MIHALKA, Joseph Alexander, 1918-
SEQUENTIAL TESTING - A SYSTEMS APPROACH.

The Ohio State University, Ph.D., 1970
Education, guidance and counseling

University Microfilms, A XEROX Company, Ann Arbor, Michigan
SEQUENTIAL TESTING - A SYSTEMS APPROACH

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Joseph Alexander Mihalka, B.A. M.A.

* * * * * *

The Ohio State University

1970

Approved by

[Signature]

Advisor

College of Education
The writer wishes to express his sincere appreciation to Dr. Herman J. Peters for his encouragement and guidance during this study.

A word of thanks is also extended to those who made this study possible through assistance in the gathering and processing of the data.
VITA

June 8, 1918 . . . . . . . Born - Shelby, Ohio

1942-1945. . . . . . . Ammunition Officer (Captain),
1948-1952 U. S. Army

1947 . . . . . . . . . . . B.A., The Ohio State University,
Columbus, Ohio

1948 . . . . . . . . . . Field Executive, Boy Scouts of America,
Parkersburg, West Virginia

1952-1954. . . . . . . Quality Control Supervisor,
Kilgore Mfg. Corp., Westerville, Ohio

1954-1958. . . . . . . Group Leader in Drafting Section
Diamond Power Specialty Co., Lancaster,
Ohio

1958-1970. . . . . . . Chief of Testing,
The Ohio Bureau of Employment Services,
Columbus, Ohio

1966 . . . . . . . . . . M.A., The Ohio State University,
Columbus, Ohio

1967 . . . . . . . . . . Consultant to The Pakistan Manpower
Council, International Labour Office (U.N)
Karachi, Pakistan

PUBLICATIONS

"Ohio 'Sells' the E.S. Through Promotional Meetings." Employment

"Ohio Serves the Rubber Industry." Employment Security Review,

GATR, Dial-A-Score (Conversion Device), Philadelphia; Specialty
VITA - continued

FIELDS OF STUDY

Major Field: Education (Guidance)

Studies in Vocational Guidance and Counseling. Professor Herman J. Peters

Studies in Vocational Education. Professor Robert M. Reese

Studies in Adult Education. Professor John F. Ohliger
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>ii</td>
</tr>
<tr>
<td>VITA</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>viii</td>
</tr>
<tr>
<td>Chapter</td>
<td></td>
</tr>
<tr>
<td>I.   INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>The Problem</td>
<td></td>
</tr>
<tr>
<td>Limitations of the Study</td>
<td></td>
</tr>
<tr>
<td>Definition of Terms</td>
<td></td>
</tr>
<tr>
<td>Organization of the Dissertation</td>
<td></td>
</tr>
<tr>
<td>II.  REVIEW OF THE LITERATURE</td>
<td>14</td>
</tr>
<tr>
<td>Sequential Testing</td>
<td></td>
</tr>
<tr>
<td>Pre-test Orientation</td>
<td></td>
</tr>
<tr>
<td>Test-taking Behavior</td>
<td></td>
</tr>
<tr>
<td>The Systems Approach</td>
<td></td>
</tr>
<tr>
<td>Summary of the Chapter</td>
<td></td>
</tr>
<tr>
<td>III. DEVELOPMENT OF A MODEL</td>
<td>40</td>
</tr>
<tr>
<td>The Setting</td>
<td></td>
</tr>
<tr>
<td>Employment Service Tests</td>
<td></td>
</tr>
<tr>
<td>Testing Procedures</td>
<td></td>
</tr>
<tr>
<td>Test Planning Worksheet</td>
<td></td>
</tr>
<tr>
<td>Test Results Worksheet</td>
<td></td>
</tr>
<tr>
<td>Test Behavior Checklist</td>
<td></td>
</tr>
<tr>
<td>Summary of the Chapter</td>
<td></td>
</tr>
<tr>
<td>IV.  EXPERIMENTAL AND DEMONSTRATION PROJECT</td>
<td>66</td>
</tr>
<tr>
<td>Experimental Procedures</td>
<td></td>
</tr>
<tr>
<td>Discussion of Findings</td>
<td></td>
</tr>
<tr>
<td>Summary</td>
<td></td>
</tr>
<tr>
<td>V.   SUMMARY AND CONCLUSIONS</td>
<td>90</td>
</tr>
<tr>
<td>Findings</td>
<td></td>
</tr>
<tr>
<td>Recommendations</td>
<td></td>
</tr>
<tr>
<td>Need for Further Research</td>
<td></td>
</tr>
<tr>
<td>Conclusion</td>
<td></td>
</tr>
</tbody>
</table>
# APPENDIX

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Glossary of Terms</td>
<td>100</td>
</tr>
<tr>
<td>B. Annotated List of Employment Service Tests</td>
<td>113</td>
</tr>
<tr>
<td>C. Forms</td>
<td>117</td>
</tr>
<tr>
<td>D. Specimen Kit: Instructions</td>
<td>129</td>
</tr>
<tr>
<td>E. Statistical Procedures</td>
<td>139</td>
</tr>
</tbody>
</table>

# REFERENCES

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIBLIOGRAPHY</td>
<td>151</td>
</tr>
</tbody>
</table>
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Categories and Use of Employment Service Tests</td>
<td>44</td>
</tr>
<tr>
<td>2. Staff Participation and Data Contribution</td>
<td>75</td>
</tr>
<tr>
<td>3. Sample Characteristics</td>
<td>77</td>
</tr>
<tr>
<td>4. Past Test Experience</td>
<td>80</td>
</tr>
<tr>
<td>5. Total Values by Categories</td>
<td>81</td>
</tr>
<tr>
<td>6. Number of Tests Administered and Appropriateness</td>
<td>83</td>
</tr>
<tr>
<td>7. Total Scores – Test Behavior Checklist</td>
<td>85</td>
</tr>
<tr>
<td>8. Total Mean Scores – Evaluation Sheet</td>
<td>86</td>
</tr>
<tr>
<td>9. Model Evaluation by Counselors and Test Examiners</td>
<td>88</td>
</tr>
</tbody>
</table>
LIST OF ILLUSTRATIONS

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>A Basic Systems Model</td>
<td>32</td>
</tr>
<tr>
<td>2.</td>
<td>A Model for Decision Making</td>
<td>33</td>
</tr>
<tr>
<td>3.</td>
<td>Subsystems and Feedback</td>
<td>35</td>
</tr>
<tr>
<td>4.</td>
<td>Model of a Testing, Counseling and Learning System.</td>
<td>36</td>
</tr>
<tr>
<td>5.</td>
<td>Model of Testing Process</td>
<td>47</td>
</tr>
<tr>
<td>6.</td>
<td>Direct Referral to Job or Training</td>
<td>48</td>
</tr>
<tr>
<td>7.</td>
<td>Processes and Services in the WIN Program</td>
<td>50</td>
</tr>
<tr>
<td>8.</td>
<td>Test Sequence for Negative Profile</td>
<td>55</td>
</tr>
<tr>
<td>9.</td>
<td>Test Sequence for Questionable Profile</td>
<td>57</td>
</tr>
<tr>
<td>10.</td>
<td>Test Sequence for Positive Profile</td>
<td>59</td>
</tr>
<tr>
<td>11.</td>
<td>Sequence for Clerical Skills Tests</td>
<td>60</td>
</tr>
<tr>
<td>12.</td>
<td>Scale Values on Test Planning Worksheet</td>
<td>76</td>
</tr>
<tr>
<td>13.</td>
<td>Scale Values on Past Test Experience</td>
<td>79</td>
</tr>
<tr>
<td>14.</td>
<td>Category and Scale Values</td>
<td>80</td>
</tr>
<tr>
<td>15.</td>
<td>Scale for Model Evaluation</td>
<td>87</td>
</tr>
</tbody>
</table>
CHAPTER I
INTRODUCTION

Although tests have been widely used for over half a century, it was only in the past decade that there has been critical examination of the area of tests and measurements. Growth in the development and use of tests began during World War I, primarily for classification and selection in the military, education and training, and in employment. Later, tests were used for clinical and guidance activities.

Developments in the early period of tests were mainly directed to (1) innovations in the kinds of tests made available, (2) expansion of their use in different settings, and (3) technical details such as validation, reliability, errors in measurement, sampling procedures, and criterion data. Evaluation of tests was frequently made by personnel in university or clinical settings. The publication of the first edition of Buros’s *Mental Measurement Yearbook* in 1938 marked the first compilation and reviews of standardized tests. Individual and commercial test publication expanded into a lucrative business that catered to schools, industry, government, and agencies.

With the phenomenal growth in the use and application of tests came two very important innovations, the separate answer sheet and machine scoring. These reduced the cost of testing through reusable
test booklets and more economical scoring procedures by key or ma-
chine. This was almost a matter of necessity due to the complexity
of tests and the increased volume of testing. Test publishers were
quick to recognize that the economy of the separate answer sheets and
scoring services would mean a new and different source of profit with
the expanded use of tests.

An important technological change in test scoring and develop-
ment was brought about by the computer. Not only was the machine able
to score the test, but it could also perform additional functions such
as converting scores, printing out interpretive data, storing up in-
formation, and doing statistical work related to research or standard-
ization studies.

Further innovations in testing included the use of tape re-
corded instructions and computer-assisted test administration such as
The Psychological Corporation's (1969) Controlled Administration of
Standardized Tests (CAST). Test information has also become a part of
several computer-assisted counseling programs. There is every in-
dication that testing and counseling will undergo many changes as a
result of technology and automation.

This study will attempt to utilize some of the basic concepts
of data processing, programming, planning, and decision-making but not
with direct application for automation or computer use. Presently, the
availability of computers is limited because of cost and requirements
for specially trained operating and programming personnel. Thus, their
use may not be feasible or economical. However, there are indications
that wider use of machine and data processing will not only be
practical but may become a matter of necessity. Development of desktop computers of relatively low cost and also systems built into large and more sophisticated equipment may result in greater adaptability.

Many improvements in tests and testing procedures failed to cover the aspects of proper selection, improvement in test administration, and adequate interpretation of results. Thus, the late 50's and 60's brought serious indictments against testing for inadequacies of some tests, invasion of privacy, discrimination in fair employment and civil rights, and a general overuse and misuse of tests. Numerous books, journal articles, and a Congressional hearing attacked testing. The furor continued in 1970.

William Whyte's book (1958), The Organization Man, was critical of personality testing, as was Gross's book, The Brainwatchers (1962). I.Q. or intelligence testing was attacked in several books, including Black's They Shall Not Pass (1963) and Hoffman's Tyranny of Testing (1962).

With the advancement of Civil Rights and Fair-Employment, Guidelines for Testing (1966) were established by the Equal Employment Opportunities Commission. The focus of attention was on the discriminatory aspects of testing. Numerous articles and books dealt with the inappropriateness of many tests being used with minority groups and the culturally or educationally disadvantaged. Culhane (1965) indicated that most standardized paper-and-pencil tests were not appropriate for persons lacking test-taking experience or who were functioning at low educational levels. Similarly, Guion (1966) wrote extensively on the use of employment tests and discriminatory
hiring. Hamburger (1964) reviewed measurement issues in the counseling of the culturally disadvantaged from the standpoint of the technological, sociological, moral, and ethical responsibilities of the counselor.

Currently, suggestions have been made to introduce moratoriums on the use of standardized tests with the disadvantaged and minority groups. Buros (1970), at a meeting of the Association for Measurement and Evaluation in Guidance, suggested such action for a period of several years, until research will have been done for the development of fairer or more appropriate tests and testing procedures. At about the same time, a resolution was introduced and passed in the 1970 American and Personnel and Guidance Association Senate meeting that a five to ten year moratorium be encouraged regarding the use of standardized tests (particularly intelligence) with the disadvantaged. The modified version of this resolution is in the nature of a position paper on the subject with recommendations for research and practice.

However, tests can be an important source of information in counseling, vocational planning, various types of selection, and employment. An objective of this study was to improve methods and procedures that would lead to better planning and administration of tests. Using some of the concepts of programming, sequential testing can be adapted to meet individual or group needs. Sequence has reference to a logical step-by-step progression in the planning and preparation for testing and the systems approach is closely related to this idea.

Essentially, the purpose of this study was to develop an
experimental model that would be used in a demonstration or pilot project. The model included a systematic method for test selection, sequence of test administration, and the development of forms for the planning, and reporting of test results and test-taking behavior. The systems approach was used in order to provide a more logical order of the tests and to avoid needless and inappropriate testing, as well as to give more valid, reliable, and meaningful results.

**THE PROBLEM**

Considerable attention has been given to the technical aspects of test development and standardized test directions to assure valid results. However, neglected areas have been the selection of appropriate tests, test readiness involving preparation or orientation, the planning for and sequence of tests, and some structured way of observing and recording test-taking behavior.

All too frequently, the selection of tests and referral to testing is a haphazard process. Tests are often given in batteries over a relatively short period of time and, in general, there is lack of planning or logical sequence. These procedures were established in the interest of economy and also to get the counselee into training or placed on a job as quickly as possible.

Similarly, preparation for testing has been limited to a few brief introductory remarks that are a part of the standardized instructions. The interviewer or counselor rarely taken time to determine the readiness of the counselee for testing nor does he provide any special help for those who lack test-taking experience. Understanding of the
reason for tests and attitude and motivation for testing are neglected too because of the time factor or because there are no specific procedures provided.

Feedback of test results usually consists only of test scores or profiles. Comments on test-taking behavior are rarely noted, the reasons being the pressure of time and the fact that most tests are administered in groups, precluding observation and recording of individual performance.

The U.S. Employment Service interviewing, counseling, and testing guidelines do allow for the establishment of procedures for test selection, preparation for testing, and methods for reporting test-taking behavior. They have not been spelled out in detail, however, nor have they been developed into any system. It was with reference to these problems that this study was undertaken.

The steps involved in testing include the gathering of information, identification of significant items, and their integration into the total counseling process. Innovative procedures and use of electronic devices may offer promise in the future for new methods in helping to solve complex test-related problems. However, the processing of data does not necessarily have to be done by a computer, which after all, can handle only information fed into it. The same information can often be treated in similar manner without the use of a machine provided there is adequate planning, preparation, and an orderly use of the data.
As a basis for this study, the following assumptions were made:

1. that test results constitute an essential source of data for the counselor and Employment Service interviewer,

2. that certain biographical and background information is necessary in making decisions with regard to the selection and administration of tests, their order, when, and how,

3. that certain segments of the population, such as the disadvantaged, lack test-taking experience and environmental exposure which influence their performance on tests and, in turn, depress test scores,

4. that test-taking can be a learned skill and that instruction and practice in the use of the separate answer sheet, development of understanding of the reasons for tests, and learning how to utilize time, tend to reduce anxiety and tension as well as establishing attitudes and motivation for maximum performance, and,

5. that distributed instruction and carefully structured preparation, including a progressive sequence of practice exercises, will eliminate needless or inappropriate testing.

Sequential testing, as used in this study, was primarily concerned with improved procedures in the proper selection and
administration of tests. Special forms that are described in detail in subsequent chapters were developed to serve as aids in this process. In light of these objectives, the following hypotheses were examined:

1. If the Test Planning and Results Worksheets are used in counseling and testing, then test selection and administration will be more appropriate.

2. If sequential testing procedures are used with disadvantaged persons, then their test-taking behavior will improve.

The significance of this study parallels the attainment of its objectives - more meaningful and appropriate testing and maximum performance as a result of improved test-taking behavior. To achieve these goals, the study-model incorporated innovative procedures and special forms which were tried experimentally with sample groups of disadvantaged persons. The benefits derived would be the assurance for the counselor and counselee that tests were carefully selected and administered in accordance with the needs and abilities of the counselee.

The primary problem dealt with in this study was the development of better procedures for testing. The model included methods for the selection of tests in accordance with individual or group needs, adequate preparation for testing, and an orderly sequence of tests. The study also gave special attention to test-taking behavior.
LIMITATIONS OF THE STUDY

The scope of this project was limited to the tests and programs in use by the Ohio Bureau of Employment Services, which cooperates with the U.S. Department of Labor and its Manpower Administration. Among the various programs are Manpower Training, Human Resource Development, Economic Opportunities, Work Incentive Programs (WIN), Neighborhood Youth Corps, New Careers, and National Alliance of Businessmen – Job Opportunities in the Business Sector (NAB-JOBS). All of these deal with problems of employability. Vital aspects of the services embrace counseling, assessment, orientation, job development, and referral to work or training.

The model developed here was experimentally tried with WIN (Work Incentive Program) enrollees. This special program is conducted in cooperation with local welfare agencies in select communities. Persons who qualify and participate receive a $30 per month incentive payment in addition to their welfare checks. Activities include orientation, assessment (in part including testing), counseling, work or skill training, basic education, and various other things that can lead to employability development or to direct referral to training or employment.

The sampling groups were limited to WIN enrollees. The experimental and demonstration project was comprised of persons participating in Orientation and Assessment. This program component consists of a 3-week structured training to prepare the enrollees for employment. Subject matter includes grooming, money management,
job-seeking strategy, transportation, communications, pre-test orientation, standardized testing, and work sampling. The last three items are a part of the assessment. The control group of this study did not participate in Orientation and Assessment but were processed for referral to training or employment according to traditional methods.

The tests and testing procedures used were limited to those developed or approved by the Employment Service. An annotated list of tests appears in appendix B. Included among the tests are those that are diagnostic (achievement and screening) to determine the person's educational level of functioning. Non-reading tests such as the Culture-Fair and non-reading parts of the General Aptitude Test Battery are also used. Special pre-test orientation practice exercises are used to develop test-taking skills and as a method of observing and diagnosing test-taking behavior.

The time span of the experimental and demonstration project was limited to a six week period. The selection of participating employment offices was determined by the fact that they were conducting Orientation and Assessment programs at this time. As a result, the model was tried in five locations, namely, Toledo, Akron, Canton, Youngstown, and Cincinnati.

In summary, the limitations of the study were: (1) the setting - The Ohio Bureau of Employment Service offices with Work Incentive Programs, (2) the sampling - enrollees in the Orientation and Assessment component (experimental) and enrollees referred directly to training or employment (control), (3) tests and testing
procedures — those developed or approved by the U.S. Department of Labor, (4) time — a six week period (March 9 to April 17, 1970). These limits for the experimental and demonstration phase of the project were established to serve as a pilot program to try out the procedures, forms, and to evaluate the effectiveness of sequential testing and the systems approach.

DEFINITION OF TERMS

A complete glossary of terms is contained in Appendix A. The following partial list of definitions is presented here for clarification and ready reference:

(1) Feedback — the process of returning a fraction of the output as an input to a previous step or stage in the system.

