

A COMPARATIVE STUDY OF TWO MEASURES
OF ADAPTIVE BEHAVIOR FOR
MENTALLY RETARDED ADULTS

by

SUSANN LOUISE MONGRAIN

B.Sc., McGill University, 1970

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF ARTS

in the Departments

of

Educational Psychology
and
Special Education

We accept this thesis as conforming to the
required standard

THE UNIVERSITY OF BRITISH COLUMBIA

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study.

I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the Head of my Department or by his representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Education

The University of British Columbia
2075 Wesbrook Place
Vancouver, Canada
V6T 1W5

Date Sept 12, 1975

Abstract

Since 1959, adaptive behavior has been defined by the American Association on Mental Deficiency (AAMD) as one of the two dimensions to be used in diagnostic classification in mental retardation. This use of adaptive behavior has received support in the literature in this field, but it has not been adopted in clinical practice. This may be the result of problems that have been associated with the conceptualization and measurement of adaptive behavior by early scales.

In recent years two new measures of adaptive behavior have been published: the AAMD Adaptive Behavior Scale (ABS) and the Vocational Rehabilitation and Research Institute's (VRRI) Adaptive Functioning Index (AFI). The ABS consists of two parts. Part I measures the extent to which an individual is able to cope with the personal and social demands of his/her environment. Part II is designed to measure maladaptive behaviors. The AFI consists of three parts. The Social Education Test (SET) is an individualized test of practical academic and related skills. The Vocational Check List (VCL) is designed to assess basic work habits and skills. The Residential Check List is used for assessing self-help skills related to independent living.

Very little information concerning the utility of these scales is available. Therefore, the purposes of this study were to provide estimates of the reliabilities of these scales; to examine their ability to discriminate among individuals at all levels of functioning; to assess their validity by examining their mutual relationship as well as the relationship of each

scale to measured intelligence; and finally, to assess the utility of these scales for programming purposes.

A sample of fifty-one mentally retarded adults were tested on the Revised Stanford Binet Intelligence Scale (RSB) and on the SET. In addition, each subject was rated by two staff members familiar with his/her daily behavior on the ABS Parts I and II, the VCL, and the RCL. Three months after this data was collected, thirty subjects were retested on the SET.

The results of this study indicated that the internal consistencies of the ABS and the AFI scales were sufficiently high and comparable to those of standardized tests such as the RSB. The stability of SET scores also appeared to be adequate. The interrater reliabilities of the ABS Parts I and II, the VCL, and the RCL were assessed using generalizability analyses. These analyses revealed great variation in the extent to which rater pairs agree in their judgments. Some inconsistencies appeared to be the result of a tendency for different raters to assign different scores to the same persons. This bias was attributed to ambiguities in the scales as well as to differences in the opportunities for observation raters have and/or use.

Both the ABS and the AFI contain items that were not useful in discriminating among members of this sample. Some items were ambiguous; others were too easy or too difficult. In general, the ABS appeared to be more useful for discriminating among more severely handicapped individuals, whereas, the AFI seemed appropriate for individuals functioning at higher levels,

The validity of these scales was assessed using a multitrait-multimethod approach. Evidence was provided for the convergent validity of sixteen of the nineteen traits that were examined. However, support for discriminant validity emerged for only seven of these traits. The practical validity of the five domains of the ABS Part II that were examined was supported. Many aspects of adaptive behavior were found to be significantly related to RSB mental age. The SET showed the most striking relationship to mental age.

Finally, the ABS appeared to be more useful for programming in institutional environments, whereas the AFI seemed more useful for more independent settings. The AFI was judged to have certain desirable psychometric characteristics, such as equal numbers of items in the subtests of a given test, that make it more desirable for programming. However, the usefulness of both the ABS and AFI was considered to be limited by their low interrater reliabilities and hence their limited validity.

Table of Contents

Abstract

List of Tables

List of Figures

Acknowledgement

Chapter

I	Introduction	1
	Statement of the Problem	3
II	Review of Related Literature	5
	Nature of Adaptive Behavior	6
	Measurement of Adaptive Behavior: Early Scales	7
	Measurement of Adaptive Behavior: New Scales	9
	AAMD Adaptive Behavior Scales	10
	VRRI Adaptive Functioning Index	11
III	Methods	28
	Design of Study	20
	Subjects	21
	Raters	21
	Subject Attrition	24
	Materials	28
	Procedure	28
	Data Collection I: Individual Testing	29
	Rating Scales	29
	Reported Incidents of Maladaptive Behavior	31

	Data Collection II	31
	Method of Analysis	32
	Test Analysis and Item Analysis	32
	Interrater Reliability	34
	Generalizability of Each Rater Pair	36
	Generalizability of All Rater Pairs	38
	Stability	40
	Validity	41
IV	Results of Analysis	44
	Results of Analysis: ABS Part I	45
	Test Analysis	45
	Item Analysis	45
	Generalizability of Ratings	55
	Results of Analysis: ABS Part II	58
	Test Analysis	58
	Item Analysis	62
	Estimated Frequencies of Occurrence for Behavior Measured by the ABS Part II	72
	Generalizability of Ratings	74
	Results of Analysis: SET	79
	Test Analysis	79
	Item Analysis	85
	Stability	96

Results of Analysis: VCL	99
Test Analysis	99
Item Analysis	103
Generalizability of Ratings	113
Results of Analysis: Difficulty Section of the VCL	117
Test Analysis	117
Item Analysis	117
Generalizability of Ratings	125
Results of Analysis: RCL	128
Test Analysis	128
Item Analysis	132
Generalizability of Ratings	147
Results of Analysis: Validity	151
Practical Academics, Language Development and Motor Development	151
Self-Help Skills	154
Vocational Skills	157
ABS Part II	160
V Discussion	168
Reliability	169
Internal Consistency	169
Stability	170
Interrater Reliability	170

Discrimination	174
Rating Scales	174
Social Education Test	178
Measure of Maladaptive Behavior	179
Validity	181
Adaptive Behavior and Measured Intelligence	182
Program Development	183
Limitations	186
Summary	187
References	190
Appendix A	192
Appendix B	203
Appendix C	224
Appendix D	236
Appendix E	242
Appendix F	260

List of Tables

Table

1	Standard deviation and IQ score ranges corresponding to measured intelligence levels	22
2	Description of sample	23
3	Description of raters	25
4	Numbers of subjects tested, by program group, test and rater	26
5	RSB mental age equivalents by level of measured intelligence for chronological age eighteen	30
6	Distribution of subjects by sex and level of measured intelligence	33
7	Test analysis information for the subtests of the ABS Part I (1)	46
8	Test analysis information for the subtests of the ABS Part I (2)	47
9	Test analysis summary for the ABS Part I	48
10	Correlations among subtest and total test scores for the ABS Part I (1)	49
11	Correlations among subtest and total test scores for the ABS Part I (2)	50
12	Generalizability coefficients by rater pair and across raters for the subtest and total test scores of the ABS Part I	56
13	Variance components contributing to universe score variance ($\sigma^2_{\mu p}$) and error variance (σ^2_{δ}) by rater pair and across raters for subtest and total test scores for the ABS Part I	57
14	Test analysis information for the subtests of the ABS Part II (1)	59
15	Test analysis information for the subtests of the ABS Part II (2)	60

16	Test analysis summary for the subtests of the ABS Part II	61
17	Correlations among subtest and total test scores for the ABS Part II (1)	63
18	Correlations among subtest and total test scores for the ABS Part II (2)	64
19	Item analysis information for the ABS Part II: <u>Violent and Destructive Behavior</u>	65
20	Item analysis information for the ABS Part II: <u>Anti-social Behavior</u>	65
21	Item analysis information for the ABS Part II: <u>Rebellious Behavior</u>	66
22	Item analysis information for the ABS Part II: <u>Untrustworthy Behavior</u>	66
23	Item analysis information for the ABS Part II: <u>Withdrawal</u>	67
24	Item analysis information for the ABS Part II: <u>Stereotyped Behaviors and Odd Mannerisms</u>	67
25	Item analysis information for the ABS Part II: <u>Unacceptable or Eccentric Habits</u>	68
26	Item analysis information for the ABS Part II: <u>Sexually Aberrant Behavior</u>	68
27	Item analysis information for the ABS Part II: <u>Psychological Disturbances</u>	69
28	Range of estimated frequencies of occurrence for items of the ABS Part II and percentage of items for which raters provided estimates	73
29	Number of items of the ABS Part II showing different degrees of overlap in the range of estimated frequency of occurrence	75
30	Generalizability coefficients by rater pair and across raters for the subtest and total test scores of the ABS Part II	76

31	Variance components contributing to the universe score variance ($\sigma^2_{\mu p}$) and error variance (σ^2_{δ}) by rater pair and across raters for the subtest and total test scores of the ABS Part II	77
32	Test analysis information for the subtests of the Set (1)	80
33	Test analysis information for the subtests of the Set (2)	81
34	Test analysis summary for the Set	82
35	Correlations among subtest and total test scores for the Set (1)	83
36	Correlations among subtest and total test scores for the Set (2)	84
37	Generalizability coefficients for subtest and total test scores of the Set	91
38	Estimates of variance components contributing to the ϵp^2 for the subtest and total test scores of the Set	98
39	Test analysis information for the subtests of the VCL (1)	100
40	Test analysis information for the subtests of the VCL (2)	101
41	Test analysis summary for the VCL	102
42	Correlations among subtest and total test scores for the VCL (1)	104
43	Correlations among subtest and total test scores for the VCL (2)	105
44	Generalizability coefficients by rater pair and across raters for the subtest and total test scores of the VCL	114
45	Variance components contributing to universe score variance ($\sigma^2_{\mu p}$) and error variance (σ^2_{δ}) by rater pair and across raters for the subtest and total test scores for the VCL	115

46	Test analysis information for the Difficulty subtests of the VCL (1)	118
47	Test analysis information for the Difficulty subtests of the VCL (2)	119
48	Correlations among Difficulty subtest and total test scores for the VCL (1)	120
49	Correlations among Difficulty subtest and total test scores for the VCL (2)	121
50	Generalizability coefficients by rater pair and across raters for the Difficulty subtest and total test scores of the VCL	126
51	Variance components contributing to the universe score variance ($\sigma^2_{\mu p}$) and error variance (σ^2_{δ}) by rater pair and across raters for the Difficulty subtest and total test scores of the VCL	127
52	Test analysis information for the subtests of the RCL (1)	129
53	Test analysis information for the subtests of the RCL (2)	130
54	Test analysis summary for the RCL	131
55	Correlations among subtest and total test scores for the RCL (1)	133
56	Correlations among subtest and total test scores for the RCL (2)	134
57	Generalizability coefficients by rater pair and across raters for the subtest and total test scores of the RCL	148
58	Variance components contributing to the universe score variance ($\sigma^2_{\mu p}$) and error variance (σ^2_{δ}) by rater pair and across raters for the subtest and total test scores of the RCL	149
59	Intercorrelations of Language Development, Practical Academic Skills and Motor Movements as measured by the SET and ABS Part I (1) and correlations of these traits with mental age	152

60	Intercorrelations of Language Development, Practical Academic Skills and Motor Movements as measured by the SET and ABS Part (2) and correlations of these traits with mental age	153
61	Intercorrelations of Self-Help Skills as measured by the RCL (1) and ABS Part I (1) and correlations of these traits with mental age	155
62	Intercorrelations of Self-Help Skills as measured by the RCL (2) and ABS Part I (2) and correlations of these traits with mental age	156
63	Intercorrelations of Vocational Skills as measured by the VCL (1) and ABS Part I (1) and correlations of these traits with mental age	158
64	Intercorrelations of Vocational Skills as measured by the VCL (2) and the ABS Part I (2) and correlations of these traits with mental age	159
65	Intercorrelations of the subtest and total test scores of the SET with RSB mental age, corrected for attenuation due to error in estimates of the internal consistencies and stabilities of these traits	161

List of Figures

Figure

1	Item analysis information for the ABS Part I (1)	52
2	Item analysis information for the ABS Part I (2)	53
3	Item analysis information for the <u>Reading</u> subtest of the SET	86
4	Item analysis information for the <u>Writing</u> subtest of the SET	87
5	Item analysis information for the <u>Numbers</u> subtest of the SET	88
6	Item analysis information for the <u>Time</u> subtest of the SET	89
7	Item analysis information for the <u>Money</u> subtest of the SET	90
8	Item analysis information for the <u>Community Awareness</u> subtest of the SET	91
9	Item analysis information for the <u>Communication</u> subtest of the SET	92
10	Item analysis information for the <u>Motor Movements</u> subtest of the SET	93
11	Item analysis information for the <u>Concept Attainment</u> subtest of the SET	94
12	Item analysis information for the <u>Basic Work Habits</u> subtest of the VCL (1)	106
13	Item analysis information for the <u>Basic Work Habits</u> subtest of the VCL (2)	107
14	Item analysis information for the <u>Work Skills</u> subtest of the VCL (1)	108
15	Item analysis information for the <u>Work Skills</u> subtest of the VCL (2)	109
16	Item analysis information for the <u>Acceptance Skills</u> subtest of the VCL (1)	110

17	Item analysis information for the <u>Acceptance Skills</u> subtest of the VCL (2)	111
18	Item analysis information for the <u>Difficulty</u> subtests of the VCL (1)	122
19	Item analysis information for the <u>Difficulty</u> subtests of the VCL (2)	123
20	Item analysis information for the <u>Personal Routines</u> subtest of the RCL (1)	135
21	Item analysis information for the <u>Personal Routines</u> subtest of the RCL (2)	136
22	Item analysis information for the <u>Community Awareness</u> subtest of the RCL (1)	137
23	Item analysis information for the <u>Community Awareness</u> subtest of the RCL (2)	138
24	Item analysis information for the <u>Social Maturity</u> subtest of the RCL (1)	139
25	Item analysis information for the <u>Social Maturity</u> subtest of the RCL (2)	140
26	Comparison of mean scores on the ABS Part II and on reported incidents of maladaptive behavior for the domain <u>Violent and Destructive Behavior</u>	162
27	Comparison of mean scores on the ABS Part II and on reported incidents of maladaptive behavior for the domain <u>Antisocial Behavior</u>	163
28	Comparison of mean scores on the ABS Part II and on reported incidents of maladaptive behavior for the domain <u>Rebellious Behavior</u>	164
29	Comparison of mean scores on the ABS Part II and on reported incidents of maladaptive behavior for the domain <u>Untrustworthy Behavior</u>	165
30	Comparison of mean scores on the ABS Part II and on reported incidents of maladaptive behavior for the domain <u>Psychological Disturbances</u>	166

Acknowledgements

I would like to thank the following persons for their assistance with this study:

Dr. Robert Conry, Chairman of the Thesis Committee, for his support and guidance during my graduate program as well as during this study.

Dr. Julianne Conry, member of the Thesis Committee, for training I received from her and for her assistance with the final draft.

Mr. Bob Poutt, member of the Thesis Committee, for his support and understanding particularly during the collection of the data and for his assistance in obtaining subjects.

Mr. Roger Bushnell and the staff and residents of Bevan Lodge without whom this study would not have been possible.

To Nancy and Chris for their encouragement.

To Larry for his understanding and patience.

Chapter I

Introduction

The concept of mental retardation that has evolved from psycho-educational theory and knowledge has been based on the assumption that retardation reflects limitations in intellectual functioning. Despite the proliferation of theories concerning the nature of intellectual functioning, the practical problems associated with mental retardation have, to a great extent influenced the definition of more specific capacities that could be used in distinguishing the retarded from the rest of the population (Robinson & Robinson, 1965).

Most traditional definitions of mental retardation have equated it with social incompetency (Robinson & Robinson, 1965). Prior to the development of standardized tests, mental retardation was defined solely in terms of the "adequacy of the individual in managing himself and his affairs" (Nihira, Foster, & Spencer, 1968). Although this definition emphasized the importance of assessing individual characteristics in terms of social and cultural standards, (Nihira et al., 1968) it implied an adult norm of performance (Robinson & Robinson, 1965). The problems posed by the diagnostic evaluation of infants and children dictated consideration of the developmental origin of mental retardation.

Definitions of mental retardation in terms of social adaptation have, implicitly or explicitly, considered this condition to be of pathological etiology, and, therefore, incurable. Although these concepts appeared to be appropriate for severely handicapped individuals, they did not apply to those cases where

physical anomalies were absent, and where retardation proved to be temporary and the product of psychological causes (Robinson & Robinson, 1965). As a result, many theoreticians in this area came to view a diagnosis of mental retardation as indicative of current abilities rather than potential capacity, regardless of etiological factors.

Although mental retardation continued to be equated with social adaptation in the body of theoretical work, the advent of mental testing changed its practical definition. Since the development of the IQ scale by Binet and Simon, measured intelligence has been considered the most salient feature of mental retardation (Benton, 1964), and, the IQ has been used as the sole criterion in diagnosis and classification (Adams, 1973).

