present psychological situation of the testee. The instructions given are effectually the following: "I am going to say a word and I would like you to tell me the first word that comes to mind." Following a preparatory example, during which instructions may be given, if deemed essential, on the necessity of responding with the first (and only one) word coming to mind, the following ten author-selected words are offered as stimuli (the norm reaction-times in seconds are in parentheses): table (2.0), river (2.5), man (2.0), food (3.0), pencil (2.5), voice (2.5), army (3.0), sickness (2.5), color (3.0), money (3.5). The authors established these reaction-times from their standardization groups of normal people and this led to scoring the responses as follows: 80 percent of the normal group stayed within the limits for words 4- 6- 8- 9- 10 (food- voice- sickness- color- money) and so any longer reaction-time is scored 1; 90 percent of the normals did likewise for words 1- 2- 3- 5- 7 (table- river- man- pencil- army) and hence any delayed response here is double-weighted and scored 2, -thus the test carries a maximum possible score of 15. Three minutes or less usually suffices to administer the Word Association Test.

Conformity Index (C.I.).- This subtest serves primarily to screen out thought disorders and is indirectly related to contact with reality. It is very rare for a normal subject to run up a score of appreciable proportions on the ten items with the possible exception of gifted creative subjects who are original in their thinking or those whose first language is other than English. The development of language is, therefore, a variable here. As a conformity measure, it is probably not valid for non-English speaking subjects. Since it is customarily administered immediately following the Word Association Test, it is usually sufficient to remind the subject of the former procedure and ask him to "write down beside each word, the first word that comes to mind when it is read". Qualitative analysis might include graphological signs and deviations. The words were selected from the well-known Kent-Rosanoff list on the basis of having "a common or popular response of the highest frequency value" and are as follows (the conforming associations are in parentheses): man (woman), cabbage (vegetable), hard (soft), eagle (bird), ocean (water), lamp (light), long (short), tobacco (smoke or smoking), scissors (cut or cutting), and blossom (flower). Again, the authors found that 90 percent of the normal standardization group answered with four or more of the selected common responses. Accordingly, the cut-off point of normality was

set at this level, i.e., four or more conforming associations merit a zero score for the testee. Because, as noted above, common responses fewer than four decreased very rapidly with normal subjects (but the reverse for psychotic patients), they established the following weighted score scale: 3 responses = score of 3, 2 = 6, 1 = 12, and 0 = 15. This subtest is generally completed in under two minutes and has a maximum possible score of 15.

Multiple Choice Rorschach (M.C.R.).- This subtest attempts to analyse the basic personality resources and stability of the person being tested and has reference to his underlying fundamental psychological condition. The subject is given the test sheet listing the one hundred items (ten for each Rorschach card) from Series A of Dr. Molly Harrower's (who generously gave permission for their inclusion in the Battery) Multiple Choice Rorschach. He then is told in substance the following: "You will be shown on the screen ten ink blots, one at a time. For each of these blots you have listed on your sheet ten possibilities which might resemble the whole or any part of the blot. Simply draw a line <u>through</u> that <u>one</u> which to you provides the most satisfactory resemblance. Check <u>all</u> ten choices carefully before you make your selection. Pick only <u>one</u> for each blot. If you have any questions, ask them now, for no questions whatsoever will be allowed once we start the slides." The questions

provide the opportunity to review the key instructions, namely, the line to be drawn through and not over or under the choice (prevents ambiguity for the scorer), only one choice per blot to be made, the reminder to neglect other things that may be "seen" but which do not appear on the listings, the essential futility of copying (in the group situation) since all answers are only useful and helpful if personal to the testee, the necessity of selecting the strongest resemblance or similarity, etc. Needless to say the tester must guard against inadvertently mentioning color or shading in his instructions and must minimize them by offhand, superficial, noncommittal or guarded replies (in the form of a repetition of an appropriate part of the basic directions) if they should arise, as appears unlikely, especially with or among the test-innocent, in the question-period. Incidentally, the problem did not appear in the school group-testing situation described herein.

In contradistinction to the scoring system employed by Harrower (scores of 1 to 5: "normal", and 6 to 10: "disturbed"), a sliding scale was devised by the same technique described in connection with the Word Association Test, i.e., the establishment of a cut-off point where a definite "break" or dichotomy in the frequencies seemed observable. This resulted in the following score norms for the ten Rorschach cards: I - 3, II - 2, III - 2, IV - 4,

V - 2, VI - 6, VII - 7, VIII - 6, IX - 8, X - 3. On the basis of the reactions of 80 per cent of the standardization normals, on cards I - II - IV - VI and IX reactions higher than the above earned a score of 1, while the results of 90 per cent of the same group merited a score of 2 for over-reactions to cards III - V - VII - VIII and X. For the individual scores to each of the one hundred items, consult the specimen test-sheet (appendix 2). This test takes about ten minutes to give, and like its two predecessors has a maximum possible score of 15. By the process of addition, then, the complete Battery has a maximum possible score of 45, which the authors have broken down into the three distinct categories of: scores 0 to 6: within normal limits, 7 to 11: suggestive of psycho-pathology, and 12 to 45: definitely abnormal and evidence of possible psychotic disturbance.

While all the statistical data reported by the authors cannot be reproduced here, certain representative figures may serve to present a reasonably adequate picture. The score designated "within normal limits" (0 to 6) screened out 83 per cent of the "final normal" group (100) and 20.6 per cent of the "final abnormal" group (350). These latter "false negatives", incidentally, were largely the less pathological cases, such as hypomania, paranoid conditions, psychopaths, alcoholics, and psychoneurotics. A chi-square test of the three possible score ranges revealed a value of 141.6, which

indicates the Battery's power to differentiate normal from psychotic people. Statistical comparison between the results of the "original" (480) and the "final" (100) groups gave the following correlation coefficients: Word Association .84, Conformity Index .87, and Multiple Choice Rorschach .90. On a projective test, this degree of reliability is rather striking.

To summarize, then, the tests of the Verdun Projective Battery are simple and brief to comprehend and administer and require very little pre-training to give and score. The quality of interpretation increases in proportion to the knowledge, skill and experience of the examiner, but the quantitative score is of fundamental worth, especially in solving the question of whether or not to make a referral, in the school situation, to the psychological or psychiatric services. The procedure is rapid, flexible (group or individual, with the obvious exception of the timed association test), and can be repeated (re-test, follow up, etc.). In addition, the constituent elements are all long-established tests of proven worth and advantage and ones with which most psychometricians are already very familiar.

3. Cornell Index - Form N2

The second measuring instrument used in the experiment was Form N2 of the Cornell Index[8]. This test, which is distributed by the Psychological Corporation of New York, is of the personality inventory type in questionnaire form and is designed for use in psychiatric screening. Included among the authors of the 1948 revision are Arthur Weider, Ph.D., Harold G. Wolff, M.D., Keeve Brodman, M.D., Bela Mittleman, M.D., and David Wechsler, Ph.D., from the staffs of New York and Bellevue Hospitals and Cornell University Medical College. Among its virtues as a rapid neuropsychiatric screening device (though differing from the Verdun Projective Battery in that it is non-projective in nature) are its brevity and simplicity of administration, scoring and testee comprehension.

The test, which requires on the average about ten minutes to complete, consists of 101 items in question form, each one of which is followed by the printed words: "Yes","No" in that order. The subject is required to encircle his choice and the necessity of answering the complete list of questions

---

8 The data of this section were taken from:
a) Weider, Arthur, Harold G. Wolff, et al., Cornell Index Manual, New York, Psychological Corporation, 1948, 8p.
b) Weider, A., K. Brodman, et al., "The Cornell Service Index: A Method for Quickly Assaying Personality and Psychosomatic Disturbances in Men in the Armed Forces", in War Medicine, Vol. 7, issue of April 1945, p. 209-13.
For a specimen copy of the test, cf. Appendix 3.

is emphasized, even to the extent of asking the testee to guess in those areas in which uncertainty is experienced. Stencil scoring requires less than a minute per paper. The items, couched in nontechnical language, are grouped in divisions as follows (typical samples are included in parentheses): 1) an introductory neutral question which is not scored ("Have you ever had a headache?"), 2) eighteen questionson defects in adjustment expressed as feelings of fear and inadequacy ("Is it always difficult for you to make up your own mind?"), 3) seven questions on pathological mood reactions, especially depression ("Are your emotions usually dead?"), 4) seven questions on nervousness and anxiety ("Do you have any unusual fears?"), 5) five questions on neurocirculatory psychosomatic symptoms ("Do you have pains in the head or chest?") 6) eight questions on pathological startle reactions ("Are you keyed up and jittery every single moment?"), 7) fifteen questions on other psychosomatic symptoms ("Is your body always in very bad condition?"), 8) seven questions on hypochondriasis and asthenia ("Are you always in poor health and unhappy?"), 9) eleven questions on gastrointestinal psychosomatic symptoms ("Do you always have stomach trouble?"), 10) six questions on excessive sensitivity and suspiciousness ("Have you usually been treated fairly?"), and finally 11) sixteen questions on troublesome psychopathy ("Do your enemies go to great length to annoy you?"). None of these descriptive

labels appear on the test sheet used by the respondent, nor is the grouping usually apparent to the average subject. It will readily be perceived that the questions orient themselves around the two poles of neuropsychiatric and psychosomatic symptoms. The screening process aims to differentiate and single out individuals with personality disturbances of a serious nature and those with significant and intensified bodily symptoms. The authors report a Kuder-Richardson reliability coefficient of .95 for one thousand military inductees (a very similar Service Form of the Index was in military use). Its degree of validity is indicated by the fact that 83 per cent of some four hundred psychiatric rejects with personality disturbances were also rejected by the test while only 20 per cent of some six hundred non-disturbed accepts were falsely classified when using a cut-off level of a score of 13 and up together with the inclusion of one or more so-called "stop" questions (these are thirteen items which refer to crucial symptoms, e.g., "Have you ever had a fit or convulsion?" or "Were you ever a patient in a mental hospital?", etc., and are intended for interview elaboration). Incidentally, this is quite comparable to the 83 per cent "correct normality" of normals and the 20.6 per cent "false normality" of abnormals reported earlier by the authors of the Verdun Projective Battery.

Two other questions concerning the Cornell Index might be raised at this juncture. The first has to do with the

wisdom of phrasing the items in such a manner that only four of the hundred are legitimately answered in the affirmative. These are statements 20: "Do you usually feel cheerful and happy?", 69: "Is your appetite good?", 82: "Have you usually been treated fairly?", 87: "Do you make friends easily?" It certainly seems to present a degree of structural imbalance that greater foresight could have prevented or later revision remedied. However, the authors claim that while the Index alone may miss a certain number of ill persons, those that it does detect <u>are</u> ill beyond reasonable doubt. The second and very germane point is the incidence of intentional falsification of responses. Mendacious replies are reported to be rarities in hospital, industrial and military situations. Consequently, there is no reason to expect that they will prove to be common with the high school students. At any rate, should this occur, which seems unlikely, glaring lack of self-consistency on the answer sheet should show up gross fabrication and/or fearful deviation from the truth.

4. Rotter Incomplete Sentences Blank-High School Form

The third and final psychometric testing device employed in the investigation, the High School Form of the Rotter Incomplete Sentences Blank[9] (also distributed by the Psychological Corporation) is a projective technique of the partially structured variety. The co-authors of this particular form are Julian B. Rotter, Ph.D. and Janet E. Rafferty, M.A., both of Ohio State University, but the test is an outgrowth and revision of previous blanks prepared and used by a variety of authors, including Willermann (in the Army), and Shor, Hutt and Holzberg (at Mason General Hospital). A major aim and intention in its development was the provision of an objective (to the extent that projection allows) screening instrument applicable to experimental purposes, but one that would preserve the advantages of economy of time and effort in administration and scoring and yet still be well within the easy and rapid understanding and comprehension of the person undergoing testing and examination.

Accordingly, the test, for which the approximate average time for administration is twenty minutes, is made up of forty sentence-beginnings, ranging in length from one word ("People", "I", "Dating") to four ("I am best when",

---

9 The data of this section were taken from: Rotter, Julian B. and Janet E. Rafferty, The Rotter Incomplete Sentences Blank Manual, New York, Psychological Corporation, 1950, iii-86 p.
For a specimen copy of the test, cf. Appendix 4.

"My greatest worry is"), followed in each case by one line of empty space (approximately 6½ inches long) on which the subject records his spontaneous response. The individual being tested is requested to complete all the sentences in a manner that expresses his real and true feelings. No stress is laid on rapid or immediate answering and the "stems" may be tackled in random order.

The completions are scored on a sliding scale of 0 to 6 for each item according to the following evaluation (illustrative samples are inclosed in parentheses): a) positive responses are based on the degree of good adjustment evidenced in the statement: P3 = 0 ("I feel happy"), P2 = 1 ("I like dancing"), P1 = 2 ("People are interesting"); b) neutral responses, which are usually simple description, are all scored 3 ("Boys will be boys"); c) conflict responses are judged for indications of severe maladjustment: C1 = 4 ("I suffer from sinus"), C2 = 5 ("I can't concentrate"), C3 = 6 ("I wish I were dead"). Omission responses and fragments, provided they be fewer than twenty in number (or the paper is unscorable), are not directly scored, but are prorated by the formula: $\left(\frac{40}{40-\text{omissions}}\right)$ × total score. Incidentally, it may be of importance to note that unusually long responses, considered signs of maladjustment, are weighted by an extra point towards the direction of "C", unless, of course, the score is already at the maximal rating.

The over-all score is derived from totalling the weights of the forty component and constituent parts. However, to facilitate rapid, accurate and symmetrical scoring, guide manuals (one for each sex) have been prepared, which list for each "stem" or sentence-beginning several characteristic or typical responses fitting the seven (0 to 6, inclusive) possible weighted score levels. Finally, each completion is to be considered on its own merits and independently of the remainder of the replies (except in the case of clear-cut cross- references) to preclude as far as practicable the pervading influence of the "halo effect" with its resulting conscious or unconscious bias in either direction on the part of the rater which tends to undermine the accuracy and lessens the reliability of the total situation.

The reported advantages of the method of sentence-completion would seem to be largely embraced under such headings as freedom of response, some disguise of the real intent and purpose (although there is the possibility of a subject of sufficient sophistication consciously falsifying or beclouding his real affect), freedom from the necessity of highly specific training for administration, scoring, and interpretation, flexibility, evident saving of time, and ease of group screening. On the debit side must be admitted the fact that the individual scoring of a large number of these tests is a very tedious process. Also

the very obvious fact must be faced that insufficiency of material negates the whole process. This state of affairs may be encountered with improperly motivated illiterate (including younger children), disturbed or downright uncooperative subjects. Certainly, the ability to write and some skill in the language arts are of the essence in a test of this nature. In the individual testing situation, oral presentation of the material will prescind from the majority of these very real difficulties and obstacles.

Since the technique is intended to serve as a gross screening instrument, some mention of the degree of reliability and validity thus far established by the authors and other investigators should be of interest. A corrected split-half (equivalent, not odd-even, halves) reliability of .84 was found for the test-results of 124 male collegians and .83 for 71 female students. Inter-scorer reliability ranged from .90 to.96 for three scorers, two of whom were author-trained and one was a clinician who merely read the manual. Validity results include the following: a) a cutting point of 135 correctly identified 78 per cent of the adjusted and 59 per cent of the maladjusted cases, where female mental hygiene students had been so classified by their instructors (be it noted also that the 50 adjusted ones had a mean score of 121 and the 22 maladjusted, one of 137); b) the same technique and cut-off level pin-pointed 69 per cent and

52 per cent respectively of the male students (mean score of 57 adjusted: 119, of 21 maladjusted: 134); c) the same cutting score of 135 identified or partialled-out 76 per cent of some forty-six self - or counselor - referred maladjusted male students (mean score: 149). These percentages compare favorably with the ones mentioned earlier for the Verdun Projective Battery and the Cornell Index. Of perhaps greater pertinence in the present study are the results that were obtained by 299 entering (that is, just out of high school) freshmen at Ohio State University. The 85 females made a mean score of 127.4 with a standard deviation of 14.4 and the 214 males an average score of 127.5 with a sigma of 14.2. The cut-off point for the male high school students of the project under presentation was set at 130 with similar satisfactory results, the publication of which must be deferred to Chapter Three.