(2) Input — information or data that is fed into the process.

(3) Model — a set of plans, a miniature representation, structure or design, example for imitation, or pattern that is a basis for future production or activity.

(4) Orientation and Assessment — a two to three week structured instructional component of the Work Incentive Program designed to prepare enrollees for referral to training or employment.

(5) Output — a total product, yield, or end result of a process over a given period of time.
(6) **Pre-test orientation** - an instructional unit of Orientation and Assessment designed to provide learning experiences in test-taking skills, to promote understanding of reasons for tests, and to develop proper attitudes and motivation to assure maximum performance on standardized tests.

(7) **Process** - a series of actions or operations definitely leading to an end or goal. A continuity, activity, or progression in time.

(8) **Sequential testing** - the planning and administration of tests over a period of time in such an order that each step is naturally derived from the preceding.

(9) **Subsystem** - a further division of a system that is broken down into elemental steps which feed into the main system.

(10) **Systems approach** - a logical step-by-step approach to planning or problem solving. A system is an assemblage of objects (units or steps) joined in regular interaction or interdependence. Anytime one or more steps are involved in reaching a goal, a system exists.

(11) **Test readiness** - the extent to which an individual has achieved a degree of maturity or has acquired certain skills and information needed to perform successfully on tests.
(12) **Test-taking behavior** - visible and observable responses and performance on tests and in testing situations.

**ORGANIZATION OF THE DISSERTATION**

Chapter I is the introduction and includes a brief background of the study, statement of the problem, limitation of the study, definition of terms, and organization of the dissertation.

Chapter II contains a review of the literature on sequential testing, pre-test orientation, test-taking behavior, and the systems approach. Sources were texts, journals, monographs, and test manuals.

Chapter III describes the development of the model for sequential testing including the various forms and procedures that were used in planning for the sequence of tests, administration, pre-testing orientation, and recording of test-taking behavior.

Chapter IV reports on the experimental and demonstration project and a discussion of the findings.

Chapter V is a summary with conclusions and recommendations for operational use and further research. Suggestions are made for modified applications of the model in other settings.

The appendix lists a glossary of terms, an annotated list of tests used in the project, copies of the various forms used, and instructions for the pilot study project.

References are listed by chapter as well as in the bibliography.
CHAPTER II
REVIEW OF THE LITERATURE

The rapid growth in the use of tests for selection, classification, and counseling has raised probing questions regarding their appropriateness and use. Goldman (1969) contended that there was entirely too much testing and that, often, students and adults were given batteries of tests without consideration for their needs or abilities. In a study of testing and fair-employment practices, Kirkpatrick and Ewen (1968) questioned the validity of many standardized tests with the culturally different when this segment of the population was not a part of the standardized sample.

Diamond (1969) indicated that a new commission of the American Personnel and Guidance Association will review "the proper and ethical use of tests." Special attention was to be given to the bias in testing as a result of cultural or environmental deprivation. Anastasi (1968) devoted an entire chapter in her revised book to the "Social Implications of Testing." The problems dealt with the differences in test performance as a result of culture, race, or education.

Briggs and Hummel (1962) found that cultural conditions affect test performance. Standardized tests may be fair for a large percentage of the population, however, they could also be unfair for others. Hill (1969) wrote that testing practices were often
substandard, with ineptness and lack of professionalism.

The criticism of testing has not gone unanswered. Goldman (1969) felt that the breakthrough in testing would be by better test selection, interpretation, and reporting of results. Testing programs should be carefully planned with administration to small groups. Rothney (1959) also supported the theory that testing is most useful when carefully planned and tailored to meet the needs of the individual. Kirkpatrick and Ewen (1968) and studies by Wallace (1966) suggested that cultural or environmental factors be an index for a moderator variable in adjusting test scores. The U.S. Department of Labor (1968) has been researching methods of measuring cultural exposure, using a questionnaire, as a way to differentiate between advantaged and disadvantaged youth. The scoring of the questionnaire by various keys would provide adjusted or corrected test scores to compensate for variables in performance that are attributed to cultural differences. The research in this area has been somewhat inconclusive and adequate measures of cultural status have not been developed or validated.

Some critics such as Deutsch (1964), Kirkpatrick and Ewen (1968), and Wallace (1966) have proposed that the use of standardized tests be limited for the culturally different or disadvantaged. Others felt that proper selection of standardized tests, along with preparation of the examinee, can be of considerable value. Anastasi (1968), Rothney, Danielson, and Heiman (1959), and Cronbach (1960) agreed that, if the examinee were prepared for the test-taking and careful selection was made of appropriate tests, the results could be both valid and useful.
Test makers, unfortunately, have not given the same care and zeal to the writing of directions either for test selection or the preparation of the examinee for testing as they have to the technical aspects of test construction. The instrument may have adequate validity and reliability, for the representatives of the standardized sample but, if it were not appropriate for the individual or if the examinee were not giving his maximum performance, the results may be invalid. Rothney, Danielson, and Heiman (1959) and others stress the need for adequate preparation of the examinee for testing.

**SEQUENTIAL TESTING**

The concept of sequential testing as a method of planning for and administering tests is not unique, but it has rarely been developed into a system. Cronbach (1960) referred to sequential testing as a method of using multifactor tests that can relate to many variables and which can be followed up with more precise or specific tests. He called attention to the economy of dividing testing further into stages, with the option to decide on continuing or rejecting further testing. The first stage should be made up of short tests that lead into longer and more costly testing as may be necessary, but, at the same time, eliminating unnecessary tests.

Thorndike (1966) pointed out that the sequence in which separate tests in a battery were given was usually governed by factors of convenience. Often, short, speed dominated tests alternated with longer and convenient breaks or rest periods are necessary for maximum performance. Variety and content of tests can also be controlled by
sequence, thereby influencing interest and motivation.

Goldman (1961) elaborated more fully on the idea of using short tests for all persons, basing the decisions as to administration of longer, more costly tests on needs and abilities. He felt that the institutional framework often limited the tests and their use. Nevertheless, the counselor should control the planning and administration procedures and use the screening approach as a part of the planning for the sequence of tests.

Super and Crites (1962) indicated that sequence of tests should be determined by interest. Preliminary tests could be buffers, preceded by warm-up exercises that are not too difficult. The examinee should develop self confidence as he progresses to longer and more difficult exercises and final tests should be short so that he ends the testing situation with a pleasant feeling. Some testing (in batteries) predetermine the need for other tests. It is important, too, that sequence be determined as far as possible by factors that make testing a desirable and meaningful experience for the examinee.

Patterson (1962), in his unpublished doctoral dissertation, developed probability models and hypothetical populations to investigate the sequential method of testing. He compared the sequential model with the traditional cumulative method and concluded that the sequential tests were more accurate with individuals at the extreme ability levels.

Rothney, Danielson, and Heiman (1959) felt that the sequence in which tests were presented could influence the examinee's performance or effort. These authors also suggested that, if more than
one test were used, those of greater interest should be presented first. Rapport between the student and the test examiner was best developed by careful planning of the sequence of tests with emphasis upon individual needs.

PRE-TEST ORIENTATION

An important factor in the planning for tests is the preparation for or orientation to the testing situation on behalf of the examinee. The idea that orientation to testing can have a significant influence on test scores is not new, but it has been only recently, with the concern for the disadvantaged, that any specific procedures have been established. Anastasi (1968), Briggs and Hummel (1962), and Doppelt (1967) have examined the theory that disadvantaged persons may lack the basic skills and understanding to function adequately on standardized tests. Factors of attitude, motivation, anxiety, poor concepts of time and accuracy, as well as lack of understanding of the need for tests can affect performance.

Anastasi (1968) referred to test-related factors, such as cultural, educational, or environmental deprivation, as possibly causing depressed scores. This reduces the validity of such scores and may invalidate criterion data. Previous test-taking experience, motivation to perform well, rapport with the test examiner can be other influencing factors. To these may be added undue anxiety and fear of tests. A slight amount of anxiety is normal and even beneficial for maintaining an alert and high energy level, but over-anxiety has a detrimental effect.
Deutsch (1964) found minority group children less verbal, less confident, and less motivated to do well in competitive situations. Their poor performance on standardized tests indicated lack of test-taking skills, a tendency to guess and make random movements, an inability to follow directions, a short span of attention, and a lack of desire to achieve. These test-related factors lowered the validity of the scores as well as resulting in lower reliability.

Diamond (1969) stated that there may be possible systematic bias in the aptitude testing of the disadvantaged. This could be due to over-anxiety caused by emphasis on concepts and information to which the disadvantaged have not been exposed. There has been also a question, as to whether those who make low scores lack test-taking skills or just have poor attitudes and motivation.

Gordon (1969) found that tests administered to disadvantaged youth without adequate orientation resulted in their not finishing the test battery, guessing and random responses, failure to show up for scheduled tests, and a general resistance by hostility to testing, if not complete rejection.

Rothney, Danielson, and Heiman (1959) felt that test designers simply assumed the presence of interest and motivation and therefore rarely did anything to generate it in preliminary directions. Frequently students had no idea what the tests were about or why they were taking them. The question of how the results would be used sometimes caused suspicion, anxiety, and even hostility towards the tests and examiner. Orientation or motivational procedures were neglected in most testing situations. This should raise serious questions as to
the validity and reliability of test results.

Sigel, Gales, and Kost (1965) in a survey of Practitioner's Perspectives on Testing, stated that standardized testing programs must continue, but those who use tests should be more sophisticated and less routine in the preparation of the examinee for testing. The mechanical following of directions and scoring prescribed in the manuals could be very limited and harmful practices. The professional test-user must have knowledge and awareness of factors that influence test performance and, in turn, affect scores. Among these could be the lack of test sophistication, problems in following complex instructions, use of the separate answer sheet, limited environmental and cultural exposure, and the overall conditions of the test-taking situation.

Super and Crites (1962) stressed the need for an introductory or motivating talk. Orientation should provide information concerning just what was to take place, how long the tests would take, reasons for the testing, and what the results mean, and how they were to be used. Thorndike (1966) made the point that the introductory statements to testing could influence motivation and attitude. The usual reaction to a talk stressing the importance of doing well and working rapidly on a test was for one to try to do one's best. Everyone, however, does not respond in this way. Some remain apathetic and unsure.

Coaching and preparation for tests are not recent innovations. For years, books such as the ARCO series, self-study guides, Huff's Score (1965), and other information have been widely available. In most instances, the information consisted of hints on test taking, how
to prepare for tests, study questions, and sample and practice items. Manuel (1956), in a little booklet called *Taking a Test*, provided a self-study guide covering the purpose of tests, meaning of test scores, attitude toward tests, and practice items that could be self scored. A similar book was written by Anderson, Katz, and Shimberg (1965).

The Psychological Corporation (Bennett and Doppelt, 1967) recently marketed "Test Orientation Procedures" (TOP) as a standardized method that can be used with persons who have had limited exposure to tests, and who suffer from anxiety. The kit consists of tape recorded instructions and two booklets which familiarize the examinee with paper-and-pencil tests and also contain illustrations of various apparatus tests. One booklet is intended for practice with the tape-recorded instructions, for self-scoring vocabulary, spelling, arithmetic, and general information tests. The second booklet is a take-home practice item.

The U.S. Department of Labor has developed a pamphlet entitled "Doing Your Best on Aptitude Tests" (1968). The contents include information and cartoons on reasons for taking aptitude tests, how to study for tests, test-taking hints, and practice exercises from the General Aptitude Test Battery with a self-scoring sheet. Another innovation of the Labor Department has been a "Pre-testing Orientation to the General Aptitude Test Battery" (1968). A standardized presentation is contained in the test administration manual which is used in conjunction with a test booklet of practice paper-and-pencil exercises using a separate answer sheet. Both items were developed in an attempt to provide better preparation and orientation of the disadvantaged for
standardized testing. The manual suggests that familiarization and practice on the tests will reduce anxieties, motivate the examinee, develop rapport with the test examiner, and give him more meaningful experience. Test scores should be more valid because certain test-related variables have been controlled or eliminated.

An important issue in regard to test orientation involves the difference between coaching to take and pass a specific test and the more general approach of providing a learning experience which would help a person take various kinds of tests. The former situation would be unethical for it is conceivable that a person could pass a test just because of practice or coaching and still fail on the job or in training for which the test had been validated. Use of practice items commonly found in the ARCO books or in the GATB Pre-Testing Orientation is approved, for these bear only a similarity to the actual test items. One objective of test orientation would be to make available public information and certain learning experiences to individuals who may have been denied such opportunities because of cultural, educational, or environmental deprivation. Culture-free or fair tests may be questionable, but removing cultural bias due to lack of experience or environmental exposure, can make tests and testing procedures more fair.

Anastasi (1968) found that conditions such as early educational experiences combined with the nature of the tests and the amount and type of instruction influenced test performance. Persons with educational and environmental deprivation profited most from special instruction, whereas persons who had taken many tests and functioned on
higher educational levels did not make significant gains in test scores. Similarity of the practice items to the tests themselves was found to be important. Elements in the development of test sophistication included overcoming strangeness, creating a better attitude toward test-taking, gaining a familiarity with items and a time concept, and using separate answer sheets.

In 1954, Haggard demonstrated that significant improvement on IQ scores resulted after only three hours of training in the reading of instructions, in test-taking skills, and the use of practice exercises. As a result, better rapport was developed, and the examinee was motivated by greater self confidence, interest, and a sensitivity to rewards.

That certain incentives could affect motivation and thereby raise test scores was reported by Cronbach (1960). Factors such as being hired, receiving a desirable assignment, ego involvement, self respect and the respect of others, and interest in the task had beneficial influence. He cautioned against incentives that were too simple and, as a result, had little or no effect. The attempt should be made to shift the concern about test scores to the examinee. The greatest change occurred when the desire to do well, to make a good impression, to work hard, to obtain praise, and to succeed was created. These are typical middle-class values and attitudes that are often lacking in the disadvantaged. Threats, fear of punishment and tension or failure, had a negative influence and tended to depress test scores.
The use of the group process was suggested by Goldman (1961) as a most helpful method of conducting test orientation. Meetings prior to the tests should be directed to understanding of the reasons for testing and to improving attitude and motivation by the lowering of undue anxiety. The most effective method used small groups that were fairly homogeneous, allowed adequate time, and avoided the lecture or pep-talk approach.

Mallory and Graham (1954) found that, if explanation of tests, question periods, and discussion took place an hour before uniform testing, test anxieties were reduced. Unfortunately, there was no specific research data on this study and, in fact, research data seems to be lacking in all the areas of test orientation.

A test orientation and tutoring course conducted by Dr. Kenneth Clark at New York University for black youth applying for apprenticeship training was reported by Kempton (1967). Four weeks of intensive instruction concentrated on test-taking skills and practice tests. The training was so successful that all students qualified on the tests. The rationale for the course was that most middle-class whites learned these skills as a result of certain educational, cultural, and environmental experiences. Providing similar opportunities for the black youth developed self confidence and attitude and motivation that took away fears, threat, and mystery of testing. The study did not report the nature of the instruction and selection of the students. Nor was there any control group.

A special orientation for Manpower Training that included instruction and preparation for testing was described by Nichols (1966).
Standardized testing was deferred until late in the program. Following this preparation, the trainees appeared to be properly motivated and to have a better understanding of the need for testing. It was felt that the change in attitude and motivation improved performance on the tests. A similar project was conducted a year later in Ohio by the State Employment Service and was the beginning of the development of a formal pre-testing orientation that is currently being researched.

A three-week pilot program of test orientation for disadvantaged youth in the Cleveland, Ohio area was reported by Ferraro (1968). It was designed for 100 black youth who were interested in preparing for entry into apprenticeship in the construction trades. Even before application for apprenticeship could be made, qualification on standardized aptitude tests was required in many instances. The youth spent three weeks in a resident camp undergoing intensive instruction and learning experiences that included test orientation, personal orientation, interview experience, and some on-the-job training. Recreational activities and individual and group counseling were also made available. Pre and post-tests were used to determine the effectiveness of instruction in test-taking. Practice was provided with samples of common employment tests, such as the Wonderlic, Otis, Minnesota Clerical Tests, and The Bennett Mechanical Comprehension Test. Emphasis was placed on test familiarization, testing methods, and use of separate answer sheets. A significant increase in test scores was found as a result of the special instruction. Questions that should be asked on this study concern the lack of a control group, whether or not alternate forms of the test were used in post-testing, and the select
nature of the group (all had to be high-school graduates). The success of the instruction was attributed to change of attitude towards tests and a reduction of fear and anxiety.

A study of the effects of training on aptitudes was made with 40 mentally retarded high school students in a vocationally oriented program (U.S. Department of Labor, 1968). It consisted of psychomotor tasks, matching exercises, and sorting operations. Tests were included for spatial, form perception, and manipulative skills. The performance on the manipulative skills increased significantly as a result of practice, whereas the least change was noted on the paper-and-pencil (form and spatial perception) tests. Effect of a single exposure was found to be as great as repeated practice. As would be expected, the similarity of the practice to the actual test was of great importance.

Wallace, Kissinger, and Reynolds (1966) propounded the theory that compensatory training in test-taking and better orientation, as well as careful administration of tests, could be of greater value than adjusted norms or culture-fair tests. The so-called culture-free or fair tests may actually be of little value because the individual still must fulfill requirements that are of a cultural nature, such as accuracy, speed, and a desire to do well or succeed. The fairer approach is to teach the examinee how to function in these situations so he can perform well on the tests and also on the job or in training.
TEST-TAKING BEHAVIOR

An important but often neglected part of testing has been the provision for adequate observation and recording of test-taking behavior. Various manuals and texts on testing have mentioned that the test examiner should note any unusual behavior such as undue anxiety, lack of motivation or interest, or inability to understand instructions. However, most directions on standardized tests stress the importance of accurate timing and precision in the reading of instructions with little time for observing test behavior. In group tests that may involve from 10 to 20 examinees, the opportunity to watch and record individual behavior has been very limited. For that reason, testing programs for the disadvantaged should be restricted to groups of 5 to 10 persons in order to afford this opportunity.

The use of tape-recorded instructions may be one way to relieve the test examiner of the details of administration and allow him to give more attention to behavioral observations. The California Test Bureau and the Psychological Corporation have put test administration directions on tape. Providing a proctor or test monitor can be another method of obtaining recorded data on test-taking behavior.