This widespread use of intelligence tests in diagnosing mental retardation has occurred for several reasons. Robinson and Robinson have cited objectivity of assessment, "...simplicity, ease of communication, and well defined normative groups for comparison" as the major advantages of this scheme. Its critics, however, have pointed out that non-intellectual factors may substantially affect intelligence-test results. Cultural differences reflect one class of these which have received considerable attention in recent years (Beerman, Bellamy, Di Rocco, Friedland, Foss and Steinbock, 1973). In addition, intelligence tests are restricted measures of abilities and disabilities. They do not "adequately describe nor predict the way an individual maintains his personal independence nor how he meets the social expectations of his environment" (Nihira, 1969a), the very information required by workers in this field.

In response to this concern over the exclusive use of intelligence tests in diagnosis and classification, the American Association on Mental Deficiency (AAMD) proposed a definition of mental retardation that reflected practical as well as theoretical concerns:

Mental Retardation refers to significantly sub-average general intellectual functioning existing concurrently with deficits in adaptive behavior, and manifested during the developmental period (Grossman, 1973).

Diagnosis and classification are based on two general dimensions: intellectual functioning, which refers to performance on one or more standardized tests developed for that purpose (Grossman, 1973; Heber, 1961; Heber, 1959); and adaptive behavior, which is defined as "the effectiveness or degree with which the individual meets the standards of personal independence and social responsibility expected of his age and cultural group" (Grossman, 1973).

This renewed interest in adaptive behavior has received much support from current writers in the field of mental retardation. It has been argued that it is the absence or impairment of the skills and behaviors that constitute adaptive behavior that marks the person as deviant or retarded in society (Wolfensberger, 1972; Benton, 1964). And, "since the vast number of behaviors identified in this sphere, can, through modern treatment and training procedures, be modified or reversed, this area becomes the soundest area for attempting a diagnostic classification" (Leland, 1969).

Statement of the Problem

Despite this emphasis on the importance of adaptive behavior in the literature, it appears to have failed to gain support in clinical practice. In

his study of the kinds of information used for diagnostic classification purposes, Adams (1973) found that psychologists rely almost totally on the IQ, and that physicians do also, when psychological reports are available.

This failure to utilize the adaptive behavior dimension may be due to problems that have been associated with the early adaptive behavior scales. Specifically, the definitions of adaptive behavior reflected in these scales have been inadequate. Their reliabilities have been lower than those of other standardized tests. Evidence for validity has been limited. The usefulness of these scales in discriminating among individuals at all levels of impairment has been questioned. And, finally, the amount of information they provide for programming purposes have been considered inadequate (Grossman, 1973; Gardner & Giampa, 1972).

The problem of this study concerns the utility of two new adaptive behavior indices, the AAMD Adaptive Behavior Scales (Nihira, Foster, Shelhaas, & Leland, 1969); and the Vocational and Rehabilitation Research Institute's Adaptive Functioning Index (Marlett, 1973). The purposes of this investigation are to provide:

1. estimates of the reliabilities of these scales
2. indications of the abilities of these tests to discriminate among individuals at all levels of functioning
3. evidence for their validity
4. indications of their usefulness for programming and progress evaluation
5. evidence regarding the relationship of measured intelligence to different aspects of adaptive behavior.

Chapter II

Review of Related Literature

In recent years the construct of adaptive behavior has evolved in at least two ways. First, variables related to adaptive behavior have been identified and incorporated into an increasingly complex definition of this construct. Secondly, methods of assessing adaptive behavior have changed both as a result of the changing definition of this construct as well as in response to the psychometric shortcomings of early adaptive behavior scales. In this section, both the nature and measurement of adaptive behavior will be discussed.

Nature of Adaptive Behavior

The authors of early measures of adaptive behavior defined it as a unidimensional construct (Adams, 1973). The most obvious indication of this is the single score Social Age or Social Quotient derived from assessment. More recent evidence suggests that adaptive behavior is multidimensional (Balthazar & English, 1971; Nihira, 1969a; Nihira, 1969b), and, as such is a "composite of many aspects of behavior, and a function of a wide range of specific abilities and disabilities" (Grossman, 1973).

The skills and behaviors associated with adaptive behavior vary according to chronological age. During infancy and early childhood, sensory-motor, communication, self-help, and socialization skills are important. In childhood and adolescence, the acquisition and application of basic academic skills, appropriate reasoning and judgement, and socialization skills become the focal points of development. In the adult years, vocational performance and social responsibilities constitute the important dimensions of adaptive behavior (Grossman, 1973).

In addition to chronological age, cultural and other environmental factors must be considered in defining the specific components of adaptive behavior. This is because the acquisition and use of these components depends in large part on opportunity and on other people in the environment (Grossman, 1973).

The inherent characteristics of the individual are also important in exploring the nature of adaptive behavior. The most widely investigated of these is measured intelligence. Many studies have revealed substantial correlations between adaptive behavior and intellectual functioning (Grossman

1973). Himmelstein (1968), in his review of the use of the Revised Stanford Binet (RSB), Form LM, with retardates, cites correlations of .72 between RSB Mental Age and the Vineland Social Maturity Scale Social Age; .69 with the Cain-Levin Social Competency Scale (CLSCS); .73 with the communication subscale of the CLSCS; and .28, .52, and .71 with the reading, spelling, and arithmetic subtests of the Wide Range Achievement Test. The problem appears to be one of identifying those specific dimensions that are most closely associated with adaptive functioning, independent of measured intelligence, as well as those most closely associated with IQ (Nihira, Foster, & Spencer, 1968).

Although it is possible to define important dimensions of adaptive behavior, an absolute standard of performance does not exist. Since each individual has a unique pattern of coping behaviors, assessment is limited to those aspects that can be commonly observed among most of the population (Nihira et al., 1968).

Measurement of Adaptive Behavior: Early Scales

The measurement of adaptive behavior poses many problems. Empirical studies of existing scales have revealed some of these.

First, the reliabilities of these scales are low, a factor attributed to the type of data collection procedures that are used. Since adaptive behavior consists largely of routine behavior, it cannot be measured in an office setting (Grossman, 1973). Because of the considerable time required for direct observation, many measures of adaptive behavior rely on the informant method of data collection. Information is obtained by a trained rater from interviews

with persons who are familiar with the subject. These persons, the informants, provide a description of the individual which can be compared with descriptions of others with known levels of competency to obtain a rating (Gardner & Giampa, 1972). Despite widespread application, research indicates that this method shows consistently lower reliability than standardized intelligence and achievement tests (Grossman, 1973). For example, the Vineland Social Maturity Scale uses this method. Although the author of this scale reports a stability coefficient of .92, other studies have reported coefficients downward to .80 and sometimes lower (Hurst, 1962).

The content validity of adaptive behavior measures has also been criticized. In their review of existing scales in the United States and Great Britain, Leland, Shelhaas, Nihira and Foster (1967) found that most tests shared a strong emphasis on independent functioning, with only a nominal sampling of items dealing with personal and social responsibility. An analysis of reasons given by parents for institutionalizing their retarded children, and a critical-incidents study of inappropriate behavior in special education classes, day care centres, and institution wards or cottages revealed aspects of adaptive behavior that existing scales (Nihira, 1969a), did not measure.

Since many scales were developed primarily for institutionalized populations they do not sample behaviors characteristic of the retarded population living in the community. The importance of this point stems from the apparent move toward integration that has occurred in recent years, and which requires the development of scales appropriate for such environments (Grossman, 1973)

Another problem, associated with content, is the failure of some scales to discriminate among individuals at all levels of retardation. For example, the Cain-Levin Social Competency Scale has been found to be unsuitable for obtaining information about low-level retarded subjects because of its failure to discriminate among individuals with severe impairments (Gardner & Giampa 1972).

Another criterion for assessing the utility of adaptive behavior indices is the level of specificity of information they provide, both for programming purposes and progress evaluation. The Vineland Social Maturity Scale provides single score results that are useful only for describing an individual's adaptive functioning in a global way and in evaluating overall changes in behavior (Gardner & Giampa, 1972).

Finally, the adequacy of the norms of existing scales has been questioned. Many scales were standardized on institutional samples, and are therefore not appropriate for assessing the performance of the retarded in the community. In addition, the norms of many scales do not reflect the performance of the "normal" population, to which the retarded are ultimately compared. It is from this type of data that standards of performance for independent functioning could be derived empirically, rather than determined on a rational basis, only.

Measurement of Adaptive Behavior: New Scales

Two new measures of adaptive behavior have been developed which attempt to overcome some of the problems associated with earlier scales. These are the AAMD Adaptive Behavior Scales (Nihira et al., 1969), and the VRRI Adaptive Functioning Index (Marlett, 1973). Specifically, these scales

reflect the multidimensional conceptualization of adaptive behavior, and, they rely on observational techniques for assessment.

AAMD Adaptive Behavior Scales. "The Adaptive Behavior Scale is a behavior rating scale for mentally retarded and emotionally maladjusted individuals". It is designed to provide objective description and assessment of "...the effectiveness of the individual in coping with the natural and social demands of his environment" (Nihira et al., 1969).

The scale consists of two parts. Part I is the product of a comprehensive review of behavioral rating scales in the United States and Great Britain. It is concerned with skills and habits important to the maintenance of personal independence in daily living (Nihira et al., 1969). The ten behavior domains are:

- | | |
|----------------------------|---------------|
| I. Independent Functioning | (Ind. Funct.) |
| II. Physical Development | (Phys. Dev.) |
| III. Economic Activity | (Econ. Act.) |
| IV. Language Development | (Lang. Dev.) |
| V. Number and Time Concept | (N TC) |
| VI. Occupation - Domestic | (Occ. - Dom.) |
| VII. Occupation - General | (Occ. - Gen.) |
| VIII. Self-Direction | (Self-Dir.) |
| IX. Responsibilities | (Res.) |
| X. Socialization | (Soc.) |

(Nihira et al., 1969)

Part II of the Adaptive Behavior Scale is based on an analysis of the reasons why parents institutionalize their retarded children, and on a critical-incidents study of inappropriate behavior in special education, day care centres, and institution wards or cottages (Foster & Nihira, 1969). It purports to measure maladaptive behavior related to personality and behavior disorders (Nihira et al., 1969). It consists of the following domains:

I. Violent and Destructive Behavior	(Viol. Des. Beh.)
II. Antisocial Behavior	(Anti. Soc. Beh.)
III. Rebellious Behavior	(Reb. Beh.)
IV. Untrustworthy Behavior	(Un. Beh.)
V. Withdrawal	(Withd.)
VI. Stereotyped Behavior and Odd Mannerisms	(St. Beh. O. Man.)
VII. Inappropriate Interpersonal Manners	(Inapp. Inter. Man.)
VIII. Inappropriate Vocal Habits	(Inapp. Voc. Hab.)
IX. Unacceptable or Eccentric Habits	(Unacc. Ecc. Hab.)
X. Self-Abusive Behavior	(Self-Ab. Beh.)
XI. Hyperactive Tendencies	(Hyper. Tend.)
XII. Sexually Aberrant Behavior	(Sex. Ab. Beh.)
XIII. Psychological Disturbances	(Psych. Dist.)
XIV. Use of Medications	(Use Med.)

(Nihira et al., 1969)

The scale consists of three types of items. The first type requires the rater-observer to select one statement from a series of hierarchically organized behavioral descriptions that best describes the subject. For example:

EATING IN PUBLIC: Check the one statement that best describes the patient's ability to use public eating facilities.

- 3 Orders complete dinner in restaurant.
- 2 Orders simple meals; e.g. hamburgers, hotdogs.
- 1 Orders soft drinks at soda fountain or canteen.
- 0 Does not use public facility at all.

The second type requires the rater-observer to indicate whether or not a particular type of behavior occurs, as well as selecting all examples of that behavior that have been observed. For example:

DAMAGES OTHERS' PROPERTY: Check "No" or "Yes". If "Yes", select all statements that are true of the person.

Yes _____ No _____

- a. Rips, tears, or chews others' clothing.
- b. Soils others' property.

c. Deliberately tears up others' magazines, books, or personal possessions.

d. Other: _____

The third type is very similar to the second, but it also permits an estimation of the frequency with which particular behaviors occur. For example:

RUNS AWAY OR ATTEMPTS TO RUN AWAY: Circle "Yes" or "No".
If "Yes" check each statement that is true of the person being considered.

Yes _____ No _____

Occasionally Frequently

- | | | |
|-------|-------|--|
| () | () | a. Runs away from hospital grounds. |
| () | () | b. Attempts to run away from hospital grounds. |
| () | () | c. Slips out of hospital ward. |
| () | () | d. Other: _____ |

Two forms of the scale exist: one for children twelve years and younger, and one for adolescents and adults thirteen years and older. Information is obtained by observation. The authors claim that the scale can be administered by any personnel who are familiar with the daily behavior of the subjects (Nihira et al., 1969).

Very little reliability data is provided for the scale. Estimates of inter-rater reliability for samples of adults are given. These coefficients range from .86 for Independent Functioning to .40 for Withdrawal, Stereotyped Behavior and Odd Mannerisms, and Inappropriate Interpersonal Manners. Mean inter-rater reliabilities for Parts I and II are .74 and .61 respectively. The mean reliability for the entire scale is .67 (Nihira et al., 1969).

There are several factors that may contribute to these rather low estimates of inter-rater reliabilities. One is the lack of standardized procedures

for assessing basic academic skills. A second problem is the lack of operational definitions of behaviors, particularly those assessed in Part II. The third is the ambiguity of criteria for assessing the frequency of occurrence of behaviors in Part II. Behaviors are considered to occur either occasionally, defined as "...once in a while or now and then," or frequently, defined as "...quite often or habitually" (Nihira et al., 1969).

Evidence for validity comes from a series of empirical investigations. Two factor analytic studies were conducted, one with children and adolescents, and one with adults. Three factors emerged from these studies. Two of these Personal Independence and Social Maladaptation, appear to be stable across ages. The third factor, Personal Maladaptation, failed to appear in the early adolescent group. The domains that define this factor in other age groups loaded significantly on Social Maladaptation (Nihira, 1969a; Nihira, 1969b).

At least two other studies indicate that parts of the scale correlate highly with previous classifications of adaptive behavior, based on clinical judgements (Nihira et al., 1969). In one of these investigations, the experimental version of Part I of the scale was used. A large number of items were found to discriminate significantly between at least two levels of adaptive functioning. In addition, at least 68.7% of the original 325 items showed a significant correlation ($\alpha = .05$) with adaptive behavior, independent of measured intelligence, on at least one adaptive behavior level (Nihira et al., 1968).

The normative data provided for the Adaptive Behavior Scale consists of mean scores for each domain by age, sex, and level of measured intelligence. The standardization sample consisted of institutionalized retardates.

VRRI Adaptive Functioning Index. "The Adaptive Functioning Index

(AFI) is a training and assessment tool to be used by those working with the developmentally handicapped adolescent and adult" (Marlett, 1973). The scale is divided into three sections. The Social Education Test (SET) is an individualized objective test, with criteria for marking correct and incorrect responses (Marlett, 1973). It consists of items from the following domains of content:

- | | |
|------------------------|-------------|
| I. Reading | (Read.) |
| II. Writing | (Writ.) |
| III. Numbers | (Num.) |
| IV. Time | (Time) |
| V. Money | (Mon.) |
| VI. Community | (Commun.) |
| VII. Communication | (Comn.) |
| VIII. Motor Movements | (Mot. Mov.) |
| IX. Concept Attainment | (Con. Att.) |
- (Marlett, 1973)

Most items are scored 0, 1. However, several items in the subtest, Concept Attainment, consist of two or more sections. If one is answered incorrectly, half marks are assigned. A test kit accompanies this scale.

At least two unpublished studies have assessed the stability of SET scores. Coefficients range from .75 for Concept Attainment to .99 for the Reading and Time subtests. The majority of coefficients are greater than .90 and the stability of total scores was assessed at .97 and .99. The stability of scores for the subtest, Motor Movements, was not evaluated in these studies.

At least one unpublished study has examined inter-rater reliability of SET scores. The values range from .83 for the subtest, Community Awareness, to .97 for the subtests, Writing and Money. The coefficient for total

scores was .99. Scores on the subtest, Motor Movements, were not evaluated.

Another study examined the relationship between SET scores and WAIS Full Scale Scores. Only scores on the first six subtests were included. All correlations were significant ($p < .05$) and the total scores for these subtests correlated .71 with WAIS scores.