Finally, in the interests of strict accuracy, it must, of course, be made abundantly clear that the statistics reported in the previous paragraph refer to the College Form of the test. The authors point out that as yet there has been no formal standardization with students of the High School Form. However, this hazard ceases to be a major one when it is borne in mind that, in the first place, the College Form has already been occasionally used with senior high school students with apparent profit; secondly, thirty-

four of the forty items remain unchanged; and lastly, the stem modifications are extremely slight for the other six beginnings. The actual changes are as follows (the College Form wording is in parentheses): 4 - "At home ("Back home"), 14 - "In the lower grades" ("In high school"), 17 - "When I was younger" ("When I was a child"), 19 - "Other kids" ("Other people"), 26 - "Dating" ("Marriage"), and 31 - "At school" ("This school"). Thus, there is every reason to believe that both the manuals and the general scoring principles outlined earlier are still valid and applicable in the hands of a reasonably flexible scorer to the situation under investigation in the present writing.

In this opening chapter, the aims of the problem have been outlined and the various tests made use of described. This constitutes the "what" of the investigation. The next chapter is concerned with the "how" of the project, i.e., the design of the experiment.

CHAPTER TWO

DESIGN OF THE EXPERIMENT

There are four fields that logically present themselves for grouping under the generic title of experimental design. First, why were the particular tests used chosen for the purpose; secondly, how were the testees designated as the subjects of the research; thirdly, what was the actual technique and mode of the procedure of administration, - a detailed discussion here is obviously necessary; and finally, some introduction to the method of analysis proposed, even though a full coverage of this field is reserved to the third chapter. This, then, is the order in which the material of this chapter will be presented.

1. Selection of the Tests

This section will not review the general descriptions already provided in the first chapter. Rather, the idea is to outline the reasons that lay behind the selection and inclusion of these (and not some other) particular testing devices in the experimental work. In addition, since certain word changes and clarifications were made, with the permission and active cooperation of the authors of the Verdun Projective Battery, for the school situation, it was deemed à propos to reserve ~~their~~ their discussion for this place in the report.

Finally, some consideration is given to the V.P.B. test sheet as it appeared in its final form in this research.

As intimated in the first chapter, the reason why the Verdun Projective Battery was chosen for investigation in the first place was the realization that it might prove sensitive to the degree of psychopathology discoverable in the average school population. If it could screen out the deviants, then a valuable addition to the armoire of the school psychologist would have been contributed. Since the trained clinician could be eliminated in the administration, scoring and quantitative interpretation (but not the qualitative depth analysis) of this group instrument, it would be of assistance, if valid, in the pupil evaluation, counselling, guidance and personnel disciplines of the educational setting. Its other technical advantages, outlined in Chapter One, will not be repeated here.

With what tests, then, could it be logically compared in a statistical analysis? As mentioned and specified earlier, the Battery is to be evaluated primarily as a screening and not a diagnostic procedure, brief and uncomplicated in content, and simple and facile to administer, score and interpret. This precluded, of course, using any of the many excellent diagnostic psychological devices available. A lengthy search of the testing literature and catalogues of the major North American distributors of psychological tests

finally brought to light two other unrestricted instruments purporting to screen out the neuropsychiatric unfit, and amenable to either group or individual administration. These were Form N2 of the Cornell Index (questionnaire inventory) and the High School Form of the Rotter Incomplete Sentences Blank (partially structured projective).

These tests, as described earlier, likewise possessed many "mechanical virtues" from the point of view of comprehension, administration, and ease of scoring and interpretation. Furthermore, since one was in the tradition of the personality inventory (but very modern in structure and use), and the other featured aspects and elements of the projective technique, opportunity was manifestly presented to consider what degree of similarity existed in the results obtained by the two different approaches, as one of the sidelights of the investigation.

No alteration in the material as received from the publishers was made in the case of the two latter tests. However, with the Verdun Projective Battery, some minor changes were necessary to adapt the test to the school situation. The Word Association and Conformity Index subtests were not changed in any way from the form used in the hospital. But a number of word amplifications in the Multiple Choice Rorschach were thought beneficial from the standpoint of greater clarity and increased ease of comprehension by the high school students. It is always well when dealing with a school population to

bear in mind the ancient educational adage not to underestimate the intelligence or over-estimate the knowledge of the group in question. The process, as will be seen shortly, was additive in nature, with one exception, a substitution in Card IV. Seven changes in all were made (one, however, was merely a repetition), affecting Cards I - II - IV - VII - VIII and X.

The changes in detail are as follows (a brief rationale for each editing will be provided in parentheses): 1) the word "crest" was appended to the first choice of Card I: "an army or navy emblem" (this is the word current in the lexicon of today's youth, especially in connection with sports, school letters and the like, and so was thought of value in clearing up any possible difficulty with the older term "emblem"); 2) also in Card I, the explanatory expression "hip bones" was placed next to the sixth choice: "a pelvis" (since the testees were all young males, it was felt that their experiential vocabulary might not include the more technical word); 3) the sixth choice of Card II, "a white top", was followed by the parenthetical "toy" (this, of course, was to prevent the interpretation of "top" in a locational or positional sense); 4) this is the only substitution, - the word "substance" replaced the final word in choice five of Card IV: "a burnt mass" (as all the subjects were Catholics, ambiguity was feared with the religious service known as the Mass, especially as this is frequently misspelled with a small "m" by schoolboys; in addition, there was some

concern that the word "burnt" might, therefore, cause them to shy away from this particular item for reasons other than those intended by the author of the test); 5) choice six of Card VII, "a white chandelier", picked up the explanatory "lights" (this was intended as a simple clarification of "chandelier"); 6) the fifth choice of Card VIII, "an emblem", again merited the appendage "crest" (the reason being the same as before), and finally 7) the word "design" was added to the third choice of Card X: "a Chinese print" (this was to assist in the interpretation of "print" as a noun and not a verb and to prevent confusion with such considerations as the symbols on a Chinese laundry ticket, etc.). In drawing up these minor modifications, the editor kept well in mind the injunction that there were to be no questions whatsoever during the testing period consequent to the appearance of the first slide.

Accordingly, with the agreement of the authors, these alterations were incorporated in a new format especially prepared for use in this project. A specimen has been inserted (Appendix 2) for reference. Several typewritten forms of various arrangements were prepared before this definitive design was formulated, decided on, and finally multilithed. Neither the name of the test nor any other indication of its real intent appear anywhere on the sheet. The full name (both sides), age (in years and months at the time of testing) and

class (e.g., 1D, 2C, 3A, etc.) of the subject were typed in prior to the testing, as was the letter "X" next to the letter "N" ("normal") or "A" ("abnormal") on the reverse side on the basis of the result of the teacher-selection poll. The obverse embraces the Multiple Choice (note the deletion of the word "Rorschach") and the Conformity Index tests, while the reverse includes the Word Association test and scoring summary and also blank space for any commentary deemed appropriate by the examiner. Five hyphens in the numbers heading the M.C. lists and five asterisks in the W.A. test, identify the double-weighted items (described in the first chapter) for the convenience of the scorer.

2. Selection of the Subjects

This section provides a description of those doing the selection and classifying (in this case, high school teachers), those being so screened and gridded (in this instance, male high school pupils), and the specific technique by which the former designated the latter.

The research was conducted at The D'Arcy McGee Boys' High School, a large departmentalized institution (college preparatory and general courses) of some six hundred pupils and located in central Montreal. While operated by the Montreal Catholic School Commission, this day school (no boarders) is under the direction of the Toronto District of the English-Speaking Christian Brothers;

however, included on the staff are a large number of lay male teachers, who actually outnumber the Religious. The pupils of Grades Eight, Nine and Ten (First to Third High, inclusive) were chosen for the experiment. Permission to disturb the schedule of Grade Eleven (Fourth High), who are preparing for a Provincial Certificate, is practically never accorded. On the other hand, the authorities, both at the Board and in charge of the school, readily consented to the testing of the other three grades.

The population from which the preliminary selections were to be made, previous to the final screening by the testing devices, was, therefore, constituted in the following manner: Grade Eight: five classes (A to E); Grade Nine: four classes (A to D); and Grade Ten: three classes (A to C), in all, a total of some four hundred fifty male pupils to be reviewed. The mean age of the students selected was 15-1 and the range was 13-0 to 17-11 (years and months) with but three exceptions (12-5, 18-5, 18-9).

It was felt that the use of two "extreme" groups would be the most appropriate method of case selection in the investigation. With this end in view, all the teachers of these classes (thirteen laymen and six Brothers) were asked to evaluate their pupils. Since the system in vogue is for the teachers to teach particular "subjects" in several rooms (a "form master" is assigned to the administrative and record-

keeping of each class), the following multiple-rating formula (by teacher-load) resulted: 2 teachers rated only one class each, 6 two classes, 3 three classes, 6 four classes, 1 (the author) six classes, and 1 (the art teacher) nine classes. By class line-up, on the other hand, 4 classes were rated by four teachers each, 4 by five teachers, 2 by six teachers, and 2 by seven teachers. Table I presents this material in schematic form.

The teachers were all approached privately and individually to have their assistance on the selection of the testees solicited. At the time of this verbal communication, (late on a Friday afternoon), they were also given stencilled material to consider over the weekend, which consisted of a short letter asking for their cooperation and stressing the necessity of care in the selections and discretion in preventing leakages both of their choices or of the real purport of the test, a short listing of certain generally accepted "signs" of abnormality, and, of course, class-lists from which the groups were to be picked. A copy of the letter and "signs" is included as Appendix 1 in this report.

Perhaps a word justifying the use of the terms "normal" and "abnormal" might be à propos at this point. While all the teacher-raters are at the top-level in their own professional category (many of them are multiple-degree men), they are, of course, teachers and not psychologists. This consideration led to the belief that the simplest

Table 1. - Classes Screened by Each Teacher-Rater.

Teacher-Raters	Relig. or Lay	Sequence of Classes: 1A	1B	1C	1D	1E	2A	2B	2C	2D	3A	3B	3C	Totals
1	R						x				x	x	x	4
2	R	x		x	x									3
3	R		x											1
4	R	x	x	x	x									4
5	R							x						1
6	R								x		x	x	x	4
7	L	x	x		x									3
8	L					x				x				2
9	L		x				x				x		x	4
10	L					x				x				2
11	L							x	x					2
12	L		x	x								x		3
13	L	x					x		x				x	4
14	L	x	x	x	x	x	x	x	x	x				9
15	L				x		x							2
16	L										x	x		2
17	L				x							x		2
18	L					x		x		x	x			4
19	L		x		x			x			x	x	x	6
Totals		5	7	4	7	4	5	5	4	4	6	6	5	62

clearcut description of deviation (or its absence) for non-professionals would be the old standbys of "normal" and "abnormal". Of course, "abnormality" in this context refers to the behavior problem fringe of the average high school population and is not to be confused in kind or degree with that encountered on the wards of a mental hospital. In any event, these terms are certainly sufficiently "extreme", and the inclusion of the "signs" to be on the lookout for should have at least channelled the teachers' thinking from completely random or subjective choice into lines approaching some degree of adequacy from the psychological point of view. The items included in the listing of "signs of abnormality" were compounded of suggestions offered by the junior author of the test (Dr. Herbert Dorken, jr., a practicing clinical psychologist) and items culled from Wickmann's survey as reported in Shaffer's text.[1]

The teachers' ratings were returned (sometimes following an additional discussion conference for purposes of further clarification) a few days afterwards and from these the two "extreme" groups were compiled by simple counting. There seemed to be no external evidence of any collusion in the choices and, indeed, all gave every sign of making their decisions in a mood and spirit of calm deliberation with a

---

1 Wickmann, E.K., Children's Behavior and Teachers' Attitudes, New York, Commonwealth Fund, 1928, (no pagination), quoted by Laurance Frederic Shaffer, The Psychology of Adjustment, Boston, Houghton Mifflin, 1936, p. 174-176.

clear appreciation of the seriousness of the intent behind the project. Certainly no levity or flippancy could be discovered in the environs nor any betrayal of the secrecy requisite for the success of the testing that was shortly to follow, for a "tip-off" slip of the tongue to the pupils might have vitiated the worth of the whole process. While approximately ten individuals per class in each category had been requested of each teacher, this number was neither ironclad nor rigidly adhered to in the actual selections, - it was clearly better to have a few above or below this point, but sincerely selected, rather than achieve uniformity at the price of forced choices.

As a result of the teacher-screening for normality and abnormality of the pupils of these twelve classes, the two following "extreme" groups were differentiated: a) 50 "normals", b) 45 "abnormals" (the "roundness" of the numbers is a strictly accidental result of the enumeration of cases). The breakdown by grade and group is provided in Table II. The basis of the selection was as follows: 1) no cases with any "negative" votes were included (this means, for example, that if a boy was cast as "normal" - or the reverse - by six out of seven teachers, or any other combination of numbers, but the other one(s) voted him the contrary of the majority, he was <u>not</u> included in the testing; this is severe elimination, possibly, but under the circumstances seemed eminently justifiable); and

Table II. - Breakdown by Grade of the "Normal" and "Abnormal" Groups.

Grade	Class	"Normal"	"Abnormal"	Totals
Eight	1A	2	3	5
	1B	3	1	4
	1C	4	3	7
	1D	11	2	13
	1E	0	4	4
Totals		20	13	33
Nine	2A	3	5	8
	2B	8	2	10
	2C	6	2	8
	2D	0	6	6
Totals		17	15	32
Ten	3A	4	8	12
	3B	3	5	8
	3C	6	4	10
Totals		13	17	30
Grand Totals		50	45	95

2) at least three raters had to agree on the classification (that is, where four teachers rated and split two-two, for example, the student was "dropped" from the investigation). It will be recalled that only the ten outstanding cases per class in each category were to be marked out and the remainder left blank, hence there was provision for the possibility of less-than-total balloting occurring for any particular individual in any component class in the survey. Table III lists the two groups anonymously by number and indicates the grade, age and teacher-selection ratio (e.g., 6/7 indicates seven teachers rated the boy and six classified him as "normal" or "abnormal") for each individual.

In this second section, then, the subjects selected for the actual testing procedure have been described, together with their scholastic environment, how their teachers screened them for normality or its counterpart, and the technique and safeguards that were worked out to maximize the efficiency of the process and achieve an optimal level of practical progress. The next section describes the manner in which the tests selected for use were administered to the two "extreme groups" of pupils chosen by the teachers.

Table III. - The Class, Age and Teacher-Selection Ratio of the "Normal" and "Abnormal" Groups.