Over thirty years ago, Bingham (1937) stressed the importance and value of observing test behavior and recording it for convenient reference and filing. The standard test situation or orientation period both provide good opportunities for this type of activity. The test score is only a partial picture of the individual. It can be made more complete by using additional information on test behavior,
thus making interpretation of results more meaningful and accurate. Bingham stated that "the numerical score tells only half the story." He suggested looking at method of attack, output of energy, industry and tenacity, relative interest, personal reaction to difficulties, awkwardness or skill, and emotional control. His examiner's checklist was based on one developed by Baumgarten and translated by Keller (1935) for detailed observation and recording. The unwieldy form provided not only for recording but also for interpretation of the behavior. Five main areas of observation were covered, including preliminary instructions (8 items), execution of task (42 items), attitude toward performance (5 items), conduct at end of test (5 items), and post-testing period (5 items). He emphasized the importance of noting the behavior actually seen, with appropriate interpretation. This information would be reported by the test examiner to the counselor.

Super and Crites (1962) proposed an additional time for observation of behavior, during informal rest periods between tests. Observation of behavior has been most frequent in clinical testing and often neglected in educational or vocational testing. Elaborate forms, such as Bingham's are rarely used because the examiner in group testing has little time to observe detailed behavior, let alone recording and interpreting it. As a rule, only unusual or significant behavior has been reported. Super and Crites agreed that this type of information can be very helpful and valuable, but they cautioned against misinterpretation or possible lack of validity of the observations.

Froehlich and Hoyt (1959), in a discussion of observation in general, also warned against superficial and misleading information.
Several techniques should be developed in gathering and verifying observational data, to be used according to the various situations. Four proffered principles for reliable observations are: (1) observe the whole situation, (2) select one student to observe at a time, (3) observe in regular activities, and (4) make observations over a period of time. The limitations of observations mentioned were: (1) small sampling of behavior, (2) observer may be biased, (3) he may be inaccurate, and (4) there is difficulty in communicating exactly what is observed. Rating scales based on clearly defined judgments could be helpful, as developed in some graphic form. Ways of recording are the descriptive method, written comments, and evaluation, as on rating scales.

The staff of the University of Maryland Counseling Center (1960) and Kirk (1961) developed a checklist for recording test-taking behavior and attempted to study the reliability of the observations. The task was found to be complicated and difficult, particularly in the interpretation phase. Kirk stated that reports such as these should be considered tentative and should not be fully accepted until proven in the counseling process. Some suggested points to watch for were the person's approach to the tests, his persistence, carefulness, accuracy, pride, his dependency upon the examiner, fear or anxiety, the number of demands, efficiency, and organization. Both the counselee and counselor can gain valuable information on test behavior despite the questions about the lack of validity or reliability.

A type of test behavior was reported by Cronbach (1950) as "response-set," related to test design. He claimed to have extensive
evidence to show that response-set can be a bias in favor of particular alternatives or patterns of responding to test items. On personality or attitude tests, the examinee may fake his answers in the direction he believes would be favorable, rather than providing an accurate pattern. Or he may resort to a tendency to guess, working for speed versus accuracy or vice versa, skipping around on the items, and marking only one type of answer or pattern of answers. Response-set can be altered by change in directions or coaching. It should be avoided if possible for it reduces the validity of the test. One way to help to control this type of behavior is to administer tests to small groups. Some tests provide a method of determining response-set scores through a comparison of items attempted and number of correct answers.

**THE SYSTEMS APPROACH**

Churchman, in his book, The Systems Approach (1968), wrote that the development of a system should be based on a philosophy of efficiency and exploration for a best way of doing a task. Underlying this can be time-motion studies as a way to gain greatest economy. A system is a set of parts coordinated to accomplish a set of goals. The parts must be defined, as well as coordinated, spelling out in detail the whole system, including the environment, objectives, and the series of "thinking steps."
The five basic considerations in the systems approach are as follows:

1. The total system's objectives and performance—measure of the whole system,
2. The system's environment—the fixed constraints,
3. The resources of the system,
4. The components of the system, their activities, goals and measures of performance, and
5. The management of the system.

The environment outside the system's control determines in part how it performs. The resources inside the system provide the "general reservoir" out of which specifications can be shaped. Specific actions are taken by the component parts or subsystems. The separation of the system into these parts gives the analyst information needed to tell if it is operating properly.

Churchman referred to the use of simulation in developing a model in order to proceed in a feasible manner, with each step being taken in turn. Models and simulation are methods of developing a prototype. Essentially, a model of a system is an input, a process, and an output with the process as the heart of the concept or the "Black Box." An illustration of this was reported by Veri and Mocker (1969) in the use of the system approach for the development of basic adult education. It is illustrated in figure 1.
Systems develop into complex networks that usually are depicted as linear or circular flow charts. Multi-stage thinking provided for various modifications of the plan. Churchman considered that the way to look at a whole system was in terms of a plan that offered courses of action to follow to a desired goal. The planner or decision-maker should be in a position to assess alternatives before selecting a route or course of action. Planning in a systematic manner is essentially breaking down into component parts concerned with multistage decision-making.

Osipow (1968) illustrated the systems approach in career development with decisions as links on a chain, each made prior to action on the next. Minor decisions influence the major decisions. Potential avenues offer various alternatives, some are logical and can be successfully followed to reach the goal, while others may not be effective and hinder the efficient attainment of the final
Osipow described the process as a sequence of information collection, assessment, selection, and predicted outcome in terms of probability and desirability. Figure 2 illustrates a model for decision-making.

Ramo (1969) suggested that the systems approach can be a method of curing social problems. The basic ideas come from the field of engineering that provides techniques for assembling and processing data, comparing alternatives, possible compromises, qualitative analysis, predictions, seeking out judgments based on past experiences, and developing creative innovations. The impetus for the use of systems has been influenced by the availability of computers, but this is not essential except in extremely complex situations.
The concentration on the analysis and design of the whole is distinct from that of the components. Examination of the entire problem requires an understanding of the sub-problems and their relationships to the whole. Ramo also stressed the importance of simulation and modeling in the development of a system.

Rabow, in his book, *The Era of The System* (1969), defined it as an assembly of components that perform together in an organized manner. A component may be a smaller system or a subsystem. Actually, it is a method of looking at a problem from an overall viewpoint and dividing it into a set of subsidiary problems which together help to solve the original problem. Modeling or simulation is characterized by the systems engineer who constructs block diagrams to achieve the best configurations. Feedback becomes a part of the system in the evaluation process. Simulation is a technique of duplicating a situation to verify prediction of behavior and outcome when it may be impossible or impractical to make a direct study of a problem. Furthermore, certain experimental controls can be introduced that might not be possible in a real situation.

An interesting application of the systems approach was used by Veri and Mocker (1969) in basic adult education. One of their subsystems related to diagnostic testing and counseling. The procedure was to determine the educational or achievement level. Tests were introduced so the learner understood that, by finding his operational level, the curriculum could be planned to meet his individual needs. The model of subsystems presented by the authors illustrated the feedback and linear progression of steps. (See Figure 3)
A circular model of the testing, counseling, and learning system is shown in figure 4.
Additional models of various subsystems and processes were used to graphically demonstrate the various steps and their relationships within the program. Testing was used to diagnose the level of ability and educational need as well as providing a method of determining behavioral change pre- and post-testing during the educational process. If the learner revealed a difference in performance as
measured on tests after the teaching or learning situation, the goals may have been achieved.

Findley (1963) identified a testing program as a small independent system. He described it as "consisting of the systematic use, at more than one level and in more than one area of one or more tests where selection, administration, recording, and interpretation of the tests are all planned and conducted in close accord with the major objectives of the educational program."

Linn and Rock (1968), in an article on programmed tests, referred to a sequential system of branching. The purpose of this was to reduce testing time with the selection of appropriate items within the scope of the examinee's abilities. This must be accomplished in accordance to acceptable levels of reliability and validity. The authors were of the opinion that programmed testing could be improved with computer assistance.

In a monograph on automated data processing and programming in testing, Stoker (1968), characterized decision-making as a step-by-step process as it is found in the systems approach. The first level of activity was the analysis of the problem with identification of input data, the required processing steps, and the types of output desired. The second level was the formulation of the problem, which would include flow charts and modeling. The third step was the actual programming.

The systems approach could be a very sophisticated and technical method that, in its more complex forms, would rely heavily on data processing, programming, and computer assistance. On the other
hand, there have been attempts to apply some of the basic concepts of system analysis, planning, and processing of data without the use of computers. This study will utilize these methods as they apply to the decision-making process, test selection, planning for testing, steps or sequence in test administration, and final integration of total test and test-related information.

**SUMMARY OF THE CHAPTER**

Current interest in the selection of appropriate tests and their proper use has focused on the problems of the disadvantaged. Numerous writings have dealt with the technical aspects of testing, whereas operational procedures concerned with the planning and selection of tests, preparation of the examinee for testing, and observation of test-taking behavior have been neglected. This chapter reviewed the literature on sequential testing, pre-test orientation, test-taking behavior, and the systems approach as it may apply to the subject of this study.

Sequential testing is a procedure for selecting, administering, and utilizing tests in a logical, orderly manner. The progression is from simple screening or practice exercises on to more difficult, complex tests that are still within the realm of the examinee's abilities and skills. An essential aspect of this method is the orientation to testing, which is especially important for disadvantaged persons who lack test-taking experience. Limited research indicates that the necessary skills can be learned through special instruction, resulting in significant improvement of test scores. Related variables that
influence test performance are anxiety, attitude, and motivation. These, in turn, can also be affected by training. The emphasis should not be placed on coaching to pass specific tests, but rather upon learning experiences that provide a change of attitude toward tests in general and an opportunity to acquire test-taking skills. Hopefully, this will inspire a maximum performance and produce more valid and reliable test scores.

The study of test-taking behavior has also been generally ignored and neglected. Test scores are only half the story, and, particularly with the disadvantaged, the supplemental information gained from noting test behavior would make test interpretation more complete. Few studies, however, have been made in this area in spite of the recognition of its importance. There are many problems connected with objective observation of test behavior, but the need for developing the means to do it remains critical.

With the advent of the computer, data processing, and programming has come the realization that systematic procedures are also needed in fields of testing that are now complex, costly, and time consuming. The principles of the systems approach, used in engineering and efficiency planning, have been incorporated into this study with practical applications designed for sequential testing and procedures.
CHAPTER III
DEVELOPMENT OF A MODEL

The purpose of this chapter was to develop a model that would serve as a foundation for an experimental and demonstration project. Although the basic concepts and procedures may be applicable in a variety of settings, certain conditions may determine the feasibility of this model to other situations.

Essentially, a model is a method of describing verbally or graphically a miniature representation or facsimile. It may be a style or structure, a design, an example, or a pattern. Modeling has become a popular way of illustrating ideas or procedures. An advantage of this type of presentation is that it organizes complex information into a more clearly understood form that can be utilized and adapted in many different ways.

This model may be adapted to different settings but for purposes of this study, it was primarily designed for use in employment or vocational counseling programs. In keeping with the times, special emphasis was given to activities related to the disadvantaged, the culturally different, and the educationally and economically deprived. The several programs currently active will be described as they pertain to this study.

The tests used in the model were those developed or approved
by the Manpower Administration of the U.S. Department of Labor. They are used by all the State Employment Services and some have been made available to non-profit educational, vocational, charitable, religious, and community organizations. A detailed description of the tests may be found in the Appendix.

The type of tests and the way they are used are largely determined by the nature of the agency and the ongoing programs. In some instances for example, certain tests are restricted as to their use because of special requirements, such as qualifications of the test examiner or the counselor who interprets the results. Personnel who participated in this study were specially selected and trained in the administration and use of the tests described.

The various forms and checklists were developed as aids to the counselor and test examiner, providing a method of recording information to be inserted into the individual counselee's case record folder. Every effort was made to keep the records brief and simple as well as avoiding duplication.

Diagrams and flow charts were used to graphically illustrate different aspects of sequencing or modeling of the systems approach. This procedure has been frequently used in programming and systems analysis, from which the basic concepts for this study were drawn.

THE SETTING

The setting for which this model was devised was the Ohio Bureau of Employment Services. This agency is part of a national organization of a State-Federal system under the Manpower
Administration of the U.S. Department of Labor. It has three main functions: (1) the collection of contributions (taxes) from employers, (2) payment of claims to unemployed persons who qualify, and (3) the operation of an employment service. In recent years, the role of the Employment Service has changed from a labor exchange activity to a total community manpower agency (Cassell, 1968). The late 50's and early 60's were periods of persistent and high unemployment due to post-war readjustment and technological change. At the same time, there was a growing awareness that the characteristics of the unemployed were lack of employable skills, low levels of education, poor attitudes and work habits, and low motivation. Special legislation such as the Manpower Development Training Act (1962), the Vocational Education Act of 1963, the Civil Rights Act of 1964, and the Economic Opportunity Act (1964) focused attention on the unemployed, the disadvantaged, and the deprived.

The main emphasis in Employment Service programs now is upon employability development, training, counseling and assessment, as well as job placement. Employability development, for example, includes interviewing, counseling, testing, orientation, and preparation for training or referral to a job.

A recent program, apart from the regular work of the Employment Service, is Human Resource Development (HRD), bringing particular attention to the vocational or employment needs of the disadvantaged. Included in this group are out-of-school youth (ages 16 to 22), older workers (over 45 years of age), welfare recipients, minority groups, and the educationally disadvantaged (less than 9th grade), socio-
economically, and culturally deprived. Each category is carefully identified in terms of qualification for special programs or services.

EMPLOYMENT SERVICE TESTS

Employment Service tests are used as a part of the assessment process in obtaining information about the individual's level of skill, knowledge, ability, aptitude, and interests. Essentially, the tests are classified in three ways:

1. **proficiency and trade knowledge tests** - that measure acquired skills, achievement, or knowledge, all of which may be learned either through special training or experience,

2. **aptitude tests** - measure learning potential to acquire skill or ability through education or training, and,

3. **other** - a general category that includes interest checklists, color vision tests, pre-test orientation, and work-sampling, a new experimental type of assessment.

Paper-and-pencil tests, apparatus, work-sampling, and checklists are used chiefly to obtain a sample of the applicant's behavior. These measure knowledge, achievement, ability, or skills as they may relate to education, training, or prediction of success on a particular job. All tests have been standardized and validated on studies of employed workers or trainees. Extensive research studies are an ongoing process, with individual states participating in national or
local projects.

Tests are used to supplement information about the individual applicant that has already been obtained through interviewing, counseling, or case records. Test data may be used for four main purposes; (1) selection and placement (into training or job), (2) classification, (3) counseling, and (4) research. Table 1 classifies the various Employment Service tests and their use. The tests are described in greater detail in Appendix B.

**TABLE 1**

**CATEGORIES AND USE OF EMPLOYMENT SERVICE TESTS**

<table>
<thead>
<tr>
<th>Category</th>
<th>Test</th>
<th>Use</th>
<th>Type of Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Proficiency and trade knowledge</td>
<td>A. Clerical skills</td>
<td>Determine level of skill for classification or referral to job or training</td>
<td>Speed and accuracy</td>
</tr>
<tr>
<td></td>
<td>1. Typing</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2. Dictation</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. Spelling</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>B. Oral Trade Questions</td>
<td>Verify knowledge or experience</td>
<td>Amount of information about a trade or job</td>
</tr>
<tr>
<td></td>
<td>(OTQ)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>C. Reading and arithmetic achievement tests.</td>
<td>Determine level of basic skills for diagnosis or referral to basic education</td>
<td>Grade level in particular skill</td>
</tr>
<tr>
<td>Category</td>
<td>Test</td>
<td>Use</td>
<td>Type of Information</td>
</tr>
<tr>
<td>------------------</td>
<td>------------------------</td>
<td>----------------------------------------------------------</td>
<td>--------------------------------------------------------------</td>
</tr>
<tr>
<td>2. Aptitude</td>
<td>B. Screening Exercises</td>
<td>Determine ability to understand instructions and do parts of GATB</td>
<td>Number of correct items and test-taking behavior.</td>
</tr>
<tr>
<td>(continued)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>C. Non-reading measure of &quot;G&quot; (mental or learning ability)</td>
<td>Determine learning ability as it relates to basic or remedial education.</td>
<td>Standard scores for predicting educable level.</td>
</tr>
<tr>
<td></td>
<td>D. Non-reading parts of GATB</td>
<td>Used with C. above for predicting ability in fields of work (OAP's)</td>
<td>Standard scores normed for fields of work (OAP's) or specific occupations (SATB's).</td>
</tr>
<tr>
<td></td>
<td>E. General Aptitude Test Battery (GATB) B-1001.</td>
<td>With individuals who have had limited test experience</td>
<td>Same as above.</td>
</tr>
<tr>
<td></td>
<td>F. GATB, Forms B-1002A B-1002B</td>
<td>With individuals capable of using separate answer sheets</td>
<td>Same as D above.</td>
</tr>
<tr>
<td></td>
<td>G. SATB's</td>
<td>Selection for Specific jobs.</td>
<td>Cutting scores &quot;H&quot; - High &quot;C&quot; - Band of consideration &quot;L&quot; - Low</td>
</tr>
<tr>
<td>3. Other</td>
<td>A. Pre-test orientation (relates to various types of tests)</td>
<td>With persons who have had limited or no test-taking experience.</td>
<td>Test-taking behavior and ability to perform on various practice exercises.</td>
</tr>
</tbody>
</table>

TABLE 1 - Continued
TABLE 1 - Continued

<table>
<thead>
<tr>
<th>Category</th>
<th>Test</th>
<th>Use</th>
<th>Type of Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>3. Other (continued)</td>
<td>B. Interest Checklist (ICL).</td>
<td>Counseling aid for area of interests.</td>
<td>Preference for various types of work arranged for fields of work.</td>
</tr>
<tr>
<td></td>
<td>C. Color vision test.</td>
<td>With certain specific aptitude test batteries.</td>
<td>Ability to differentiate between colors as may be required for specific jobs.</td>
</tr>
</tbody>
</table>

TESTING PROCEDURES

The decisions as to when to test and what kinds of test to use are determined by various Employment Service staff members depending on the nature of the individual applicant's characteristics and background, and by types of services or programs available. Essentially, the inputs for decisions, selection, and scheduling of tests are based on information obtained in the interview, counseling, or case records. Information about an applicant can come from several sources, but a primary one has been the initial interview during which, for one thing, the application form is completed. Different types of records, military, employment, education, welfare, and supportive services, for example, can also be valuable as biographical facts. In addition,
special forms developed for specific programs such as Manpower Training and Work Incentive are used. The combination of the initial information and assessment data gathered during services such as counseling, testing, orientation, training, and employability development would determine the sequence of activities. Figure 5 illustrates this aspect of the model.