The second part of the scale is a Vocational Check List (VCL). It consists of ten equal training sequences, each with five goals and three difficulties. Its domains are:

- | | |
|-----------------------------|---------------|
| I. Basic Work Habits | (BWH) |
| A Independence | (Ind.) |
| B Making Decisions | (Mak. Dec.) |
| C Use and Care of Equipment | (U.C. Equip.) |
| D Taking Direction | (Tak. Dir.) |
| II. Work Skills | (Wk. Sk.) |
| A Speed | (Sp.) |
| B Following Instructions | (Fol. Ins.) |
| C Task Analysis | (Ta. An.) |
| III. Acceptance Skills | (Acc. Sk.) |
| A Appearance | (App.) |
| B Punctuality | (Punc.) |
| C Self-Expression - Talking | (Self. Ex.) |
| D Relations with Co-workers | (Rel. Cow.) |

(Marlett, 1973)

The VCL uses a three point scale (0, 1, 2) to assess performance:

- 0 - if the trainee does not show or have a chance to show the behavior
- 1 - if the trainee is making a concerted effort but has not quite mastered the skill, or requires assistance or prompting
- 2 - if the trainee has mastered the skill or behavior and does it on his own

(Marlett, 1973)

In addition the VCL includes a Task Analysis Sheet. This permits the assessment of performance on specific work tasks. These tasks are classified, by the rater, according to skill level. The five skill levels are:

1. One step tasks with few decisions and/or gross dexterity, e.g. one part of a car wash, one task on assembly line, wiping tables, peeling potatoes, sanding, etc.
2. Simple routines, with few decisions and/or gross dexterity, e.g. two or more step packaging, basic janitorial routines, table clearing, etc.
3. Tasks requiring judgment or exacting precision, e.g. spot welding, set-up work, precision assembly, etc.
4. Complex routines requiring judgment and high skill, within an established work routine, e.g. quality control, stock control, lead hand, etc.
5. Complex routines requiring judgment, high skill, flexibility, and an ability to make decisions on his/her own, e.g. pay clerk, work foreman, librarian, etc.

(Marlett, 1973)

Competency of performance is assessed using a four point scale (1, 2, 3, 4).

1. Work needs to be checked several times each day.
2. Work O.K. if checked once a day.
3. Good but needs checking 1-2 times per week.
4. Quality good and stays good.

(Marlett, 1973)

The Residential Check List (RCL) is an index of 150 skills considered necessary for independent functioning in the community. It consists of the following domains:

I	Personal Routines	(Per. Rou.)
A	Cleanliness	(Clean.)
B	Appearance and Eating	(App. & Eat.)
C	Room Management	(Rm. Man.)
D	Time Management	(Tm. Man.)
E	Health	(Heal.)
II	Community Awareness	(Com. Aw.)
A	Transportation	(Trans.)
B	Shopping	(Shop.)
C	Leisure	(Lei.)
D	Budgeting	(Bud.)
E	Cooking and Home Management	(Ck. Hm. Man.)
III	Social Maturity	(Soc. Mat.)
A	Communication	(Commun.)
B	Consideration	(Consid.)
C	Getting Friends	(Get. Fr.)
D	Keeping Friends	(Kp. Fr.)
E	Handling Problems	(Hand. Prob.)

(Marlett, 1973)

The RCL uses the same three point scale as the VCL to assess performance. At the present time there is no information on the reliabilities of the current forms of these two scales.

Some evidence of the validity of the AFI is available. From a factor analysis of the AFI, three factors emerged: home-living, work, and education (Marlett, 1974). SET subtests showed the strongest loading for the third factor, education. Although the RCL and VCL subtests did emerge as separate factors, a few subtest scores did load on factor 3. This overlap represents vocational and residential behavioral descriptions of SET items (Marlett, 1974). For example the subtest, Shopping, of the RCL, loaded on factor 1, home-living, and factor 3, education.

Normative data is not available, at the present time, for the AFI.

Although these scales (the ABS and the AFI) appear to have overcome some of the problems of early measures of adaptive behavior by reflecting a multidimensional conceptualization of adaptive behavior and by relying on observational techniques for assessment, the information available to date suggests that the ABS and AFI have other problems associated with them. First the interrater reliabilities of the ABS are low, indicating that a great deal of measurement error contributes to scores on this test. However, the method of analysis used by the authors of the ABS does not permit the isolation of particular sources of error. Secondly, the validity of both the ABS and the AFI has been assessed using sophisticated techniques such as factor analysis. Neither scale, however, has been correlated (at least in the published literature) with other measures of adaptive behavior. This type of analysis would indicate how much of what is measured by these scales is generalizable from one test to another and how much is an artifact of particular test. At the present time there is insufficient information available for assessing the utility of the ABS and the AFI. This study will attempt to provide some of this information by

1. providing estimates of the internal consistencies of the ABS Parts I and II, the SET, the VCL, and the RCL
2. assessing the interrater reliabilities of the ABS Parts I and II, the VCL, and the RCL using generalizability analysis; this analysis will help to isolate specific types of error that appear to be influencing scores, at least on the ABS
3. estimating the stability of SET scores¹
4. assessing the validity of these scales by examining their mutual relationship in a multitrait, multi method framework as well as the relationship of each scale to measured intelligence

5. determining the ability of these scales to discriminate among individuals at all levels of adaptive functioning
 6. evaluating the usefulness of these tests for programming purposes
-

¹It was not possible to assess the stability of the ABS Parts I and II, the VCL and RCL (see page 32)

Chapter III

Methods

Design of Study

In order to evaluate the two measures of adaptive behavior the following procedures were used. A sample of mentally retarded adults were selected. During the first phase of data collection each subject was assessed individually on the Stanford-Binet Intelligence Scale and on the Social Education Test of the AFI. In addition, each subject was rated independently by two staff members who were familiar with the daily behavior of the subjects on the ABS Parts I and II and on the VCL and RCL of the AFI. Three months after this data was collected, a random sample from the initial group of subjects was retested on the Social Education Test.

Subjects

The subjects used in this study were selected from Bevan Lodge, a facility for mentally retarded adults located near Courtenay, B.C. At the time of this study, there were fifty-nine adults in residence and an additional thirty-six who attended on a day basis. Only those adults who lived at Bevan Lodge were considered for this study. From this initial pool of fifty-nine (thirty-five males and twenty-four females), four individuals were excluded because of profound sensory handicaps. These included two deaf males, one deaf female, and one blind male. An additional four subjects, three males and one female were not included in the final sample because their levels of measured intelligence exceeded the upper limit of the mildly retarded range, as defined by the AAMD classification system. Table 1 summarizes this system. All four subjects had been assessed within the two years prior to this study by the psychologist at the Department of Mental Health, Courtenay, B.C. Their full scale WAIS IQ's ranged from 76 to 115.

The final sample consisted of fifty-one subjects, twenty-nine males and twenty-two females. All subjects participated in one of eight programs offered by Bevan Lodge, from which raters were selected. Table 2 provides a description of these groups of subjects and the total sample by age, sex, and other biodemographic variables.

Raters

Two staff members from each of the eight program groups were selected to complete the adaptive Behavior Scales and the Vocational and

TABLE 1

Standard deviation and IQ score ranges corresponding to AAMD measured intelligence levels¹

Level	Range in 50 units	Revised Stanford Binet scores	Wechsler scores
Mild	-2.01 to -3.00	67-52	69-55
Moderate	-3.01 to -4.00	51-36	54-40
Severe	-4.01 to -5.00	35-20	39-25
Profound	< -5.00	<u>≤</u> 19	<u>≤</u> 24

¹Taken from Grossman (1973)

TABLE 2
Description of sample

Program Group	N			Mean Ca. (Yrs.)	Mean Time at Bevan (Yrs.)	Mean Time in other Institution (Yrs.)	Mean Total Time in Institutions (Yrs.)
	M	F	T				
Maintenance	9	0	9	32.03	3.63	7.58	11.21
Housekeeping	4	6	10	33.57	3.55	6.29	9.84
Woodworking	2	0	2	26.96	2.96	0.00	2.96
Community Work	2	2	4	39.50	4.58	9.54	14.12
Arts and Crafts	1	4	5	27.42	4.88	7.55	12.43
Activity I	4	2	6	28.43	2.56	4.58	7.14
Activity II	4	2	6	26.85	3.43	4.97	8.40
Cafeteria	3	6	9	28.16	2.31	4.03	6.34
Total	29	22	51	30.55	3.40	5.90	9.30

Residential Check Lists of the Adaptive Functioning Index. Since a number of staff members were involved in the Cafeteria program, rater selection was based on the consistency of exposure and the recommendations of the program director. Thus, casual, part-time, and rotating staff were excluded. For programs that usually involved only one full-time staff person, that is, Laundry and Housekeeping, Woodworking, and Arts and Crafts, a second rater was chosen on the basis of past and anticipated exposure to this group.

During the first data collection period, three staff members left Bevan Lodge, and a fourth changed from full-time to casual part-time status and assumed new responsibilities. Three other staff members left for annual summer vacations. As a result of these changes, five additional raters were selected to complete both Adaptive Behavior Scales and Adaptive Functioning Index check lists, for the Cafeteria, Arts and Crafts, Housekeeping and Laundry, and Community Work Programs. Two other staff members were asked to finish rating assignments. One completed Adaptive Behavior Scales for the Activity Group I; the second was responsible for AFI check lists for the Cafeteria Program. Of these seven additional raters, two had other rating assignments as well. Table 3 provides a description of the raters who participated in this study.

Subject Attrition

Despite attempts to ensure that each subject was tested on both the RSB and the SET, as well as rated twice on the ABS, the VCL and the RCL, not all this data was obtained. Table 4 summarizes the data that was obtained,

TABLE 3
Description of Raters

Age	N	Sex	N	Education	N	Length of Employment Yrs.	Job Classification	N
21-30	5	M	7	Graduate School	0	0-1 3	Aide	2
31-40	3	F	11	Four Years College	0	1-2 2	Asst. Program Worker	3
41-50	3			Two Years College	4	2-3 2	Program Worker	4
51-60	4			High School	12	3-4 0	Senior Program Worker	1
61-65	2			Elementary School	1	4-5 2	Cook	3
Unknown	1			Unknown	1	5-6 8	Maintenance	2
						Unk. 1	Housekeeping	1
							Other	2

TABLE 4

Numbers of subjects tested by program group, test and rater

Sample		RSB	SET	ABS		VCL		RCL	
Program	N			R1	R2	R1	R2	R1	R2
Maintenance	(9)	9	9	9	9	8*	8*	8*	8*
Housekeeping	(10)	10	9*	10	10	10	0*	10	0*
Woodworking	(2)	2	2	2	2	2	0*	2	2
Community	(4)	4	4	4	4	4	4	4	4
Arts & Crafts	(5)	5	5	5	5	5	5	5	5
Activity I	(6)	6	6	5*	5*	6	6	6	6
Activity II	(6)	6	6	6	6	6	6	6	6
Cafeteria	(9)	9	9	9	5*	9	5*	9	0*

*indicates that one or more subjects were not assessed

by test and program group. It indicates the numbers of subjects included in the initial sample, by program group, and the numbers of subjects who were assessed on each test and by each rater.

In the Maintenance Program, one subject was not included in the sample to be rated on the VCL and RCL because he had left Bevan Lodge.

One subject in the Laundry and Housekeeping Program was away on summer holidays when the SET was administered. In addition, one rater for this program did not complete the practice RCL and VCL before leaving on vacation. Because this rater's check lists were returned after the first data collection period had ended and were completed incorrectly, they were excluded from this study.

Only one staff member was responsible for the Woodworking program. As a result, a second rater was not available who considered himself/herself able to do an adequate assessment of these subjects' vocational performance in the detail demanded by the VCL.

Additional changes in the composition of program groups, subsequent to the developments that occurred just prior to Data Collection I resulted in two subjects (one by each rater) being excluded from the rating lists for Activity Group I.

With regard to the Cafeteria program, one rater had worked with only five of the nine subjects, because of the organization of the shifts. In addition, only one rater felt able to complete the RCL for this sample.

Materials

The standard protocols and test kits for the Revised Stanford Binet Intelligence Scale, Form LM, (Terman & Merrill, 1960) and the Social Education Test (Marlett, 1973) were used. Two packages were prepared for each staff rater. The first, the ABS package contained a brief description of the present study, a copy of the adult version of the AAMD Adaptive Behavior Scales, answer forms, a list of scoring instructions for this scale, and, a list of subjects that a given staff member would be required to rate. The scoring instructions included those contained in the AAMD Adaptive Behavior Scales Manual (Nihira et al., 1969), as well as additional ones required for this study. The second, the AFI package, contained copies of the Vocational and/or Residential Check Lists, as well as instructions for completing these scales. These instructions included those which were part of the scale, those suggested by the Administration Manual, and those required for this study.

The hand scoring sheets for the Adaptive Behavior Scales and the Vocational and Residential Check Lists were also used.

A questionnaire, based on the Description of Raters contained in the Manual of the AAMD Adaptive Behavior Scales (Nihira et al., 1969) was developed in order to obtain biodemographic information from raters.

Procedure

Data was collected in two sessions separated by an interval of three months. The first session was of nine weeks' duration, the second required two weeks.

Data Collection I: Individual Testing. During this first phase, the Revised Stanford Binet Intelligence Scale (RSB) was administered to all subjects by the experimenter. A modification of standard procedures occurred at the end of each individual testing session, when each subject was asked to define the terms handicapped and mentally retarded. This information was requested by the Bevan Lodge staff. Individuals, whose current level of intellectual functioning was suspected to be in the mildly retarded range on the basis of information from their files, were assessed first to ensure that no subject exceeded the upper limit of this range. Mental age scores for Chronological Age eighteen were used to define the IQ ranges for the AAMD levels of retardation of intellectual functioning. These are presented in Table 5.

The Social Education Test was administered to each subject after the IQ testing was completed. Because of ambiguities in certain administration and scoring instructions, a supplementary list of guidelines was prepared. This is presented in Appendix A.

Rating Scales. The raters selected for this study did not begin their training and testing sessions until three weeks after the beginning of the first data collection period. This delay was the result of recent program developments that involved changes in the composition of the groups.

Each staff member participated in two individual training sessions. During the first, the rater received the AAMD Adaptive Behavior Scale package. After these materials were reviewed, each person completed a practice scale for a subject from their program area who was not

TABLE 5

RSB mental age equivalents by level of measured intelligence for
chronological age eighteen

Level	IQ Range	M.A. Range
Mild	67 to 52	11-0 to 8-6
Moderate	51 to 36	8-5 to 6-1 to
Severe	35 to 20	6-0 5-3
Low Severe and Profound	≤ 19	$\leq 5-2$

participating in this study. This was scored by the experimenter and discussed with the rater, who was then asked to complete the remainder of his/her assessments. These were also scored by the Experimenter; and, any omissions or questions were discussed with the rater. This was not possible in two cases. One rater left Bevan Lodge; a second did not return from her annual summer leave before the end of the first data collection period.

After completing the Adaptive Behavior Scales, raters received the Adaptive Functioning Index packages and began the second training session. The procedures were the same as those for the Adaptive Behavior Scales, with one addition. Items for the Task Analysis Sheet of the Vocational Check List were selected for each program group by the experimenter and the appropriate raters.

It was decided that raters would complete the Adaptive Behaviors Scales first since the more specific items of the Vocational and Residential Check Lists might bias the ABS ratings.

Critical Incidents. During the last part of Data Collection I, the daily reports for the year June 1, 1973 to June 1, 1974 were reviewed. All reported incidents of maladaptive behavior on the part of members of the sample were retrieved. These were to be used in a practical validation of Part II of the ABS.

Data Collection II. During the second data collection period, a stratified random sample of thirty subjects were retested on the Social Education test. Four subjects of each sex and of each modified level of retardation

of measured intelligence were selected.¹ Table 6 summarizes the distribution of the total sample of subjects according to these variables.

Retesting on the Adaptive Behavior Scales and Vocational and Residential Check Lists was not feasible. A total of seven staff members had either left Bevan Lodge or assumed new responsibilities there since the beginning of the first phase of data collection. In addition, the opening of a new day program facility, and the hiring of new staff members left only nineteen subjects with the original raters and program groups.

Method of Analysis

Because two ratings were available for the majority of subjects on the ABS, RCL, and VCL, two "analysis samples" were constructed for these tests. Subjects, within each program group, were randomly assigned to the two analysis samples so that one-half of a given rater's tests were in each sample and so that no subject was included more than once in each sample. These analysis samples were used for test and item analyses, as well as for the correlational analysis that was used to assess the validities of these tests.

Test Analysis and Item Analysis. These analyses were performed on all tests, except the RSB, using the LERTAP test analysis package (Nelson, 1974), available in the Faculty of Education at the University of British Columbia.

¹The total sample included only three mildly retarded males and three severely retarded males.

TABLE 6

Distribution of subjects by sex and level of measured intelligence.

Level	Male	Female	Total
Mild	3	5	8
Moderate	14	6	20
Severe	3	6	9
Low Severe and Profound	9	5	14

Cronbach's alpha for stratified tests (Cronbach, Schönemann, & McKie, 1965) was used to assess the internal consistency of total test scores. The formula for this coefficient is provided in Rajaratnam, Cronbach, & Gleser, 1965)

$$\alpha_{\text{strat}} = 1 - \frac{\sum_h k_h MS_{r(h)}}{s_t^2}$$

where k = the number of items in a given subtest or stratum, h .