"Normals"				"Abnormals"			
No.	Class	Age	Ratio	No.	Class	Age	Ratio
1	1B	12- 2	6/7	51	1D	12-11	6/7
2	1D	13- 0	5/7	52	1C	13- 4	4/4
3	1B	13- 2	5/7	53	2A	13- 6	4/5
4	1C	13- 2	3/4	54	1A	13- 8	4/5
5	1D	13- 3	6/7	55	1D	13-10	4/7
6	1D	13- 4	4/7	56	1C	14- 0	3/4
7	1C	13- 4	3/4	57	1E	14- 1	3/4
8	1D	13- 5	4/7	58	2A	14- 1	4/5
9	1D	13- 6	4/7	59	1A	14- 2	4/5
10	1D	13- 7	4/7	60	1E	14- 2	3/4
11	1C	13- 7	3/4	61	2A	14- 5	4/5
12	1D	13- 9	5/7	62	2D	14- 7	3/4
13	1C	13-10	3/4	63	3B	14- 8	4/6
14	1B	14- 0	5/7	64	3A	14- 8	4/6
15	1A	14- 2	4/5	65	1C	14- 9	4/4
16	2C	14- 3	3/4	66	2C	14- 9	3/4
17	3A	14- 4	5/6	67	3A	14-10	5/6
18	2C	14- 4	4/4	68	2D	14-11	4/4
19	1D	14- 5	5/7	69	2D	14-11	4/4
20	2A	14- 6	5/5	70	2A	14-11	3/5
21	1D	14- 7	5/7	71	2B	14-11	4/5
22	2C	14- 8	3/4	72	3B	15- 0	4/6
23	2B	14- 8	4/5	73	2B	15- 3	5/5
24	1A	14- 9	3/5	74	2D	15- 4	3/4
25	2C	14- 9	3/4	75	3C	15- 5	3/5
26	3A	14- 9	4/6	76	2D	15- 5	4/4
27	1D	14- 9	4/7	77	2D	15- 5	3/4
28	1D	14-10	4/7	78	1A	15- 6	3/5
29	2A	15- 0	5/5	79	3A	15-10	5/6
30	2B	15- 1	3/5	80	3B	15-11	5/6
31	3B	15- 1	4/6	81	1B	15-11	7/7
32	2A	15- 1	5/5	82	2C	16- 0	4/4
33	2C	15- 3	3/4	83	1E	16- 1	3/4
34	2B	15- 3	3/5	84	1E	16- 1	3/4
35	2B	15- 3	4/5	85	3C	16- 1	3/5
36	3B	15- 5	4/6	86	3B	16- 2	6/6
37	2B	15- 6	4/5	87	3A	16- 2	5/6
38	2B	15- 9	4/5	88	3B	16- 6	4/6
39	3A	15-10	4/6	89	3A	16-11	6/6
40	3C	15-10	3/5	90	2A	16-11	3/5
41	3C	16- 1	4/5	91	3C	17- 3	3/5
42	3C	16- 2	5/5	92	3A	17- 7	5/6
43	3B	16- 4	5/6	93	3A	17-10	4/6
44	2B	16- 6	4/5	94	3C	18- 5	3/5
45	3C	16- 9	5/5	95	3A	18- 9	4/6
46	3A	16- 9	4/6				
47	2B	17- 2	4/5				
48	2C	17- 9	3/4				
49	3C	17-11	5/5				
50	3C	17-11	3/5				

3. Administration Procedure

The intention here is to epitomize the preliminary organization (including the necessary equipment and arrangements) that preceded the actual application of the tests to the subjects and then briefly to describe the highlights of the administration of each test and subtest in turn, concluding with a terse commentary on the entire procedure.

About a month ahead, a tentative date of Thursday, February 25, 1954 (the following day was the annual mid-term Staff Holiday) was set for the test administration. Incidentally, this was the height of the school hockey season and furthermore the annual operetta was being presented; conflict, therefore, had to be skilfully avoided with such events as practises, games, rehearsals and performances. During this period, nevertheless, such essential preliminaries as the following were encompassed: the permission and approval of the authorities in the Department of Studies of the Montreal Catholic School Commission and the direction of the school itself was secured, the copies of the Cornell Index and the Rotter Incomplete Sentences Blank were ordered and received from the Psychological Corporation, the printing of the special format of the Verdun Projective Battery was done, the check-lists were explained to and received back from the teachers and the two groups compiled from them, clearances

assuring free and unhampered use that day of the gymnasium (morning) and auditorium (afternoon) were obtained, the co-operation of all the teachers was again evidenced by their ready consent to the inevitable class disturbance involved in the whole process (in addition, some had to sacrifice their gym period that morning), and arrangments were made for the procuring of such incidental equipment as tables, chairs, stopwatches, etc., and above all, the services of a school monitor for messages, doorman duties and the like.

The overall plan was to devote the morning to the administration of the timed Word Association test in the spacious gymnasium and in the afternoon to assemble the subjects in the large auditorium for the remainder of the Verdun Projective Battery and the Cornell Index. The following week the Rotter test could be administered at ease in the actual classrooms. A few days before, a preliminary examination of the suitability of the auditorium (which seats about 700 people) for the Rorschach screening established certain area limits where neither the color nor form vision was apparently adversely affected and these designated seats and rows were the ones subsequently occupied by the pupils chosen for the experiment.

These, then, were the general introductory, though fundamental, preliminaries; the actual specific administration of each of the testing instruments will now be described in

detail. However, the wording of the instructions, already presented in the first chapter, will not be repeated here, - it is obvious they were given as the occasion arose to each successive individual or group concerned.

Since communication (even innocently) of the matter of the Word Association test was considered deleterious to the purpose of the project (absolutely spontaneous responses being the aim), it obviously had to be completed in one and only one school session. Accordingly, the author enlisted the services of Dr. Dorken (one of the clinicians who designed the test), Miss Margaret Calvert, M.A. (his assistant at the hospital), and Mr. Nathan Wisebord, M.A., a trained psychologist. Tables and chairs were set up in each of the four quarters of the large gymnasium (the fear of being overheard rapidly proved groundless) and the monitor brought the selected pupils (using typed lists) of each class in turn, who were then individually tested. Each W.A. response, together with the time taken in tenths of a second, was recorded on the test sheet. Presumably none of the testees knew they were to be tested, prior to being summoned to the gymnasium for what purpose they knew not. As an extra precaution, since the morning quarter-hour recess period might accidentally provide an occasion for the sharing of the word-list, thus frustrating spontaneity, secrecy was enjoined on the testees as they left the gym "to avoid spoiling the test for the other boys".

They seemed at least appreciative of the seriousness of the purpose and the reasonableness of the request.

The author had pre-determined the cases for each of the examiners; basically, the following division was adhered to: the "abnormals" were done by the two male psychologists and the "normals" by the rest of the testing team. The morning proved ample for the Word Association test, the entire procedure taking from nine o'clock to eleven-thirty, approximately. Incidently, as the testees left, they were informed to be present in the auditorium that afternoon (at either one-fifteen or two o'clock, depending on their class-level) and to come equipped with a pencil and a large book to serve as a writing base.

In the afternoon at the appointed time the first platoon (the entire group was divided into sections on a grade-level sectioning, not on the "normal" versus "abnormal" dichotomy) assembled and were seated and arranged in the areas previously delimited in the preview screening of a few days earlier. Their sheets (unscored from the morning, of course) were returned to them and the instructions for the Multiple Choice Rorschach given. As practically no questions seemed forthcoming, or very few at any rate, the directions were repeated and emphasized two or three times. The boys, however, seemed to be eager to get at the test itself, so the lights were dimmed for the showing of the first slide. To facilitate the mechanics of supervision (not only to prevent possible

over-exuberance but also by their mere presence to militate against the "copying" of choices) and presentation, the assistance of Dr. Dorken (possibly mistaken by the boys for a teacher from some other school) to operate the projector and more or less time the showing of each slide was procured (he also supplied through the courtesy of the Verdun Protestant Hospital the screen, Harrower slides and projector used), and Mr. Vincent Patton, B.A., B.Ed., (a very popular teacher and highly successful football and hockey coach) helped with the auditorium lighting and overseeing of the first group, as did another member of the McGee staff, Mr. John Lying, B.A., B.Ed., with the second seating. In addition, the distribution and collection of the test papers was expedited by the ready aid of one of the most popular pupils in the school, the professional hockey player, Reggie Fleming, who was himself selected for testing from one of the Grade Ten classes. The monitor mentioned earlier (Joseph Taddeo of 1A, not included in the testing report) contributed further assistance here. In addition to the services of the people listed above, the cooperation of the testees was solicited by their being informed that McGee (their school) had been selected for the experiment (there is tremendous spirit and pride of school here), and that what they were being asked to do, if done properly, would prove of real benefit to future pupils everywhere. Their fears (if any) were calmed by the announcement that their responses

were merely means to this end and would not be reported to the authorities or held against them in any personal way (promotion, references, extracurricular activities, etc.). Finally, they were told that the sheets would be checked by the author (also a teacher at the school); this was to spur them to their maximum effort (certainly to full and careful completion of all items at least) and to try to preclude any intentionally mischievous activities. It must be reported that good will prevailed fortunately throughout every aspect of the administration and the "outsiders" to the school were quick to note and report their extremely favorable impression of the courtesy and cooperation of all the adolescents encountered. Not one sour note was struck in the entire day, - certainly a sincere and valued tribute in these days of juvenile ebullience and extroversion! Each slide was displayed about a minute, its number and place on the sheet were pointed out, the study of all ten choices was encouraged, and as the time elapsed they were asked to make their selection just before the screen went blank. The author and some of the other assistants circulated around the group during the screening and neither whispering nor neck-craning (for copying) seemed in evidence. The pupils seemed to enjoy the experience. The screen had been set up at the front and center of the stage in full and easy view of everyone and the neutral-shaded draw curtains closed behind it to provide an indifferent

background. Sufficient light was maintained in the auditorium to enable the choices to be easily legible without at the same time dimming the projected slides. The projector was about thirty feet or so down the center aisle on a heavy table to ensure steadiness and constancy of distance and illumination. Exactly the same procedure was gone through again with the second group, who were scheduled for the two-fifteen showing.

The other three tests were relatively uncomplicated and presented no particular administration problem. Immediately after the Multiple Choice Rorschach (with each group), the Conformity Index was administered. The subjects were reminded of the procedure they had gone through that morning and were asked to apply it themselves to the new list of words offered as stimuli. This task was rapidly executed and the Verdun Projective Battery sheets were collected from the group. Then Form N2 of the Cornell Index was quickly distributed and explained. The only items of general information at the top they were requested to complete were to fill in their name and class (the remainder either did not apply or could be found from the V.P.B. sheets). The "Yes"-"No" responses were soon circled and these forms were gathered in. This terminated the testing activities for the first day. Early in the following week (beginning of March), the High School Form of the Rotter Incomplete Sentences Blank was administered, for the most part in the classrooms themselves during actual school

periods (the author was fortunate enough to be scheduled for time in seven of the twelve classes concerned in the project), with the remainder of the students being assembled after hours in their own room or in the library for the short testing time required. In one or two instances, cooperative teachers permitted some of the boys to work on the test during their teaching time, after the instructions had been explained to them for transmission to the pupils. One slight change in the administration of the test was effected in a definite effort to secure the maximum number of responses. This was to insist on all forty items being completed; in other words, no omissions were allowed. This seemed to work no hardship on the pupils, and assured uniformity of response total for scoring and comparison purposes.

All the collected test sheets were sorted, classified and checked for incompletions. Of the relatively small number of these that were found, most were on the Rotter test (as might have been expected from a study of its nature and the added inadmissibility of omissions), a few occurred (probably through oversight) on the Cornell, and only two on the V.A.B. (two boys each marked one of the M.C.R. choices in a way that was difficult to decipher which response was the intended one). Individual and personal interviews rapidly cleared up these few lacunae. Incidentally, the very few unwitting absentees on the testing day of those designated in the teacher-selection

procedure for examination were tested individually by the author as soon as they returned to school (three in the "normal" group and four in the "abnormal", making seven in all, - specifically, subjects 33 - 41 - 44 - 68 - 81 - 90 - 92).

Possibly a note on the general reaction of the pupils to the whole procedure might not be amiss. Generally speaking, reactions were minimal or absent. Perhaps they are becoming test-sophisticated, (e.g., earlier in the same year one of the Brothers had prepared a thesis, also for the University of Ottawa, which had involved testing some of these boys in a different setting) or the myriad extra-curricular activities (or maybe even study for the term examinations) soon diverted their attention. There were a few inquiries afterwards as to the purpose of the tests. However, it was not considered judicious to explain the real nature of the test out of consideration for the feelings of all concerned (the idea of testing the test and not the individuals had been stressed during the administration). The ones not tested, by the way, sometimes wished to know the reason for their exclusion. When it was pointed out that time was at a premium, that no stigma was cast by omission, and that the teachers had run their eyes down the class lists picking some at each level (a euphemistic but technically accurate description of the selection procedure), their minds seemed at ease.

As to the tests themselves, the least popular seemed to be the Rotter, probably because in its form it bears a close resemblance to ordinary school exercises and also perhaps because its administration took place in the classroom during the school hours (or even after sometimes), whereas the others were very different from the usual routine and also were administered in settings different from everyday humdrum scholastic existence. The cinema-like atmosphere surrounding the administration of the Multiple Choice Rorschach seemed to intrigue their interest and fascinate their attention to the greatest extent. Another likely contributor to their affect concerning the tests was the fact that the formulation and writing of the Rotter sentences represented a far greater expenditure of effort than the mere checking or encircling of responses and pronunciation or writing of a few words required by the other procedures.

4. Method of Analysis

This concluding section on the general method of analysis to be employed in the handling of the data obtained will be brief, for the detailed analysis of the data is the property and prerogative of the entire third chapter. Their actual discussion and interpretation is reserved for the final chapter. Here and now, some passing attention will be paid to the scoring of all the tests collected and a brief descriptive

listing of the statistical tools and devices to be applied to the data in the ensuing parts of the report.

The scoring of the various tests was completed by late August of 1954 and was uneventful but fatiguing. The directions of the different authors were adhered to faithfully and eventually the entire mass of material had been evaluated. Preliminary to any analysis, the levels theoretically separating normality from maladjustment had to be determined and ascertained for the test material. This resulted in the dichotomies described in the balance of this paragraph. On the Verdun Projective Battery, the cut-off point of a score of 6 was followed; on the Cornell Index, the scores of 13 and up designated disorder or disturbance (the so-called "stop questions" were not especially utilized, as their references to drink, dope, arrests, mental hospital incarceration and the like did not seem very applicable to an adolescent school population); finally, the count of 130 was adopted on the Rotter test (this perhaps is somewhat arbitrary but no official cut-off point has been set as yet for the High School Form, - in any event, it is five points lower, i.e., stricter, than for the College Form with the idea in mind that the younger and more sheltered high school boys should be encountering less stressful life situations). With these points of division in mind, the statistical analysis of the results was initiated in September and brought to completion some three months later in early December 1954.

For the breaking-down of the accumulated data into coherent and comparable forms that would yield the maximum amount of verification with the greatest degree of illustrative clarity, the value of the techniques of "t" ratio study, chi square analysis and correlation procedure seemed of paramount contribution. Accordingly, for the endeavor of attempting to ascertain the degree of validity possessed by the instruments, these processes were applied to the facts gathered. As far as the status of the reliability was concerned, Rulon's method of estimating internal consistency appeared to offer the greatest possibilities of optimal service in that rôle. It, therefore, was the device adopted. Lastly, the element of relationship of results with academic standing (after all, the testing milieu was a school and it seemed almost wilful to neglect such an obvious source of information) offered another avenue or channel to be explored, possibly with profit. The final scholastic record for the school year 1953-54 of each individual tested in the survey was, therefore, secured and the study made of the significance of these findings in terms of reference to the data procured by the psychological devices. The description and discussion of all this matter constitute the content of the two succeeding chapters.

With this abbreviated and condensed summation in concentrated form of the general method of analysis pursued in the investigation, this second chapter on the experimental design, the "how", as was remarked before, of the project is concluded. The selection of both the tests and the population to whom they were applied and, in addition, the administration procedure followed in the application have all been described in the earlier sections. The time has now arrived to analyze the results in order to unearth and bring to light the recondite interrelationships contained therein. The results of the analysis of the facts obtained will be reported in the ensuing third chapter.

CHAPTER THREE

ANALYSIS OF THE DATA

The report of the process of statistical analysis constitutes the content of this third chapter. The analytic crucible will be scrutinized, so to speak, from four different approaches. First of all, an overview of the general validity will be attempted. Next in line, some attention will be paid to the relationship borne to the results by the matter of academic standing. The third topic concerns the effect of type of administration, i.e., group versus individual procedure. Fourthly, some analysis of the degree and extent of reliability will be essayed. It may be pertinent to point out here that this chapter is concerned solely with analysing the data obtained. The discussion of their meaning and significance, an interpretative endeavor, is assigned to the concluding chapter.