One possible sequence would be referral of the applicant directly to a job or training, in which case, the information would indicate that no special services or activities are necessary. An
illustration could be that of the applicant who is trained or already has the necessary experience and meets the employer's specifications. In other instances, tests may be required such as proficiency (Clerical Skills), Oral Trade Questions, Specific Aptitude Test Batteries, GATB for multi-battery scoring, or the Color Vision Test. If they are used, the results are returned to the interviewer who interprets them along with other information and makes the referral. Figure 6 illustrates the model for this situation.

Figure 6 - Direct Referral to Job or Training
If the applicant lacks experience or employable skills, has no specific job choice, or has some special problems of employability, he is usually sent to the counselor. He in turn may refer him to certain programs or services designed for individuals or groups with special needs such as the Work Incentive Program (WIN) for which the model was devised. This type of program, with services for welfare recipients, disadvantaged youth and adults, and the handicapped, has been more extensive and complex than most others. The information gathered has been more detailed and the assessment process, including testing, and other services are very carefully planned. Figure 7 illustrates the various services in the WIN program as they may relate to the counseling and testing process.

Essentially, the decisions with regard to the sequence of testing and other services are usually made by the counselor after analyzing the information available and determining the readiness of the counselee. In the process of developing this model, the various sources and types of information were carefully reviewed. Certain items were selected on the basis of face validity and empirical evidence as being of significance to the counselor in the decision-making process.
Figure 7 - Processes and Services in the WIN Program
For purposes of this study, a Test Planning Worksheet (TPW) was organized into three areas: namely, biographical data, socio-economic and cultural information, and previous testing experiences as stated in the counseling interview. The worksheet was designed to provide the counselor with a method of profiling information about the counselee and graphically determining whether a full range of tests would be needed in some specific sequence or if select tests were to be used in accordance with the individual's characteristics or requirements of a program.

A number of factors have been known to affect performance in a testing situation and must be seriously considered. For example, there has been evidence to indicate a relationship between both age and education and the ability to perform on tests. Older workers (45 years and above) have often had limited or not very recent test-taking experiences and, to them, tests have appeared threatening. Older persons tend to react less favorably to highly timed tests because of slower response sets that are a result of physiological change with aging. Equally frustrating is the fact that most standardized tests are heavily loaded with factors that relate to educational skills such as reading, arithmetic, problem-solving, and comprehension. Thus, those with less education have an initial disadvantage in addition to the fact that they have had fewer test-taking experiences. A similar rationale is found with persons who have had limited work or military experience, where tests are frequently used.
It is obvious that persons with physical handicaps have problems with tests. Test examiners have been trained to be alert for clues that may indicate a physical problem which may appear to influence performance. Perhaps they notice a visual, hearing, or psycho-motor disability. This is usually recorded in comments and reported to the counselor who will interpret the test results.

Notation of interests, hobbies, leisure-time activities, and reading patterns can also give hints as to possible test-taking behavior. This information can supplement or verify that already obtained.

Economic status has also been found to have a relationship to environmental exposure and experience, especially with lower income groups who may have had few or no contacts with educational situations which place emphasis on culture, knowledge, education, and middle-class values. Standardized tests, on the other hand, have been recognized as definitely being culturally oriented and economic income has had a direct influence on this fact. Social status, involving family structure, size, and housing, was also included as a variable. Descriptive suggestions for interpretation are provided in the directions found on the reverse side of the form.

Finally, direct information on test-taking experiences and attitudes are to be obtained through the counseling interview and entered on the worksheet. In general, evidence has shown that persons who have the most varied, recent, and greatest amount of experiences are most likely to do well on tests. Such factors as attitude, motivation, anxiety, and understanding of the reasons for tests have
also had an influence on performance.

The Test Planning Worksheet profile scale was developed with a five-point rating (Likert Scale) for each item. The interpretation was based on three categories. These were: (1) **positive**, which meant that certain tests may be used without any special preparation or sequence, (2) **questionable**, indicating that caution should be exercised and preparation for and orientation to testing, as well as sequence should be planned, and, (3) **negative**, which signifies that the appropriateness of prospective tests should be seriously reviewed and extreme care used in the preparation for and orientation to testing, which, in itself should be limited.

In situations that show a positive profile on the Test Planning Worksheet, direct referral may be made to testing, as indicated in Figure 6. This may include Clerical Skills Tests, Oral Trade Questions, The General Aptitude Test Battery, Specific Aptitude Test Batteries, or the Color Vision Test. If the evidence implied that testing were questionable or negative, careful planning and preparation should be a part of the counseling and testing process.

The key point to be stressed here is that when counseling is involved, the counselor makes all decisions about testing and the results would be reported back to him for interpretation. Counselees, in some instances, may be entered into various programs without the use of tests. However, at an appropriate time, certain tests can be administered in accordance with the counselee's needs or readiness, as determined jointly with the counselor. An aspect of orientation or employability development has been pre-test instruction, described in
Chapter II. This structured preparation for testing was based on a series of sequential experiences in the learning process of developing test-taking skills and of reducing anxieties through a better understanding about tests and test-taking. In the case of severely disadvantaged persons, all testing should be deferred until some orientation and pre-test instruction was provided. This group would include persons whose profiles were predominantly in the negative area of the rating scale. Tests that may be taken by such persons could include the Elementary Form of the Metropolitan Achievement tests (for diagnosis of reading and arithmetic abilities), the GATB screening exercises (to determine ability to take the full GATE), the non-reading measure of "G" (mental ability), and the non-reading parts of the GATB. The flow chart for this type of testing is depicted in Figure 8.

If the test planning profile is in the middle or questionable area, the sequence would utilize similar procedures, but the orientation would be modified. Sequence would be directed toward diagnosis and referral to other standardized tests, such as the Metropolitan or GATB (Form B-1001 or B-1002). The intermediate Metropolitan Achievement Tests (MAT) and GATB Screening Exercises can be used by the counselor to determine the counselee's ability to read and understand verbal parts of the GATB. There are two forms of this battery. The B-1001 is the earlier edition that uses an expendable booklet in which the answers are written and arithmetic problems worked. It was designed for persons who have had limited test-taking experience and may not be familiar with the use of separate answer
Figure 8 - Test Sequence for Negative Profile
sheets. Included in this group may be older or handicapped persons. The B-1002 has forms A (used for regular testing) and B (alternate form used for retesting and release to outside agencies). A new edition, to be published sometime in 1970, will have a non-reading content. The sequence of tests for persons whose profile may be in the questionable area appears in Figure 9.

The counselee would be referred back to counseling after each stage of testing for an interpretation and determination of the next step. An individual may skip certain stages if they were not appropriate for him. For example, a person with a negative profile may have taken the Achievement Tests and scored above the 6th grade level, which would be an adequate functioning level to take the B-1002 form of the GATE. The Screening Exercises, orientation, and B-1001 would therefore not be administered. If the person were found to function below the 6th grade level, the sequence should be followed with feedback at each stage for interpretation and determination of the next step in testing.

The categories of the Test Planning Worksheet profiles are not all inclusive, nor are they precise enough to spell out specifically the tests in the exact order in which they should be given. The counselor must exercise considerable judgment and rely on his knowledge of tests, as well as assessing the attitudes, motivation, and demonstrated abilities of the counselee. The latter are determined in the feedback sessions and are extremely important in making decisions as to the sequence of tests.
Figure 9 - Test Sequence for Questionable Profile
In situations that resulted in a positive profile, the full range or select tests, with the exception of the orientation and diagnostic tests (achievement and screening exercises), may be given. However, in a borderline case where there would be some doubt about the person's abilities or attitude toward tests, the orientation and diagnostic tests may be utilized. The intermediate form of the Metropolitan Achievement Tests may be administered to persons who are between the 6th and 8th grade attainment. The performance tests (Clerical Skills and Oral Trade Questions) should be used only when there is evidence that the person has had special training or experience. Figure 10 illustrates the sequence of tests for a positive profile.

The Interest Checklist (ICL) is a counseling aid and is not scored. It may be used along with the GATB in counseling if there were a need for vocational or training choice. Occasionally, the GATB may be given for multi-battery scoring (several Specific Aptitude Test Batteries) for selection and referral. If just one occupation or job were considered, only the SATB is given. Several batteries require the Color Vision Test (such as Cable Splicer, Painter, and Decorator).

The Clerical Skills Tests and Oral Trade Questions are usually administered to persons with experience or training and are used in selection, referral, or classification. The SATB's, Color Vision, Clerical Skills Test, and Oral Trade Questions may be requested by an interviewer or training selection officer without the services of counseling.
Figure 10 - Test Sequence for Positive Profile
The Clerical Skills Tests include typing, dictation, and spelling. Typing tests offer a choice of a manual or electric machine and plain or statistical copy, the job-order specifications determining the appropriate test. The dictation tests are administered at 60, 80, or 100 words per minute. The usual procedures are to administer a short practice exercise at one of the slower speeds and let the examinee decide what speed is preferred. Several forms of typing and dictation allow retesting. Spelling tests consist of clerical, legal, and medical forms. Again, the selection would be determined by the experience or training of the examinee or the requirements of the job specification. Figure 11 shows the sequence for Clerical Skills tests.

Figure 11 - Sequence for Clerical Skills Tests
The Oral Trade Questions (OTQ) are used by an interviewer to verify knowledge an applicant has about an occupation. Some 250 occupations are covered by sets of questions, but many of these are now out-dated. Thus, the OTQ's are not in wide use and have had limited applications. Research is in progress to up-date some of the lists of questions.

TEST RESULTS WORKSHEET

Test scheduling, results (scores), and comments are recorded on the Test Results Worksheet (TRW). A copy of this form (#2) is found in Appendix C. Test results are reported by the local office test examiner on a special Employment Service form (537) that is a duplicate of an appointment form provided the examinee. The Testing Section maintains a file of test records (Test Record Card, E.S. Form 538) for each examinee. Copies of these forms are also found in Appendix C. The advantage of the Test Results Worksheet over the existing forms is the fact that it provides a structured sequence of tests and scheduling and a method of recording all test results and comments on a single sheet. As the title of the form indicates, this was intended to be a worksheet for quick and ready reference, as a part of the counselee's case-record folder.

TEST BEHAVIOR CHECKLIST

The last phase of this model involved a procedure for observing, recording, and reporting test-taking behavior. Bingham (1937) stated that a test score is only half the story. Information regarding the
examinee's test-taking behavior will help the counselor make a more accurate interpretation of results. Form #3, Test Behavior Checklist (see Appendix C), was developed after a review of the literature, discussions with test examiners and counselors, observations of test-taking behavior, and several experimental forms.

Most Employment Service tests are administered to groups (10 to 20) and, up to now, the test examiner has had little time for observing and recording test behavior. For the disadvantaged, however, tests are usually given to smaller groups (5 to 10). With the examiner's problem in mind, attention was given to the development of the form by selecting only a limited number (10) of the most important items.

Each item is accompanied by a brief description to assist in identifying visible behavior. The checklist covers observations of ability to follow instructions, use of the separate answer sheet, accuracy, motivation, concentration, understanding, confidence, and doing own work. The rating scale has three levels, high, medium, and low with two intermediate marks, resulting in a possible five-point (Likert) range. Additional comments may be made concerning unusual test behavior on the bottom or reverse side of the sheet. The name of the person administering the test and the one who made the referral to testing are entered on the form. The use of the rating scales was intended to simplify the reporting procedures, with additional provision for comments if desired.

In situations where a proctor or monitor was available, this person would complete the form while the test examiner concentrated on
administration of the exercises. The checklist may be used in any test situation, but would be most valuable when observation time is longer and varied, such as during the administration of the General Aptitude Test Battery which involves $2\frac{3}{5}$ to 3 hours. Space was provided for recording of the GATB scores. This information would be routed to the counselor making the test interpretation and would then be placed in the counselee's case record folder.

In addition to the Test Behavior Checklist, the Test Record Card (E.S. Form 538) and Record of Apparatus Tests (see Appendix C), may also have comments on test behavior. These forms are maintained in the Testing Section, but the remarks may be transferred to the Behavior Checklist for reporting purposes.

**SUMMARY OF CHAPTER**

In this chapter, a model for sequential testing was developed and various procedures utilizing the systems approach were illustrated with flow charts. Test content and process were included so that application of the model could be made in other settings or situations. The model also served as a basis for the experimental and demonstration project reviewed in chapter IV.

The Employment Service and its programs were described as the setting for the model and the study. The various E.S. tests were listed by category, use, and the type of information provided.

Essentially, the model consists of three phases: (1) test planning, selection, and scheduling, (2) preparation for and administration of tests in sequence and according to a plan, and (3) a method
of reporting and recording test results and test behavior. Appropriate forms have been developed, including the Test Planning Worksheet, Test Results Worksheet, and the Test Behavior Checklist.

The Test Planning Worksheet profiles biographical, socio-economic, and interview information about the counselee. By using rating scales, the profile is charted in three ways: positive, questionable, and negative. If the information is predominantly in the positive area, the full range of tests may be administered without great concern for any special preparation for testing or specified sequence of tests but rather according to job or training specifications. The questionable profile alerts the counselor to exercise care in the planning, selection, preparation for, and administration of tests. The negative profile suggests the necessity for considerable caution in the test planning and sequence procedures. A number of situations and activities were described and illustrated with flow charts.

The Test Planning Worksheet is filled out by the counselor with information provided by the interviewer from various forms and records, and the counseling interview. The plotting of the profile determines the sequence of tests or procedures and range, with feedback to the counselor of the results and test behavior. Pre-test orientation and diagnostic testing (achievement and screening exercises) are essential parts of the negative and questionable profiles. Information and results are given to the counselor at each step before proceeding to the next. If deemed necessary, testing should stop or be delayed at points which reflect the maximum abilities or motivation of the counselee.
The Test Behavior Checklist is completed by the test examiner or proctor and supplements the test scores. The three forms are considered worksheets and are not necessarily formal reports. The information becomes a part of the counselee's case record folder to be used by the counselor or referral person in the total assessment and evaluation process.

The model presented in this chapter may be adapted to other settings or situations. The variables would be the test, examinee characteristics, and programs or reasons for testing. The procedures would essentially be the same, with some modification of the Test Planning Worksheet and Test Results Form to accommodate the differences.

Sequential Testing is based on careful planning, selection of tests in an orderly step-by-step manner in accordance with the needs and abilities of the counselee, and special preparation or orientation for testing. The systems approach provides a method of accomplishing this, both in a descriptive and graphic way through the use of flow charts.

There has been a recognized need for allowing more time for the development of test-readiness on the part of the counselee and for improved feedback of information to him in small understandable segments. Hopefully, the model does accomplish this, but it needs to be tried out in the experimental and demonstration project that is reported in the next chapter.
A very important aspect of research or theory development is the experimental phase in which the theory is tried scientifically in order to evaluate its usefulness and decide how it may be put into practice. Counseling and testing research have become quite sophisticated in techniques and methodology, but a common complaint seems to be the lack of research by the practitioner. The purpose of this chapter was to introduce the model for sequential testing experimentally and determine its value for operational use.

A frequently used type of research today is the experimental and demonstration (E & D) project. Sometimes referred to as field or action research, the activities incorporate laboratory studies into a real situation. Cornell and Monroe (1950) described it as a method in which one group (experimental) is treated according to the procedures of the design and a matched group (control) reacts to the conventional procedures. An attempt should be made to identify the variables and determine their effect through observable differences between the two groups. One of the critical problems in this type of study is the identification and limitation of the variables. Another condition that may exist is the "Hawthorne effect" which influences the performance of the experimental group because of motivational factors resulting
from special attention on the part of the researchers.

A relatively recent term in educational studies is "action research" and it was described by Corey (1953) as "research undertaken by practitioners in order that they may improve their practices." The techniques are not unique other than that they use systematic evaluation and applications of theory in actual operational situations. The motivation of the researcher is not only to find the answer to "generalized questions," but also to improve practices in individual situations.

Research has for years been an integral part of the U.S. Department of Labor. It was not until the 60's, however, with the "Manpower Challenge Years," that experimental and demonstration projects were given special attention (Wolfbein, 1967). The Manpower Training Act (1962) included programs for experimental, developmental, demonstration, and pilot projects. Contracts were granted to public and private non-profit organizations for improving techniques and demonstrating the effectiveness of specialized methods to meet manpower needs and solve the problems of unemployment or related factors. Particular emphasis was given to the problems of youth, older workers, and the disadvantaged. Similarly, research appropriations were made under the Vocational Education Act of 1963, the Economic Opportunity Act (1964), and other Federal programs. It is very apparent that this method of research has both a popular and utilitarian appeal. Briefly stated, Cassell (1968), former Director of the U.S. Employment Service, saw this method as "translating research into operations: the potential for change."
The E. and D. method of research was selected for this study for practical and economic reasons as well as for its flexibility in a short-term pilot project that could be replicated at a later date in other situations or agencies. The setting for this study was previously described in Chapter III as the Ohio Bureau of Employment Services and its current programs for the disadvantaged.

The model developed was experimentally tried with enrollees in the Work Incentive Program (WIN). This activity was initiated in 1967 to provide special supportive services and assistance to welfare recipients in developing employability, various types of training, and placement into productive employment. It is jointly administered by local welfare agencies and offices of the Employment Services. Participants receive an incentive payment of $30 a month and other possible training allowances, in addition to their regular welfare grants.

All enrollees are given program orientation in which the services, activities, and conditions of the program are described. Participants who are not ready for referral to work or training because of personal problems, lack of proper attitude, motivation, work habits, or no employable skill are assigned to WIN teams on a case-load basis. The teams are comprised of a counselor, a work-training specialist, a job-development specialist, an enrollee advisor or coach, and a clerk-typist. The team members, with the assistance of other local office staff, conduct the Orientation and Assessment (O. & A.) component.

The Orientation and Assessment consists of structured program activities including personal development (grooming, money management,
and legal aid), job-finding strategy (the interview, job hunt, the
application, supervision, transportation, and job requirements), pre-
testing orientation (test-taking skills, practice on tests, and actual
taking of standardized tests), and finally, work-sampling, a new ex-
perimental form of assessment. Within this schedule, both individual
and group counseling are provided.

The portion of the O. & A. dealing with counseling and testing
(standardized) was the area of exploration for this study. The pre-
test orientation unit was described in detail in Chapter III as five
to seven and a half hours of preparation for and the sequential ad-
ministration of tests. The three-week O. & A. component allows ade-
quate time for careful planning, preparation, and selective adminis-
tration of tests that are not normally found in other Employment Ser-
vice programs.