$MS_{r(h)}$ = the residual mean square from the persons-by-items matrix for a given subtest or stratum, h .

s_t^2 = variance for the total test.

A criterion of .40 for item-total test correlations was selected in order to examine discriminating and non-discriminating items (Nunnally, 1967). The stringency of this criterion is justifiable since these scales are behaviorally rather than cognitively oriented.

Interrater Reliability. Traditional methods of assessing the interrater reliability of rating scales have been based on the correlational model of test theory. Ordinarily, two or more raters judge subjects on a list of items. A correlation coefficient is computed for these sets of scores. This coefficient is taken as an indication of the extent to which raters agree in their judgments. According to this model, raters are considered to be parallel forms of the measuring operation, and, the error of measurement, or, more specifically, disagreement among raters, is assumed to be random and undifferentiated. These assumptions may be appropriate where carefully equated parallel forms of a test are used. However they are less descriptive

of rating scales, where

1. raters are not necessarily equivalent
2. error may be due to
 - a. differences in the central tendency of values that raters assign;
 - b. differences among raters in overall judgments of specific subjects; and
 - c. differences due to the qualities (e.g. items) to which different raters attend.

(Cronbach, Gleser, Nanda, & Rajaratnam, 1972)

Cronbach, Rajaratnam, and Gleser (1963) have proposed a theory of generalizability which utilizes an analysis of variance approach to test theory. Within the framework of this theory, an observation is described in terms of conditions.¹ The general term referring to conditions of a certain kind is facet.² Facets alone or in combination define universes (Cronbach et al., 1972). Practical and theoretical concerns determine the universe to which an observation is to be generalized, and, hence, the plan or design for collecting this data.

Generalizability theory assumes that

1. conditions of observation are not necessarily parallel;
2. conditions (especially items) may be randomly sampled from a universe, or in accord with a stratified plan; and
3. two or more facets may be analyzed simultaneously.

(Cronbach et al., 1972)

¹Similar to levels in traditional ANOVA terminology.

²Similar to factors in traditional ANOVA terminology.

From the analysis of data collected in a given generalizability study, estimates are obtained for the components of variance, each of which is attributed to one facet or combination of facets represented in the experimental design. These variance estimates, "the hypothetical components of an observed score" (Cronbach et al., 1972), are derived from the equations for Expected Mean Squares (EMS) (Cornfield & Tukey, 1956). From these components, a coefficient of generalizability can be calculated. This is defined as the ratio of universe score variance or "wanted variance" (Gleser, Cronbach, & Rajaratnam, 1965) to expected observed score variance. "It expresses on a 0 to 1 scale how well the observation is likely to locate individuals relative to other members of the population.... For each of several possible universes of generalization, there is a corresponding variance ratio" (Cronbach et al., 1972).

This study was concerned with the accuracy of generalizing from

1. the judgments of each rater pair; and
2. the judgments of all rater pairs

over the universe of judges for each subtest of the ABS, the VCL, the RCL, and for the total scores of these scales. To assess this, two types of analyses were carried out.

Generalizability of each rater pair. To estimate the generalizability of the judgments of each rater pair, a three way ANOVA was carried out for each program group for each subtest and the total tests. The three facets were persons (p), judges (j), and items (i). The design was completely crossed, pxjxi. The analysis was for a mixed model, with

persons and judges treated as random, and items as fixed. The equations for the Expected Mean Squares for the mixed model are:

$$\text{EMSp} = n_i \sigma^2(p_j, \bar{e}/I^*) + n_i n_j \sigma^2(p/I^*)^1$$

$$\text{EMSj} = n_i \sigma^2(p_j, \bar{e}/I^*) + n_p n_i \sigma^2(j/I^*)$$

$$\text{EMSi} = \sigma^2(p_{ij}, \bar{e}/I^*) + n_p \sigma^2(ij/I^*) + n_j \sigma^2(pi/I^*) + n_p n_j \sigma^2(i/I^*)$$

$$\text{EMSpj} = n_i \sigma^2(p_j, \bar{e}/I^*)$$

$$\text{EMSpi} = \sigma^2(p_{ij}, \bar{e}/I^*) + n_j \sigma^2(pi/I^*)$$

$$\text{EMSji} = \sigma^2(p_{ij}, \bar{e}/I^*) + n_p \sigma^2(ij/I^*)$$

$$\text{EMSpij} = \sigma^2(p_{ij}, \bar{e}/I^*)$$

The variance components obtained for each of the eight facets and combinations of facets were used in the following calculations:

1. Universe Score Variance ($\epsilon \sigma^2_{\mu p}$) or wanted variance

$$\epsilon \sigma^2_{\mu p} = \sigma^2 p + (1/n'_i) \sigma^2(pi/I^*)$$

where, $\sigma^2(p) =$ variance component for persons

$\sigma^2(pi/I^*) =$ variance component for the persons-by-items interaction for a fixed set of items

$n'_i =$ number of conditions for facet i (items)

2. The error ($\epsilon \sigma^2_{\sigma}$):

$$\epsilon \sigma^2_{\sigma} = (1/n'_j) \sigma^2(pj/I^*) + (1/n'_i n'_j) \sigma^2(p_{ij}, \bar{e}/I^*)$$

where $\sigma^2_{pj} =$ variance component for the persons-by-judges interaction for a fixed set of items

$\sigma^2_{p_{ij}, \bar{e}} =$ variance component for the persons-by-judges-by-items interaction for a fixed set of items

$n'_i =$ number of conditions for facet i (items)

$n'_j =$ number of conditions for facet j (judges)

¹The notation "I*" is used to indicate that the number of conditions of facet i, that is, the number of items, is fixed.

3. The Expected Observed Score Variance ($\epsilon\sigma^2(x)$):

$$\epsilon\sigma^2(x) = \sigma^2(\mu p) + \sigma^2(\delta)$$

4. The coefficient of Generalizability (ϵp^2)

$$\epsilon p^2 = \frac{\sigma^2(\mu p)}{\epsilon\sigma^2(x)}$$

Generalizability of all rater pairs. To estimate the generalizability of the judgments of all raters, a four facet design with persons (p) crossed with judges (j), both nested in program group (c), and all crossed with items (i) was used. The analyses were performed on a sample of observations consisting of four subjects randomly selected from each program group. For the ABS, the Woodworking program was excluded ($N < 4$). For the VCL, the Woodworking and Laundry and Housekeeping programs were excluded because only one rating was available for each subject. The Woodworking program was excluded from the RCL analysis ($N < 4$). In addition the Cafeteria and Laundry and Housekeeping programs were excluded because only one rating was available for each subject. The analysis was for a mixed model with persons, judges, and program treated as random and items as fixed. An analysis was performed for each subtest and total test.

The equations for the Expected Mean Squares for this design are

$$\begin{aligned} \text{EMSc} &= (n_p n_j n_i) \sigma^2(c/I^*) + (n_j n_i) \sigma^2(p(c)/I^*) + \\ &\quad (n_p n_i) \sigma^2(j(c)/I^*) + (n_i) \sigma^2(pj(c)/I^*) \end{aligned}$$

$$\begin{aligned} \text{EMSi} &= (n_c n_p n_j) \sigma^2(i/I^*) + (n_p n_j) \sigma^2(c_i/I^*) + \\ &\quad (n_j) \sigma^2(p_i(c)/I^*) + (n_p) \sigma^2(j_i(c)/I^*) + \\ &\quad \sigma^2(p_j i(c) \bar{e}/I^*) \end{aligned}$$

$$\text{EMSp}(c) = (n_j n_i) \sigma^2(p(c)/I^*) + (n_i) \sigma^2(p_j(c)/I^*)$$

$$\text{EMSj}(c) = (n_p n_i) \sigma^2(j(c)/I^*) + (n_i) \sigma^2(p_j(c)/I^*)$$

$$\begin{aligned} \text{EMSci} &= (n_p n_j) \sigma^2(c_i/I^*) + (n_j) \sigma^2(p_i(c)/I^*) + \\ &\quad (n_p) \sigma^2(j_i(c)/I^*) + \sigma^2(p_{ij}(c) \bar{e}/I^*) \end{aligned}$$

$$\text{EMSpj}(c) = (n_j) \sigma^2(p_j(c)/I^*)$$

$$\text{EMSpi}(c) = (n_j) \sigma^2(p_i(c)/I^*) + \sigma^2(p_{ij}(c) \bar{e}/I^*)$$

$$\text{EMSji}(c) = (n_p) \sigma^2(j_i(c)/I^*) + \sigma^2(p_{ij}(c) \bar{e}/I^*)$$

$$\text{EMSp_{ij}}(c) = \sigma^2(p_{ij}(c) \bar{e}/I^*)$$

The variance components obtained for each of the nine facets and combinations of facets were used in the following calculations:

1. The Universe Score Variance ($\epsilon \sigma^2_{\mu p}$) or wanted variance:

$$\epsilon \sigma^2_{(\mu p)} = \sigma^2(p(c)/I^*) + (1/n'_i) \sigma^2(p_i(c)/I^*)$$

where $\sigma^2(p(c)/I^*)$ = variance component for persons, nested in program group, for a fixed set of items.

$\sigma^2(p_i(c)/I^*)$ = variance component for the persons-by-items interaction, nested in program group for a fixed set of items.

n'_i = number of conditions for facet i (items).

2. The Error ($\epsilon \sigma^2(\delta)$):

$$\epsilon \sigma^2(\delta) = (1/n_j) \sigma^2(p_j(c)/I^*) + (1/n'_i n'_j) \sigma^2(p_{ij}(c) \bar{e}/I^*)$$

where $\sigma^2(p_j(c)/I^*)$ = variance component for the persons-by-judges interaction, nested in program group, for a fixed set of items.

$\sigma^2(\text{pij}(c)\bar{e}/I^*)$ = variance component for the persons-by-items-by-judges interaction, nested within program group for a fixed set of items.

3. The Expected Observed Score Variance ($\epsilon\sigma^2(x)$):

$$\epsilon\sigma^2(x) = \sigma^2(\mu p) + \sigma^2(\delta)$$

4. The Coefficient of Generalizability (ϵp^2):

$$\epsilon p^2 = \frac{\epsilon\sigma^2(\mu p)}{\epsilon\sigma^2(x)}$$

Stability. Generalizability theory was also used to assess the stability of the SET subtest and total test scores. Occasions were treated as conditions of a facet and were considered to be randomly sampled from a three month time span. Generalization was over the universe of occasions, for this time span. A three way ANOVA was performed for each subtest as well as the total test. The three facets were persons (p), occasions (o), and items (i). The analysis was for a mixed model with persons and occasions treated as random, and items as fixed. The equations for the Expected Mean Squares for this design are:

$$\text{EMSp} = n_o n_i \sigma^2(p/I^*) + n_i \sigma^2(po/I^*)$$

$$\text{EMSo} = n_p n_i \sigma^2(o/I^*) + n_i \sigma^2(po/I^*)$$

$$\text{EMSi} = n_p n_o \sigma^2(i/I^*) + n_o \sigma^2(pi/I^*) + n_p \sigma^2(oi/I^*) + \sigma^2(poi\bar{e}/I^*)$$

$$\text{EMSp}o = n_i \sigma^2(po/I^*)$$

$$\text{EMSpi} = n_o \sigma^2(pi/I^*) + \sigma^2(poi, \bar{e}/I^*)$$

$$\text{EMSoi} = n_p \sigma^2(oi/I^*) + \sigma^2(poi, \bar{e}/I^*)$$

$$\text{EMSp}oi = \sigma^2(poi, \bar{e}/I^*)$$

The variance components obtained for each of the eight facets and combinations of facets were used in the following calculations:

1. Universe score variance ($\epsilon\sigma^2(\mu p)$) or wanted variance.

$$\epsilon\sigma^2(\mu p) = \sigma^2(p/I^*) + (1/n'_i)\sigma^2(pi/I^*)$$

where $\sigma^2(p/I^*)$ = the variance component for persons, for a fixed set of items.

$\sigma^2(pi/I^*)$ = the variance component for the persons-by-items interaction for a fixed set of items.

n'_i = number of conditions of facet i (items).

2. The Error ($\epsilon\sigma^2(\delta)$):

$$\epsilon\sigma^2(\delta) = (1/n'_o)\sigma^2(po/I^*) + (1/n'_on'_i)\sigma^2(poi\bar{e}/I^*)$$

where $\sigma^2(po/I^*)$ = the variance component for the persons-by-occasions interaction, for a fixed set of items.

$\sigma^2(poi\bar{e}/I^*)$ = the variance component for the persons-by-occasions-by-items interaction, for a fixed set of items.

n'_o = number of conditions of facet o (occasions)

n'_i = number of conditions of facet i (items).

3. The Expected Observed Score Variance ($\epsilon\sigma^2(x)$):

$$\epsilon\sigma^2(x) = \sigma^2(\mu p) + \epsilon\sigma^2(\delta)$$

4. The Coefficient of Generalizability (ϵp^2)

$$\epsilon p^2 = \frac{\epsilon\sigma^2(\mu p)}{\epsilon\sigma^2(x)}$$

Validity. In order to provide some evidence of the validity of the ABS and the AFI, two approaches were used.

The first involved the comparison of Part I of the ABS to the AFI in multitrait-multimethod matrices (Campbell & Fiske, 1959). Traits were

defined by matching the domains and/or subdomains of the ABS with the appropriate subtests of the SET, VCL, and RCL. These traits fell into three categories: Practical Academic Skills and Motor Development (ABS with SET); Vocational Skills (ABS with VCL); and, Self-Help Skills (ABS with RCL). Each test was treated as a separate method of measuring these traits. A correlation matrix was obtained for each category of traits as measured by the two methods. Convergent and discriminant validity were assessed using the following criteria:

1. The validity coefficients, i.e., correlations for the same trait as measured by two methods are significantly different from zero ($\alpha = .05$).
2. The validity coefficients are greater than correlations where neither trait nor method is the same.
3. The validity coefficients are greater than correlations where traits are different but the method is the same.
4. The pattern of correlations is the same for different traits using the same method as for different traits using different methods.

(Campbell & Fiske, 1959)

Test-retest reliability coefficients for the subtests of the SET were used in the multi-trait multi-method matrices. Inter-rater reliability coefficients were used for the ABS Part I, the VCL, and the RCL. These were selected as being closest to Campbell and Fiske's (1959) definition of reliability as agreement between two measurements of the same trait using the same method.

The second validation procedure used the reported incidents of maladaptive behavior as a practical criterion in assessing the validity of Part II of the ABS. The recorded incidents were sorted, by subject, into the thirteen behavioral

domains defined by ABS Part II. All incidents were sorted in this way. For each domain of maladaptive behavior, three subgroups were defined. The first, the "No Incidents" group (NIN), included those subjects for whom no instances of maladaptive behavior had been recorded. The second, the "Low Incidents" group (LIN), consisted of subjects for whom there were very few recorded incidents of maladaptive behavior. The last, the "High Incidents" group included subjects for whom a relatively large number of maladaptive behaviors had been reported. The latter two groups were defined in terms of the mean number of incidents for subjects for whom such incidents were recorded. Subjects in the LIN group fell below the mean; whereas, subjects in the HIN group fell above the mean. The means for these three groups of subjects were computed for each domain of the ABS Part II, and compared to those for the reported incidents data.

Chapter IV

Results of Analysis

In this section the results of the test analyses, item analyses, and generalizability analyses will be presented for each test: the ABS Part I, the ABS Part II, the SET, the VCL, and the RCL. The results of the multi-trait multi-method analyses of the ABS Part I, the SET, the VCL, and the RCL will be presented in a section on validity. This section will also include the assessment of the practical validity of the ABS Part II.

Results of Analysis - ABS Part I

Test Analysis. Tables 7 and 8 show the mean, range, standard deviation, standard error of measurement, and, the Hoyt estimate of internal consistency for each domain of the ABS Part I, for each analysis sample. Table 9 summarizes these results for the total test. For analysis sample one, the Hoyt coefficients range from .46 for Responsibilities to .85 for Independent Functioning. These values range from .50 for Responsibilities to .88 for Economic Activity for the second sample. The internal consistencies of the ratings for the total test are .95 and .94, for the first and second samples, respectively. A restriction in the range of scores was observed for at least four domains: Independent Functioning, Physical Development, Language Development, and Socialization. This sample does not appear to represent all ranges of functioning for these domains.

Subtest and total test correlation for the ABS Part I are presented in Tables 10 and 11. For both analysis samples, ratings of Physical Development correlate lowest with total test scores; and, scores on Independent Functioning correlate highest with total test results.