1. Validity Studies

The data obtained were subjected in turn to statistical analysis by the three techniques of the t-ratio test, chi square comparison and correlation coefficient procedure. The basic and related measures of the mean, standard deviation and standard error of the mean were derived for all tests and subtests. The standard error of the difference between

the independent means was also computed to assist in testing the null hypothesis. For the chi square derivations, the cut-off levels described earlier were applied to the test totals. Suitable sectioning of the Verdun Projective Battery total was performed for the computation of the subtest comparisons. In this report on the validity analyses, the raw data will be presented first, followed in succession by the "t" ratios, the chi squares and the series of intercorrelations.

The results obtained by each of the testees, dichotomized by group, on the Word Association Test, Multiple Choice Rorschach, Conformity Index, and total Verdun Projective Battery, Cornell Index and Rotter Incomplete Sentences Blank are reported in Table IV under the same anonymity of identification by number introduced in Table III of the second chapter. In each instance, the total has been appended at the foot of the column. Remarkable in the "normal" group are the Conformity Index scores of subjects 8 - 26 - 27 - 45, which stand out as apparent anomalies. Subject 5 is interesting with a high V.P.B. total of 10 (brought about by the W.A. score of 8) coupled with no Cornell score at all and the lowest and only two-figure Rotter result. Cases 12 - 13 - 19 - 44 are also noteworthy, but for diametrically opposite reasons. In the "abnormal" group, subjects 54 - 67 - 68 - 70 - 77 - 78 - 80 - 82 - 85 are rather striking. The Rotter score (181) of subject 77 is far ahead of the next one in line, the 145 of

Table IV. - Scores Made on All Tests and Subtests by the "Normal" and "Abnormal" Subjects.

No. "NORMALS"	W.A.	M.C.R.	C.I.	Total V.I.	Cornell	Rotter	No. "ABNORMALS"	W.A.	M.C.R.	C.I.	Total V.I.	Cornell	Rotter
1	1	0	0	1	10	111	51	8	0	0	8	14	112
2	1	2	0	3	2	128	52	7	3	0	10	6	119
3	2	0	0	2	5	114	53	0	1	12	13	4	122
4	1	4	0	5	11	121	54	0	2	0	2	8	127
5	8	2	0	10	0	99	55	11	1	0	12	3	111
6	4	3	0	7	16	125	56	4	3	0	7	8	123
7	1	2	0	3	10	125	57	2	2	0	4	10	115
8	1	1	12	14	0	114	58	10	0	0	10	13	112
9	0	7	0	7	3	120	59	6	0	0	6	10	105
10	0	1	0	1	6	123	60	2	11	0	13	14	114
11	1	1	0	2	9	111	61	2	3	0	5	20	126
12	7	1	3	11	20	118	62	5	4	3	12	10	132
13	0	0	0	0	5	120	63	9	1	0	10	9	119
14	0	2	0	2	16	110	64	9	7	3	19	8	145
15	0	3	0	3	0	119	65	6	0	0	6	4	126
16	3	3	0	6	6	125	66	2	2	0	4	25	132
17	0	6	0	6	2	123	67	0	2	0	2	7	125
18	2	1	0	3	1	123	68	0	0	0	0	9	137
19	0	0	0	0	7	118	69	2	2	0	4	5	124
20	1	5	0	6	14	125	70	0	1	0	1	3	124
21	2	2	0	4	32	134	71	1	3	0	4	13	134
22	0	2	0	2	11	123	72	1	0	3	4	5	140
23	1	3	0	4	3	121	73	3	5	0	8	8	123
24	2	0	0	2	3	121	74	5	2	0	7	8	128
25	1	4	15	20	7	126	75	2	1	0	3	12	122
26	0	0	15	15	5	118	76	9	3	0	12	26	112
27	0	3	0	3	11	127	77	1	1	0	2	16	181
28	4	3	0	7	9	118	78	0	2	0	2	0	123
29	3	2	0	5	2	123	79	3	4	0	7	18	135
30	1	0	3	4	14	127	80	1	1	0	2	14	123
31	1	1	0	2	13	119	81	2	6	0	8	12	123
32	3	1	0	4	4	116	82	0	1	0	1	13	123
33	2	0	3	5	25	120	83	4	4	0	8	31	140
34	0	0	0	0	9	115	84	2	5	0	7	16	120
35	3	2	0	5	12	118	85	0	2	0	2	2	120
36	0	2	0	2	11	126	86	2	3	0	5	12	123
37	0	2	0	2	4	115	87	0	3	3	5	4	121
38	0	0	0	0	5	122	88	9	0	0	9	11	146
39	5	3	0	8	18	122	89	8	7	0	15	16	126
40	1	1	0	2	10	117	90	1	7	0	8	8	120
41	1	2	0	3	33	140	91	3	0	0	3	2	122
42	1	4	0	5	7	112	92	8	5	0	13	24	128
43	2	1	0	3	13	127	93	3	5	0	8	11	138
44	0	8	6	14	2	117	94	7	4	0	11	0	122
45	3	5	12	20	10	126	95	4	3	0	7	4	114
46	2	3	0	5	3	115							
47	0	1	0	1	7	114							
48	0	3	0	3	10	124							
49	1	4	0	5	4	107							
50	3	1	0	4	0	103							
	75	107	69	251	440	5985		164	121	24	309	476	5657

case 64. The highest "normal" Rotter is the 140 of subject 41, who also obtained the greatest Cornell score (33) of either group. The second highest Cornell score (case 21's 32) is also to be found in the "normal" group.

These results have been grouped by frequency distribution in Table V for the V.P.B. subtests. For the "normals", the great cluster around the low scores is easily observable, especially on the W.A. test. The odd high C.I. score must also be remarked. In general, the "abnormals" have a greater spread in their scores. Their relative absence of C.I. scores is worthy of notice. An attempt has been made, in the three-part structure of Table VI, to line up frequencies on the V.P.B., Cornell and Rotter totals, even though these are not directly comparable because of the varying maximum scores. As can be noted at a glance, the trend of score concentration is at the lower end of the V.P.B. (0 to 8, primarily) and Cornell (0 to 14, approximately) tests and around the middle (111 to 128, more or less) of the Rotter instrument.

Could these results have arisen by chance fluctuations alone or is there a real difference between the findings for the two groups? Do the tests actually discriminate between "normal" and "abnormal" subjects and if so, to what degree? To try to answer these and similar questions that naturally arise in such an experimental situation, some investigation of the criteria of significance or confidence levels or

Table V. - Frequency Distribution of Verdun Projective Battery Subtest Scores for the "Normal" and "Abnormal" Groups.

	Word Assoc.			Mult. Choice Rors.			Conform. Index		
Score	N.	A.	Total	N.	A.	Total	N.	A.	Total
0	17	9	26	10	8	18	42	40	82
1	15	5	20	11	8	19	--	--	--
2	7	9	16	11	9	20	--	--	--
3	6	4	10	9	7	16	3	4	7
4	2	3	5	4	4	8	--	--	--
5	1	2	3	2	4	6	--	--	--
6	0	2	2	1	1	2	1	0	1
7	1	2	3	1	3	4	--	--	--
8	1	3	4	1	0	1	--	--	--
9	0	4	4	0	0	0	0	0	0
10	0	1	1	0	0	0	--	--	--
11	0	1	1	0	1	1	--	--	--
12	0	0	0	0	0	0	2	1	3
13	0	0	0	0	0	0	--	--	--
14	0	0	0	0	0	0	--	--	--
15	0	0	0	0	0	0	2	0	2
Totals	50	45	95	50	45	95	50	45	95

Table VI. - Frequency Distribution of V.P.B., Cornell and Rotter Total Scores for the "Normal" and "Abnormal" Groups.

V.P.B.			Cornell			Rotter		
Score	N.	A.	Score	N.	A.	Score	N.	A.
0	4	1	0- 1	5	2	99-101	1	0
1	3	2	2- 3	8	4	102-104	1	0
2	9	6	4- 5	7	6	105-107	1	1
3	8	2						
4	5	5	6- 7	6	2	108-110	1	0
5	7	3	8- 9	3	8	111-113	3	4
6	3	2	10-11	9	5	114-116	7	3
7	3	5						
8	1	6	12-13	3	6	117-119	9	2
9	0	1	14-15	2	3	120-122	8	8
10	1	3	16-17	2	3	123-125	10	10
11	1	1						
12	0	3	18-19	1	1	126-128	7	6
13	0	3	20-21	1	1	129-131	0	0
14	2	0	22-23	0	0	132-134	1	3
15	1	1						
16	0	0	24-25	1	2	135-137	0	2
17	0	0	26-27	0	1	138-140	1	3
18	0	0	28-29	0	0	141-143	0	0
19	0	1	30-31	0	1	144-146	0	2
20	2	0	32-33	2	0	147-180	0	0
						181	0	1
Totals	50	45		50	45		50	45

fiducial limits of the data obtained must be attempted. Accordingly, the various essential components were determined and the calculations made leading to the establishment of the "t" ratios reported in Table VII. The Word Association Test screens the groups to the greatest degree, closely followed in turn by the Rotter and the total Verdun Projective Battery. The Multiple Choice Rorschach, Conformity Index and Cornell tests fail to reveal anything of worth in splitting the high school students. Accordingly, the sensitivity of the V.P.B. in this situation is practically wholly attributable to the W.A. subtest.

In a further effort to measure the validity of the testing devices with these groups, chi square analysis was carried out for all the tests and subtests. These results are presented in Table VIII. Yates' correction for continuity[1] was applied throughout. As mentioned earlier, in dealing with the subtests of the Verdun Projective Battery, some form of satisfactory cut-off level had to be devised and applied to the data. The following was the one adopted: since six is the highest tolerable score in the "within normal limits" area, a division of zero to three was arbitrarily set as the boundaries for classification in the "normal" group for the Word Association and Multiple Choice

---

1 Guilford, J. P., Fundamental Statistics in Psychology and Education, Second Edition, New York, McGraw-Hill, 1950, p.278.

Table VII. - Means and Other Statistics Leading to the "t" Ratios for All Tests and Subtests.

Test	Group	M	σ	σ_M	$D_M \div \sigma_{d_M}$	"t"	Prob.	Interpretation
..A.	N.	1.50	1.76	0.25	2.14 / 0.55	3.71	0.005	significant
	A.	3.64	3.28	0.49				
M.C.R.	N.	2.14	1.84	0.26	0.55 / 0.44	1.25	0.25	insignificant
	A.	2.69	2.36	0.36				
C.I.	N.	1.38	3.76	0.54	0.85 / 0.61	1.39	0.20	insignificant
	A.	0.53	1.93	0.29				
V.P.B. total	N.	5.02	4.61	0.66	1.85 / 0.92	2.01	0.05	significant
	A.	6.87	4.23	0.64				
Cornell	N.	8.80	6.82	0.97	1.78 / 1.42	1.25	0.25	insignificant
	A.	10.58	6.89	1.04				
Rotter	N.	119.70	7.07	1.01	6.01 / 2.10	2.86	0.005	significant
	A.	125.71	12.20	1.84				

Table VIII. - Obtained Frequencies[1] Leading to Chi Squares for All Tests and Subtests.

	Obtained freq. N.(by test)A.		Chi Sq.	Prob.	Interpretation
W.A.	0-3	4-15			
N. (by teachers)	45	5	10.04	0.01	significant
A.	27	18			
M.C.R.	0-3	4-15			
N.	41	9	1.04	0.30	insignificant
A.	32	13			
C.I.	0	1-15			
N.	42	8	0.13	0.70	insignificant
A.	40	5			
V.F.B. total	0-6	7-45			
N.	39	11	8.65	0.01	significant
A.	21	24			
Cornell	0-12	13-100			
N.	39	11	1.02	0.30	insignificant
A.	30	15			
Rotter	0-130	131-240			
N.	48	2	6.60	0.01	significant
A.	34	11			

1 Yates' correction for continuity was applied to all data.

Rorschach subtests, and because only disturbed people as a rule miss out on the Conformity Index, any score at all here was considered justification for inclusion under the "abnormal" heading. This serves to explain and account for the cut-off points found in the Table. It will be observed the results for the W.A., total V.P.B. and Rotter tests are again worthwhile from the point of view of personality screening devices, while the others remain nugatory and insignificant. There is a considerable degree of consistence in these statistical findings.

Finally, all possible intercorrelations among the tests and subtests were worked out from the original measurements by the formula:

$$r^2_{xy} = \frac{[N\Sigma XY - (\Sigma X)(\Sigma Y)]^2}{[N\Sigma X^2 - (\Sigma X)^2][N\Sigma Y^2 - (\Sigma Y)^2]}$$

and are listed in Table IX. It is important to note that the final step in finding r_{xy} by this formula is to extract the square root of the number obtained from the above procedure. Both the "normal" and the "abnormal" groups are included in the same illustration, the former constituting the upper right-hand half of the table above the diagonal and the latter found in the lower left-hand section below the dividing line. Lines have been drawn to make the Verdun Projective Battery stand out from the other tests; by coincidence, these also serve to highlight the major correlations found from the data.

Table IX. - Test and Subtest Intercorrelations for the "Normal" (upper right above diagonal) and "Abnormal" (lower left below diagonal) Groups.

	W.A.	M.C.R.	C.I.	V.P.B. total	Cornell	Rotter
W.A.		-.05	-.03	.33	.12	-.21
M.C.R.	-.02		.10	.41	-.12	.12
C.I.	-.15	-.07		.84	-.09	.06
V.P.B. total	.72	.54	.30		-.07	.02
Cornell	.16	.25	-.21	.16		.51
Rotter	-.18	-.01	.06	-.12	.06	

Reference to the Wallace-Snedecor tables[2] will show that the "normal" M.C.R. and C.I. "r" values of .84 and .41 are significant at the .01 level and the W.A. of .33 just below it, while the "abnormal" W.A. (.72) and M.C.R. (.54) are so at the .01 and the C.I. (.30) at the .05 levels of confidence. In addition, the Rotter-Cornell "normal" r of .51 is also significant at the .01 level. All the other intercorrelations are insignificantly positive or negative.

This report on the various statistical analyses made of the data with a view to ascertaining the degree of validity possessed by the instruments will suffice for the present. The discussion of the implications contained therein is reserved to the next chapter.

2. Academic Standing Comparisons

It is no startling disclosure to affirm that because of their variegated training, outlook and experience there will be a wide divergence between the ways in which professional teachers and professional psychologists rate and evaluate people. The standpoints are obviously different. With the end in view of attempting to ascertain to what degree the teachers were consciously or unconsciously influenced by scholastic standing and achievement in their rating and classification of the pupils into the two "extreme" groups, the final June 1954 results of all the latter

---

2 Guilford, J. P., op. cit., p. 610.

were obtained in September from the school archives. These are listed in Table X, together with the annotation as to whether the pupil passed or failed his year and also for the grades concerned the membership list in the"recommended" group (this term will be explained and clarified shortly). The statistical procedures applied to the test data were again utilized.

The outcomes of this analysis are not without interest. For the "normal" group, the mean general average is 75.31, the standard deviation is 9.48 and the standard error of the mean is 1.35; these figures for the "abnormal" subjects are 63.59, 9.76 and 1.49, respectively. This leads to a "t" ratio value of 5.93 which is significant at the 0.005 level of confidence. Furthermore, since the previous data seemed to indicate the Word Association Test to be the most sensitive screen, correlations were calculated for W.A. versus total June results (93 cases, two boys having left prior to the closing of the school), and those of the "normal" (49 subjects) and "abnormal" (44 students) groups. The coefficients obtained were -0.10 for the full complement of testees, 0.33 for the "normals" (almost the .01 level) and 0.14 for the "abnormals, using the same formula as before.

Two other dichotomies were possible (pass-fail and recommended-non-recommended) and chi square analysis was effected here. The entire ninety-five cases were used in

Table X. - Academic Results[1] in June 1954 of the "Normal" and "Abnormal" Subjects.