The organization of this chapter was threefold: (1) to pre-
sent the experimental procedures used, (2) to provide information and
tabulation on data collection, and (3) to analyze and discuss the
findings. In essence, an experimental tryout was made of the model
to refine the procedures, evaluate the practical application of the
various concepts of sequential testing and the systems approach, and
determine what improvements or procedural changes might be suggested.
The experimental and demonstration approach placed more emphasis on
empirical rather than statistical evidence of the practical and opera-
tional value of the model. These are fully discussed in subsequent
parts of this study.
EXPERIMENTAL PROCEDURES

The experimental and demonstration project or model was installed in five local employment offices (Akron, Canton, Cincinnati, Toledo, and Youngstown) that had WIN programs and were conducting Orientation and Assessment components. A six-weeks period was established for the tryout of the model and collection of the data.

The WIN teams were instructed in the study procedures through a training session of approximately an hour and a half. Each office was provided with a specimen kit describing the project, procedures, sample forms, and check lists. A copy of the instructions and training outline may be found in Appendix D. Forms referred to are contained in Appendix C.

The training or instruction session included the entire WIN team. Each team is assigned a case-load of 250 enrollees (on an annual basis). Cincinnati and Toledo have several teams, whereas Akron, Canton, and Youngstown have only one team each.

The presentation to the local office staff included the purpose of the project, procedures, a review of the various forms, and data-collection details. It was felt to be imperative that the participating staff not only understand the procedures but also see merit in the application of the model for operational use. Although some of the concepts and methods were not entirely new to the staff, the organization of the material, the forms, and definite goals or outcomes that could be evaluated by the various team members constituted a new approach.
The WIN program has been typically over-burdened with paper work in the nature of record-keeping and reporting. In light of this, extreme care was taken to see that the various forms and checklists were essentially work-sheets designed to assist the counselor and not a part of formal records or reports. It was stressed that the real value of the forms would be determined through actual use and evaluation by the staff.

Another point emphasized was in regard to the experimental and demonstration nature of the study. Reference to basic research was avoided so there would be no feeling that the staff or enrollees were being manipulated in any way. Rather, weight was given to the fact that this was a pilot or tryout of the model containing some procedures already in use, along with a few innovations, and, thus, this was essentially a revision or reorganization.

The concepts of sequence and systems were explained, although not in great detail. Elementary flow charts were used as visual aids, but every effort was made to keep the information simple and practical without confusing the staff with many theories or technical details. This was necessary in that the backgrounds of the team members varied considerably.

The response to these presentations was good and it was not necessary to "sell" the project in order to gain cooperation of the local office staff. It appeared that the purpose and procedures of the study were clear to them, based on the nature of questions asked and discussion of the project.

Each local office was requested to provide an experimental
group of 10 to 15 persons and a control group of approximately the same number. If Orientation and Assessment were to be conducted more than once in the specified six-weeks period, the sampling could be enlarged or duplicated. The time limitation was imposed by the dissertation requirements.

The experimental group consisted of WIN enrollees who were scheduled for Orientation and Assessment. This activity permitted use of the sequential testing procedures with the various forms as a basis for study data collection and for the case record folder and counseling information.

The control group was made up of WIN enrollees who were not referred to Orientation and Assessment and who were processed according to traditional employment service procedures with reference to counseling and testing. The key factor in the Orientation and Assessment was the pre-test orientation which contains the main elements of sequential testing. These were preparation for testing and step-by-step tests from simple diagnostic to more difficult and demanding batteries, such as the General Aptitude Test Battery.

The Test-Planning Worksheet was completed on each enrollee from existing sources of information and that obtained during the counseling interview where test-taking experiences or attitude toward tests would be brought out. The profile was to be used by the counselor and test examiner as a basis for planning test-selection for the experimental group using the sequential method. In the case of the control group, the Planning Worksheet was to be completed but not used for test-planning or selection. Instead, it was a source of data
for comparison of the experimental and control groups.

The Test Results Worksheet was also filled out for each examinee in both groups, and served as a device for reporting test results to the counselor and as a record of testing for the case record folders. For the purpose of this study, it would indicate the comparative appropriateness of the tests administered according to the Test Planning Worksheet profiles and those given in the traditional way.

The Test Behavior Checklist was completed for each person in the sample groups who took the Metropolitan Achievement Test or the General Aptitude Test Battery. This form revealed test readiness and provided information for comparison of the performance of the experimental and control groups. The assumption was that persons receiving the special attention of pre-test orientation and sequential testing would perhaps exhibit more favorable test-taking behavior than those who didn't and, in turn, would result in improved performance on tests.

A final Evaluation Sheet (see Appendix C) was completed by the counselor on each enrollee using a five point scale on each of five items that related to his test-readiness, test-taking skills, attitude and motivation, understanding, and maximum performance. A duplicate Evaluation Sheet was made out using information obtained from the examinee. The five point scale was based on the Likert procedures (Sax, 1960).

In addition, a three item Interview Questionnaire (see Appendix C), was used by just the counselor and test examiner of each team for an overall appraisal of the model, as to the value of the
procedures, adequacy and practicality of the forms, and the overall effectiveness of the model in the WIN program.

The specimen kit included an outline of the training presentation of the project, a schedule guideline, summaries of procedures and disposition for each form, as well as sample copies. A supply of forms was provided for each team.

Follow-up contacts were made by personal Central Office staff visitations or periodic telephone calls to determine status and progress of the project. A few minor problems were encountered concerning clarification of the forms, supply of additional material and the coordination of various phases of the study based on the proposed time-table.

**DISCUSSION OF FINDINGS**

The completed forms and data sheets for 135 participating enrollees were collected, screened, and sorted into their respective groups. Of these, 16 sets of forms (11.8%) were withdrawn because of incomplete data. The balance included 86 enrollees in the experimental group and 33 in the control. Table 2 shows local offices and the number of staff persons who assisted with the project as well as their contribution to the sample groups.
### TABLE 2

**STAFF PARTICIPATION AND DATA CONTRIBUTION**

<table>
<thead>
<tr>
<th>Local Office</th>
<th>Local Office Staff</th>
<th>Test Counselor</th>
<th>Examiner</th>
<th>Experimental group</th>
<th>Control group</th>
<th>Incomplete data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Akron</td>
<td>1</td>
<td>1</td>
<td></td>
<td>14</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>2. Canton</td>
<td>1</td>
<td>1</td>
<td></td>
<td>19</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>3. Cincinnati</td>
<td>1</td>
<td>3</td>
<td></td>
<td>15</td>
<td>17</td>
<td>4</td>
</tr>
<tr>
<td>4. Toledo</td>
<td>1</td>
<td>1</td>
<td></td>
<td>32</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>5. Youngstown</td>
<td>1</td>
<td>1</td>
<td></td>
<td>6</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Totals</td>
<td>5</td>
<td>7</td>
<td></td>
<td>86</td>
<td>33</td>
<td>16</td>
</tr>
</tbody>
</table>

Canton and Toledo referred all enrollees to Orientation and Assessment and therefore did not contribute to the control group which consisted of persons not taking part in this phase. Instead, the control group was treated according to traditional employment service procedures without the use of the project forms, orientation, and sequential testing.

Numerical values were assigned to items on the forms to facilitate the statistical treatment of the data. These were posted on worksheets and, with the use of a calculator, the sums and the squares of the scores were obtained. From these, the means, and t-ratio were computed, the latter for the purpose of determining the significance of differences between the means. Only select
factors were used from the various forms in an attempt to evaluate key aspects of the study. The statistical procedures are reviewed in Appendix E.

The Test Planning Worksheet provided information for a comparison of matching characteristics of the experimental and control groups. Each of 14 items was rated on a five point scale from 1 (low) to 5 (high). Of these, three items, (1) age, (2) education (grade attainment), and (3) work experience (years) were used to compare the two groups.

Figure 12 illustrates the part of the worksheet used for data collection, the scale values, and the identification of the various points on the scales.

Scale Values

1. Age (years)  
   \[
   \begin{array}{ccccc}
   & 1 & 2 & 3 & 4 & 5 \\
   56+ & 46-55 & 36-45 & 26-35 & 16-25 \\
   \end{array}
   \]

2. Education (grade attainment)  
   \[
   \begin{array}{ccccc}
   0 & 1-5 & 6-8 & 9-12 & College \\
   \end{array}
   \]

3. Work Experience (years)  
   \[
   \begin{array}{cccc}
   0 & 1 & 2 & 3 & 4+ \\
   \end{array}
   \]

Figure 12 - Scale Values on Test Planning Worksheet
Table 3 indicates that the experimental and control groups were fairly well matched. There were no significant differences between the groups as measured by the three factors. The small size of the control group was a result of the fact that most enrollees are referred to orientation and the limited time allotted to the study did not permit any delays.

**TABLE 3**

SAMPLE CHARACTERISTICS

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Sex (M/F)</th>
<th>Age (M)</th>
<th>Sex %</th>
<th>Value Grp. Education (M)</th>
<th>Value Grp. Work Experience (M)</th>
<th>Work Experience Years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>86</td>
<td>M-2 F-84</td>
<td>97.6%</td>
<td>4.023</td>
<td>3.860 9-12</td>
<td>3.012</td>
<td>2</td>
</tr>
<tr>
<td>Control</td>
<td>33</td>
<td>M-2 F-31</td>
<td>93.9%</td>
<td>3.939</td>
<td>3.878 9-12</td>
<td>3.303</td>
<td>2</td>
</tr>
<tr>
<td>t-ratio</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.444</td>
<td>.286</td>
<td>.967</td>
</tr>
</tbody>
</table>

The ratio of males to females in the groups was quite similar, reflecting statewide averages which show that most welfare referrals have been females on Aid to Dependent Children (ADC) programs. Three other characteristics also compared at about the same level within the group. These were age (approximately point 4, 26-35 years), education
(slightly below point 4, grades 9-12), and work experience (point 3, 2 years).

The age group (26-35) is indicative of the fact that many of the women have school age children, thus making them available for work now, although the lack of day child care for after-school hours and child illness is one of the major reasons for absence or dropout in the program. The mean for education is just short of point 4 on the scale, which would mean that many of the enrollees were educationally deficient according to the U.S. Department of Labor standards (less than a 9th grade education). This condition was supported by achievement test scores (reading and arithmetic). The need for basic and remedial education for many of the enrollees was all too apparent and this also pointed to the fact that tests such as the General Aptitude Test Battery might not be appropriate because of the heavy loading with reading and math skills (above the 6th grade level).

The mean years of work experience was about two. Most enrollees had intermittent work and usually of an unskilled or semi-skilled nature. This was not surprising since lack of employable skill has been correlated with low educational attainment.

Although other factors on the Planning Worksheet were given consideration, in the total use of the form, the second area evaluated was in regard to previous test-taking experience. Figure 13 is a portion of the worksheet that shows the items, the scale values, and describes the various points on the scale.
Table 4 is a tabulation of the numerical values for the two groups. Again, a similarity was reflected in the fact that no significant differences were found. The first mean value was approximately 2, which signified that 1 to 2 tests had been taken by the enrollees about a year ago (point 4+). The attitude was close to point 4 on the scale, denoting an "indifferent" (plus) attitude towards tests. The data on this part of the worksheet also supported the fact that the experimental and control groups were fairly well matched.
### TABLE 4

**PAST TEST EXPERIENCE**
*(Test Planning Worksheet)*

<table>
<thead>
<tr>
<th>Group</th>
<th>No. Tests Taken (M)</th>
<th>Years Ago Taken (M)</th>
<th>Value Yrs. Ago (M)</th>
<th>Attitude Value Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>1.872</td>
<td>1-2</td>
<td>4.560</td>
<td>1+</td>
</tr>
<tr>
<td>Control</td>
<td>2.062</td>
<td>1-2</td>
<td>4.865</td>
<td>1+</td>
</tr>
<tr>
<td>t-ratio</td>
<td>.118</td>
<td>1.333</td>
<td>1.227</td>
<td></td>
</tr>
</tbody>
</table>

The three categories on the Test Planning Worksheet, Negative, Questionable, and Positive, were assigned values of 1, 2, and 3. Figure 14 illustrates the categories with a part of the scales that were assigned values within them of 1 to 5. These procedures were used to facilitate the computation and statistical treatment of the data.

<table>
<thead>
<tr>
<th>Category values</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Questionable</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scale values</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>56+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>46-55</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>36-45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26-35</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16-25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Figure 14 - Category and Scale Values*  
*(Test Planning Worksheet)*
Table 5 shows the totals for the Test Planning Worksheet based on the category values and also the classification of the profiles as determined by visual inspection. The procedure used was to screen the worksheet for the greatest number of scale values marked within a particular category. Values in each category, assigned 1 to 3 points, were added and the three categories then totaled. From these figures, means were obtained with which to compare the two groups.

**TABLE 5**

**TOTAL VALUES BY CATEGORIES**
(Test Planning Worksheet)

<table>
<thead>
<tr>
<th>Group</th>
<th>Total Values by Category</th>
<th>% Category Classifications</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Negative</td>
</tr>
<tr>
<td>Experimental (N-86)</td>
<td>25.360</td>
<td>14%</td>
</tr>
<tr>
<td>Control (N-33)</td>
<td>23.121</td>
<td>42%</td>
</tr>
</tbody>
</table>

| t-ratio | * 3.20 |

* Significant at the .01 level

The t-ratio (Ferguson, 1966) was used as a method for determining whether or not the differences between the means were significant, (see formula and table reference in Appendix E). This statistical procedure is frequently used when sample sizes are relatively small but may be used with sample sizes of 50 to 100
(Yuker, 1958). The means of the total scores on the Test Planning Worksheet were found significantly different at the .01 level. The interpretation of this data somewhat surprisingly reveals that the control group, which was assumed to be test ready, scored lower in the profile categories and was actually in greater need of careful test planning than the experimental group. This result might be due in part to the small sample size. Nevertheless, the category classification showed 42% in the negative area for the control group and only 14% for the experimental. As far as the questionable category was concerned, the percentages were 58% to 85%. In the positive, the classifications were 0 and 1%. This may point to the fact that more of the disadvantaged are being processed by the Employment Services who need special help in testing than is being offered by traditional procedures.

The criteria of test appropriateness consisted of the types of tests administered and their sequence. Pre-test orientation and screening as well as diagnostic tests were considered appropriate for enrollees in the negative and questionable categories. For persons in the positive area, no special order or selection of tests was necessary other than job order specifications or past experience or training of the applicant. Table 6 is a summary of the number of tests administered and the percentages of appropriate and inappropriate testing.
A significant difference (at the .01 level) was found in the number of tests administered. The experimental group was provided pre-test orientation including, in most cases, a sequence of screening and diagnostic tests. Fewer aptitude tests were given and, when they were, wider use was made of the B-1001 form of the General Aptitude Test Battery which uses expendable test booklets rather than separate answer sheets. The control group, on the other hand, was most frequently given the B-1002 (separate answer sheet form) of the GATE. Too often, this proved to be inappropriate due to low educational attainment or poor reading skills, resulting in extremely low scores. The Interest Checklist was also administered to part of each group (to nearly all of the control), but it too is heavily loaded with verbal factors and requires special reading skills. No use was made of the screening exercises or the non-reading measure of "G" for

### Table 6

**NUMBER OF TESTS ADMINISTERED AND APPROPRIATENESS**

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Number of E.S. Tests Administered</th>
<th>(M)</th>
<th>% Appropriate</th>
<th>% Inappropriate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>86</td>
<td>3.244</td>
<td></td>
<td>96.5%</td>
<td>3.5%</td>
</tr>
<tr>
<td>Control</td>
<td>33</td>
<td>1.242</td>
<td></td>
<td>75.8%</td>
<td>24.2%</td>
</tr>
</tbody>
</table>

\[ \chi^2 \] * 9.88

*Significant at the .01 level
(Yates correction applied)
the control group.

Tests for 96.5% of the experimental enrollees were found to be appropriate whereas 3.5% were not appropriate. For the control group, 75.8% of the individuals were given appropriate tests and 24.2 inappropriate. The differences were significant at the .01 level. Due to the small and uneven sample size, the Yates correction for continuity (Ferguson, 1966) was applied in the statistical calculations. (see Appendix E).

The data on the Test-Behavior Checklist was treated in a manner similar to that on the Planning Worksheet. The points on the scale were given values of one to five (low to high) so that, with ten items on the list, the possible range of scores would be from 5 to 50. Table 7 is a tabulation of the means for the total scores on the checklist.

The experimental group was scored significantly higher (at the .025 level) on test behavior than the control, as measured by the Test Behavior Checklist. This indicated that pre-test orientation and the proper sequence of tests is reflected in improved test-taking behavior.
TABLE 7
TOTAL SCORES - TEST BEHAVIOR CHECKLIST

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>(M) Total Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>86</td>
<td>36.205</td>
</tr>
<tr>
<td>Control</td>
<td>33</td>
<td>32.968</td>
</tr>
<tr>
<td>t-ratio</td>
<td></td>
<td>*2.160</td>
</tr>
</tbody>
</table>

* Significant at the .025 level

The fact that the control group had a higher percentage of inappropriate testing was possibly a contributing factor to the differences between the two groups.

The Evaluation Sheet (Form #4) consisted of five items with a five-point scale similar to that used on the Test Behavior Checklist, providing a possible range of 5 to 25 for a total score. The data on these forms were incomplete for the counselees and there may be some question about the value of this part of the study. Table 8 shows the means of the total scores on the evaluation sheets.
TABLE 8
TOTAL MEAN SCORES - EVALUATION SHEETS

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Total Scores</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Counselor</td>
<td>Counselee</td>
</tr>
<tr>
<td>Experimental</td>
<td>86</td>
<td>19.225</td>
<td>18.879</td>
</tr>
<tr>
<td>Control</td>
<td>33</td>
<td>16.937</td>
<td>Incomplete data</td>
</tr>
<tr>
<td>t-ratio</td>
<td></td>
<td>*2.009</td>
<td></td>
</tr>
</tbody>
</table>

* Significant at the .025 level

The Evaluation Sheet was designed to assess test-readiness as determined after testing by both the counselor and the counselee. The differences between the mean scores of the counselor and the counselee in the experimental group were not significant. Incomplete data from the counselees in the control group prevented this type of comparison. However, there was a significant difference (at the .025 level) in the mean scores of the counselors in the two groups. The counselors in the experimental group rated their enrollees higher on test-readiness than did the counselors in the control group. This is not surprising because the variable for the experimental group was orientation and special instruction in test-taking which was not provided the enrollees in the control group.

A final evaluation of the model was made by interviewing each
test examiner and counselor who participated in the study. Three questions were asked and it was requested that responses be made on a five point scale, as illustrated in figure 15.

Scale Values

1 2 3 4 5
low median high

Figure 15 - Scale for Model Evaluation by Test Examiners and Counselors

The questions were as follows:

1. How would you assess the value of the procedures in making tests and testing more valid and meaningful to the counselor and counselee?
2. How practical and useable were the forms?
3. How would you rate the overall model of sequential testing?