Item Analysis. Figures 1 and 2 present the results of the item analyses for the ABS Part I for analysis samples one and two respectively.¹ For analysis sample one, thirty-nine of the sixty-nine items correlate .40 or greater with total test scores. At least one item from each domain

¹In these and subsequent similar figures, each point represents an item and indicates the item difficulty (ordinate) and item-total score correlation (abscissa). The difficulty and correlational indices for the items of the ABS Part I are presented in tabular form in Appendix B.

TABLE 7

Test analysis information for subtests of the ABS, Part I(1)

Subtest	Mean	Range	St.Dev.	Hoyt	SEM
Ind. Funct.	112.60	82 - 125	10.21	.85	3.86
Phys. Dev.	24.02	19 - 26	1.82	.53	1.14
Econ. Act.	10.88	3 - 17	4.35	.80	1.70
Lang. Dev.	31.71	17 - 43	6.93	.81	2.84
NTC	10.56	0 - 14	3.74	.75	1.51
Occ. Dom.	12.52	1 - 19	4.26	.74	1.99
Occ. Gen.	10.85	3 - 13	2.23	.57	1.20
Self-Dir.	26.85	9 - 33	5.86	.86	2.04
Res.	4.33	1 - 6	.97	.46	.50
Soc.	23.25	11 - 31	4.26	.64	2.37

TABLE 8

Test analysis information for subtests of the ABS, Part I(2)

Subtest	Mean	Range	St.Dev.	Hoyt	SEM
Ind. Funct.	111.83	89 - 126	9.42	.80	4.15
Phys. Dev.	23.79	19 - 26	2.05	.52	1.29
Econ. Act.	10.56	1 - 18	4.69	.88	1.42
Lang. Dev.	31.08	18 - 43	6.52	.75	3.09
NTC	10.65	1 - 14	3.41	.63	1.70
Occ. Dom.	11.71	4 - 19	4.67	.76	2.08
Occ. Gen.	10.85	3 - 13	2.35	.62	1.18
Self-Dir.	25.98	11 - 33	5.62	.83	2.16
Res.	4.15	1 - 6	1.13	.50	.56
Soc.	23.48	15 - 32	4.37	.71	2.18

TABLE 9

Test analysis summary for the ABS, Part I

	Mean	Range	St.Dev.	α Strat	SEM
ABS (1) Pt. I	267.58	184 - 317	31.78	.95	7.93
ABS (2) Pt. I	264.08	186 - 322	31.57	.94	8.16

TABLE 10

Correlations among subtest and total test scores for the ABS, Part I(1)¹

Subtest	1	2	3	4	5	6	7	8	9	10	11
1 Ind. Funct.	1.00										
2 Phys. Dev.	.23	1.00									
3 Econ. Act.	.64	-.03	1.00								
4 Lang. Dev.	.51	.23	.70	1.00							
5 NTC	.64	.22	.74	.78	1.00						
6 Occ. Dom.	.34	.17	.30	.25	.24	1.00					
7 Occ. Gen.	.38	.09	.31	.17	.25	.38	1.00				
8 Self-Dir.	.54	.18	.36	.33	.32	.38	.79	1.00			
9 Res.	.46	-.12	.36	.19	.33	.26	.68	.60	1.00		
10 Soc.	.42	.18	.55	.48	.52	.20	.28	.47	.31	1.00	
11 ABS 1 Part I	.85	.29	.79	.76	.79	.51	.56	.71	.54	.66	1.00

¹N = 48; $r_{\text{critical}} = .282$ ($P \leq .05$)

TABLE 11

Correlations among subtest and total test scores for the ABS, Part I(2)¹

Subtest	1	2	3	4	5	6	7	8	9	10	11
1 Ind. Funct.	1.00										
2 Phys. Dev.	.19	1.00									
3 Econ. Act.	.65	.25	1.00								
4 Lang. Dev.	.40	.19	.70	1.00							
5 NTC	.40	.10	.55	.72	1.00						
6 Occ. Com.	.44	.29	.49	.33	.15	1.00					
7 Occ. Gen.	.59	.14	.35	.27	.34	.40	1.00				
8 Self-Dir.	.66	.24	.55	.29	.22	.48	.72	1.00			
9 Res.	.59	.11	.44	.10	.31	.36	.47	.55	1.00		
10 Soc.	.40	.30	.39	.55	.42	.32	.39	.43	.31	1.00	
11 ABS Pt. I(2)	.84	.35	.83	.72	.62	.63	.66	.76	.57	.66	1.00

¹N = 48 $r_{\text{critical}} = .282$ ($P \leq .05$)

appears to be discriminating among members of this sample. For the second analysis sample, forty-one items correlate .40 or greater with total test scores. Similarly, each domain contains some discriminating items. The non-discriminating items will be described for each domain.

For the domain Independent Functioning eleven items correlate .40 or greater with total test scores for the first sample. Only eight items reach this criterion for sample two. A number of items from this domain, do not appear to be discriminating among members of these samples. These are the items measuring the skills of Drinking, Table Manners, Cleanliness, Appearance, Care of Clothing, and Dressing and Undressing. These items appear to be easy for these samples.

For the subtest, Physical Development, only those items measuring Body balance and Hearing are discriminating among members of sample one and two respectively. Thus five items fail to discriminate among members of each sample.

All items in the domain, Economic Activity, correlate .40 or greater with total test scores.

The items measuring Ability to Speak, Clarity of Speech and General Language Development do not distinguish between high and low scores on this test. They appear to be easy for these samples. The remaining six items correlate .40 or greater with total test scores for both samples. All items measuring Number and Time Concepts correlate .40 or greater with total test scores.

FIGURE 1

Item analysis information for the ABS Part I (1)

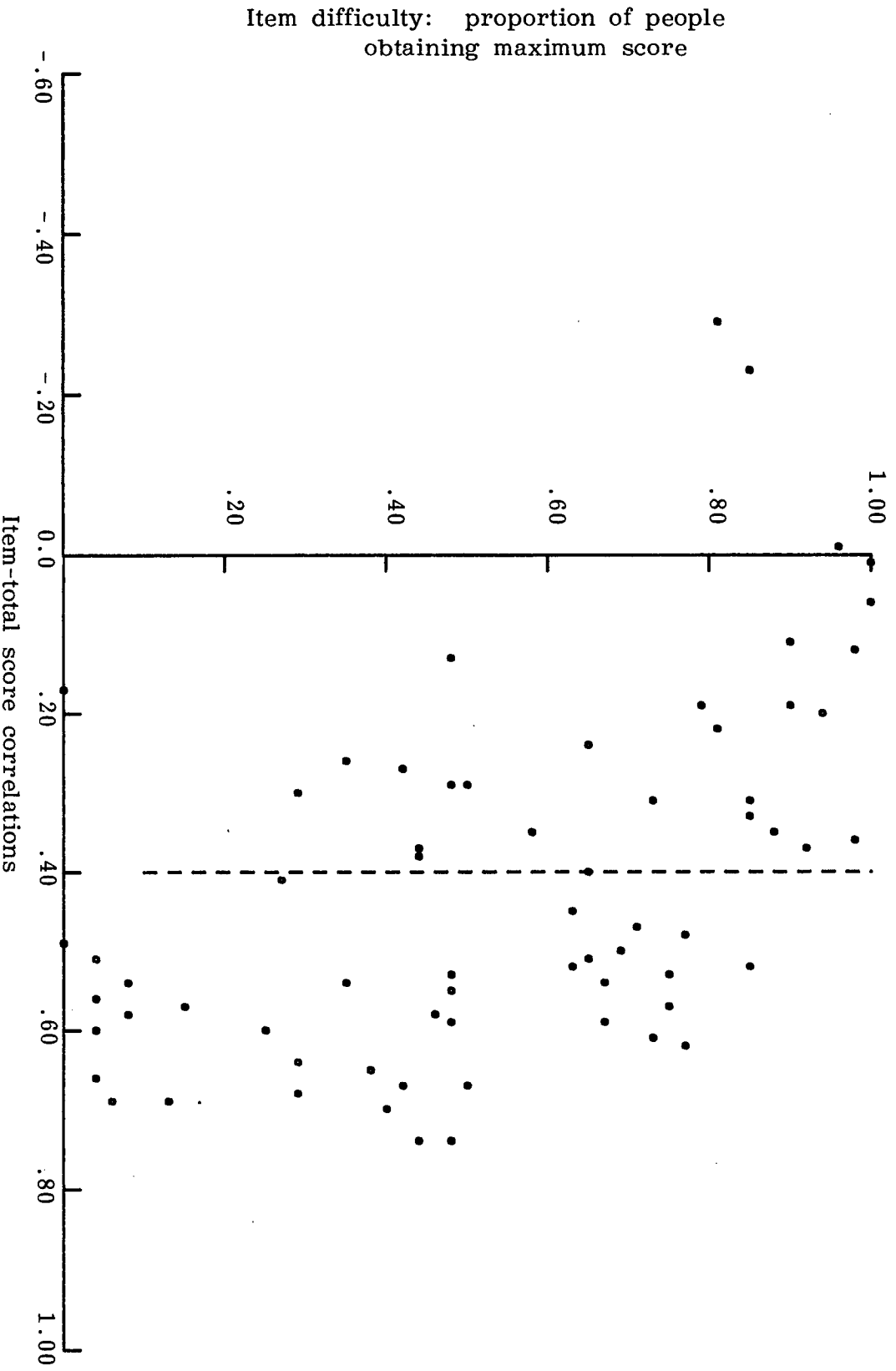
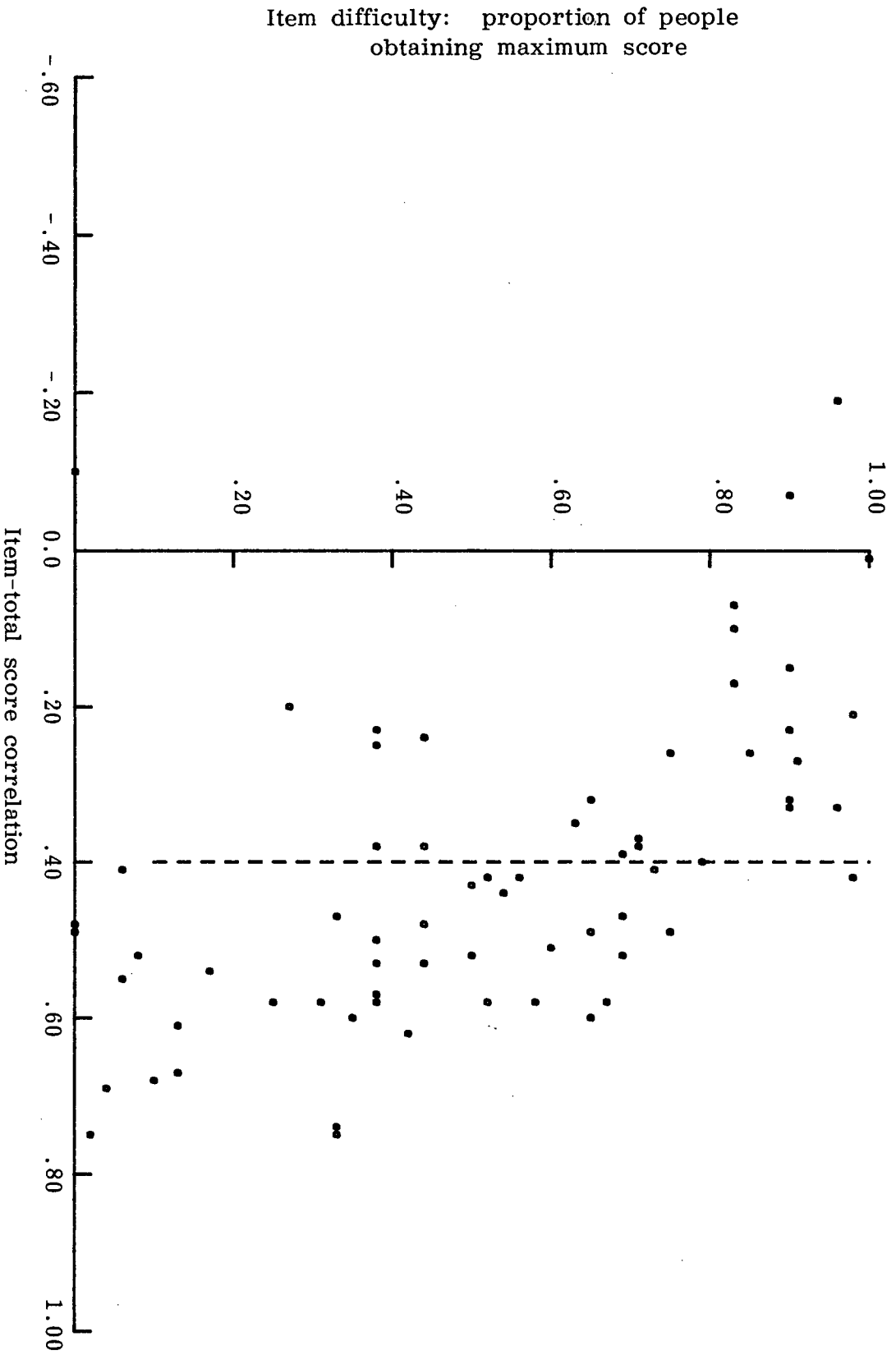


FIGURE 2
Item analysis information for the ABS Part I (2)



For the domain, Occupation Domestic, the items measuring cleaning, table clearing, and general domestic skills do not discriminate among members of these samples. Table clearing appears to be an easy item. Thus only three items from this subtest discriminate among members of these samples.

The item measuring Work Fitness in the subtest, Occupation General, does not discriminate among members of analysis sample one. It appears to be difficult for this group. All three items discriminate among members of the second sample.

For the subtest, Self-Direction, the items pertaining to Sluggishness in Movements and Perseverance do not correlate well with total test scores although Perseverance approaches .40. Six items from this domain clearly discriminate among members of these samples.

The item measuring Responsibility for Belongings, from the domain Responsibilities, does not discriminate well among members of analysis sample one. Both this item and the item measuring other responsibilities discriminate among members of the second sample.

Finally, for the subtest, Socialization, the items measuring Interaction with Others and Selfishness fail to discriminate among members of either sample. In addition, the items Cooperation and Other Immaturities in Socialization correlate less than .40 with total test scores for sample one. The item Knowledge of Others fails to discriminate among members of the first sample. Only the items measuring Consideration and Participation distinguish among high and low scorers for both samples.

Generalizability of Ratings. Estimates of the generalizability of ratings over the universe of raters for the ABS Part I are presented in Table 12. Table 13 provides estimates for the variance components that were used in determining these coefficients.

In examining these results for the domains, it can be seen that nine coefficients are equal to zero. For these coefficients, the universe score variance equals zero. Of the ten domains, there are five for which at least one coefficient equals zero. Responsibilities has three coefficients equal to zero. For rater pairs 3, 5, and 8,¹ the generalizability coefficient for at least one domain equals zero. Rater pair three has five coefficients equal to zero.

Of the remaining seventy-one coefficients, seventeen are greater than or equal to .80. For the subtest Language Development, four are greater than or equal to .80. The subtests, Occupation Domestic and Occupation General have no coefficients greater than or equal to .80. The remaining seven subtests have at least one coefficient greater than .80. For rater pair 7, six coefficients are greater than or equal to .80. Rater pairs 2 and 8 have no ratings where generalizability is greater than or equal to .80. The remaining five pairs of raters have at least one coefficient greater than or equal to .80.

In examining the fifty-four coefficients that are greater than zero but less than .80, it can be seen that the persons-by-judges (pj) variance component is the largest or equal to the largest for twenty-one coefficients.