No.	Class	Gen. Aver.	Pass or Fail	Recommended?	No.	Class	Gen. Aver.	Pass or Fail	Recommended?
1	1B	85	P	yes	51	1D	69	P	no
2	1D	86	P	yes	52	1C	66	P	no
3	1B	82	P	yes	53	2A	82	P	--
4	1C	64	P	no	54	1A	70	P	no
5	1D	87	P	yes	55	1D	68	P	no
6	1D	87	P	yes	56	1C	66	P	no
7	1C	62	P	no	57	1B	63	F	no
8	1D	88	P	yes	58	2A	73	P	--
9	1D	88	P	yes	59	1A	67	P	no
10	1D	77	P	yes	60	1B	55	F	no
11	1C	--	P	no	61	2A	76	P	--
12	1D	77	P	yes	62	2D	56	F	--
13	1C	74	P	no	63	3B	77	P	no
14	1B	72	P	no	64	3A	72	P	yes
15	1A	82	P	no	65	1C	63	P	no
16	2C	83	P	--	66	2C	68	P	--
17	3A	54	F	no	67	3A	74	P	no
18	2C	64	P	--	68	2D	32	F	--
19	1D	81	P	yes	69	2D	58	F	--
20	2A	85	P	--	70	2A	67	P	--
21	1D	84	P	yes	71	2B	65	P	--
22	2C	79	P	--	72	3B	59	F	no
23	2B	56	F	--	73	2B	72	P	--
24	1A	73	P	no	74	2B	53	F	--
25	2C	83	P	--	75	3C	61	F	no
26	3A	63	P	no	76	2D	64	P	--
27	1D	86	P	yes	77	2D	56	F	--
28	1D	82	P	yes	78	1A	71	P	no
29	2A	90	P	--	79	3A	53	F	no
30	2B	75	P	--	80	3B	70	P	no
31	3B	71	P	no	81	1B	57	F	no
32	2A	90	P	--	82	2C	45	F	--
33	2C	63	P	--	83	1B	74	P	no
34	2B	66	P	--	84	1B	65	P	no
35	2B	77	P	--	85	3C	71	P	no
36	3B	69	P	no	86	3B	63	P	no
37	2B	70	P	--	87	3A	64	P	no
38	2B	71	P	--	88	3B	75	P	no
39	3A	55	F	no	89	3A	45	F	no
40	3C	85	P	yes	90	2A	71	P	--
41	3C	74	P	no	91	3C	50	F	no
42	3C	76	P	yes	92	3A	--	F	no
43	3B	67	P	no	93	3A	61	P	no
44	2B	71	P	--	94	3C	50	F	no
45	3C	78	P	yes	95	3A	61	P	no
46	3A	72	P	no					
47	2B	72	P	--					
48	2C	76	P	--					
49	3C	69	P	no					
50	3C	70	P	no					
Totals		3690	P=47	yes=16			2798	P=30	yes=1

1 Taken from the archives of The Thomas D'Arcy McGee Boys' High School, Montreal, Que.

the pass-fail categorizing for of the two boys who left, one did so to leave with his family to join his father in the Canadian Armed Forces stationed in Germany and on the verbal report of his teacher was a student who would most certainly have passed his examinations had he remained and the other was a boy who never evinced the slightest interest in matters scholastic (actually his departure back to Prince Edward Island practically coincided with the ending of the hockey season) and whose failure for lack of effort and concentration (or homesickness) was a foregone conclusion. The "normal" group split 47 to 3 and the "abnormal" 30 to 15 on a pass-fail basis. This resulted in the significant (0.01) chi-square value of 9.91.

The question of being "recommended" or not (that is, of being excused from writing the June final examinations and of being promoted on the results of the year's work alone under certain specified percentage restrictions which are of no concern here) applies only to Grades Eight and Ten. The pupils of Grade Nine (Second High) all must write for one of the junior Provincial Certificates that lead to the senior and specialized high school course. Automatically, they are excluded from the following data. Incidentally, the percentages in Table X for the "recommended" pupils represents their general average for the year and are not, strictly speaking, June examination results. On the "recommended" or not division,

the groups separated as follows: "normals" 16 to 17, "abnormals" 1 to 29, - a striking difference, resulting in a chi square value of 14.07 which is significant at the 0.01 level. The discussion of all these data are reserved for the next chapter.

3. Group Versus Individual Administration

It has already been specified that the school population underwent group administration of the V.P.B. (with the obvious timed W.A. exception), while the hospital cases (both normal and psychotic groups) were individually tested. Some investigation of possible differences arising from the manner of administration seemed worthwhile as an adjunct or secondary branch of the main scope of the research. Accordingly, hitherto unpublished data from the individual testing of normal subjects (both boys and girls) within the same age range as the school groups was made available from the research files of the hospital and chi-square comparisons were drawn between these and both the "normal" and "abnormal" school subjects. The results are presented in Table XI for consideration and analysis. Incidentally, the youngsters mentioned were all enrolled at a summer camp and, therefore, on vacation.

The first three sets of calculations represent a brief illustration of the apparent lack of sex differentiation reported in Chapter One. As no appreciable differences could be observed between the results of the individually-tested

Table XI. - Obtained Frequencies[1] Leading to Chi Squares for the Sex Difference and Campers vs. Students Comparisons.

	Obtained freq. N.(by test) A.		Chi Sq.	Prob.	Interpretation
W.A.	0-3	4-15			
boys	21	4	0.06	0.80	insignificant
girls	13	4			
M.C.R.	0-3	4-15			
boys	12	13	0.14	0.70	insignificant
girls	10	7			
C.I.	0	1-15			
boys	20	5	0.06	0.80	insignificant
girls	15	2			
W.A.	0-3	4-15			
camp N.	34	8	0.92	0.75	insignificant
school N.	45	5			
M.C.R.	0-3	4-15			
camp N.	22	20	8.06	0.01	significant
school N.	41	9			
C.I.	0	1-15			
camp N.	35	7	0.15	0.70	insignificant
school N.	42	8			
W.A.	0-3	4-15			
camp N.	34	8	3.68	0.06	insignificant
school A.	27	18			
M.C.R.	0-3	4-15			
camp N.	22	20	2.54	0.12	insignificant
school A.	32	13			
C.I.	0	1-15			
camp N.	35	7	0.18	0.65	insignificant
school A.	40	5			

1 Yates' correction for continuity was applied to all data.

boys and girls, they were lumped together into one group of forty-two cases for the subsequent comparisons with the high school students. There seems to be no significant difference between group and individual administration of the Conformity Index for either the "normal" or the "abnormal" subjects (of course, while the setting may be a group one, the test itself is fundamentally an individual procedure).

The Word Association (an individual test for all by its very nature) results seem to offer one item of particular interest, namely, the 3.68 value (approximate 0.06 level of confidence) found in comparing the summer camp youngsters to the school "abnormals". There is very properly no significant difference with the "normal" testees. Such, however, is definitely not the case with the Multiple Choice Rorschach test. Here the "normals" show up adversely with a significant (0.01) chi square of 8.06, while no major variation was recorded for the "abnormal" group. Possibly the group administration of this test is less sensitive with normal people. Incidentally, unpublished normative data of the authors show that high scores are more common on this test in adolescence.

4. Item Analysis for Reliability

The method chosen to elicit the amount of reliability possessed by the instrument in question, the Verdun Projective Battery, was that of the internal consistency analysis formulated and devised by P. J. Rulon[3]. This is, of course, a variation of the split-half procedure; but one which does not, however, require a Spearman correction as the reliability of the total test scores and not the halves is given by the formula. The other methods of assaying reliability were not applicable to the present situation. As there is only one edition of the test, use of the equivalent forms device was clearly impossible. The test-retest technique is apparently considered by the leading statistical authorities of somewhat dubious value in personality testing. Garrett, for example, has this to say concerning reliability methods: "Likewise, many personality tests cannot be given in alternate form nor repeated because of radical changes in the subject's attitude and interests when taking such tests for the second time."[4]

Since Rulon's is one of the split-half devices, the first method of half splitting for reliability used in the investigation will now be described. The percentage of "wrong" answers (i.e., ones meriting a score on the V.A. and M.C. tests

---

3 Guilford, J. P., op. cit., p. 497.
4 Garrett, Henry E., Statistics in Psychology and Education, Third Edition, New York, Longmans, 1947, p. 383.

or not being included on the list of conforming associations for the C.I. instrument) was obtained for each of the thirty items, ten on each subtest, of the Verdun Projective Battery. These were then divided into two balanced groups of fifteen items each, five from each subtest, on the basis of equivalent total "failure" percentage. The procedure was repeated for the "normal" and "abnormal" school groups independently. This resulted in the two percentage matching halves for the "normal" group as follows: a) W.A. items 1 - 2 - 3 - 4 - 8, M.C.R. 1 - 2 - 7 - 9 - 10 and C.I. 3 - 5 - 8 - 9 - 10 versus b) W.A. 5 - 6 - 7 - 9 - 10, M.C.R. 3 - 4 - 5 - 6 - 8 and C.I. 1 - 2 - 4 - 6 - 7. For the "abnormal" group, the similarly-obtained split was the following: a) W.A. items 1 - 3 - 4 - 5 - 7, M.C.R. 1 - 4 - 5 - 7 - 10 and C.I. 3 - 5 - 8 - 9 - 10 versus b) W.A. 2 - 6 - 8 - 9 - 10, M.C.R. 2 - 3 - 6 - 8 - 9, and C.I. 1 - 2 - 4 - 6 - 7. All this material is summarized in Table XII.

The next step was the application of Rulon's estimation of internal consistency according to the formula: $1 - \left(\frac{\leq d^2}{N} \div \sigma_t^2\right)$, where d = difference between two scores of half-tests for one individual. For the "normal" group, this gave a reliability estimate of 0.88 and for the "abnormal" one of 0.77; however, both are probably underestimates because of the fact that the intercorrelations are low and the items are definitely not of equal difficulty. For example, in the W.A. test, the words "food" (4), "voice" (6) and "color" (9) show practically no

Table XII. - Percentages of "Failure" on Each Item of the V.P.B. Subtests Leading to the Establishment of the Reliability Split Halves.

	Word Assoc.				Mult.Choice Rors.				Conform. Index			
Item	N. %	sub-group	A. %	sub-group	N. %	sub-group	A. %	sub-group	N. %	sub-group	A. %	sub-group
1	6	a	38	a	14	a	33	a	34	b	44	b
2	2	a	18	b	36	a	33	a	66	b	64	b
3	18	a	31	a	34	b	38	b	56	a	55	a
4	20	a	22	a	22	b	24	a	40	b	20	b
5	2	b	18	a	8	b	13	a	38	a	44	a
6	24	b	27	b	10	b	13	b	30	b	22	b
7	4	b	13	a	10	a	9	a	52	b	49	b
8	12	a	40	b	2	b	9	b	64	a	53	a
9	16	b	16	b	4	a	11	b	50	a	47	a
10	12	b	24	b	10	a	7	a	16	a	22	a
Totals N.	a 58		b 58		a 74		b 76		a 224		b 222	
Totals A.	a 125		b 122		a 104		b 86		a 199		b 221	

Grand Totals: "Normals": a= 356; b= 356; "Abnormals": a=428, b=429

difference between the "normal" and "abnormal" subjects, while the allegedly most neutral words ("table" (1), "river" (2) and "pencil" (5), for instance) show the greatest degree of difference. On the M.C.Rorschach, Card I separates the best, followed far behind by Cards V, VIII and IX, while II - III - IV - VI - X show minimum differences. On the Conformity Index, "eagle", "tobacco", "man" and "lamp" in that order screen the groups most adequately, while the results with the remaining six words are practically equivalent as far as present grouping is concerned. Incidentally, applying Rulon's formula to the more customary odd-even dichotomy (as a second method of half splitting) resulted in almost identical results, namely, .83 for the "normals" and .77 for the "abnormals".

This brings to a close the presentation of the analysis of the data. The statistical techniques that were applied in an endeavor to ascertain the degree of validity and reliability, the influence and relationship of academic standing and scholastic status, and the possible interplay of group versus individual administration have all been described. They were primarily the "t", "X^2" and "r" procedures. In the next and final chapter, the data will be examined for significance and these results discussed at somewhat greater length from the point of view of the primary and associated aims and objectives of the whole investigation.

CHAPTER FOUR

DISCUSSION OF THE DATA

A statement of the general aims of the whole research problem was set forth early in the first chapter. These were said to be, in short, the assessment of the validity and reliability of the Verdun Projective Battery for high school adolescents, as the primary project, and the analysis of group versus individual administration procedures, the examination of the degree of relationship of academic standing to test and rater evaluation and the investigation of the extent of similarity in the results obtained by the projective and non-projective test material, as topics of related interest. The data relating to these specific aims were analyzed in the preceding chapter. It is the intention in this concluding part of the thesis to discuss and interpret, where the evidence would seem to warrant such an undertaking, these results following and according to the pattern of the aims enumerated above. The sub-sections will naturally be very similar in scope and theme to those in Chapter Three. Specifically, attention will be devoted, in turn, to an appraisal of the significance and worth of the validity found, some estimate of the scholastic standing relationship, the extent of the influence of the administrative technique employed, and an evaluation of the degree of reliability established.

1. Validity Appraisal

Even a cursory examination and mere inspectional analysis of the data secured for both groups of subjects reveals the great divergence among the various tests and subtests used in the experiment. However, for the most part, the degree of discrimination found, even when low, is at least in the right direction. For certain of the measurements, namely, the Word Association subtest and because of it the total V.P.B., and the Rotter test, the results seem to indicate a dependable degree of discriminative power and therefore these seem to possess considerable validity as personality screening instruments. In any event, in this section an attempt will be made to account for the strengths and shortcomings revealed by the statistical analysis.

In the first place, teachers and psychologists are inevitably going to rate people from different bases and points of view. This fundamental fact may be the foundation of the apparent worth of the Word Association test and seeming nugatory value of the Multiple Choice Rorschach with these groups. Incidentally, the observable relationship between personality rating and scholastic standing would seem to confirm this hypothesis. This will be discussed further in the second section of this chapter. The poor report from the Conformity Index was, perhaps, to be expected. Major scores on this test

are usually made by definite psychotics; from the very nature of the environmental set-up, these are not going to be found, in the main, in a modern high school. The possible exception might be the odd latent schizophrenic, who, being withdrawn and causing no disturbance, might be considered a "good" pupil by the teachers. This might account for the four high Conformity Index scores in the "normal" group. Again, highly creative and original thinking subjects also produce false positives.

Why does the Word Association test discriminate so well and so constantly, as a glance at the "t", "X^2" and "r" values shows? Perhaps the fact that it assays mental control is the key. The teachers by their training, experience and professional interest would be expected to be able to rate and grade pupils accurately from the intellectual and academic proficiency standpoint. Very likely, this unintentional and probably unconscious bias affected to some extent their ratings. This situation, of course, would have most effect on the more or less related Word Association test. Thus the teachers and this subtest would concur very well in their respective judgments. Of the three aspects measured by the V.P.B.: mental control, personality structure and commonness of thought content, only the former bears any strong relationship to capacity to function in an academic setting, for the extreme cases are probably already out of school. The fact that teachers rating their

pupils as teachers (and not as psychologists) do nevertheless achieve a sufficiently high validity and trustworthiness that ensures entrusting the test to them a safe, practical and feasible procedure may prove to be the most vital finding of all, if the Battery is to be placed in service at the educational level.