The purpose of this final evaluation was not only to get responses to the questions from the persons most closely related to the study but also to obtain opinions about the various aspects and details of the model. These comments are reported in the Summary and Conclusions chapter.
Table 9 is a summary of the mean values of the responses of the counselors and test examiners. The latter responded with higher ratings to all three questions than did the counselors, but the differences were not significant. Both groups rated the forms as having the greatest value. The overall rating for the project was slightly above the median point on the rating scale.

**TABLE 9**

**MODEL EVALUATION BY COUNSELORS AND TEST EXAMINERS**

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>1-Procedures</th>
<th>2-Forms</th>
<th>3-Overall</th>
<th>(M) Total Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Counselors</td>
<td>7</td>
<td>3.428</td>
<td>3.571</td>
<td>3.428</td>
<td>3.475</td>
</tr>
<tr>
<td>Test Examiners</td>
<td>5</td>
<td>3.600</td>
<td>3.800</td>
<td>3.800</td>
<td>3.766</td>
</tr>
<tr>
<td>t-ratio</td>
<td></td>
<td>.514</td>
<td>.319</td>
<td>.378</td>
<td></td>
</tr>
</tbody>
</table>

**SUMMARY**

In summary, the significant findings of the experimental and demonstration project were as follows:

1. Select comparison factors were age, education, work experience, and test-taking experience which indicated that the experimental and control groups were closely matched. Total evaluation of the
Planning Worksheet, however, revealed that the control group, (much smaller than the experimental) had greater need for careful planning and preparation for testing. There was a significantly lower mean score and a greater percentage of negative sheets in the control group. The special import of these facts appears to be that enrollees for whom orientation is not considered necessary, according to standard Employment Service procedures, actually do need special help in testing.

2. The number of inappropriate tests administered to the control group (24%) was significantly greater than the number to the experimental group (4%). The mean of the number of tests administered to the experimental enrollees was significantly greater than that (mean) for the control, but they were used more diagnostically and were provided in a more logical sequence.

3. There was a significant difference on test behavior as assessed by the Checklist. The experimental group performed more favorably than the control. This was probably because of the pre-test orientation and careful scheduling and administration of appropriate tests.

4. Test Examiners gave the model a higher evaluation than counselors, although the differences were not significant. The overall value of the model was slightly over a median rating.
CHAPTER V

SUMMARY AND CONCLUSIONS

The idea of the need for this study was generated by the mounting criticism against the misuse of standardized tests with disadvantaged persons. The inadequate selection of appropriate tests, lack of preparation for testing, and general overuse and misuse of tests are only a few of the serious indictments of tests and testing procedures. Some professional organizations have suggested that the use of standardized tests be severely curtailed, if not abandoned, until more adequate instruments are developed along with better testing procedures and greater competence in the interpretation of test results. In an attempt to supply some answers for improving this situation, methods, forms, and procedures were developed during this study to assist the counselor and test examiner in making better selection and use of tests.

The dissertation research was undertaken within the limitations of Employment Service tests, programs, and staff, with six weeks allowed for data collection in the experimental and demonstration project. Emphasis in the study was placed on disadvantaged persons and their exposure to tests.

One basic assumption was that information obtained from tests can be important to counselors and that certain background and
biographical data can also be of value in decisions about testing. Relative to this, it was noted that disadvantaged persons often lack test-taking skills or experiences, but these could be learned if proper and adequate instruction were provided. Finally, it was assumed that the administration of tests in sequence and a logical order would result in improved test-taking behavior and maximum performance on tests.

A review of the literature revealed that, although the concept of sequential testing was not new, it had not been fully explored or developed. Likewise, pre-test orientation had been used only to limited degree with the disadvantaged, but, in most cases, showed a favorable effect on test performance. Test-taking behavior was studied as early as 1934, according to several writers. Still, little or no provision had been made for the recording and reporting of this type of information. As for the systems approach, it has only recently been applied in settings other than engineering and management. For example, the use of the computer and data processing has developed a new technology for its application in the areas of education, the behavioral sciences, and general problem-solving. Essentially, it consists of input, process (or procedures), and output. Flow charts are commonly used to graphically illustrate various models and activities. A system consists of a logical step-by-step progression for product development, problem solving, or for the attainment of specific goals.

The model developed for sequential testing was for disadvantaged persons in the Work Incentive Program, using Employment
Service Tests and procedures. A key factor was the Orientation and Assessment component which included pre-test orientation. Special forms were developed to facilitate the recording of essential data. The first was a Test Planning Worksheet designed to assist the counselor in profiling certain characteristics of the individual that would help in the selection of appropriate tests for him, their sequence, and procedures to be used. The Test Results Worksheet was a sequential listing of all Employment Service tests with spaces for inserting results, dates, persons recommending tests, and comments. This was later to be available for quick reference in the enrollee's case record folder. A third form, The Test Behavior Checklist, provided the counselor with additional information from observations noted in writing by the test examiner of the counselee during testing. This data supplemented the test results and helped the counselor in test interpretation.

FINDINGS

The model was introduced for tryout in five local offices in the form of an experimental and demonstration project. Data was collected for a sample of 86 enrollees in the experimental group. They were treated with the sequential procedures and were exposed to pre-test orientation. The control group consisted of 33 persons who did not participate in orientation and for whom sequential procedures were not used, although data was collected on the same forms. For them, the traditional methods of test selection and administration were followed.
Statistical analysis of the data (means and t-ratio) indicated that the two groups were closely matched as far as four select characteristics were concerned. However, the total values on the Test Planning Worksheet reflected a significant difference on the mean category scores. The lower score of the control group could be interpreted to mean that these enrollees were in as great a need (or even greater) of test orientation, screening, and careful test selection as the others. Unfortunately, these were not considered necessary for them. The percentage of persons tested inappropriately was 24% for the control group contrasted to only 4% for the experimental. This is probably the most significant aspect of this study. With the use of the forms and sequential testing procedures, inappropriate testing is less likely to occur.

Another positive outcome was the fact that the experimental group performed significantly better, as assessed by the Test Behavior checklist, than the control group. The project evaluation by counselors of both groups was a median (value 3 on a five-point rating scale) score, with no significant differences. The test examiners, however, rated the model slightly higher on the scale.

On the basis of both empirical and statistical evidence, the following hypothesis (as stated in Chapter I) were confirmed:

1. If the Test Planning and Results Worksheets are used in counseling and testing, then test selection and administration will be more appropriate.

2. If sequential testing procedures are used with disadvantaged persons, then their test-taking
behavior will improve.

These conclusions are, of course, founded on the limitations of the data collected and analyzed in the experimental and demonstration project phase of this study.

RECOMMENDATIONS

Some recommendations were obtained from two sources; namely, personal interviews with the counselors and test examiners who participated in the actual tryout of the model and their opinions and remarks as recorded on an Interview Questionnaire from (see Appendix C). As might be expected, these were both favorable and unfavorable, hopefully, the positive aspects outweighing the negative.

Individual interviews with the participating counselors and test examiners brought out the following opinions and suggestions:

1. A general negative attitude toward the forms due to what they felt was excessive paperwork, already a problem with numerous required records and reports.

2. There was a duplication of information that was needed for other records.

3. Clarification of some of the descriptive data and instructions with reference to some of the forms would be advisable.

4. The Test Behavior Checklist was considered the most unique and valuable form.

5. Several counselors expressed the idea that more thorough training in the procedures and proper
use of tests would be helpful. It was suggested that this should be a part of regular Employment Service counselor training.

6. The interviews indicated a greater need for cooperation between the counselor and test examiner. Differences in background and training can, at times, have a detrimental effect on the two functions.

Based on the research findings of this study, the following conclusions and recommendations are made:

1. There is a definite need for improved procedures for the selection and administration of tests with disadvantaged persons.

2. Counselors and test examiners should be trained more thoroughly in the testing and assessment processes. This should begin with more information regarding the technical background of tests and include specific instruction in determining appropriateness of tests and in the methods of observing and recording test behavior.

3. It is essential to overcome the bias against the use of forms and records. It is true that there is at present an over-abundance of this material and, unfortunately, the data is often poorly organized and is not useable. A few well-designed, efficient forms could be a satisfactory answer to these criticisms.

4. Local office staff must recognize the need for the
procedures of sequential testing and be sold on the idea that they can work. The investment of time with the forms and procedures will pay dividends in eliminating needless or inappropriate testing. More important, it can mean improved test behavior and maximum performance.

5. Elaborate and technical details of systems theory should be avoided. Instead, emphasis should be on the simplicity of the model and its practical nature.

6. The pilot study conducted in the experimental and demonstration project appears to support the position that the model has value and is workable. There should be further research, however, for improvement and validation.

NEED FOR FURTHER RESEARCH

A study of this type would not be complete without observations regarding suggested additional research. In effect, what was developed here was merely a pilot study of a limited nature. Replication of the study or parts of it should be undertaken with larger samples and perhaps in a variety of settings. The following recommendations are made for molecular research on specific areas or parts of the model:

1. The forms should be carefully reviewed and item analysis of the Test Planning Worksheet would validate the various parts of the worksheet.

2. A correlational study could be made with the U.S.
Department of Labor's Measure of Cultural Exposure (1968) and the Test Planning Worksheet.

3. The Test Behavior Checklist should be refined with improvements of the descriptive scales and instructions for use. Item analysis could perhaps reduce the number of factors to the most significant ones.

4. More detailed research is needed on pre-test orientation to actually determine its effectiveness on test performance and scores.

5. Various types of orientation should be researched to provided for the different needs of individuals.

6. New methods and procedures for screening should be developed to help make the decisions on test-readiness, not only to assist the counselor and test examiner but the counselee.

7. A method should be devised to determine the type and amount of orientation that would best meet the requirements of individuals in different categories.

8. Counselors and test examiners should be encouraged to participate in research on a local level. They should be recognized openly for their suggestions and ideas, which, in turn, should be made available to other staff persons.
CONCLUSION

It was hoped that the counselors and test examiners who participated in this study would recognize the real need for improved procedures in testing, such as the ones described in this model for sequential testing. Although they gave the project a slim margin of approval, there was a disappointing lack of enthusiasm for what it could do. They seemed to be unaware of the universally poor testing procedures being carried on with the disadvantaged in the local offices and also of their own ineptness and lack of professionalism. It is becoming apparent that the misuse of tests with disadvantaged and minority groups has come under close scrutiny. Failure on the part of counselors and test examiners to heed the warnings that are coming from very responsible persons in the field could well mean the abolishment of standardized testing with the disadvantaged.

As tests have become one of the most important techniques for assessment to the counselor, they should be worthy of studied and proper use. The true mark of any craftsman is the way in which he uses his tools and the care and respect he gives them. In a sense, the counselor is a craftsman and the tests are a part of the tools of his trade. He should strive to improve his skill and competence in using them so the product of his efforts reflects his pride and knowledge.

The pursuit of culture-free or fair tests for the disadvantaged has been all but abandoned as impossible attainments. In their place, proper selection of tests, adequate orientation, and sequence
of tests may be a partial answer to fairer testing. Education and testing cannot escape technological change and the idea of the systems approach, as applied to this study, may provide some of the solutions to the problems of testing.
APPENDIX A

GLOSSARY OF TERMS

**Achievement Tests** (reading and arithmetic) - tests that measure the extent to which a person has acquired certain information, knowledge, or skill, usually as a result of special instruction. Often reported as scholastic (grade level) performance.

**Apparatus Tests** - tests that utilize pieces of equipment or instruments to measure performance. In the case of the GATB, finger (rivets and washers) and peg boards measure finger and manual dexterity.

** Appropriateness of Test Selection** - the suitability or compatibility of a test in terms of the examinee's abilities, test readiness, and need or purpose of testing.

**Aptitude Tests** - tests designed to measure an individual's potential for acquiring specified skills and knowledge. A measure of present characteristics that have been found to be predictive of capacity to learn.

**Bennett Mechanical Comprehension Tests** - designed to measure ability to understand mechanical relationships, these tests consist of drawings with simply phrased questions about them.

**Blackbox** - a formal model used in formulating hypothetical constructs: given a certain input, what must be hypothesized as taking place in the blackbox to account for the output.
Case Record Folder — a collection of data (forms, records, reports) that may be used for the evaluation, planning, disposition of the individual or his referral to training, treatment, or job placement.

Checklist — a record form on which to indicate how often certain items have been observed. The list is prepared beforehand and serves as a reminder of what is to be watched for.

Clerical Skills Tests — proficiency tests that are standardized and validated for clerical occupations and measure skills acquired through experience or special training. Included are typing, dictation, and spelling.

Coaching — to train someone intensively by instruction, demonstration, and practice.

Color Vision — the person's ability to differentiate between various colors as measured by tests (see appendix B).

Control Group — a group as closely as possible equivalent to an experimental group, and exposed to all the conditions of the investigation except the experimental variable or treatment being studied.

Controlled Administration of Standardized Tests (CAST) — a method of administering tests with specialized mechanical equipment using recorded instructions, console controls, multiple administration, automatic timing, and machine recording and scoring of results. Developed and leased by the Psychological Corporation.

Cultural Bias — factors of race, ethnic origin, environmental exposure, religion, or patterns of social behavior that may influence performance on a test but were not characteristic of the sample population upon which the tests were standardized.
and are standardized and validated for clerical occupations. Included are typing, dictation, and spelling tests.

**Color Vision Test** - American Optical's pseudoisochromatic color plates. Background of various colored dots form geometric designs when there is ability to differentiate between colors. This measure is important for specific occupations, such as cable splicer, painter and decorator, or in the printing trades.

**Dictation Tests** - clerical skills tests that measure ability to take and transcribe dictation. Administration is at predetermined speeds of 60, 80, and 100 words per minute by "live" or standardized recordings. A short practice exercise is provided to determine level of ability to take exercises at various speeds. Examinees mark a true-false answer sheet in transcribing that yields an accuracy score.

**General Aptitude Test Battery** - a multifactor test measuring aptitudes (potential ability to learn) as they relate to fields of work (occupational aptitude patterns) or specific occupations (specific aptitude test batteries). Twelve exercises measure the following nine basic aptitudes: "G" - mental or learning ability, "V" - verbal aptitude, "N" - numerical aptitude, "S" - spatial aptitude, "F" - form perception, "Q" - clerical perception, "K" - motor coordination, "P" - finger dexterity, and "M" - manual dexterity. Parts 1 to 7 are paper-and-pencil exercises and 8 to 12 are apparatus (finger and peg boards). The B-1001 (earlier edition) uses expendable booklets, in which the answers are recorded. The B-1002 uses separate answer sheets and reusable booklets. B-1002 consists of two forms, A used for initial testing and B (alternate) used for retesting and release
Culture Free (fair) Tests — a test of general ability from which have been eliminated, as far as possible, all items depending upon experiences that are more commonly found in one culture than another. Such tests must eliminate language and the information or skills selectively employed in one culture more than others. Cattell's test used as a non-reading measure of "G" (see appendix B).

Cutting Score — a dividing point in a series of scores that have been ranked in order of magnitude so that, for a stated purpose, all above that point receive one treatment, all below another.

Depressed Scores — test scores that have been reduced because of the influence of some variable that may not have been considered in the standardization of the test.

Diagnostic Tests — are designed to locate the particular source of a person's difficulties in learning, especially school subjects, thus providing clues to what further measures of instruction, guidance, or study are needed.

Disadvantaged — a classification (U.S. Department of Labor) of individuals in need of special services or referral to programs. Criteria may include race, age, education, culture, socio-economic and employment status. Specific guidelines have been established for each area.

Employability Development — specific services or programs that include counseling, assessment, training, and referral to a job. Emphasis is on the factors of attitude, motivation, work habits, vocational choice, interest, job skills, and other items that may lead to obtaining and holding an appropriate job.
Employment Service (Federal-State) - a network of public local offices operating under a state-federal (U.S. Department of Labor) guidelines. Emphasis in programs is on job placement, training, and employability development.

Enrollee (Counselor, Examinee) - an individual who has been formally registered and accepted for a program or services provided by the Employment Service.

Environmental (Cultural) Exposure - a sum of the external conditions, factors, and experiences that influence or are capable to influencing the functioning of an individual.

Evaluation Sheet (E. & D. Project) - a form developed for the experimental and demonstration project to determine the effectiveness of the model in terms of test selection, preparation for testing, and adequacy or reporting test behavior and results.

Experimental and Demonstration Study - a research method of trying out and implementing a model through simulation or actual operation to determine its effectiveness and how the procedures may be modified or improved.

Experimental Group - those subjects who are exposed to the experimental variable and whose performance will therefore reflect the influence, if any, of the variable.

Feedback - a principle or method of reporting or returning information of results in a system to a previous step. The return to the input of a part of the output of a system or process, the partial reversion of the effects of a process to its source to a preceding stage.
Flow Chart - a diagram or outline showing progress of material or activity through a process or a complicated system of activities.

GATB Screening Exercises - a short series of practice items relating to two parts of the GATB that are individually administered by a counselor. Used to determine if the applicant can take the full General Aptitude Test Battery.

General Aptitude Test Battery (GATB) - a multifactor aptitude test battery validated to predict success in training or learning situations for fields of work or specific occupations (see appendix B).

Group Process - a general term referring to the procedures through which a group approaches, attacks, and solves a common problem. Used as a method in pre-test orientation.

Input - the energy entering a system from without. Information fed into a system or process.

Interest Checklist - an interviewing or counseling aid used to determine interest by preferences for work activities. It is not scored but information about areas with most checks is explored (see appendix B).

Likert Procedure - a method of constructing and scoring rating scales using a three or five step range.

Mean (arithmetic) M or X - a measure of central tendency calculated by dividing the sum of all the values by the number of cases in a statistical series.

Minnesota Clerical Test - a commercial test consisting of number comparison and name comparison, both of which emphasize perceptual speed. Norms established are based on studies of clerical workers and stu-
Model - a small copy of the real thing. That which is to be copied (an ideal or perfect form). A method of showing how something works and displays the relationship of parts to a whole. May be a system of relationships and operating principles of a given empirical theory that may be copied.

Moderator Variable - factors such as culture, race, education, ethnic origin, environmental exposure, or experience may influence test performance and, because of differences, would require adjustments in the interpretation of scores.

Multi-battery Scoring - a method of scoring a test, such as the GATB, for several specific aptitude test batteries. The GATB can be scored for 36 occupational aptitude patterns and for over 500 specific occupations.

Multifactor Test - a test, usually in a battery, that measures several factors. Test of this type are developed through factor analysis, which is a statistical procedure of determining what each part of a test measures and to what degree.