¹Rater pairs are not equivalent from test to test

TABLE 12

Generalizability coefficients by rater pair and across raters for the subtest and total test scores of the ABS Part I

Domain	Rater Pair								Across Raters
	1	2	3	4	5	6	7	8	
Independent Functioning	.90	.66	.36	.04	.74	.81	.68	.12	.74
Physical Development	.83	.62	0.0	.63	.39	.11	.87	.15	.74
Economic Activity	.52	.72	.76	.06	.95	.94	.85	0.0	.62
Language Development	.79	.55	.95	.42	.91	.83	.83	.35	.76
Number and Time Concepts	.78	.70	0.0	.63	.76	.21	.87	.04	.69
Occupation: Domestic	.11	.49	0.0	.42	.16	.67	.37	.78	.53
Occupation: General	.58	.05	0.0	.14	.70	.54	.56	0.0	.33
Self-Direction	.88	.65	.38	.16	.04	.55	.86	.59	.54
Responsibilities	.58	.54	0.0	.80	0.0	.67	.49	0.0	.26
Socialization	.35	.33	.55	.88	.76	.50	.81	.34	.50
Adaptive Behavior Scales Pt. I	.85	.73	.13	.18	.79	.80	.94	.40	.74

TABLE 13

Variance components contributing to universe score variance (σ^2_{up}) and error variance (σ^2_{δ}) by rater pair and across raters for subtest and total test scores of the ABS Part I¹

Subtest		Rater Pair								Across Raters
		1	2	3	4	5	6	7	8	
		A B	A B	A B	A B	A B	A B	A B	A B	
Independent Functioning	σ^2_{up}	.12+.02	.10+.01	.004+.003	.004+.00	.04+.01	.40+.01	.05+.01	.002+.00	.09+.01
	σ^2_{δ}	.01+.01	.04+.02	.002+.01	.09+.02	.01+.01	.09+.02	.01+.01	.01+.01	.04+.01
Physical Development	σ^2_{up}	.06+.02	.08+.02	0.0+0.0	.04+.01	.02+.00	.01+.00	.03+.02	.001+.00	.07+.003
	σ^2_{δ}	.01+.02	.03+.04		.01+.03	.02+.04	.03+.05	.01+.01	.004+.009	.01+.04
Economic Activity	σ^2_{up}	.38+.03	.65+.03	.66+.07	.02+.02	1.21+.05	2.31+.00	.53+.01	0.0+0.0	.52+.004
	σ^2_{δ}	.23+.15	.22+.04	.12+.13	.59+.09	.03+.04	.09+.06	.05+.05		.25+.01
Language Development	σ^2_{up}	.26+.06	.19+.01	1.20+.05	.08+.00	.83+.06	.57+.08	.33+.05	.09+.01	.37+.02
	σ^2_{δ}	.05+.03	.12+.04	.01+.06	.07+.04	.06+.02	.11+.02	.06+.02	.12+.06	.09+.02
Number, Time Concepts	σ^2_{up}	1.40+.11	.90+.13	0.0+0.0	.26+.06	1.30+.12	0.0+.14	.73+.08	.03+0.0	.58+.01
	σ^2_{δ}	.30+.12	.25+.19		.06+.13	.16+.19	.41+.11	.04+.08	.51+.21	.17+.02
Occupation Domestic	σ^2_{up}	0.0+.02	.12+.03	0.0+0.0	.14+.02	0.0+.03	.32+.00	.07+.00	.20+.08	.13+.02
	σ^2_{δ}	.12+.03	.10+.06	.01+.05	.19+.03	.14+.03	.07+.09	.07+.05	.04+.03	.08+.05
Occupation General	σ^2_{up}	.29+.02	0.0+.01	0.0+0.0	.09+.00	.15+.06	.32+.11	.27+.00	0.0+0.0	.13+.00
	σ^2_{δ}	.13+.09	.10+.05	1.38+.18	.38+.17	.02+.07	.26+.10	.07+.13	.07+.09	.15+.11
Self-Direction	σ^2_{up}	.44+.04	.18+.02	.39+.08	.03+.02	.01+.00	.38+.00	.66+.04	.02+.02	.15+.04
	σ^2_{δ}	.03+.04	.08+.04	.74+.02	.18+.03	.15+.04	.26+.05	.07+.03	.01+.01	.11+.08
Responsibilities	σ^2_{up}	.16+.08	.05+.06	0.0+0.0	.17+.08	0.0+0.0	.17+.00	.15+.00	0.0+0.0	.04+.00
	σ^2_{δ}	.08+.09	.06+.03	.78+.03	.06+.00	.11+.04	.04+.04	.08+.06	.01+.01	.07+.04
Socialization	σ^2_{up}	.05+.01	.07+.01	.22+.05	.22+.05	.27+.03	.08+.03	.46+.03	.11+.00	.13+.04
	σ^2_{δ}	.06+.05	.11+.05	.21+.02	.02+.02	.07+.02	.07+.03	.08+.03	.11+.09	.10+.12
ABS Part I	σ^2_{up}	.11+.01	.08+.01	0.0+.01	.01+.002	.11+.01	.27+.01	.15+.003	.01+.001	.09+.003
	σ^2_{δ}	.02+.003	.02+.04	.04+.01	.05+.04	.03+.03	.06+.04	.01+.003	.01+.004	.03+.004

¹For σ^2_{up} $A=\sigma^2_{pi}/I^*/n_i$; for σ^2_{δ} , $A=\sigma^2_{pj}/I^*/n_j$ and $B=\sigma^2_{pij}/I^*/n_{ij}$. Equations are found on page

All subtests have at least one coefficient where the p_j component is the largest. Occupation Domestic has three coefficients where the p_j component contributes most to the observed score variance. All pairs of raters have at least one coefficient where the p_j component is the largest. Rater pairs 2 and 4 have five coefficients each to which the p_j component contributes most to the observed score variance. Thus, rater bias appears to be influencing the ratings on all subtests and the judgments of all rater pairs.

The generalizability coefficients, for total test scores, by rater pair range from .13 to .94. For rater pairs 3, 4, and 8, the p_j component contributes most to observed score variance.

The generalizability coefficients across all raters range from .26 for Responsibilities to .76 for Language Development. The p_j component contributes most to the observed score variance for the subtests Occupation General and Responsibilities. The generalizability coefficient for total test scores is .74, a value comparable to that cited by the authors of this scale.

Results of Analysis - ABS Part II

Test Analysis. Tables 14 and 15 show the mean, range, standard deviation, standard error of measurement, and the Hoyt estimates of internal consistency for each subtest of the ABS Part II for each analysis sample. Hoyt coefficients are not provided for the subtests, Inappropriate Interpersonal Manners, Unacceptable Vocal Habits, Self-Abusive Behavior, Hyperactive Tendencies, and Use of Medication since these subtests consist of only one item. Table 16 summarizes these results for the total tests. The Hoyt coefficients range from 0.0 for Stereotyped Behavior and Odd Mannerisms

TABLE 14

Test analysis information for subtests of the ABS, Part II (1)

Subtest	Mean	Range	St. Dev.	Hoyt	SEM
Viol. Des. Beh.	1.42	0 - 22	3.64	.67	1.86
Antisoc. Beh.	3.33	0 - 28	6.18	.77	2.68
Reb. Beh.	3.46	0 - 26	5.58	.65	3.00
Un. Beh.	.92	0 - 16	2.97	.71	1.14
Withd.	.79	0 - 7	1.73	.53	.97
St. Beh. O. Man.	.40	0 - 3	.76	0	.58
Inapp. Inter. Man.	.40	0 - 3	.79	-	-
Unacc. Voc. Habits	.52	0 - 4	1.07	-	-
Unacc. Ecc. Habits	.54	0 - 4	1.15	.17	.91
Self-Ab. Beh.	.08	0 - 2	.35	-	-
Hyper. Tend.	.38	0 - 8	1.33	-	-
Sex. Ab. Beh.	.83	0 - 14	2.50	.51	1.51
Psych. Dist.	5.04	0 - 46	8.84	.91	2.50
Use of Med.	.83	0 - 4	1.10	-	-

TABLE 15

Test analysis information for the subtests of the ABS Part II, (2)

Subtest	Mean	Range	SD	Hoyt	SEM
Viol. Des. Beh.	1.23	0 - 8	2.19	.34	1.59
Antisoc. Beh.	3.65	0 - 18	4.99	.60	2.87
Reb. Beh.	3.10	0 - 16	4.54	.63	2.52
Un. Beh.	.69	0 - 8	1.65	0	1.20
Withd.	1.38	0 - 10	2.13	.38	1.37
St. Beh. O. Man.	1.00	0 - 14	2.32	.46	1.21
Inapp. Inter. Man.	.52	0 - 8	1.43	-	-
Unacc. Voc. Hab.	.44	0 - 6	1.11	-	-
Unacc. Ecc. Hab.	.58	0 - 4	1.16	.04	.99
Self-Ab. Beh.	.13	0 - 2	.44	-	-
Hyper. Tend.	.67	0 - 8	1.75	-	-
Sex. Ab. Beh.	.67	0 - 9	1.63	.26	1.22
Psycho. Dist.	6.33	0 - 44	9.55	.87	3.14
Use of Med.	.96	0 - 4	1.22	-	-

TABLE 16
Test analyses summary for the ABS Part II

	Mean	Range	St.Dev.	α Strat	SEM
ABS(1) Part II	18.94	0 - 142	27.00	.95	6.90
ABS(2) Part II	21.33	0 - 100	23.90	.92	7.06

to .91 for Psychological Disturbances, for analysis sample one. For the second analysis sample, these values range from 0.0 for Untrustworthy Behavior to .87 for Psychological Disturbances. The internal consistency estimates for total test scores are .95 and .92 for analysis samples one and two respectively.

Subtest and total test correlations for ABS Part II are presented in Tables 17 and 18. For analysis sample one, the subtests, Stereotyped Behavior and Odd Mannerisms, Self-Abusive Behavior, and Use of Medication correlate lowest with total test scores ($r = -.05, -.01, \text{ and } -.01$, respectively). Violent and Destructive Behavior, Antisocial Behavior, and Psychological Disturbances correlate highest with total test scores ($r = .90$). For the second analysis sample, this pattern is somewhat altered with Self-Abusive Behavior, Unacceptable Vocal Habits, and Use of Medication showing the lowest correlations with total test scores ($r = .18, .27, \text{ and } .28$, respectively). Psychological Disturbances, Antisocial Behavior, and Rebellious Behavior have the highest correlations with total test scores ($r = .88, .83, \text{ and } .82$, respectively).

Item Analysis. The results of the item analyses for each domain of the ABS Part II are presented in Tables 19 to 27. For analysis sample one, twenty-six of the forty-four items correlate .40 or greater with total test scores. All domains contain discriminating items except

Stereotyped Behavior and Odd Mannerisms
 Unacceptable or Eccentric Habits
 Inappropriate Interpersonal Manners
 Self-Abusive Behavior
 Use of Medication

TABLE 17
Correlations among subtest and total test scores of the ABS Part II(1)¹

Subtest	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Viol. Des. Beh.	1.00														
Anti-Soc. Beh.	.70	1.00													
Reb. Beh.	.67	.82	1.00												
Un. Beh.	.35	.58	.59	1.00											
With.	.15	.23	.26	.13	1.00										
St. Beh., O. Man.	-.04	-.06	-.03	-.13	.03	1.00									
Inapp. Interper. Man.	.14	.36	.36	.14	.33	.09	1.00								
Unacc. Voc. Hab.	.48	.54	.54	.31	-.02	.08	.33	1.00							
Unacc. Ecc. Hab.	.17	.22	.33	.18	.19	.14	.23	-.06	1.00						
Self-Ab. Beh.	.06	-.11	-.02	-.08	.10	.27	-.12	-.12	.31	1.00					
Hyper. Tend.	.22	.40	.40	.15	-.03	.06	.04	.16	.39	.16	1.00				
Sex. Ab. Beh.	.12	.47	.42	.39	.06	-.04	.002	-.02	.37	-.08	.62	1.00			
Psycho. Dist.	.41	.71	.69	.75	.34	-.18	.21	.9	.45	-.04	.43	.64	1.00		
Use of Med.	-.11	-.15	.05	-.10	.09	.18	.10	-.07	.31	-.15	.10	-.06	-.08	1.00	
ABS Pt.II(1)	.66	.90	.90	.73	.35	-.05	.34	.45	.44	-.01	.50	.60	.90	-.01	1.00

¹N=48; r_{critical} = .282 (p ≤ .05)

TABLE 18
Correlations among subtest and total test scores of the ABS Part II(2)¹

Subtest	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1 Vio. Des. Beh.	1.00														
2 Antisoc. Beh.	.43	1.00													
3 Reb. Beh.	.23	.53	1.00												
4 Un. Beh.	.17	.54	.54	1.00											
5 With.	-.02	.20	.57	.03	1.00										
6 St. Beh., O. Man.	.12	.11	.29	-.09	.40	1.00									
7 Inapp. Inter. Man.	.01	.17	.47	.03	.29	.22	1.00								
8 Unacc. Voc. Hab.	.12	.12	.23	-.01	.42	.83	.16	1.00							
9 Unacc. Ecc. Hab.	.22	.42	.58	.23	.43	.36	.20	.52	1.00						
10 Self-Ab. Beh.	.34	.11	.18	-.12	.20	.27	-.07	.23	.35	1.00					
11 Hyper. Tend.	.02	.23	.26	-.04	.23	.55	.33	.37	-.22	-.11	1.00				
12 Sex. Ab. Beh.	.34	.51	.70	.59	.18	.04	.25	.13	.41	.06	-.04	1.00			
13 Psych. Distur.	.35	.77	.61	.64	.27	.07	.18	.02	.36	.03	.31	.54	1.00		
14 Use of Med.	.04	.35	.06	-.04	.10	.08	.001	.09	.14	.01	.33	-.09	.24	1.00	
15 ABS Pt.II(2)	.44	.83	.82	.59	.48	.37	.36	.27	.59	.18	.43	.65	.88	.28	1.00

¹N=48; $r_{\text{critical}} = .282$ ($p \leq .05$)

TABLE 19

Item analysis information for the ABS Part II: Violent & Destructive Behavior

Item	Mean		Standard Deviation		r_{subtest}		r_{total}		p^1	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
1	1.71	1.56	1.71	1.17	.68	.30	.78	.19	.75	.75
2	1.04	1.02	.20	.14	.04	.19	-.04	-.01	.96	.98
3	1.10	1.02	.37	.14	.85	.19	.48	-.01	.92	.98
4	1.13	1.02	.44	.14	.62	.40	.21	.18	.92	.98
5	1.33	1.60	.98	1.44	.81	.27	.59	.50	.81	.77

TABLE 20

Item analysis information for the ABS Part II: Antisocial Behavior

Item	Mean		Standard Deviation		r_{subtest}		r_{total}		p	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
1	2.15	2.23	2.17	2.34	.78	.65	.75	.73	.69	.67
2	1.77	2.27	1.48	2.27	.64	.76	.58	.64	.73	.60
3	1.13	1.13	.44	.64	.38	.05	.29	.29	.92	.96
4	1.31	1.29	1.10	.62	.49	.41	.51	.64	.90	.79
5	1.23	1.21	1.06	.71	.47	.03	.66	.10	.92	.90
6	1.67	1.54	1.67	.90	.51	.22	.72	.49	.73	.67

¹For this and tables 20 to 29, p refers to the proportion of subjects receiving a score of zero.

TABLE 21

Item analysis information for the ABS Part II: Rebellious Behavior

Item	Mean		Standard Deviation		r_{subtest}		r_{total}		p	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
1	1.31	1.29	.69	.65	.50	.41	.54	.33	.79	.81
2	2.15	1.98	2.08	1.86	.77	.47	.84	.60	.67	.65
3	1.29	1.23	1.07	.75	.28	.08	.48	.41	.90	.90
4	2.10	2.02	2.17	2.16	.56	.42	.68	.56	.67	.71
5	1.06	1.15	.32	.58	.20	.37	.09	.28	.96	.94
6	1.46	1.50	1.41	1.13	.22	.59	.32	.75	.79	.79

TABLE 22

Item analysis information for the ABS Part II: Untrustworthy Behavior

Item	Mean		Standard Deviation		r_{subtest}		r_{total}		p	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
1	1.89	1.19	.92	.89	.77	-.03	.69	.04	.94	.92
2	1.60	1.50	1.70	1.41	.76	-.03	.69	.67	.85	.83

TABLE 23

Item analysis information for the ABS Part II: Withdrawal

Item	Mean		Standard Deviation		r_{subtest}		r_{total}		p	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
1	1.15	1.38	.51	1.16	.52	.35	.40	.37	.92	.88
2	1.29	1.35	.94	.84	.24	.16	.19	.48	.90	.81
3	1.35	1.65	.89	1.16	.38	.18	.24	.17	.79	.69

TABLE 24

Item analysis information for the ABS Part II: Stereotyped Behavior
and Odd Mannerisms

Item	Mean		Standard Deviation		r_{subtest}		r_{total}		p	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
1	1.44	1.56	1.09	1.18	-.71	.62	.02	.39	.79	.75
2	1.06	1.29	.32	.65	-.09	.49	-.06	.25	.96	.81

TABLE 25

Item analysis information for the ABS Part II: Unacceptable or Eccentric Habits

Item	Mean		Standard Deviation		r_{subtest}		r_{total}		p	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
1	1.17	1.23	.56	.75	.21	.003	.35	.49	.92	.90
2	1.29	1.27	.82	.79	.17	.07	.37	.29	.89	.88
3	1.00	1.02	.001	.14	0.0	.19	0.0	.21	1.00	.18
4	1.08	1.06	.40	.32	-.09	-.09	.02	.19	.96	.96

TABLE 26

Item analysis information for the ABS Part II: Sexually Aberrant Behavior

Item	Mean		Standard Deviation		r_{subtest}		r_{total}		p	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
1	1.02	1.04	.14	.29	.13	-.06	.01	-.06	.98	.98
2	1.13	1.04	.87	.29	.56	-.06	.42	.21	.98	.98
3	1.44	1.19	1.40	.61	.31	.37	.17	.50	.90	.90
4	1.25	1.40	1.06	1.27	.34	.31	.85	.56	.94	.90

TABLE 27

Item analysis information for the ABS Part II: Psychological Disturbances

Item	Mean		Standard Deviation		r_{subtest}		r_{total}		P	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
1	1.77	2.02	1.63	1.77	.72	.68	.68	.62	.73	.65
2	2.06	2.23	2.08	1.79	.81	.74	.87	.68	.60	.52
3	2.15	2.10	2.04	1.98	.75	.85	.75	.78	.65	.60
4	1.81	1.81	1.58	1.75	.65	.78	.58	.71	.73	.71
5	1.38	1.50	1.27	1.43	.85	.65	.79	.59	.88	.83
6	1.27	1.35	.96	1.06	.62	.44	.64	.51	.88	.81
7	1.69	2.17	1.60	1.99	.81	.66	.80	.81	.73	.58

For the second analysis sample, twenty-two items correlate .40 or greater with total test scores. All domains contain discriminating items except

Stereotyped Behavior and Odd Mannerisms
 Inappropriate Interpersonal Manners
 Unacceptable Vocal Habits
 Self-Abusive Behavior
 Use of Medication

For the subtest Violent and Destructive Behavior, three items correlate .40 or greater with total test scores for the first sample; only one item discriminates among members of the second sample. The items measuring whether an individual damages personal and public property (2 and 4) do not discriminate among members of both samples. In addition, the items measuring tendencies to threaten or do physical violence (1) and to damage others' property (3) correlate less than .40 with total test scores for the second analysis sample.