The reverse side of the picture holds up on examination also. The Multiple Choice Rorschach assays the personality resources. Here the teachers are on unfamiliar ground. Their classroom and extracurricular contacts with the pupils, even including the advantageously placed athletic coaches, are of a definitely limited nature in comparison with the full span of activities carried on by the students in their daily lives. In this area of personality rating, where the professional psychologist would be very much at home, the teachers may well have completely foundered. Is this the reason for the apparent futility of the Multiple Choice Rorschach as a screening device for these pupils? Certainly, the hypothesis does not seem to be beyond the bounds of credibility. If this subtest is of such little worth (in teachers' hands) at the school level, it might be advisable to extirpate it from the Battery altogether for that stratum or reserve its use to the referral stage by the psychologist. On the other hand, it may not be that the M.C.R. is no good or that teachers would be unable to use the test, but rather that they cannot rate the pupils

personality-wise and hence their divisions are unrelated to test results.

Although all the statistical devices show the total Verdun Projective Battery to be valid and discriminative, it is evident that the source of its apparent strength is the very high incidence of validity possessed by the Word Association subtest in this situation. It will be noted, by the way, that all the subtests, while correlating poorly or even negatively with each other, do have significant "r" values with the total test, and this situation maintains for both the "normal" and "abnormal" groups. This is a very worthwhile finding, i.e., that the subtests are all independent measures. As it happens, two of the "things" measured have no relevance to this situation, but they are all separate, nonetheless. The summ ry for the pivotal V.P.B. instrument, then, is that the Word Association subtest and the Battery as a total survive the "t", "χ^2" and "r" probing and seem possessed of a reasonable degree of validity, but the Conformity and Rorschach subtests, perhaps for the reasons adduced earlier, fail to reveal any worthwhile and dependable discriminative power between these groups, possibly due to the method of group selection.

What of the Cornell and Rotter tests? How do they stand the strain of statistical analysis? Briefly, the former fails on practically all counts in this project. The "t" ratio and chi square results are insignificant and the correlations

are low, - in fact, they are almost completely negative for the "normals" with the V.P.B. total and subtests and not very much better with the "abnormal" group. The only respectable figure is the "r" of .51 for the "normals" with the Rotter test. This contrasts rather sharply with the .06 for the "abnormals", however. What gave rise to this can only be speculated upon. Does this test for adults just not lend itself adequately and realistically to the screening of an adolescent high school population? Again, it may be that only one test is sensitive to abnormality as teachers see it. The Rotter would logically be the one. Thus, for the normals there is nothing to note as the tests correlate. For the abnormals, however, the Rotter scores are sensitive, and so the tests fail to correlate. Whatever the reason, Form N2 of the Cornell Index contributed nothing positive to this investigation.

With the High School Form of the Rotter Incomplete Sentences Blank, on the other hand, very different results were obtained. Very significant discrimination was brought to light by the "t" ratio and chi square techniques. It is true the correlation values, with the exception of the above-mentioned "normal" Cornell (.51), are all low and sometimes negative. Nevertheless, the test seems able to screen out "abnormals" from "normals" to a very considerable degree. The source of this differentiating power may be inherent in the test itself, which was, after all (and unlike the Cornell), especially

prepared for high school students. It is interesting to note that while the Word Association and Rotter tests discriminate between the students almost equally well, they are actually uncorrelated. Therefore, the students differ on more than one criterion. At any rate, the Verdun Projective Battery and the Rotter test are both effective (of course, they are related in that both are projective devices) in their approach to and discriminative power with these adolescents at the secondary school level. The services of both could well be enlisted as worthy adjuncts by the school counselling and guidance personnel.

2. Estimate of Scholastic Standing Relationship

The comparison of the personality rating of the pupils made by the teachers in January 1954 with the scholastic results secured by these identical students six months later on the June final examinations is certainly striking. Even with the stencilled list of some of the observable signs of abnormality to guide them, apparently it was impossible for the teachers to rate the students other than as teachers. Their mental set, it seems, was too powerful to overcome. The examination of the high degree of accuracy achieved by the "mental-control" oriented Word Association test, where the teachers would be expected to possess the greatest amount of clarity and insight into the personality structure of the students, confirms this hypothesis. The concomitance of group rating and academic

result reinforces this supposition.

As corroborative evidence, consider the pass-fail dichotomy: only three "normals" failed (and of these the V.P.B. classed one as "abnormal" and another on the boundary with the borderline score of six), whereas five times as many of the fewer "abnormals" did so (actually the fifteen cases constitute one-third the group) and the V.P.B. test concurred on eight of them. The question of being "recommended" or not presents an even more striking illustration. Of the thirty "abnormals" eligible for recommendation, only one did in fact receive the accolade (probably because he was repeating his year); this individual stands out, incidentally, as the one who made the highest V.P.B. in the entire "abnormal" group. On the other hand, fifty per cent of the thirty-three eligible "normals" achieved the honor and the V.P.B. agreed on nine of the sixteen cases.

The mean general averages of the two groups were so diverse that the very significant "t" ratio of 5.93 resulted. Coupled with the extremely high chi square values (9.91 and 14.07), this makes the relationship even more apparent. In addition, the "r" of .33 between the W.A. scores of the "normals" and their June results is almost significant at the 0.01 level of confidence, according to the Wallace-Snedecor tables. The fact that reaction-times to the ten stimulus words (or only seven actually, for three failed completely to differentiate

the groups) of the W.A. subtest can apparently predict year end academic results to some degree is of possible great significance in educational circles and is an aspect of the results of this research that would certainly seem to merit very considerable further investigation.

Be this as it may, the incontrovertible fact is that teachers in rating pupils are going to be heavily influenced by their knowledge or opinion of the latter's scholastic proficiency. Furthermore, it would seem that "alert" pupils are the ones whom teachers consider "normal". Does this situation vitiate the results and is it going to prejudice their successful use of the Verdun Projective Battery at the educational level? It is felt the answer to both inquiries is to be made in the negative. Variations will certainly enter in and must be provided and accounted for in the new setting, but the consistency, as it were, of the deviation should run "true" after the initial start. That is to say, all teachers will be professionally and personally inclined to gear their rating to the same criteria (in which intellectual acuity and academic endeavor are inevitably going to loom large) and once this is accepted as a constant, there will be logic and validity and self-consistency in the operations of the mechanism thereafter. This is presuming the staff of instructors of the school utilized in the survey (The Thomas D'Arcy McGee Boys' High School of Montreal) is representative of Canadian high school

teachers throughout the country, and there certainly seems no particular reason from the point of view of their training, experience and professional outlook and practice why this premise and postulate should not be valid and acceptable. The pupil sampling likewise in no way appeared unique or specialized.

In the school use setting, the major emphasis in interpreting the situation will have to be accorded to the Word Association results, with the other two subtests serving in reserve as auxiliary adjuncts to the over-all screening. Any score on the Conformity Index will still be viewed with suspicion, while the Multiple Choice Rorschach criteria may best be left, perhaps, to the trained psychologist (the quantified score will, of course, be preserved and used in the total rating). The study really throws no light on the M.C.R. or C.I. effectiveness, so that their value in a school situation remains in doubt. To evaluate them properly, some means of selecting a truly "abnormal" group would have to be devised.

3. Influence of Type of Administration

As rapid scanning of the chi square values, all insignificant and above the .05 level of confidence, will reveal, there seems no observable appreciable difference between the results the boys and girls at the summer camp obtained on the Battery. It was this state of affairs, of course, that naturally led to the suggestion that they be grouped together to

constitute and form one larger body for purposes of comparison, from the point of view of the effect and influence of type of administration procedure followed, with the results of the high school students in each of the two categories into which they were classified. The youngsters of the camp group were all theoretically "within normal limits" and so the Verdun Projective Battery found most of them to be, as a glance at the listings in Table XI will show. The Word Association and Conformity Index subtests reveal nothing too unusual; however, the incidence (almost fifty per cent) of apparent abnormality on the Multiple Choice Rorschach seems rather startlingly high for such a group of average adolescents and will be dealt with later.

The two sets of "normal" testees synchronized very well on their Word Association returns, as, indeed, they certainly should, for this test is of its essence (having to be timed with a stop watch) limited to the individualized procedure of administration. The school "abnormals" also lived up to their teacher-rating in comparison with the summer-camp people, achieving the 3.68 chi square referred to in the preceding chapter. This is almost significant at approximately the .06 level of confidence and gives further credence to the apparently well-founded belief that the Word Association subtest is far and away the most dependable and discriminating instrument for use in the school situation.

The number of individuals obtaining a zero score, that is, who gave at least four conforming associations on the Conformity Index subtest, does not seem to vary much from group to group with these young subjects. But, as was pointed out earlier, psychotic propensities (the discovery of which is this test's especial forte) are happily rare at these age-levels and rarer still in educational institutions of this type. Whatever ones there are in the groups concerned have probably earned the few high C.I. scores, but they are not numerous enough to affect the statistical significance (or, rather, the absence of it). The fact must be admitted also that it may be possible that these young people (irrespective of their grouping) lack the experiential background necessary to produce to certain of the stimulus words the responses that have been designated as conforming associations on the basis of adult reactions. The verification of this hypothesis would be well worthwhile as a form of additional research on the Battery.

Finally, the spotlight returns to the Multiple Choice Rorschach, where most divergence was observed. It is on this subtest that both the "abnormal" and "normal" groups contrast rather strangely with the comparison data of the campers. Consider, first, the findings of the school "normal" versus camp "normal" inquiry. Here the significant value of 8.06 was derived from the chi square analysis. In sharp contrast, however, the school "abnormals" have no significant M.C.R. difference from their camp cousins.

What can possibly account for this seeming paradox? Two tentative hypotheses may be put forth at this juncture. The first is simply the frank admission that perchance the group administration is, after all, less sensitive in its screening potential than the original individual technique. Maybe visual power is partially affected by use of the slides instead of the customary standardized cards, perhaps the colors are less bright or the shading more vague and diffuse or the minor details lost or obscured, possibly distance and the angle of the vision play a part, or a host of other similar factors could influence the selections made, - and all this in spite of the many precautions taken during the preparations and preview screening of the school auditorium. These things may be inherent in and intrinsic to the group administration setting, no matter what the locale.

The second thought that occurs, and one that seems to embrace a good deal of commonsense in it, is the possibility that these testees are all too young for their personalities to be sufficiently stabilized for accurate screening by the Multiple Choice Rorschach. Since they are not adults, they can neither logically nor justly be judged and evaluated in terms of adult standards. Incidentally, among the authors' as yet unpublished data is evidence that a fair number of high M.C.R. scores are "normal" among adolescents and consequently the test is probably not suitable for them.

Perhaps these considerations account for the unusual twenty-two to twenty split on the campers' scores and the other anomalies, statistical and otherwise, observable in the M.C.R. results. Again, it must be brought to the fore of attention that teachers and psychologists are inevitably going to screen people by different channels of approach and that the estimation of a person's personality resources is not the particular prerogative of the average pedagogue. It is a moot question and one requiring detailed study and analysis whether anything more than the quantified score should be extracted from the Multiple Choice Rorschach in the school setting. Differential interpretation of this instrument is far more safely entrusted to the hands of the trained and experienced clinician. Even its retention at all in the Battery for school use could be a subject for discussion.

In short, then, the influence of the type and the procedure of administration seems to be limited to the Rorschach area of the total test and even there has not been definitively decided. It is apparently negligible elsewhere. This is important in the average school testing situation where the group, rather than the individual, administrative technique will expectedly be employed, except in the case of the few individual "emergencies". This in no way negates the practical value of brief individual tests, such as the Word Association, where five minutes' study of a particular student may reveal very worthwhile results.

4. Reliability Evaluation

While there are three general methods for estimating reliability from empirical data, namely, internal-consistency (split-half), test-retest (degree of stability on repetition), and alternate or parallel forms (leading to a coefficient of equivalence), the first type was the one selected for use in this investigation. The lack of equivalent forms for the tests employed precluded any consideration of the alternate-parallel method. Also, as pointed out in the preceding chapter, the test-retest device is of very limited use and value. It requires double the time and may be subject to a host of extraneous influences. Garrett speaks of its lesser accuracy and goes on to report that "the test-retest method is used less generally than are the other two methods".[1] Guilford, too, points out that the changes in the examinee himself, e.g., fatigue, boredom, emotional state and the like, are determiners that contribute to retest errors.[2] In addition, Adkins flatly states: "Further, it has been pointed out that the test-retest coefficient is affected not only by the unreliability of the test but also by any lack of stability of the function being tested. These

---

1 Garrett, Henry E., Statistics in Psychology and Education, Third Edition, New York, Longmans, 1947, p. 381.

2 Guilford, J.P., Fundamental Statistics in Psychology and Education, Second Edition, New York, McGraw-Hill, 1950, p.482.

difficulties are so serious that the test-retest method is rarely used".[3] These affirmations assume all the greater importance when it is recalled that the tests to be evaluated were personality instruments. Certainly, if fluctuations are going to occur anywhere, this is the area "par excellence" for their occurrence, with devastating effects on the impartial evaluation of the true degree of reliability possessed by the tests.

These considerations led to the selection of the internal-consistency method for the purposes of this project. The particular single administration device utilized was Rulon's method of reliability estimation. It has already been fully described, as have the techniques of obtaining the two halves into which all the Verdun Projective Battery items were dichotomized. In addition, some results of the very general item analysis have been enunciated for the three subtests in turn. Possibly, the whole question of reliability is encapsulated in the two, no doubt underestimated, sets of indices derived: .88 for the "normals" and .77 for the "abnormals" of the school population by the "failure percentage" split and .83 and .77, respectively, by the more usual odd-even grouping. Perhaps, in connection with these figures, it might be à propos to review a very pertinent comment of Guilford's on requisite reliability levels: "This is a dramatic demonstration of the

---

3 Adkins, Dorothy C., et al., Construction and Analysis of Achievement Tests, Washington, D.C., U.S. Government Superintendent of Documents, 1947, p. 150.

fact that the requirement of reliabilities of .90 and above is very unrealistic and that if a selection test proves to be valid we can tolerate its low reliability. Standards for evaluating both validity and reliability coefficients must be relative. No universal limits can be applied".[4] The figures obtained are at least satisfactory, for on applying Fisher's formula[5] for "t" in testing a coefficient of correlation $\left(t = r\sqrt{\frac{N-2}{1-r^2}}\right)$, the following "t" values, all significant at the .01 level, are the result: a) .88 = 12.72, b) .83 = 10.33, c) .77 = 7.89.

It is of some interest to note in the item study made that in the case of all three subtests, the first stimulus encountered screens the subjects by group exceedingly well. Thus, on "table" in the W.A., the "normals" record only 6 per cent failure as opposed to 38 per cent for the "abnormals"; similarly, the initial Rorschach blot divides them 14 to 33; and "man" on the C.I. makes a 34 - 44 split. Are the "abnormals" especially experiencing "initial shock"? Incidentally, the blots bearing color show very meager discriminative power with these subjects. The reason can only be a matter for conjecture. Conforming associations for the two most divergent words ("eagle" and "tobacco") are seemingly more difficult for the "normal" group to produce. An unusual result on the W.A. subtest is the fact that the expectedly "neutral" words ("table", "river", "pencil",

---

4 Guilford, J.P., op. cit., p. 485.
5 Guilford, J.P., op. cit., p. 227.

"man", etc.) show the most marked differences between the groups, while the more "abstract" ones ("color", "food", "voice", - the exception was "sickness") fail to differentiate them. Many of these and similar findings could well provide subjects for further study and investigation.

With this brief overview of the inherent reliability of the Battery for the average high school population, the discussion of the data is concluded. Attention has been given in this fourth chapter to some study of the validity appraisal and estimation of relationship of scholastic standing to test and teacher personality rating. The influence of the group versus individual type of administration was examined and the test reliability evaluated. The final step will be to essay a short summary together with a few brief conclusions arrived at on the basis of the evidence of the data gathered, collated and analyzed.

SUMMARY AND CONCLUSIONS

The purpose of these brief concluding remarks is to resume finally the overall process of investigation, analysis and interpretation carried on throughout the project and to collect together and sum up the significant conclusions reached and arrived at as a result of the research. In short, then, it will be recalled that the pupils of D'Arcy McGee Boys' High School were classified by their teachers into the two categories designated, for want of a more easily grasped nomenclature, as "normal" and "abnormal", respectively. Three personality screening devices, the Verdun Projective Battery (three subtests), Form N2 of the Cornell Index and the High School Form of the Rotter Incomplete Sentences Blank, were subsequently administered to these "extreme" groups.