Negative Profile - a profile category, as may be charted on the Test Planning Worksheet, indicating that existing standardized tests, may not be appropriate for an individual or that testing should be limited and administered with careful planning and extreme caution.

Non-reading Tests - tests or exercises that do not require reading in order for the person to do them. Directions are read to the individual by the test examiner. Illustrations are Non-reading "G" and
parts 3, 5, 7, and 8 through 12 of the GATB (see appendix B).

**Occupational Aptitude Patterns (OAP's)** - a grouping of occupations into families based on aptitudinal requirements, as determined by studies of the General Aptitude Test Battery and Specific Aptitude Test Batteries.

**Oral Trade Questions** - a series of questions that measure the amount of information or knowledge that an experienced or trained person may have about a specific occupation (see appendix B).

**Orientation and Assessment (O. & A.)** - standardized instructional component of the Work Incentive Program that offers learning experiences that may help the enrollees in employability development. These include grooming, transportation, money management, job-finding strategy, pre-test orientation (including standardized testing), counseling, and referral to training or a job.

**Otis Test** - a test that measures general learning ability (intelligence). It is used in some employment and educational settings.

**Output** - that which, or the amount of that which, may be produced by a system in a given period of time. The information or end product which is fed out by a system or process.

**Paper-and-pencil Tests** - exercises that require responses to material contained in booklets with the recording in the booklet or separate answer sheet.

**Positive Profile** - a category on the Test Planning Worksheet that indicates the person's ability to handle a full range of employment tests without any special preparation or need for any specific order
of tests.

Practice Exercises – short simple test items that are a sample of parts of tests that may be used to acquaint a person with test content and also to determine if the person comprehends directions and is capable of handling the test.

Pre and Post-Testing – tests administered before and after intervening variables are introduced (such as special instruction). The purpose is to determine if there is any change in performance. Alternate forms should be used in post-testing to avoid practice effect.

Pre-test Orientation – a structured instruction or preparation for persons who may lack test-taking experience or are lacking in test-taking skills. The main objectives are to reduce anxieties and develop understanding of reasons for tests, thereby affecting attitude and motivation. An attempt is made to remove certain biases and provide more valid and meaningful performance on standardized tests.

Process – something going on marked by change that leads toward a particular result. A series of actions or operations conducing to an end.

Proctor – an individual designated to assist in test administration, checking practice exercises, time-keeping, and in maintaining standard procedures. May be assigned responsibility of observing and recording test-taking behavior.

Proficiency and Trade Knowledge Tests – measure acquired skills or knowledge (through training or experience) as they may be important for performance in specific jobs or fields of work.

Profile – a representation of something in outline by a curved or
irregular line, such as a chart.

**Questionable Profile** - a category that may be charted on the Test Planning Worksheet indicating an area between negative and positive. Tests should be administered with caution, careful planning, and sequence in accordance with the individual's abilities and needs.

**Rating Scales** - a device by which a rater can record, for the case in question, the estimated magnitude of the trait or quality rated. An attempt is made to provide procedures designed to make the estimate more careful and objective.

**Referral** - the act of sending or directing, as to training or a job.

**Reliability** - the extent to which a test is consistent in measuring whatever it does measure; dependability, stability, relative freedom from errors in measurement. It usually is estimated but may also be determined statistically.

**Response-Set** - a set to follow a certain pattern in responding: e.g. to alternate responses to true and false or to produce a pattern of responding according to what an examinee thinks is expected or desired.

**Sampling (sample)** - persons who, for a specific reason, are considered as a group or set, upon which a procedure or experimental program is tried for purposes of study.

**Sequential Testing** - a method of determining the effect of further data upon the level of statistical significance, in order to decide when the point has been reached where the addition of further data is probably unnecessary. The following of one test by another in a
logical order.

**Specific Aptitude Test Battery (SATB)** - a group of tests (with the GATB as the "core" battery) developed by experimental research and designed to measure the aptitudes and amounts needed to learn to perform the tasks involved in a given occupation or a small group of related occupations.

**Standard Deviation** - a measure of the dispersion or variability of a whole distribution. A statistic used to express the extent of the deviations from the mean for the whole distribution.

**Standard Error of Measurement** - indicates how closely the individual score compares with the true score. A measure or an estimate of the sampling errors affecting a statistic.

**Subsystem** - a further division of a system that is broken down into elemental steps which feed into the main system.

**System** - the set of orderly and persisting interrelations between parts of a whole. All the elements that work together to perform a given function. Generally made up of a number of interlocking theories putting sharp emphasis upon orderly and logical structure or arrangement.

**t ratio** - a statistical method for determining the significance of the differences between two means. Usually used when the sample size is small (between 30 to 50).

**Test** - a systematic procedure for comparing (measuring) the behavior of two or more persons. Any measurement that yields quantitative data for judgment.
Test Behavior Checklist – a ten-item checklist utilizing a five-point scale (low to high) for recording of test-taking behavior by the test examiner or proctor (see appendix C). Also provides space for recording and reporting test scores.

Test Orientation Procedures (TOP) – a familiarization procedure with two sets of practice tests to help inexperienced job applicants learn how to take tests. Marketed by Psychological Corp.

Test Planning Worksheet – a ten-item worksheet for charting information of biographical, cultural, socio-economic, and test-taking experience. Five-point scales provide profiling into three categories: negative, questionable, and positive. Reverse side of sheet has instructions for item recording and interpretation.

Test Point – a point or stage in system that is used to evaluate or measure status in terms of continuance or return to a previous stage in the system.

Test Readiness – preparedness to respond or react to tests and testing. A state or condition of the person where he is ready to perform at his maximum ability because of factors such as attitude, motivation, and test-taking skills.

Test Results – test scores or interpretation thereof in terms that are meaningful to the individual (counselor, interviewer, or examiner). May be supplemented with additional information in comments.

Test Results Worksheet – a single page (10 items) for the recording and reporting of test results. Worksheet becomes a part of case-record folder with complete record of tests on a single sheet. Entries
are arranged in sequence in which they may normally be administered.

Test-taking Behavior — observable (visible) performance of examinee
during pre-test orientation or during standardized testing. Checklist
is designed to be used with standardized tests to supplement test
scores with additional information.

Validity — the extent to which a test measures what it purports to.
Can be determined through statistical treatment of data. Various
types of validity include face, empirical, content, item, predictive,
construct, and concurrent.

Wonderlic Personnel Test — a shortened form of the Otis-self adminis­
tering test for mental (learning) ability. Frequently used for
employment testing.

Work Incentive Program (WIN) — a program funded by the U.S. Department
of Labor for employability development and job placement of welfare
recepients. Enrollees receive a $30. a month incentive payment while
participating in a program of Orientation and Assessment, on-the-job
training. Referrals are made by local welfare agencies.

Worktable — form used in tabulating data for statistical treatment.
APPENDIX B

ANNOTATED LIST OF EMPLOYMENT SERVICE TESTS

Employment Service Testing Program (E.S. Manual, Part II, 9010)
E.S. tests are a standardized means of measuring an individual's possession of, or ability to acquire, job skills and knowledge. Tests which measure job potential are called aptitude tests and used primarily in connection with jobs which require no previous training or experience; those which measure job proficiency are called proficiency tests and are used in connection with jobs open only to workers who are qualified by previous experience or training. E.S. tests supplement information obtained in interviews concerned with classification, selection, placement, or employment counseling.

Achievement Tests (Reading and Arithmetic) – measure levels of ability for basic skills in reading and arithmetic. The Ohio agency uses the Metropolitan Achievement tests in the Elementary form (grades 3-4) and Intermediate (grades 5-6). These tests are used to determine grade level at which the enrollee is functioning and possible need for basic or remedial education, as well as ability to handle tests that require specific reading skills.

Clerical Skills Tests – trade knowledge or proficiency tests that measure acquired skills (through experience or special training)
to outside agencies.

**Interest Checklist** - an interviewing and counseling aid used to obtain information on the range of vocational interests of the counselee. It is particularly useful with persons who have no definite work interests or who have limited knowledge of the variety of jobs and occupational fields. Consists of 173 sample tasks grouped in accordance with the Worker Trait Arrangement of the Dictionary of Occupational Titles. There is no scoring of the list but a review of preferences checked is used as a guide for discussion of interests.

**NR (non-reading) Measure of "G"** - an interim non-reading measure of "G", mental or learning ability. Exercises consist of three parts of Cattell's "Culture Fair" test and part 7 (form matching) of the GATB. May be used to determine if the person would profit from basic or remedial education. Or, along with other non-reading parts of the GATB, it may be scored with other aptitudes for occupational aptitude patterns as used in counseling. This measure will soon be replaced by a complete non-reading edition of the GATB which will be called the NATB (Nonreading Aptitude Test Battery).

**Oral Trade Questions** - sets of 14 to 18 simple questions phrased in language of workers and covering essential parts of occupations. Questions are indirect measures of skill or knowledge about a job. There are approximately 240 occupations covered by the OTQ. The questions are administered during the interview and are scored as the questions are asked. Scores are interpreted as "well informed," "some information," or "little information." Cut-off scores were based on studies of employed workers who also served for standardi-
zation of the questions.

Pretest Orientation Exercises for GATB - designed to assist in the preparation and orientation of disadvantaged persons to take the General Aptitude Test Battery. A series of practice exercises and instructions, with separate answer sheets, is intended to teach test-taking skills and how to follow instructions, and to give practice in actually doing exercises.

Screening Exercises for GATB - are used by the counselor to determine if the examinee (who may be educationally deficient) is capable of understanding directions and able to perform on sample or practice exercises. The screening exercises consist of verbal and non-verbal parts. Interpretation of performance is based on the counselor's judgment of the examinee's handling of the items and ability to make correct responses.

Specific Aptitude Test Batteries (SATB) - are used to measure and predict potential ability to learn a specific job, as based on aptitudes found to be significant through test development and research. The GATB is used as the core battery, with two to four of the most significant aptitudes selected for greatest predictive efficiency. Multiple cutting scores are interpreted as: "H" - high, having met or placed above the minimum scores for each aptitude, "C" - band of consideration (scores fall within one standard error of measurement for each aptitude), and "L" - low, one or more scores below the band of consideration (SEM).

Spelling Tests - are clerical skills tests measuring the person's ability to recognize correct or incorrect spelling of words. A
general clerical form consists of 90 items on a separate answer sheet. Medical and legal tests consist of 40 items each which are dictated and spelled out by the examinee. Norms for accuracy have been established for various levels of skill.

Typing Tests - are clerical skills tests that measure speed (words per minute) and accuracy (errors). Several forms of the test are for both plain copy and statistical (tabular) typing. Short practice exercises are provided. Plain copy is a 5-minute test, whereas statistical takes 10 minutes. Separate norms have been established for manual or electric typewriters. Alternate forms provide for retesting.
**APPENDIX C**

**TEST PLANNING WORKSHEET**

(see back for instructions and interpretation)

<table>
<thead>
<tr>
<th>Name _________________</th>
<th>Sex _____</th>
<th>Soc. Sec. No. ____________</th>
<th>Profile Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>last</td>
<td>first</td>
<td></td>
<td>Profile Scale</td>
</tr>
<tr>
<td>Date ____ Counselor ____</td>
<td></td>
<td></td>
<td>(circle point on scale)</td>
</tr>
</tbody>
</table>

1. **Age** (years)
   | Neg. | Quest. | Pos. |
   | 56+  | 46-55  | 36-45 | 26-35 | 16-25 |

2. **Education**
   a. **Grade attainment**
   | 0    | 1-5   | 6-8   | 9-12  | Coll. |
   b. **Spec'l. Ed or Trng.**
   | none | self  | Voc HS | post | Coll. |
   | trained | M.D. | H.S. |

3. **Work Experience**
   a. **Years**
   | 0    | 1     | 2     | 3     | 4+    |
   b. **Type**
   | none | unsk. | semisk. | skilled | prof. |
   c. **Military (years)**
   | 0    | 1     | 2     | 3     | 4+    |

4. **Physical Condition**
   | poor | fair  | excell. |

5. **Interests and Hobbies**
   | none | med.  | high |

6. **Economic Status**
   (income)
   | public | $2000. | $3000. | $4000. | $5000.+ |
   | assistance |

7. **Cultural Status**
   (race or ethnic)
   | minority | borderline | majority |

8. **Family** (structure, status, size, housing, etc.)
   | negative | neutral | positive |

9. **Test Taking Experience**
   a. **No. of tests**
   | 0    | 1-2   | 3-4   | 5-6   | 7+    |
   b. **When taken**
   (years ago)
   | 5+   | 4     | 3     | 2     | 1     |
   c. **Attitude & Motivation**
   | negative | indifferent | positive |

10. **Remarks**
    (record on reverse side)
Form #1 (back)

Instructions

A. Secure information from existing forms such as application, WIN records, case record folder, or direct interview.

B. 1. **Age** - circle point on scale that includes enrollee's age.
   2. **Education** - grade attainment from interview, forms, or records.
   3. **Work** - total years, **Type** - per D.O.T., **Military** - years service.
   4. **Physical** - Excel. - no evidence of defects; Fair - one disability of a minor nature; Poor - two or more minor or major defects. Information from interview, medical or case records.
   5. **Interests and Hobbies** - High - two or more that involve high levels of skill or ability; Medium - one activity of an elementary level, Low - no interests or hobbies.
   6. **Economic** - state or recorded information.
   7. **Cultural Status** - Majority - white, middle - class; Borderline - white appalachia, black or ethnically different; Minority - Black.
   8. **Family** - Positive - stable, medium size (4-5), live in own home or rent (3 or more rooms); Neutral - live with one parent, divorced or separated, large size (6 or more), rent (1-2 rooms); Negative - no family, divorced or separated, boards in one room.
   9. **Test Taking** - information from interview or records.
   10. **Remarks** - record in space on bottom of this page.

Interpretation

Plot profile and connect points. Predominance of items in:

1. **Positive Area** - indicates that little or no special preparation or orientation is required for testing. Various tests may be administered as needed.
2. **Questionable** - exercise caution in the planning and preparation for testing. Need for special orientation and some diagnostic testing. Administer only appropriate tests as needed.
3. **Negative** - exercise extreme caution in the planning, selection, preparation, and administration of tests. Need for special preparation and orientation. Only administer tests of a diagnostic nature. Delay testing or omit is not appropriate.

Remarks: ______________________________________
_____________________________________________
_____________________________________________
_____________________________________________
_____________________________________________
Group: Exp. ___ Control ___ TEST RESULTS WORKSHEET

Name ______________________  Sex ___ Soc. Sec. No. ______________

1. Pre-testing Orientation: Date _____ Adm. by ____________________
   Remarks ______________________________________________________

2. Metro. Ach. Tests: Date _____ Form ___ Adm. by ________________
   Remarks ______________________________________________________

3. GATB Screening Exercises: Date _____ Adm. by ________________
   Remarks ______________________________________________________

4. NR "G" and NR parts of GATB: Date _____ Adm. by ______________
   Standard Scores
   NR "G" | S | P | K | F | M
   Remarks ______________________________________________________
   OAP's ___________________________ SATB's _______________________

5. GATB Form B-1001: Date ____________ Adm. by ________________
   Standard Scores
   G | V | N | S | P | Q | K | F | M
   OAP's ___________________________ SATB's _______________________

6. GATB Form B-1002A: Date ____________ Adm. by ________________
   Standard Scores
   G | V | N | S | P | Q | K | F | M
   OAP's ___________________________ SATB's _______________________

7. GATB Form B-1002B: Date ____________ (retest) Adm. by ________________
   Standard Scores
   G | V | N | S | P | Q | K | F | M
   OAP's ___________________________ SATB's _______________________

8. Proficiency:
   Typing: Man. _ Elec. ___  Dictation: Date ___ Form ___
   Plain ___ Stat. ___  Score _____ WPM _____ Errors
   Date ___ Adm. by ___________ Spelling: Date ___ Form ___
   Score _____ WPM _____ Errors  Score ______
   OTQ: Date _____ Form ___ Score ______

9. Other:  ICL: Date _____ Remarks ______________________________
           Color Vision: Date _____ Score ______

10. Additional remarks on back of sheet.
Group:  
Exp.  
Control  

TEST BEHAVIOR CHECKLIST  
(to be completed during or immediately after testing)  

Name  
Sex  
Soc. Sec. No.  

See back of sheet for instructions  

Performance Rating  
(circle point on scale)  

1. Followed instructions.  
2. Used separate answer sheet properly.  
3. Did practice exercises properly.  
4. Worked rapidly.  
5. Worked accurately.  
6. Appeared interested and motivated.  
7. Concentrated on task.  
8. Understood reason for tests.  
9. Anxiety during testing.  
10. Did own work.  

Remarks:  

Test Results: Screening Exercises  
Ach. Tests  
Other  

GATB: Form Date Tested by  
G V N S P Q K F M  

St's. Sc.  
Referred to testing by  

See back of sheet for additional remarks  
Descriptions for observed behavior:

1. Read instructions or listened and responded appropriately.

2. Located and made marks in proper places according to instructions. Positioned answer sheet for efficient marking.


4. Started immediately on signal and moved ahead on task without delays or stopping.

5. Made few or no errors which were quickly corrected.

6. Looked interested and alert. Moved ahead well with task.

7. Did not look around or appear distracted.

8. Felt the tests were going to provide information that would help him. Obtained from pre or post-test interview.

9. Did not appear nervous, apprehensive or anxious.

10. Did not copy from others.

Additional Remarks: ___________________________________________________________

__________________________________________________________________________

__________________________________________________________________________

__________________________________________________________________________

__________________________________________________________________________

__________________________________________________________________________

__________________________________________________________________________

__________________________________________________________________________

__________________________________________________________________________

__________________________________________________________________________

__________________________________________________________________________

__________________________________________________________________________
<table>
<thead>
<tr>
<th>Section</th>
<th>Rating</th>
<th>Remarks:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Test Readiness</td>
<td>Oriented</td>
<td>not poorly question- almost ready ready prepared able ade- adequ.</td>
</tr>
<tr>
<td>for taking tests</td>
<td>prepared</td>
<td>quacy.</td>
</tr>
<tr>
<td>2. Test-Taking Skills</td>
<td>Knows how to</td>
<td>none poor fair good excell.</td>
</tr>
<tr>
<td>take tests, use of answer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sheet, time, organization of</td>
<td></td>
<td></td>
</tr>
<tr>
<td>material</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Attitude - Motivation</td>
<td>Feelings</td>
<td>negative poor fair good excell.</td>
</tr>
<tr>
<td>about tests, approach to</td>
<td>hostile</td>
<td></td>
</tr>
<tr>
<td>exercises, level of activity.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Understanding</td>
<td>Reason for</td>
<td>none poor fair good excell.</td>
</tr>
<tr>
<td>tests, ability to follow</td>
<td></td>
<td></td>
</tr>
<tr>
<td>instructions.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Maximum Performance</td>
<td>Works rapidly,</td>
<td>below min question- less max.</td>
</tr>
<tr>
<td></td>
<td>accurate, top</td>
<td>able than max.</td>
</tr>
<tr>
<td></td>
<td>test ability.</td>
<td></td>
</tr>
<tr>
<td>Remarks:</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Form #5

INTERVIEW QUESTIONNAIRE

Local Office ______________________

Position - Counselor ______________ Test Examiner ______________

1. How would you assess the value of the procedures in making tests and testing more valid and meaningful to the counselor and counselee?