The item measuring a tendency to disrupt others' activities, (3) in Antisocial Behavior does not discriminate among members of the first analysis sample. The items measuring a tendency to disrupt others' activities and disrespect for others' property (5) from this subtest correlate less than .40 with total scores for the second analysis sample. Thus, five items discriminate among members of the first sample and four items discriminate among members of the second sample.

For the subtest Rebellious Behavior four items discriminate among members of each sample. The item measuring the tendency to "run away" (5) does not discriminate among members of either sample. In addition, the items

measuring the tendencies to ignore regulations and routines (1) and to misbehave in group settings (6) correlate less than .40 with total scores for analysis sample one and two respectively.

For the subtest, Untrustworthy Behavior, both items discriminate among members of the first sample. The item measuring whether an individual takes another's property without permission (1) does not discriminate among subjects in the second analysis sample.

For the subtest, Withdrawal, the items measuring a tendency to be withdrawn but active (2) and withdrawn or shy (3) fail to discriminate among members of the first sample. The latter item also correlates less than .40 with total test scores for the second sample. In addition, the item measuring a tendency to be withdrawn and inactive (1) does not discriminate among members of the second sample.

None of the items in the subtest Stereotyped Behavior and Odd Mannerisms correlate .40 or greater with total test scores for both samples.

Only one item in the subtest, Unacceptable or Eccentric Habits correlates greater than .40 with the total scores for analysis sample one. This item measures tendencies to display strange and unacceptable habits (1). No items in this subtest discriminate among members of the second analysis sample.

For the subtest, Sexually Aberrant Behavior, the item indicating a tendency to engage in inappropriate masturbation (1) does not discriminate among members of either sample. The items measuring the tendencies to expose the body improperly (2) and to display homosexual tendencies (3)

do not discriminate among members of analysis samples two and one respectively. Thus two items correlate .40 or greater with total test scores for each sample.

All items in the subtest, Psychological disturbances correlate .40 or greater with total test scores for both samples.

Of the five one item subtests, only Unacceptable Vocal Habits and Hyperactive Tendencies discriminate among members of analysis sample one. Hyperactive Tendencies also correlates greater than .40 with total scores for the second sample (see Tables 17 and 18).

Estimated Frequencies of Occurrence for Behaviors Measured by ABS Part II. For the items of the ABS Part II, raters are required to indicate whether or not a behavior occurs, and if so, to state whether it occurs "occasionally" defined as "...once in a while or now and then", and "frequently", defined as "...quite often or habitually" (Nihira et al, 1969). For this study, raters were also asked to provide numerical estimates, e.g. 1/da, 2/wk, etc., for behaviors they considered their ratees to display. Table 28 presents the range of estimated frequencies provided by raters for each item for each category: occasionally occurring or frequently occurring behaviors. For many items, these ranges overlap for occasional and frequently occurring behaviors. For some items this overlap is complete. For example, for item eleven,

Uses Profane or Hostile Language

raters who judge this behavior as occurring occasionally provide frequency estimates ranging from once a day to four times a year. Raters who judge

TABLE 28

Range of estimated frequencies of occurrence for items of the ABS Part II and percentage of items for which raters provided estimates

Item	Estimated Frequency			
	Occasional		Frequent	
	Range	%	Range	%
1	1/da - 1/4yr	72	1/da - 1/wk	100
2	3/yr - 1/yr	100	-----	--
3	3/mo - 1/yr	100	4/mo	100
4	1/wk - 1/2yr	100	1/wk - 1/mo	100
5	1/wk - 2/yr	48	3/wk - 1/wk	42
6	1/da - 1/mo	82	1/da - 2/wk	39
7	1/da - 1/mo	70	2/da - 1/wk	60
8	3/mo - 1/mo	100	1/da - 5/wk	60
9	1/da - 1/mo	100	5/da - 3/wk	64
10	2/wk - 4/yr	100	1/da - 5/wk	75
11	1/da - 4/yr	66	1/da - 2/mo	62
12	1/da - 2/yr	71	1/da - 3/wk	50
13	1/da - 1/mo	75	1/da - 3/mo	66
14	2/mo - 1/2mo	33	1/da - 1/wk	75
15	1/da - 1/mo	45	1/da - 1/wk	75
16	1/mo	100	2/wk - 1/wk	100
17	2/mo - 4/yr	38	1/da - 2/wk	76
18	1/mo - 2/yr	80	5/wk - 1/wk	57
19	4/wk - 2/yr	89	1/da - 2/wk	71
20	1/da - 6/yr	89	1/da - 3/wk	88
21	1/wk - 6/yr	66	1/da - 2/wk	64
22	1/da - 1/mo	61	3/da - 4/wk	83
23	2/wk - 1/mo	63	8/da - 3/wk	82
24	1/da - 1/mo	100	2/da - 1/da	66
25	1/da - 2/mo	69	1/da - 2/mo	66
26	1/da - 1/mo	78	4/da - 1/da	69
27	3/mo	100	1/da - 5/wk	33
28	-----	--	4/da - 1/wk	71
29	4/mo	100	-----	--
30	-----	--	1/da	100
31	1/da - 1/mo	66	1/da	100
32	1/da - 1/wk	75	4/da - 5/wk	43
33	-----	--	4/da	100
34	-----	--	3/wk	33
35	1/mo - 1/yr	50	1/da	30
36	1/wk - 1/mo	100	5/wk - 1/wk	83
37	1/da - 4/yr	64	1/da - 1/wk	45
38	1/da - 4/yr	61	1/da - 2/wk	19
39	1/da - 1/2mo	64	5/wk - 1/wk	13
40	2/da - 1/mo	100	1/da - 1/wk	38
41	1/da - 1/mo	90	1/da - 2/wk	50
42	1/da - 3/yr	66	1/da - 1/mo	100
43	2/wk - 1/mo	62	1/da - 1/wk	63
44	1/da	33	4/da - 1/da	79

that subjects display this behavior frequently, provide estimates ranging from once a day to twice a month. For other items, the ranges of estimated frequencies of occurrence for occasional and frequently occurring behaviors showed partial overlap. For example, for item one

Threatens or does Physical Violence,

raters who judge this behavior as occurring occasionally provide frequency estimates ranging from once a day to once every four years. Raters who judge that subjects display this behaviour frequently, provide estimates ranging from once a day to once a week.

Table 29 provides a summary of items showing different degrees of overlap in the range of estimated frequencies of occurrence. It would appear from these results, that raters use different criteria in assigning behaviors to the occasional and frequent categories.

Generalizability of Ratings. Estimates of the generalizability of ratings over the universe of raters for the ABS Part II are presented in Table 30. Estimates of the variance components used in calculating these coefficients are presented in Table 31.

In examining these results for the subtests or domains, it can be seen that fifty-six coefficients are equal to zero. For thirty of these, the universe score variance is equal to zero. For the remaining twenty-six coefficients all components are equal to zero. All domains have from one to seven coefficients equal to zero. All rater pairs have from four coefficients (rater pairs 2 and 7) to eleven coefficients (rater pair 3) that are equal to zero.

TABLE 29

Number of items of the Adaptive Behavior Scales, Part II showing different degrees of overlap in the range of estimated frequency of occurrence.

Degree of overlap in frequency estimates			
No Overlap	Some Overlap	Complete Overlap	Undetermined
8	26	1	9

TABLE 30

Generalizability coefficients by rater pair and across raters for the subtest and total test scores of the ABS Part II

Domain	Rater Pair								Across Raters
	1	2	3	4	5	6	7	8	
Viol. Des. Beh.	.87	.68	0.0	.77	.22	0.0	.06	.80	.45
Antisoc. Beh.	.90	.18	0.0	.91	0.0	0.0	.54	.36	.53
Reb. Beh.	.48	.78	0.0	.73	.40	.03	.70	.42	.30
Un. Beh.	0.0	.15	0.0	.72	0.0	0.0	.58	0.0	.55
With.	.61	.13	.32	0.0	.74	.16	.68	.40	.57
St. Beh., O. Man.	0.0	0.0	.62	.17	.58	.13	.29	.89	.41
Inapp. Inter. Man.	0.0	.31	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Unacc. Voc. Hab.	0.0	0.0	.49	0.0	0.0	0.0	0.0	0.0	0.0
Unacc. Ecc. Hab.	0.0	.53	0.0	.44	.98	0.0	.75	0.0	.70
Self-Ab. Beh.	0.0	0.0	0.0	0.0	.89	0.0	0.0	0.0	.57
Hyper. Tend.	0.0	.08	0.0	0.0	0.0	0.0	1.00	0.0	.48
Sex. Ab. Beh.	0.0	0.0	0.0	0.0	.47	0.0	0.0	0.0	.12
Psych. Dist.	.62	.48	0.0	.53	.53	.21	.18	.70	.52
Use of Med.	.37	.89	0.0	0.0	1.00	1.00	1.00	0.0	.87
ABS Part II	.78	.45	.61	.93	.54	.17	.49	.62	.58

TABLE 31

Variance components contributing to universe score variance ($\sigma^2_{\mu\rho}$) and error variance (σ^2_δ) by rater pair and across raters for subtest and total test scores of the ABS Part II

Subtest		Rater Pair								Across Raters
		1	2	3	4	5	6	7	8	
		A B	A B	A B	A B	A B	A B	A B	A B	
Violent and De- structive Behavior	$\sigma^2_{\mu\rho}$.11+.0002	.07+.01	0 + 0	.05+.05	.03+.01	0 + 0	0+.02	1.39+.42	.04+.01
	σ^2_δ	.01+.01	.01+.04	.05+.02	.01+.02	.12+.04	.003+.01	.15+.13	.38+.06	.04+.02
Rebellious Behavior	$\sigma^2_{\mu\rho}$.01+.003	.45+.10	0 + 0	.09+.08	.002+.02	0 +.02	.62+.13	.16+.16	.04+.04
	σ^2_δ	.01+.01	.02+.08	.05+.11	.01+.05	.03+.01	.41+.11	.08+.25	.33+.10	.11+.08
Antisocial Behavior	$\sigma^2_{\mu\rho}$.13+.06	.08+.02	0 + 0	.31+.01	0 + 0	0 + 0	.24+.23	.20+.03	.11+.03
	σ^2_δ	.02+.004	.25+.19	.06+.04	.01+.02	.12+.06	.06+.04	.25+.14	.38+.02	.07+.06
Untrustworthy Behavior	$\sigma^2_{\mu\rho}$	0 + 0	0 +.09	0 + 0	3.00+1.00	0 + 0	0 + 0	3.45+ 0	0 + 0	.40+.12
	σ^2_δ	0 + 0	.27+.27	.50+.50	1.00+ .56	0 + 0	0 + 0	1.19+1.36	.07+.08	.24+.18
Withdrawal	$\sigma^2_{\mu\rho}$.03+.05	.03+ 0	0 +.17	0 + 0	.28+.16	.09+.05	.42+ 0	.06+.10	.22+.07
	σ^2_δ	.03+.02	.13+.10	.35+.01	0 + 0	.12+.04	.59+.17	.02+.18	.13+.10	.14+.08
Stereotyped Behavior and Odd Mannerisms	$\sigma^2_{\mu\rho}$	0 + 0	0 + 0	0 +.50	.08+ 0	.48+.13	.04+ 0	0 +.13	.10+.10	.10+.06
	σ^2_δ	.03+.04	.10+.10	.03+.28	.27+.10	.23+.11	.14+.14	.16+.16	.01+.01	.13+.09
Inappropriate Inter- personal Manners	$\sigma^2_{\mu\rho}$	0	.62	0	0	0	0	0	0	0
	σ^2_δ	.01	1.40	0	1.23	.08	1.00	.18	.13	.36
Unacceptable Vocal Habits	$\sigma^2_{\mu\rho}$	0	0	3.00	0	0	0	0	0	0
	σ^2_δ	0	0	3.08	1.06	.25	.33	.40	.85	.38
Unacceptable or Eccentric Habits	$\sigma^2_{\mu\rho}$	0 + 0	.03+.07	0 + 0	0+.06	.10+.03	0 + 0	.06+.08	0 + 0	.05+.06
	σ^2_δ	0 + 0	.06+.04	0 + 0	.04+.03	.01+.01	.07+.03	.02+.03	.003+.003	.03+.02
Self-Abusive Behavior	$\sigma^2_{\mu\rho}$	0	0	0	0	.40	0	0	0	.07
	σ^2_δ	.03	0	.50	0	.05	.25	0	0	.05
Hyperactive Tendencies	$\sigma^2_{\mu\rho}$	0	.20	0	0	0	0	2.67	0	.62
	σ^2_δ	.03	2.25	4.50	1.00	0	.06	0	0	.67
Sexually Aberrant Behavior	$\sigma^2_{\mu\rho}$	0 + 0	0 + 0	0 + 0	0 + 0	.06+.09	0 + 0	0 + 0	0 + 0	.01+.01
	σ^2_δ	0 + 0	.31+.10	0 + 0	.02+.02	.08+.10	.04+.06	.16+.10	0 + 0	.12+.05
Psychological Disturbances	$\sigma^2_{\mu\rho}$.09+.003	.45+.05	0 + 0	.26+.05	.25+.002	.10+ 0	.39+.03	.30+.01	.33+.01
	σ^2_δ	.04+.02	.43+.11	.02+.02	.25+.03	.16+.06	.18+.19	1.87+.06	.09+.07	.25+.08
Use of Medication	$\sigma^2_{\mu\rho}$.14	1.42	0	0	1.20	1.00	1.07	0	.64
	σ^2_δ	.24	.18	0	.25	0	0	0	0	.09
ABS Part II	$\sigma^2_{\mu\rho}$.01+.002	.06+.007	.03+.001	.15+.01	.04+.004	.02+ 0	.15+.02	.13+.01	.06+.04
	σ^2_δ	.003+.001	.07+.01	.01+.01	.004+.01	.03+.01	.09+.01	.15+.02	.08+.01	.04+.01

Of the remaining fifty-six coefficients, twelve are greater than or equal to .80. Besides the subtest Use of Medication, for which four coefficients are greater than .80, the domains Violent and Destructive Behavior and Antisocial Behavior each have two coefficients that are greater than or equal to .80.

In examining the forty-four coefficients that are greater than zero but less than .80, it can be seen that the persons-by-judges variance component is the largest or equal to the largest component for twenty-three of these. The subtests Unacceptable or Eccentric Habits, Self-Abusive Behavior, and Sexually Aberrant Behavior have no coefficients where the p_j component is the largest, although these subtests have four to seven coefficients that are equal to zero. The remaining subtests have from one to four coefficients where the p_j component contributes most to the observed score variance. Thus rater bias influences ratings on most domains. All rater pairs have from one (rater pair 4) to five (rater pair 3) coefficients where the p_j component is the largest or equal to the largest component of variance.

The generalizability coefficients for total test scores, by rater pair, range from .17 to .93. For rater pairs 2, 6, and 7 the p_j component contributes the most to the observed score variance.

The generalizability coefficients across all raters range from 0.0, for Inappropriate Interpersonal Manners and Unacceptable Vocal Habits to .70 for Unacceptable or Eccentric Habits. ϕ^2 for Use of Medication is equal to .87. The p_j component contributes the most to the observed score

variance for the domains, Violent and Destructive Behavior, Rebellious Behavior, Stereotyped Behavior and Odd Mannerisms, Hyperactive Tendencies and Sexually Aberrant Behavior. The generalizability coefficient for total test scores is .58.