The statistical techniques of the "t" ratio, chi square analysis and correlation procedure were then applied to the raw data gathered. This was done primarily in an endeavor to ascertain the degree of validity and reliability the pivotal Verdun Projective Battery possessed as an adequate personality screening instrument at the high school level. But other matters also received consideration and attention. These included an examination of the extent of influence of the type (group or individual) of administration procedure employed, an estimation of the relationship of academic proficiency and

eminence to the personality rating of test and teacher, and lastly, a more or less superficial overview of the interrelationship of the projective and non-projective approaches to the whole problem of personality discrimination and differentiation.

What definite conclusions have logically been reached as a result of all these analytic procedures? The first and foremost of these is, surely, the statistical evidence that the Verdun Projective Battery does, indeed, do the job in question, that is, does achieve a creditable degree of accuracy in screening high school students on a personality basis. The key point of paramount importance is that in the hands of teachers - and this would ordinarily be the case in its use at the educational level - the V.P.B. retains a sufficient measure of its clinical significance to be useful and worthwhile. At the same time, the degree of reliability is likewise found to be more than satisfactory.

In the school setting, the three subtests have uneven capacity. The strongest and most clearcut performance of all was achieved by the Word Association test; this, of course, by its nature, bears some resemblance to the typical school situation and may even be predictive of scholastic results. Any Conformity Index score of major proportions requires investigation, as is the case in the clinical setting. The Multiple Choice Rorschach, however, apparently is unable to measure the pupils' personality resources, possibly because

of the teachers' seeming inability to rate their pupils accurately on a personality basis. With the progress of time and the accumulation of experience, no doubt, gradual changes[1] and minor modifications will be made and take place in the items of the various V.P.B. subtests, in order to strengthen its general performance in any particular setting, including the educational one. These improvements will gradually exert a beneficial influence on the future form and use at all levels of the Verdun Projective Battery.

What were the findings recorded for the two companion testing devices? The non-projective Cornell Index proved to bear no clear relationship to the data obtained from the other sources. It was, however, designed for use with a slightly older group. The Rotter test, on the contrary, revealed very significant results and ones quite comparable to those obtained by the Battery. The remaining matters may be rapidly

---

1 As a matter of fact, a personal communication (January 4, 1955) from the junior author lists the following changes already decided upon: a) W.A.: "pencil" down to 2.0, "sickness" up to 3.0 and "money"down to 3.0; b) M.C.R.: CardIII up to 3, Card V up to 4 and Card VI down to 4; c) C.I.: "cabbage" and "ocean" are deleted and double service is made of "table" and "pencil" from the W.A. subtest; d) the new cut-off point is 7, resulting in the two total score groups of 0 - 7 and 8 - 45, with the former middle group disappearing altogether, - all these alterations have resulted from statistical evaluation of normative data now accumulated for the age groups from 10 - 79, inclusive.

itemized as follows: the M.C.R. subtest shows the only reaction to administration variations, possibly for the reasons adduced earlier; pupil scholastic standing and teacher personality rating show a rather close relationship; and the projective results seem to be roughly similar and comparable.

From these and other considerations enumerated and described in the main body of the thesis, it would seem permissible to believe that the Verdun Projective Battery has a definite contribution to make as an effective personality screening and eveluation device at the high school level, - the investigation of which hypothesis constituted this particular research project.

ANNOTATED BIBLIOGRAPHY

Abt, Lawrence Edwin and Leopold Bellak, Projective Psychology, New York, Knopf, 1950, xv - 435 - xiv p.
Eclectic contribution to the theory and practice of projective methods; Harrower describes the Rorschach group techniques and the sentence completion device also receives considerable attention.

Adkins, Dorothy C., et al., Construction and Analysis of Achievement Tests, Washington, D.C., U.S. Government Superintendent of Documents, 1947, xvii - 292 p.
Practical text on the technical problems and methods of examining; points out weakness of test-retest method as an estimate of test reliability.

Allen, Robert M., Introduction to the Rorschach Technique, New York, International Universities Press, 1953, 126 p.
A useful introductory manual on the administration and scoring of the test; some consideration is given to both the theory and the mechanics of the Rorschach approach.

Allen, Robert M., Elements of Rorschach Interpretation, New York, International Universities Press, 1954, 242 p.
An introduction to the complexities of evaluating and interpreting the Rorschach record; includes a very extensive bibliography of 1130 references; a well-organized attempt to initiate the neophyte on the road to competence.

Anderson, Harold G. and Gladys L. Anderson, Editors, An Introduction to Projective Techniques, New York, Prentice-Hall, 1951, xxiv - 720 p.
A recent contribution to the understanding of the dynamics of human behavior by a general survey of this specialized field; extensive material on the Rorschach, word association, and sentence completion methods is provided.

Bell, John Elderkin, Projective Techniques, New York, Longmans, 1948, xvi - 533 p.
Comprehensive picture of the whole field of projective technique theory and practice; contains material on the word association, incomplete sentences and Rorschach tests.

Canadian Mental Health Association, "Let's Stop This Human Waste Now!", (no publisher), (no place), (no date), 4 p.
Leaflet inviting membership in the Association; gives facts and figures, including those for children, of occurrence of mental illness in Canada.

Coleman, James C., Abnormal Psychology and Modern Life, Chicago, Scott, Foresman, 1950, xi - 657 p.
A bio-psycho-social overview of psychological abnormality against the developmental background of present-day society; provides statistics on the incidence of mental illness in U.S.

Dorken, Herbert, jr., "Projective Tests and the Consistency of the Personality Structure: a Pilot Study", in The Journal of Abnormal and Social Psychology, Vol. 48, No. 4, 1953, p. 525-531.
Analysis of Rorschach and fingerpainting test correlations for normal and psychotic subjects; latter had high intra- but low inter-test "r's", while substantial intertest "r" of former considered a result of the consistency of the normal personality structure.

Freeman, Frank S., Theory and Practice of Psychological Testing, New York, Holt, 1950, xxi - 518 p.
Descriptions of major psychological instruments together with their theoretical principles and basic assumptions; offers considerable matter on the Cornell, Rorschach, word association and sentence completion devices.

Froehlich, Clifford P., and John G. Darley, Studying Students, Chicago, Science Research Associates, 1952, xviii - 411 p.
Guidance methods for the individual analysis and understanding of students; contains a useful outline of the technique of identifying and diagnosing various types of students' problems and difficulties.

Garrett, Henry E., Statistics in Psychology and Education, Third Edition, New York, Longmans, 1947, xii - 487 p.
Practical applications of basic statistical theory; emphasizes superiority of split-half technique for personality test reliability.

Greene, Edward B., Measurements of Human Behavior, Rev. Ed., New York, Odyssey, 1952, xxiv - 790 p.
An analytical introduction to measuring instruments, involving extensive statistics; abundant Cornell, Rorschach, single-word association and sentence-completion descriptive material.

Guilford, J. P., Fundamental Statistics in Psychology and Education, Second Edition, New York, McGraw-Hill, 1950, xiii - 633 p.

Practical textbook on statistical method; good discussion of testee and environmental variations vitiating retest method of reliability; illustrates Rulon's split-half technique; includes adaptation of Wallace-Snedecor tables.

Hahn, Milton E. and Malcolm S. MacLean, General Clinical Counseling in Educational Institutions, New York, McGraw-Hill, 1950, xi - 375 p.

The principles, techniques and tools of counseling in a good balance of theory and practical applications with many useful concrete examples; word association, sentence completion and Rorschach ink-blots are among the projective devices outlined.

Harrower, Molly, Appraising Personality, New York, W. W. Norton, 1952, xvii - 197 p.

A step-by-step description and study of certain of the objective, rapid and accurate tools and techniques used to appraise personality; presents Rorschach and sentence completion matter in discussion form.

Harrower-Erickson, M. R. and M. E. Steiner, Large Scale Rorschach Techniques, Springfield, Ill., Thomas, 1945, xi - 419 p.

Rorschach procedure adapted for group administration by those without extensive clinical training; "group" and "multiple choice" techniques described.

Klopfer, Bruno and Douglas McGlashan Kelley, The Rorschach Technique, Yonkers, World Book, 1942, x - 475 p.

One of the earliest detailed accounts of the whole field, comprising the history, administration, scoring and interpretation of the test; provides the essence of what is necessary to a preliminary understanding of the technique; the 1946 Supplement is included.

Klopfer, Bruno, Mary D. Ainsworth et al., Developments in the Rorschach Technique, Vol. 1, Yonkers, World Book, 1954, x - 726 p.

A recent extended treatment of Rorschach technique and theory; contains some discussion of the problems relating to short-cut methods, including Harrower's Multiple Choice Test.

Lasagna, Louis and Johm M. von Felsinger, "The Volunteer Subject in Research", in Science, Vol. 120, No. 3114, issue of September 3, 1954, p. 359-361.

Forty-five per cent of 56 college students volunteering as "normal" controls in a drug experiment were found to be suffering from severe psychological maladjustment; selected "extreme group" school subjects expectedly of greater statistical discrimination.

Lehmann, Heinz and Herbert Dorken, jr., "The Clinical Application of the Verdun Projective Battery", in the Canadian Journal of Psychology, Vol. 6, No. 4, issue of December 1952, p. 164-172.

Describes the clinical use of the V.P.B. as a quantitative screening device; gives consideration to some of its limitations; discusses the problem of "false negatives" and rejection failures.

Lehmann, Heinz and Herbert Dorken, jr., "The Administration, Scoring, and Percentile Standardization of the Verdun Projective Battery", in the Canadian Journal of Psychology, Vol. 7, No. 2, issue of June 1953, p. 69-80.

Description of the background work in the preparation of the V.P.B.; absolutely essential to any intelligent and successful usage of the test.

Mursell, James L., Psychological Testing, New York, Longmans, 1949, xvi - 488 p.

Basic theories of mental measurement and their practical application; discussion of word association and Rorschach tests is included.

Paterson, Donald G., Gwendolen G. Schneidler and Edmund G. Williamson, Student Guidance Techniques, New York, McGraw-Hill, 1938, xviii - 316 p.

Diagnostic techniques to assay the personality characteristics, etc., of modern youth; good discussion of students' emotional, educational and other personal problems.

Rapaport, David, Merton Gill and Roy Schafer, Diagnostic Psychological Testing, Vol. 2, Chicago, Year Book Publishers, 1946, xi - 516 p.

The systematic exploration of the diagnostic potentialities of seven psychological tests used as a battery; lengthy treatments are afforded the Word Association and Rorschach tests.

Rosenzweig, Saul and Kate Levine Kogan, Psychodiagnosis, New York, Grune & Stratton, 1949, xii - 380 p.
An introduction to the integration of tests in dynamic clinical practice; illustrative protocols of the Word Association and Rorschach techniques are valuable aids to understanding.

Rotter, Julian B. and Janet E. Rafferty, The Rotter Incomplete Sentences Blank Manual, New York, Psychological Corporation, 1950, iii - 86 p.
Rationale, administration, scoring, normative data and interpretation given; male scoring examples used as patterns.

Strang, Ruth, Counseling Technics in College and Secondary School, Rev. Ed., New York, Harper, 1949, xi - 302 p.
The theory and practice of educational counseling; special attention is given to the problem of interpreting the reliability and validity of the various techniques; useful chapter on the projective method; contains Rorschach and incomplete sentences material.

Super, Donald E., Appraising Vocational Fitness, New York, Harper, 1949, xxi - 727 p.
An extremely comprehensive evaluation of the materials for testing vocational and other capacities; illustrates the use of the Rorschach and incomplete sentences devices with "normal" subjects.

Weider, Arthur, Editor, Contributions Toward Medical Psychology, Vol. 2, New York, Ronald, 1953, xi - (457 to) 885 p.
Comprehensive summary of the salient features of general, child, developmental, personality, social and cultural psychology; considerable material on the word association, Rotter, and group and multiple choice Rorschach tests is provided.

Weider, A., K. Brodman, et al., "The Cornell Service Index: A Method for Quickly Assaying Personality and Psychosomatic Disturbances in Men in the Armed Forces", in War Medicine, Vol. 7, issue of April 1945, p. 209-213.
Description of test administration, scoring, etc.; stresses its ability to screen out those who are really ill and severely disturbed.

Weider, Arthur, Harold G. Wolff, et al., Cornell Index Manual, New York, Psychological Corporation, 1948, 8 p.

Purpose, development, administration, scoring, uses and statistical data provided; cutoff levels utilized; has an excellent bibliography on the test.

Wickmann, E. K., Children's Behavior and Teachers' Attitudes, New York, Commonwealth Fund, 1928, (no pagination), quoted by Laurance Frederic Shaffer, The Psychology of Adjustment, Boston, Houghton Mifflin, 1936, p. 174-176.

The seriousness of children's behavior problems as judged differentially by teachers and clinical psychologists; supplies comparative table with pertinent discussion.

APPENDIX 1

RATING INSTRUCTIONS TO TEACHERS

A. M. D. G.

January 22, 1954

Dear Friend:

Will you help me? An important assignment in my University work is the examination of the worth of the personality test called the Verdun Projective Battery. This involves administering the test to two very distinct groups, one "normal" and the other "abnormal".

Will you help me select two test groups by choosing ten 'guaranteed' "normals" and ten "abnormals" (in your honest opinion) in each of the enclosed class lists. To assist you in your selection, I include a list of items usually accepted as being indicative of evidence of some degree of "abnormality". As you can readily understand, unless the groups are really representative, my work is hopeless, - therefore, I earnestly ask you to think your choices over carefully.

One last word, - the entire procedure is in the strictest confidence (for example, your listings will never be made public), so I ask your cooperation in preventing any word of the true nature of the test leaking out to the pupils or, above all, their parents.

Thank you most sincerely for your generous assistance in this matter of prime importance in my present degree studies.

Yours truly,

P.S. This project has the official approval of the M.C.S.C. and Bro. Ambrose.

SIGNS OF ABNORMALITY

1) "NERVOUS" CHILDREN (including speech defects, tics, nail-biters).

2) BEHAVIOR PROBLEMS (discipline, truancy, lying, cheating, bullying, cruelty, rebellion, over-critical attitude, "broken homes").

3) EMOTIONAL IMMATURITY (marked excitability, childishness, suggestibility; or below-average intelligence).

4) RETIRING GROUP (shy, timid, seclusive, day-dreaming, suspicious, depressed, fearful, over-sensitive, resentful, discouraged, withdrawn).

5) NON-CONFORMISTS (take no voluntary part in normal extra-curricular activities of a group, social or "school spirit" nature).

N.B. An "N" on your lists will denote "normal" and an "A" "abnormal", - the numbers above are merely for your convenience and do not need to appear as such on the class lists.

AJW:ID

L. D. S.

APPENDIX 2

VERDUN PROJECTIVE BATTERY

A. M. D. G.

LAST NAME: FIRST NAME: AGE: CLASS:

TEST B: MULTIPLE CHOICE

NOTE: In each set of ten choices, draw a line THROUGH the ONE you choose.