   1  2  3  4  5  

   low     median     high

2. How practical and usable were the forms?

   __ __ __ __  

   low     median     high

3. How would you rate the overall model of sequential testing?

   __ __ __ __  

   low     median     high

Additional Comments: ____________________________________________

_________________________________________________________________

_________________________________________________________________

_________________________________________________________________

_________________________________________________________________

_________________________________________________________________
TEST APPOINTMENT FORM

ADMIT: Name ___________________________________________ Soc. Sec. Number ______________________

Address ______________________________________________________

Sex __________________ Classification _____________________________

TO: Testing unit _____________________________________________ (Address)

Date of tests ______________________ Time _______________________

TESTS: Typing: Manual ☐ Electric ☐; Dictation ☐ 60, 80, 96, 120 w.p.m.; Spelling ☐

Aptitude test battery S-____, S-____, S-____, S-____

S-____, S-____, S-____, S-____

GATB for Counseling B-1001 ☐ B-1002 ☐

Other tests: ________________________________________________

______________________________ EMPLOYER NAME ________________________________

From ______ Date ______

E. S. Form 537 - Test Appointment Form (Front)
TEST RESULTS

COMMENTS: ____________________________________________________________
_____________________________________________________________________
_____________________________________________________________________

TYING: Form: ______ Speed: ______ w.p.m., A, B, C, D, E  Accuracy: A, B, C, D, E

DICTATION: Form: ______ Speed: 60, 80, 96, 120 w.p.m.  Accuracy: A, B, C, D, E

SPELLING T-_____: A, B, C, D, E

GATB FOR _______________________
S-____H, L; S-____H, L; S-____H, L.
MULTI-BATTERY SCORING: S-____H, L; S-____H, L; S-____H, L.

OTHER TESTS: __________________________________________________________________________

COLOR VISION: ________________________________________________________________________

<table>
<thead>
<tr>
<th>INDIIVIDUAL APTITUDE PROFILE</th>
</tr>
</thead>
<tbody>
<tr>
<td>B-1001</td>
</tr>
<tr>
<td>B-1002</td>
</tr>
</tbody>
</table>

OAP* OAP* OAP* B-*
0 15 30
1 16 31
2 17 32
3 18 33
4 19 34
5 20 35
6 21 36
7 22 37
8 23 38
9 24 39
10 25 40
11 26 41
12 27 42
13 28 43
14 29 44

*Encircle number for qualifying scores.
Cross out number for non-qualifying scores.
**Name**

**Code**

**Address**

**Social Security Account Number**

### TYPING

<table>
<thead>
<tr>
<th>Date</th>
<th>Form</th>
<th>SPEED</th>
<th>ACCURACY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Words Per Minute</td>
<td>Grade</td>
</tr>
</tbody>
</table>

### DICTATION

<table>
<thead>
<tr>
<th>Date</th>
<th>Form</th>
<th>SPEED</th>
<th>ACCURACY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Words Per Minute</td>
<td>Errors</td>
</tr>
</tbody>
</table>

### OTHER

<table>
<thead>
<tr>
<th>Date</th>
<th>Form</th>
<th>Test</th>
<th>Score</th>
<th>Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Spelling</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Legal Spelling</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Medical Spelling</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### V score

<table>
<thead>
<tr>
<th>KUDER PERCENTILES</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>OUT</td>
<td>MEC</td>
</tr>
</tbody>
</table>

**TEST RECORD CARD**

**OHIO BUREAU OF EMPLOYMENT SERVICES**

---

_E. S. Form 538 - Test Record Card (Front)_
### GATB DATA

<table>
<thead>
<tr>
<th>Date</th>
<th>Comments:</th>
<th>Research data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Age</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Education</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Part</th>
<th>Raw Score</th>
<th>G</th>
<th>V</th>
<th>N</th>
<th>S</th>
<th>P</th>
<th>Q</th>
<th>K</th>
<th>F</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**OAP* | SATB NUMBER**

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>13</th>
<th>26</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>14</td>
<td>27</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>15</td>
<td>28</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>16</td>
<td>29</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>17</td>
<td>30</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>18</td>
<td>31</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>19</td>
<td>32</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>20</td>
<td>33</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>21</td>
<td>34</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>22</td>
<td>35</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>23</td>
<td>36</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td>24</td>
<td>37</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td>25</td>
<td>All</td>
</tr>
</tbody>
</table>

H—Met Cutoff Scores.
C—Met Cutoff Scores with standard error of measurement.
L—Did Not Meet Cutoff Scores.

---

E. S. Form 538 - Test Record Card (Back)
# Record of Individual Student Test Performance

<table>
<thead>
<tr>
<th>Applicant Name of</th>
<th>Applicant Number (From reverse side)</th>
<th>Did not follow test directions</th>
<th>Had difficulty working practice exercises</th>
<th>Had difficulty marking answer sheet</th>
<th>Had difficulty doing test exercises</th>
<th>Excessive anxiety</th>
<th>Tended to waste time during test</th>
<th>Additional Comments on Applicant Performance on Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX D
SPECIMEN KIT - INSTRUCTIONS
SEQUENTIAL TESTING - A SYSTEMS APPROACH
An Experimental and Demonstration Project

I. Purpose - To determine the feasibility, value, and practical nature of the model developed for a systems approach to sequential testing. This phase of the project is a pilot or demonstration study of a tryout of the model, forms, and operational aspects.

II. Procedures - The model will be introduced in five to seven local employment offices which have on-going WIN programs and are using the Orientation and Assessment (Q.&A.) component. The E. and D. project will be initiated by Central Office Testing Supervisors in a training session of approximately one hour. An outline for the training is contained in this folder.

Participating local office staff will include the team leader, the counselor, the person conducting Orientation and Assessment, and the test examiner. In some instances, staff may have combined duties.

Of utmost importance is the timing of the demonstration project. Data collection should start immediately with the planning for the O. and A. sessions. The final cut-off data will be April 17th, at which time all data, forms, and material are to be forwarded to the Central Office Testing Section.
III. **Forms and Data Collection** - These will be carefully explained by the staff person installing the procedure. Attention should be given to obtaining accurate, complete information. An attempt was made to develop forms and worksheets that were short, concise, and not time-consuming for entries. Additional comments from participating staff are welcome.

IV. **Project Schedule** - March 2 to 13, installation of study in local offices. Collection of data should begin immediately, contingent upon availability of enrollees and staff services and the schedule of the O. & A. component. April 17 is the final cut-off date. Data may be forwarded prior to this date if completed.

V. **Questions or Requests for Assistance** - The following Central Office Testing staff can be contacted for information, forms or inquiries regarding the study: Joe Mihalka, Elwood Ziegler, John Showalter, or Penny Purviance. Call Columbus (614) 469-2303 or 469-4919.

**Outline for Presentation to L.O. Staff**

I. **Introductory Remarks** -

A. Current issues and problems in use of tests.

1. Selection of appropriate tests.

2. Preparation of the examinee for testing (orientation).

3. Obtaining valid test results through maximum performance.

4. Avoiding needless testing or the administration of
tests the examinee cannot handle.

5. Reporting and profiling of test results (scores and test behavior) for the counselor.

6. Improved test interpretation as a part of total information available on the counselee.

II. The Study Project —

A. Experimental and Demonstration (E. & D.) as a method of study.
   1. Not research in the strict sense of the word.
   2. Rather, a tryout for a new method or revision and improvement of existing methods and procedures.
   3. Less emphasis on statistical treatment of data and more attention to empirical evidence with practical applications.

B. The Systems Approach —
   1. Origin from engineering (systems analysis), data processing, programming, and management.
   2. Essentially, an input, process, and output through the use of flow charts or steps.
3. There are subsystems within systems. Movement may be forward as in referral or backward as feedback for evaluation and interpretation.

4. The procedure consists of a logical step-by-step activity with provision for orderly, planned, and evaluated increments of a sequential nature.

C. Sequential Testing -

1. Rationale
   a. Testing should be planned as a step-by-step procedure in accordance with the individual's needs and abilities.
   b. Tests should be administered in order from simple diagnostic forms (may be practice) to more complex, difficult, standardized forms.
   c. Maximum performance should assure valid and reliable results.
d. Needless or inappropriate tests and testing may be harmful to the examinee and also uneconomical.

2. Steps in Sequential Testing -
   a. Planning - Gather and assess data and information on counselee. Plot information on Test Planning Worksheet. Evaluate profile.
   b. Decision Making - Based on Planning Worksheet, develop a sequence or plan for test preparation and administration of standardized tests (Test Results Worksheet may be used).
   c. Feedback - Check test results at each point to determine next step in process, delay or termination of testing.
   d. Evaluation - Evaluate supplemental information, such as test-taking behavior and all other data available on counselee.
   e. Counseling - Integrate all data and information during test interpretation.

D. Study Procedures -
   1. Experimental Sample Group from each local office will consist of seven to fifteen WIN enrollees who are referred to Orientation and Assessment. With participation of 5 to 7 offices, the total experimental sampling will be approximately 50 enrollees.
2. The Control Group will consist of a matched number of WIN enrollees who do not participate in O. & A. and also will not be processed with the sequential procedures.

3. Experimental Procedures –
   a. The Experimental Group –
      (1) The Test Planning Worksheet will be completed on enrollees prior to referral for O. & A. It may recommend test procedures particularly for pre-test orientation and administration of standardized tests.
      (2) Feedback to the counselor should take place during the process of pre-test orientation and after test administration. The Test Results Worksheet should be utilized during this phase.
      (3) The Test Behavior Checklist should be utilized at the point that involves the longest testing period (GATB or Metro. Ach.).
      (4) The Procedures Evaluation Sheets are to be worked on independently by the counselor, the enrollee (in the counseling interview), and the test examiner. The counselor should complete his form prior to the interview with the enrollee. The enrollee's may be done during test interpretation with the counselor.
The test examiner's sheet should be finished as soon as possible after testing.

b. The Control Group will not be exposed to O. & A. and will not utilize the procedures outlined above but will have data gathered as follows:

(1) The Test Behavior Checklist and Test Results Worksheet will be used but not forwarded to the counselor (for data collection only).

(2) The Test Planning Worksheet will be filled in after testing has been completed (for data collection only).

(3) The Procedure Evaluation Sheet will be completed as item a. (4) above.

c. In essence, the experimental group will follow the model design, using the forms and procedures as outlined. The control group will follow traditional testing and counseling methods. Forms are used only to gather data for the comparison of the experimental and control groups and not for operational purposes.

(1) All data, forms, test answer sheets, and other material are to be completed and sent to the Central Office Testing Section on or before April 17th. Any additional comments, observations, or recommendations concerning the study as a whole will be welcome. A
summary report of the study will be provided all cooperating offices upon the completion of the project.

III. Summary and Conclusions -
A. Stress time limitations of study.
B. Request as large a sample of each group as may be available.
C. Follow procedures exactly as outlined.
D. Forward all forms, test answer sheets, and procedural material as well as additional comments or observations on or before April 17th.
E. Further information or questions should be sent to C.O. Testing Section, Attention - Research.
F. Summary of study will be forwarded.
G. Express appreciation and thanks for cooperating.

IV. Questions and Discussion -

Checklist for Forms and Procedures

1. Test Planning Worksheet
A. Complete one form on each enrollee in sample groups.
   (1) Experimental group (make two copies) prior to referral to C. & A. and testing. Use according to procedures.
   (2) Control group (make one copy) after testing or test interpretation (do not use procedurally, only for data information). Please code sheet "E" for Experimental and "C" for control.
   (3) Items 1 to 8 are obtained from various biographical data, other forms, records, or counseling interview.
B. Disposition of forms -

(1) Place originals in case record folder for experimental group, and forward carbon copy with other data to Central Office.

(2) For control group, send the one copy to C.O.

2. Test Results Worksheet

A. Complete one form for each enrollee.

(1) Experimental group (make two copies) may be stated by counselor but maintained and recorded also by test examiner or orientation coordinator (pre-test orientation).

(2) Control group (make one copy) kept by test examiner and not forwarded to counselor or entered into case record folder.

B. Disposition of forms -

(1) Place originals in case record folder for experimental group, and forward carbon copy with other data to Central Office.

(2) For control group, send the one copy to C.O.

3. Test Behavior Checklist

A. Complete one form on each enrollee.

(1) To be completed by test examiner.

(2) Experimental group (make two copies) return one to counselor and carbon to C.O.

(3) Control group (make only one copy) forward to C.O.
B. Disposition of forms -

(1) Experimental group, place original in case record folder and send carbon to C.O.

(2) Control group, send only copy to C.O.

4. Evaluation Sheets (after testing is completed)

A. To be complete independently by each person.

(1) Counselors - fill in prior to evaluation by enrollee and without reference to test examiner's evaluation.

(2) Enrollee - to be completed with the assistance of the counselor during the counseling session on test interpretation.

(3) Test examiner - finish as soon as possible after testing and completion of Test Behavior Checklist.

B. Disposition of forms

(1) Send all copies to C.O. Testing Section.
APPENDIX E

STATISTICAL PROCEDURES

The various forms (with the exception of the Test Results Worksheet) were provided with a rating scale for each item to which values from 1 to 5 (Likert procedures) were assigned. Low values were to the left and became higher toward the right end of the scale. Moreover, the Test Planning Worksheet scales were divided into 3 categories (negative, questionable, and positive) which were given values of 1 to 3. This method was used to determine total scores on this worksheet.

The means of the values as recorded on the individual item scales were calculated. Significance of differences between the means was determined by t-ratios (tests). This procedure is often used with small samples (30 or less), but some statisticians are of the opinion that the numbers may be 50 to 100 (Yuker, 1958).

The t-test for significance of differences between means assumes equality of the population variances. Approximate methods to use when variances are unequal have been suggested by Aspen and Welch that make adjustments in the number of degrees of freedom on special tables (Ferguson, 1966).

On the basis of data available for the experimental and control groups, the assumption was made that the characteristics of
the samples were equal and there was normality insofar as the WIN enrollees were concerned. This position was supported by information on the Test Planning Worksheet from which select factors, such as sex, age, education, work experience, and previous test experience, indicated no significant differences between the two groups.

In most experimental situations, the subjects are randomly assigned to the groups and the experimental sample is given special treatment. Comparison is then made to the control group which is not given this treatment. The random assignments were not possible in this study because the groups were already made up in accordance with standard procedures in the WIN program. With these facts in mind, the formula used for the t-ratio (test) was one for determining the differences between two independent means. The basic computation formula was from Bruning and Kintz (1968), as follows:

\[ t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left(\frac{\sum X_1^2}{N_1} + \frac{\sum X_2^2}{N_2} - \frac{(\sum X_1)^2}{N_1} - \frac{(\sum X_2)^2}{N_2}\right) \cdot \left(\frac{1}{N_1} + \frac{1}{N_2}\right)}} {\sqrt{\frac{(N_1 + N_2) - 2}{N_1 N_2}}} \]
where

\[ X_1 = \text{the mean of the first group of scores} \]
\[ X_2 = \text{the mean of the second group of scores} \]
\[ \Sigma X_1^2 = \text{the sum of the squared score values of the first group} \]
\[ \Sigma X_2^2 = \text{the sum of the squared score values of the second group} \]
\[ (\Sigma X_1)^2 = \text{the square of the sum of the scores in the first group} \]
\[ (\Sigma X_2)^2 = \text{the square of the sum of the scores in the second group} \]
\[ N_1 = \text{the number of scores in the first group} \]
\[ N_2 = \text{the number of scores in the second group} \]

The table for critical values was Table III of Fisher and Yates (Ferguson, 1966), using the level of significance for one-tailed test. This table provided for degrees of freedom up to 120.

The calculations of the significance of differences for the "appropriateness" and "inappropriateness" utilized the Yates correction for continuity (Ferguson, 1966). The formula computed \( \chi^2 \) from a 2 x 2 table to incorporate this correction as follows:

\[
\chi^2 = \frac{N (1AD - BCL - N/2)}{(A + B) (C + D) (A + C) (B+D)}
\]

where \( N = \text{total number of cases} \)
\( A, B, C, \text{&} D = \text{various cases in cells of a 4-fold table} \)
This correction was suggested when any of the expected frequencies were less than 5 (some writers suggest 10). As reported in the study, the number of inappropriate tests in the experimental group (N=86) was 3 and in the control (N=33) it was 8. The $\chi^2$ with the correction was 9.88, which was significant at the .01 level.

Further interpretation of the data was made in percentages and descriptive information on the different scales. Conclusion in the study were based on both empirical and statistical evidence, with perhaps some serious reservations about the statistical treatment of the data. The main emphasis in the study was on the development of the model on a tryout in a pilot or demonstration, and, finally, on some limited statistical processing of experimental data.
REFERENCES FOR CHAPTER I


REFERENCES FOR CHAPTER II


REFERENCES FOR CHAPTER III


Cattell, A.K.A. and Cattell, B.B., Test of "g": Culture Fair, Champaign, Ill., The Institute for Personality and Ability Testing, 1957.


Staff, Testing Section, ES Division, Pre-Testing Orientation, Columbus: The Ohio Bureau of Employment Services, 1969.


Forms - The Ohio Bureau of Employment Services, Columbus:

E.S. 537, Test Appointment and Results, 1968
E.S. 538, Test Record Card, 1968
Record of Apparatus Tests, 1967
E.S. 511, Application Card, 1968
REFERENCES FOR CHAPTER IV


BIBLIOGRAPHY


Staff, Central Office Testing Section, *Pre-test Orientation*, Columbus; The Ohio Bureau of Employment Services, 1969.


U. S. Department of Labor, U.S.E.S. Pre-testing Orientation Exercises to the GATB (manual and exercises), Washington D. C.: Manpower Administration, 1968.

Veri, Clive C. and Mocker, Donald W., The University of Missouri Adult Basic Education System, St. Louis: University of Missouri, 1969.