No. 1		No. 2		No.-3	
An army or navy emblem (crest)	3	A bug somebody stepped on	8	Two birds fighting	2
Crumbling cliffs	9	Nothing at all	10	Meat in a butcher's shop	8
A bat	2	Two Scottie dogs	2	Two men pulling	1
Nothing at all	10	Little faces on the sides	4	Part of my body	6
Two people	1	A bloody spinal column	6	Just colored blots	9
A pelvis (hip bones)	3	A white top (toy)	5	A colored butterfly	3
An X-ray picture	7	A bursting bomb	8	Spots of blood and paint	8
Pincers of a crab	4	Two elephants	2	Monkeys hanging	2
A dirty mess	9	Two clowns	1	A red bow tie	3
Part of my body	6	Red and black ink	9	Nothing at all	10

No. 4		No.-5		No. 6	
Head of an animal	4	A bird's beak	4	Two kings' heads	3
Lungs and chest	6	Something squashed	9	An X-ray picture	7
A nasty dirty mess	9	A ballet dancer	1	Parts of the body	6
A pair of boots	4	Nothing at all	10	A totem pole	3
A burnt substance	9	A map	7	A fur rug	3
Nothing at all	10	Sugar tongs	4	Mud and water	9
A giant in a fur coat	1	A moth	2	A polished post	4
An animal skin	3	Shoulders	6	Nothing at all	10
A big gorilla	2	Smoke	7	A turtle	2
An X-ray picture	7	A rabbit's head	4	A landslide	9

No.-7		No.-8		No. 9	
Smoke	7	An orange or pink butterfly	3	Sea horses	3
Two women talking	1	Shoulders, lungs and stomach	6	Just spilt paint	9
Parts of the body	6	Nothing at all	10	Flowers	3
Animals	2	Just colors	9	Parts of the body	6
Nothing at all	10	An emblem (crest)	3	Smoke and flames	8
A white chandelier (lights)	5	A pretty flower	3	Deer or horns of deer	2
Burning fragments	9	Heaven and Hell	8	Nothing at all	10
Lambs' tails	4	Two blue cushions	3	Two witches	1
An X-ray picture	7	Two bears climbing	2	Bloody clouds	8
Bookends	3	Colored clouds	7	A candle	3

No.-10	
Two people	1
Spilt paint	8
A Chinese print (design)	3
An X-ray picture	7
Just colored ink spots	9
Spiders, caterpillars, insects	2
Parts of my insides	6
A colored map of California	3
Nothing at all	10
A flower garden	3

TEST C: CONFORMITY INDEX

MAN	LAMP
CABBAGE	LONG
HARD	TOBACCO
EAGLE	SCISSORS
OCEAN	BLOSSOM

TEST A: WORD ASSOCIATION

WORD:	TIME:	RESPONSE:	REMARKS:
TABLE *			
RIVER *			
MAN *			
FOOD			
PENCIL *			
VOICE			
ARMY *			
SICKNESS			
COLOR			
MONEY			

TOTAL: W.A. M.C. C.I. FULL NAME: N. A.

L. D. S.

APPENDIX 3

CORNELL INDEX - FORM N2

C. I.—FORM N2

Name_____ Age_____ Today's Date_____
(Last) (First) (Middle)

Home Address_____ Are You Married?_____
(Street or RFD) (City) (State)

Occupation_____ Last School Grade Reached_____

Directions: Put a circle around **(YES)** if you can answer YES to the question asked. Put a circle around **(NO)** if you have to answer NO to the question asked. Answer all questions. If you are not sure guess.

1. Have you ever had a headache? Yes No
2. Do you frequently feel faint? Yes No
3. Do you have hot or cold spells? Yes No
4. Have you fainted more than twice in your life? Yes No
5. Do strange people or places make you afraid? Yes No
6. Do you often have spells of dizziness? Yes No
7. Do you get all nervous and shaky when approached by a superior? Yes No
8. Does the sight of blood make you want to drop down in a faint? Yes No
9. Does your work fall to pieces when the boss or a superior is watching you? Yes No
10. Are you scared to be alone with no friends near you? Yes No
11. Do you feel nervous or dizzy right at this moment? Yes No
12. Do you always get orders and directions wrong? Yes No
13. Does your thinking become completely confused when you have to do things quickly? Yes No
14. Do you always sweat and tremble a lot during inspections or examinations? Yes No
15. Do you wish that you always had someone at your side to advise you? Yes No
16. Do you have to do things very slowly in order to be sure you are doing them right? Yes No
17. Does it bother you to eat anywhere except in your home? Yes No
18. Do you have an uncontrollable need to repeat the same disturbing actions? Yes No
19. Is it always difficult for you to make up your mind? Yes No
20. Do you usually feel cheerful and happy? Yes No
21. Do you always have a bad time no matter what you are doing? Yes No
22. Do you often feel miserable and blue? Yes No
23. Does life usually look entirely hopeless? Yes No
24. Are your emotions usually dead? Yes No
25. Are you usually quiet and sad while at a party? Yes No
26. Do you often wish you were dead and away from it all? Yes No
27. Are you considered a nervous person? Yes No
28. Do you have any unusual fears? Yes No
29. Do you often have difficulty in falling asleep or staying asleep? Yes No
30. Does every little thing get on your nerves and wear you out? Yes No
31. Does worrying continually get you down? Yes No
32. Did you ever have a nervous breakdown? Yes No
33. Were you ever a patient in a *mental* hospital? Yes No
34. Do you get out of breath long before anyone else? Yes No
35. Do you have pains in the heart or chest? Yes No
36. Does your heart often race like mad for no good reason? Yes No
37. Do you often have difficulty in breathing? Yes No
38. Are you often bothered by thumping of the heart? Yes No
39. Do you often suddenly become frightened while you are thinking? Yes No
40. Do you often shake or tremble? Yes No

(OVER)

Published by The Psychological Corporation, New York, N. Y.

Printed in U.S

53-143T

41. Are you often awakened out of your sleep by frightening dreams? Yes No

42. Do you always become scared at sudden movements or noises at night? Yes No

43. Do sudden noises make you jump and shake badly? Yes No

44 Do you tremble or feel weak every time some one shouts at you? Yes No

45. Are you keyed up and jittery every single moment? Yes No

46 Do you have very disturbing or frightening thoughts that keep coming back in your mind? Yes No

47. Do you suffer badly from frequent severe headaches? Yes No

48. Do you sweat a great deal even in cold weather? Yes No

49. Are you repeatedly bothered by severe itching? Yes No

50. Are you troubled by stuttering? Yes No

51 Have you at times had a twitching of the face, head or shoulders? Yes No

52. Were you a bed wetter between the ages of 8 to 14 years? Yes No

53. Do cold hands or feet trouble you even in hot weather? Yes No

54. Do you suffer from asthma? Yes No

55. Are you a bed wetter? Yes No

56. Are you a sleep walker? Yes No

57. Have you ever had a fit or convulsion? Yes No

58. Do pains in the back make it hard for you to keep up with your work? Yes No

59 Do you sometimes find yourself unable to use your eyes because of pain? Yes No

60. Is your body always in very bad condition? Yes No

61. Do severe pains and aches make it impossible for you to perform your duties? Yes No

62. Do you get spells of exhaustion or fatigue? Yes No

63. Do you wear yourself out with worrying about your health? Yes No

64. Do weak or painful feet make you miserable every single day? Yes No

65. Do you frequently get up tired in the morning? Yes No

66. Does pressure or pain in the head make it hard for you to perform your duties? Yes No

67. Are you always in poor health and unhappy? Yes No

68. Are you constantly too tired and exhausted even to eat? Yes No

69 Is your appetite good? Yes No

70 Do you constantly suffer from bad constipation? Yes No

71. Do you often suffer from an upset stomach? Yes No

72. Do you frequently get attacks of nausea (sick to your stomach)? Yes No

73. Do you suffer from indigestion? Yes No

74. Do you always have stomach trouble? Yes No

75. Do your stomach and intestines work badly? Yes No

76. Do bad pains in the stomach double you up after every meal? Yes No

77. Do you usually have trouble in digesting food? Yes No

78. Do you suffer badly from frequent loose bowel movements? Yes No

79. Has any doctor ever told you that you had ulcers of the stomach? Yes No

80. Do people usually misunderstand you? Yes No

81. Do you have the feeling of being watched while you are at work? Yes No

82. Have you usually been treated fairly? Yes No

83. Do you have the feeling that people are watching or talking about you in the street? Yes No

84. Do people usually pick on you? Yes No

85. Are you extremely shy or sensitive? Yes No

86. Are you easily upset or irritated? Yes No

87. Do you make friends easily? Yes No

88. Do you go all to pieces if you don't constantly control yourself? Yes No

89. Were you ever sent to reform school? Yes No

90 Have you ever gotten into serious trouble or lost your job because of drinking? Yes No

91. Have you been arrested more than three times? Yes No

92. Have you ever taken dope regularly (like morphine or "reefers")? Yes No

93. Do your enemies go to great lengths to annoy you? Yes No

94 Does it make you angry to have anyone tell you what to do? Yes No

95 Do you often drown your sorrows in drink? Yes No

96. Do you always do things on sudden impulse? Yes No

97. Do people always lie to you? Yes No

98 Do you flare up in anger if you cannot have the things that you want right away? Yes No

99. Is the opposite sex unpleasant to you? Yes No

100. Do you always have to be on your guard with friends? Yes No

101 Do you often get into a violent rage? Yes No

APPENDIX 4

ROTTER INCOMPLETE SENTENCES BLANK - HIGH SCHOOL FORM

INCOMPLETE SENTENCES BLANK — HIGH SCHOOL FORM

Name _____ Sex _____ Age _____

School _____ Grade _____ Date _____

Complete these sentences to express your real feelings. Try to do every one.
Be sure to make a complete sentence.

1. I like _____
2. The happiest time _____
3. I want to know _____
4. At home _____
5. I regret _____
6. At bedtime _____
7. Boys _____
8. The best _____
9. What annoys me _____
10. People _____
11. A mother _____
12. I feel _____
13. My greatest fear _____
14. In the lower grades _____
15. I can't _____

(TURN PAGE OVER AND CONTINUE)

Copyright 1950
The Psychological Corporation
522 Fifth Avenue
New York 36 N Y

Printed in U S A 50 110 AS

JULIAN B ROTTER, AUTHOR

16. Sports

17. When I was younger

18. My nerves

19. Other kids

20. I suffer

21. I failed

22. Reading

23. My mind

24. The future

25. I need

26. Dating

27. I am best when

28. Sometimes

29. What pains me

30. I hate

31. At school

32. I am very

33. The only trouble

34. I wish

35. My father

36. I secretly

37. I

38. Dancing

39. My greatest worry is

40. Most girls

APPENDIX 5

ABSTRACT OF

THE EFFECTIVENESS OF THE VERDUN PROJECTIVE BATTERY AS A PERSONALITY SCREENING DEVICE AT THE HIGH SCHOOL LEVEL

APPENDIX 5

ABSTRACT OF

The Effectiveness of the Verdun Projective Battery As a Personality Screening Device at the High School Level.[1]

The discovery of a simple, rapid and accurate personality device that even in non-clinical hands would dependably screen those students with serious psychopathic deviation, potential or actual, from the ones with more minor and less acute disorder would prove a tremendous boon in the present educational setting. It was the major aim of this research project to attempt to ascertain if the new and as yet commercially unavailable Verdun Projective Battery, which purports to be a rapid but accurate means of screening for potential neuro-psychiatric disturbance and maladjustment, is just such a discriminating personality instrument for high school students. In other words, while the ten-year background of the V.P.B. has been chiefly concerned with the differentiation of normal adults from psychotic patients in a large mental hospital, the problem here was to test its applicability in a typical school environment. Ancillary aims of the experiment were the investigation of the influence of type of administration (group versus individual) procedure employed, an estimation of the relationship of scholastic standing and academic proficiency to test and teacher personality rating, and finally, some brief study and

---

1 Ph.D. Thesis presented by Arthur James Welbourne, in 1955, through the Institute of Psychology, to the Faculty of Arts of the University of Ottawa, 116 pages.

examination of the degree of similarity and interrelationship in the results obtained by the projective and non-projective testing approaches to the whole problem of personality discrimination and differentiation.

The design of the experiment was drawn up according to this pattern. The pupils of The Thomas D'Arcy McGee Boys' High School (Montreal, Que.) were classified by their teachers into the two categories designated, for ease of comprehension by non-clinicians, as "normal" and "abnormal", respectively. Three personality screening devices, the Verdun Projective Battery (three subtests), Form N2 of the Cornell Index (personality questionnaire) and the High School Form of the Rotter Incomplete Sentences Blank (partially structured projective test), were subsequently administered to these "extreme" groups. The statistical techniques of the "t" ratio, chi square analysis and correlation procedure were then applied to the raw data gathered to test and interpret the significance of the validity and reliability results obtained by means of these measuring instruments.

The Verdun Projective Battery, which takes approximately fifteen minutes for individual administration, comprises in its present form three subtests: 1) a ten-word Word Association Test, which is timed with a stop-watch; 2) a ten-word untimed word-association test (from the Kent-Rosanoff tables) described as a Conformity Index; and 3) series A of Harrower's Multiple

Choice Rorschach (with an amended and revised scoring procedure). It must be stressed and emphasized that this test is designed for screening and not diagnostic purposes. Form N2 of the Cornell Index is of the personality inventory type in questionnaire/form and is also designed for use in psychiatric screening. It consists of 101 items to be answered "Yes" or "No" and requires on the average about ten minutes to complete. The third screening instrument, the High School Form of the Rotter Incomplete Sentences Blank, averages twenty minutes for administration and is made up of forty sentence-beginnings or "stems" which the subject is requested to complete in a manner that expresses his real and true feelings.

The statistical analysis revealed certain definite trends. Both "t" ratio and chi square results showed the Word Association subtest, (3.71 and 10.04, respectively), the total V.P.B. (2.01 and 8.65) and the Rotter test (2.86 and 6.60) to be significant (in most cases at the 0.01 level of confidence) in distinguishing between the "normal" and "abnormal" teacher-selected student categories. On the contrary, however, the C.I. and M.C.R. subtests and the Cornell test proved insignificant in this setting. This may possibly be due to absence of severe psychosis in the first case, teacher-inability to rate pupils personality-wise in the second, and the fact that the third test was designed for use with a slightly older group. The V.P.B. subtests correlate slightly negatively with each other (the

only exception is the "normal" M.C.R. - C.I. .10) and hence are independent measures, but significantly with the total test, for both the "normal" and "abnormal" groups. The only significant correlation is the Rotter-Cornell .51 for the "normals".

The degree of reliability (.88 for the "normals" and .77 for the "abnormals") found by Rulon's split-half method of estimation is likewise very satisfactory. The half splitting was obtained by dividing the thirty items of the V.P.B. into two matched groups of fifteen items each, five from each subtest, on the basis of equivalent total "failure" percentage. This procedure was repeated for the "normal" and "abnormal" school groups independently. The results on the more customary odd-even grouping were practically identical, namely, .83 for the "normal" group and .77 for the "abnormal" one. Incidentally, all figures are probably underestimates because of the fact that the intercorrelations are low and the items are definitely not of equal difficulty.

The conclusions for the associated phases of the investigation may be summarized as follows: 1) The V.P.B. results of individually-tested youngsters (in the same age-range as the students) at a summer-camp were obtained from the hospital files and compared with the group-derived school data, - the M.C.R. subtest("normals") alone showed a reaction to administration variation, possibly for the reason specified above. It is of interest to note, in addition, that the summer-camp

testees (25 boys and 17 girls) showed insignificant sex differences in their scores and were therefore lumped together into one group for the comparisons. 2) All the subjects' academic results (June 1954), together with their final classification (pass or fail) and "recommended" status (that is, promotion by excellence of the year's record and exemption from written examinations), were taken from the school archives and compared to the test results. The pupil scholastic standing and teacher personality rating show a certain degree of relationship and confirm the hypothesis that teachers will inevitably rate pupils according to their own professional lights (and not, for example, as psychologists). However, this does not vitiate their use of the V.P.B., fortunately, as their accuracy on the "mental control" oriented Word Association test preserves the integrity of the rating. Incidentally, the apparent scholastic predictive value of the W.A. subtest seems well worth further investigation. 3) The non-projective Cornell Index scores proved to bear no clear relationship to the data obtained from the other sources, as explained above, but all the projective results seem to be roughly similar and comparable.

In short, on the basis of these and other considerations enumerated and described in the main body of the thesis, it would seem permissible to believe that the Verdun Projective Battery has a definite contribution to make as an effective personality screening and evaluation device at the high school level.