Results of Analysis - SET

Test Analysis. The mean, range, standard deviation, standard error of measurement, and Hoyt coefficient of internal consistency for each subtest of the SET for both test and retest samples are presented in Tables 32 and 33. This information for the total test is given in Table 34.

For both test and retest samples, a floor effect was obtained for the first six subtests. A ceiling effect was observed for the subtests Writing, Time, and Communication for both samples, and for Money and Concept Attainment for the retest sample. These results indicate that these subtests do not include sufficiently easy and/or difficult items for the members of this sample.

For the test sample, the Hoyt coefficients range from .65 for Motor Movements to .97 for Writing. For the retest sample, these values range from .67 for Motor Movements to .98 for Writing. Only the coefficients for Motor Movements are less than .85. The internal consistency of the total test is .99 for both samples.

Correlations among subtest and total test scores, for both test and retest samples are given in Tables 35 and 36. Motor Movements correlates lowest with total test scores for both samples.

TABLE 32

Test analysis information for the subtests of the Social Education Test 1

Subtest	Mean	Range	St. Dev.	Hoyt	SEM
Reading	6.52	0 - 16	5.93	.94	1.38
Writing	7.54	0 - 20	7.61	.97	1.22
Numbers	6.90	0 - 18	5.74	.93	1.49
Time	8.52	0 - 20	6.05	.94	1.46
Money	7.04	0 - 19	7.00	.96	1.32
Community	6.30	0 - 17	4.63	.87	1.62
Communication	14.36	5 - 20	4.32	.88	1.45
Motor Movements	11.38	5 - 16	2.75	.65	1.59
Concept Attainment	9.56	1 - 19	4.74	.85	1.77

TABLE 33

Test analysis information for the subtests of the Social Education Test (2)

Subtest	Mean	Range	St. Dev.	Hoyt	SEM
Reading	6.97	0 - 17	6.55	.96	1.31
Writing	7.43	0 - 20	7.99	.98	1.11
Numbers	7.00	0 - 17	5.78	.93	1.50
Time	8.57	0 - 20	6.52	.95	1.45
Money	8.03	0 - 20	7.04	.96	1.33
Community	6.13	0 - 15	4.17	.85	1.60
Communication	14.23	5 - 20	4.26	.89	1.41
Motor Movements	11.13	5 - 15	2.84	.67	1.60
Concept Attainment	10.83	3 - 20	5.26	.88	1.76

TABLE 34

Test analysis summary for the Social Education Test, 1 and 2

Test	Mean	Range	St. Dev.	α_{Strat}	SEM
Set I	78.12	11 - 151	43.58	.99	4.84
Set II	80.33	15 - 149	45.78	.99	4.77

TABLE 35

Correlations among subtests and total test score for
the Set 1¹

Subtest	1	2	3	4	5	6	7	8	9	10
Reading	1.00									
Writing	.95	1.00								
Numbers	.89	.90	1.00							
Time	.91	.88	.90	1.00						
Money	.88	.88	.92	.90	1.00					
Community	.80	.81	.80	.82	.82	1.00				
Communication	.70	.70	.65	.71	.65	.73	1.00			
Motor Movements	.47	.41	.43	.44	.42	.31	.45	1.00		
Concept Attainment	.77	.72	.79	.79	.77	.67	.71	.60	1.00	
Total Test	.95	.95	.95	.95	.94	.87	.79	.53	.86	1.00

¹N = 50; $r_{\text{critical}} = .276$ ($p \leq .05$)

TABLE 36

Correlations among subtests and total test scores of the Set 2¹

Subtest	1	2	3	4	5	6	7	8	9	10
Reading	1.00									
Writing	.95	1.00								
Numbers	.95	.94	1.00							
Time	.87	.86	.90	1.00						
Money	.94	.91	.95	.93	1.00					
Community	.86	.86	.88	.83	.89	1.00				
Communication	.76	.75	.74	.70	.71	.80	1.00			
Motor Movements	.45	.37	.42	.36	.44	.33	.37	1.00		
Concept Attainment	.84	.76	.82	.82	.84	.75	.71	.68	1.00	
Total Test	.97	.95	.97	.93	.97	.91	.81	.51	.89	1.00

¹N = 30; $r_{\text{critical}} = .355$ ($p \leq .05$)

Item Analysis. The results of the item analysis for the subtests of the SET are presented in Figures 3 to 11. Of the one hundred and eighty items in this test, one hundred and thirty-two were found to correlate .40 or greater with total test scores.

For the Reading subtest seventeen of the twenty items correlate .40 or greater with total test scores. The three items that do not appear to be discriminating are difficult to this sample. Only four subjects were able to correctly label all Social Sight Vocabulary words and symbols. No subjects received perfect scores on the Recall task. And, only six subjects received perfect scores on the first Comprehension task.

All items in the Writing subtest correlate .40 or greater with total test scores.

For the subtest, Number Concepts, the item measuring proportions and weights and a problem in long division do not appear to be discriminating among members of this sample. The latter is the most difficult item in this subtest ($p = .06$). The remaining eighteen items correlate .40 or greater with total test scores.

Of the twenty items in the Time subtest, only two correlate less than .40 with total test scores. The first question, "What time do you get up in the morning?", is the easiest for this sample ($p = .92$). The second, "How many working days are there in February?", is the hardest item ($p = .06$).

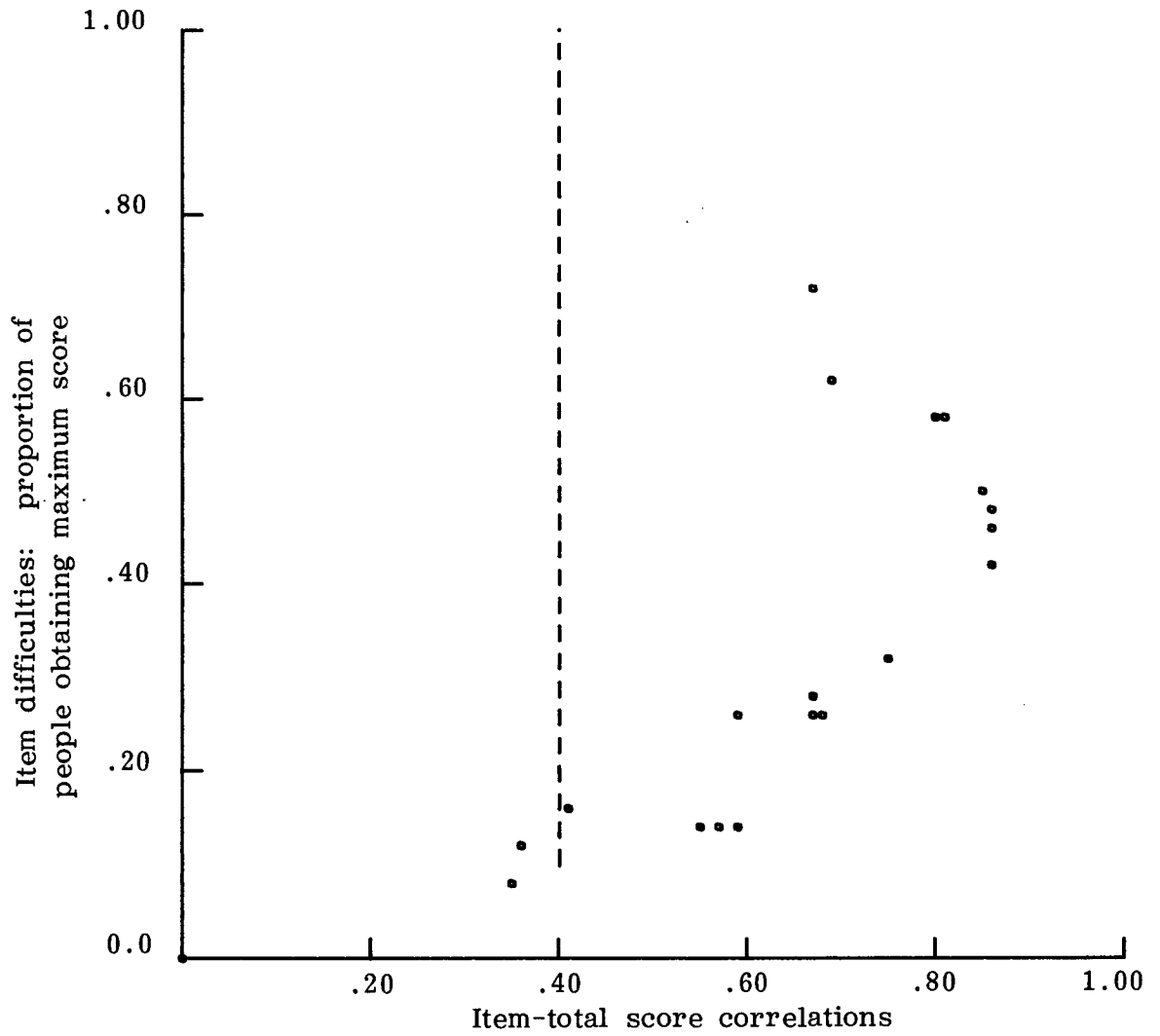


FIGURE 3

Item analysis information for the
Reading subtest of the SET

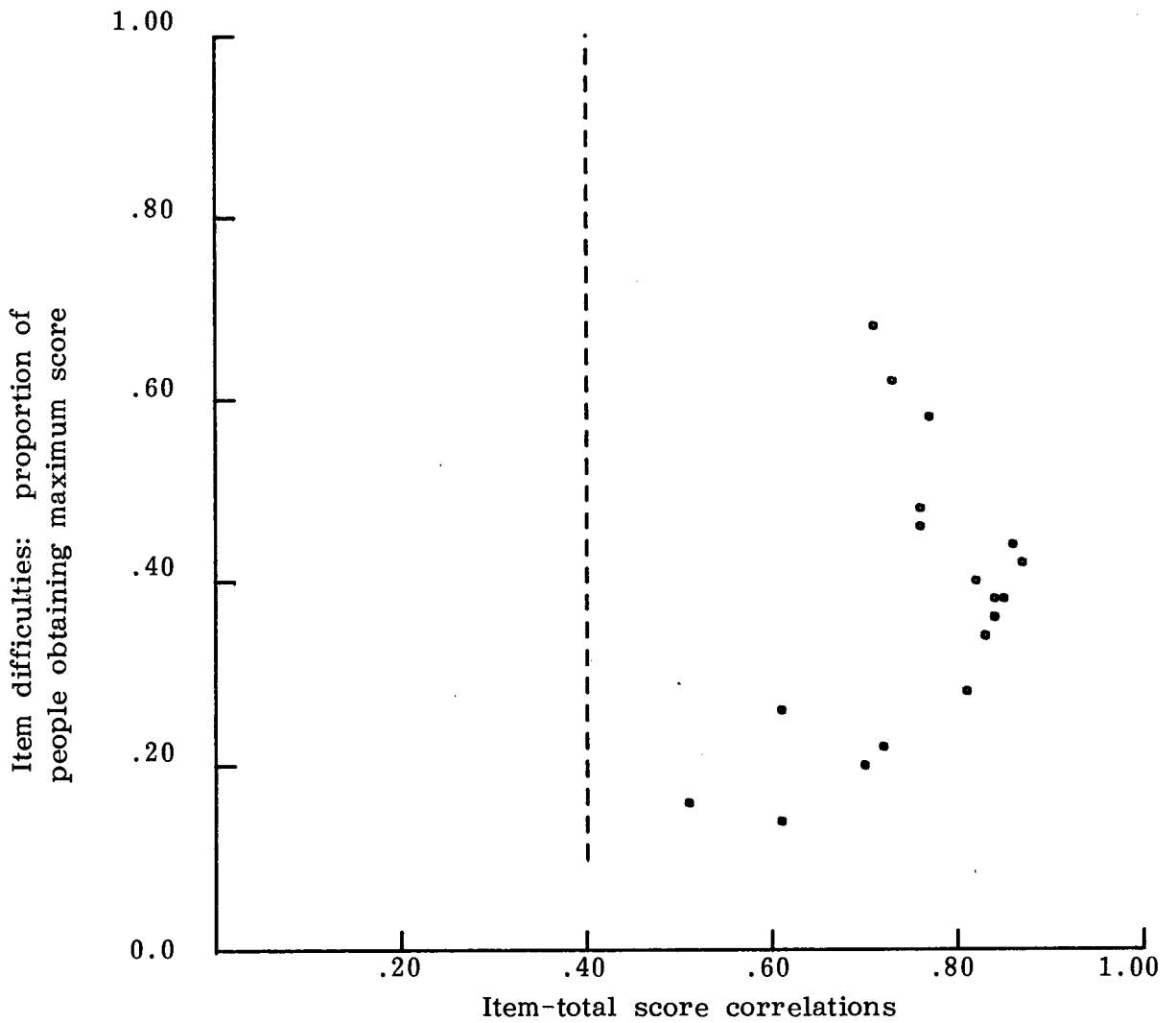


FIGURE 4

Item analysis information for the
Writing subtest of the SET

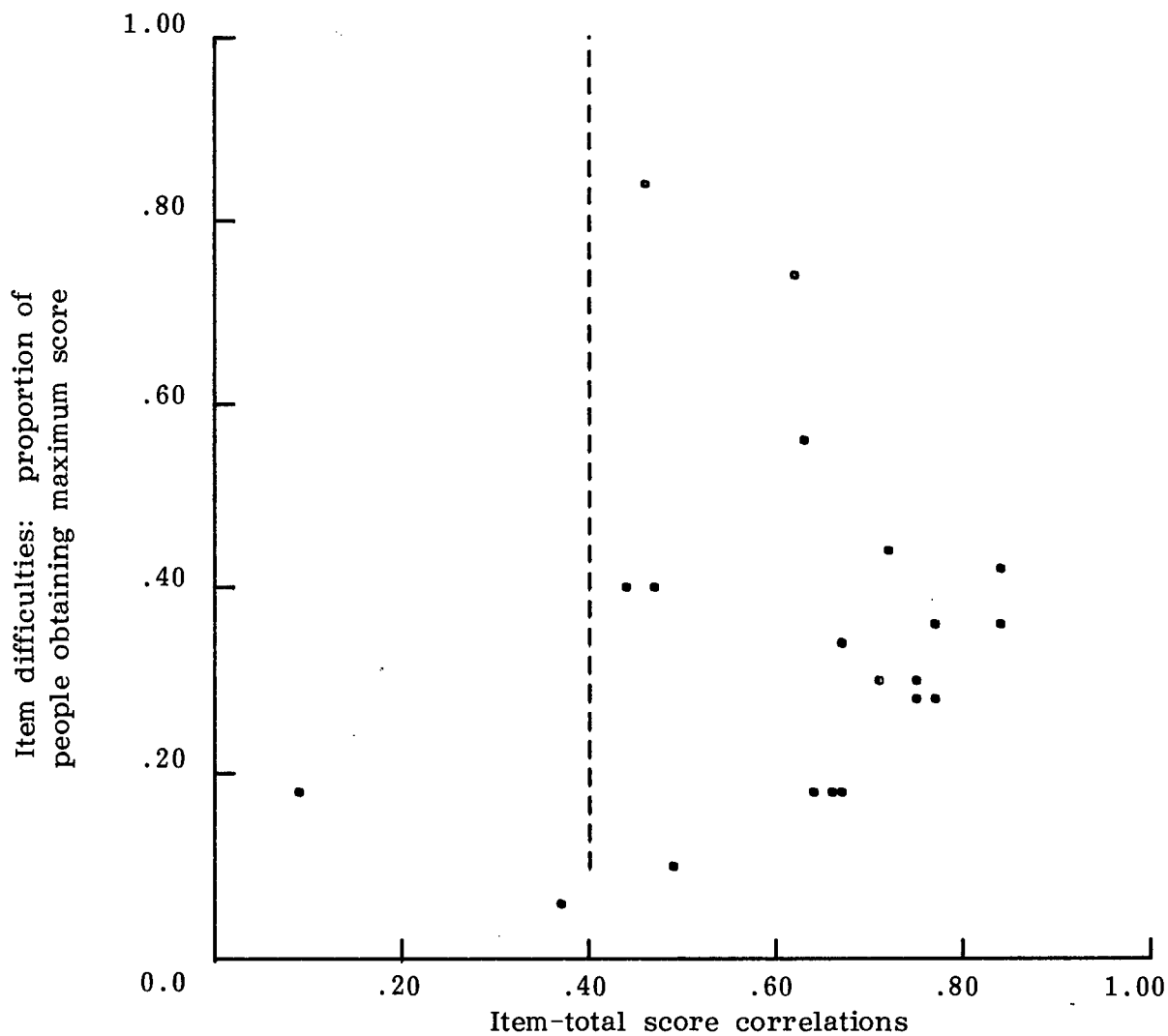


FIGURE 5

Item analysis information for the
Numbers subtest of the SET

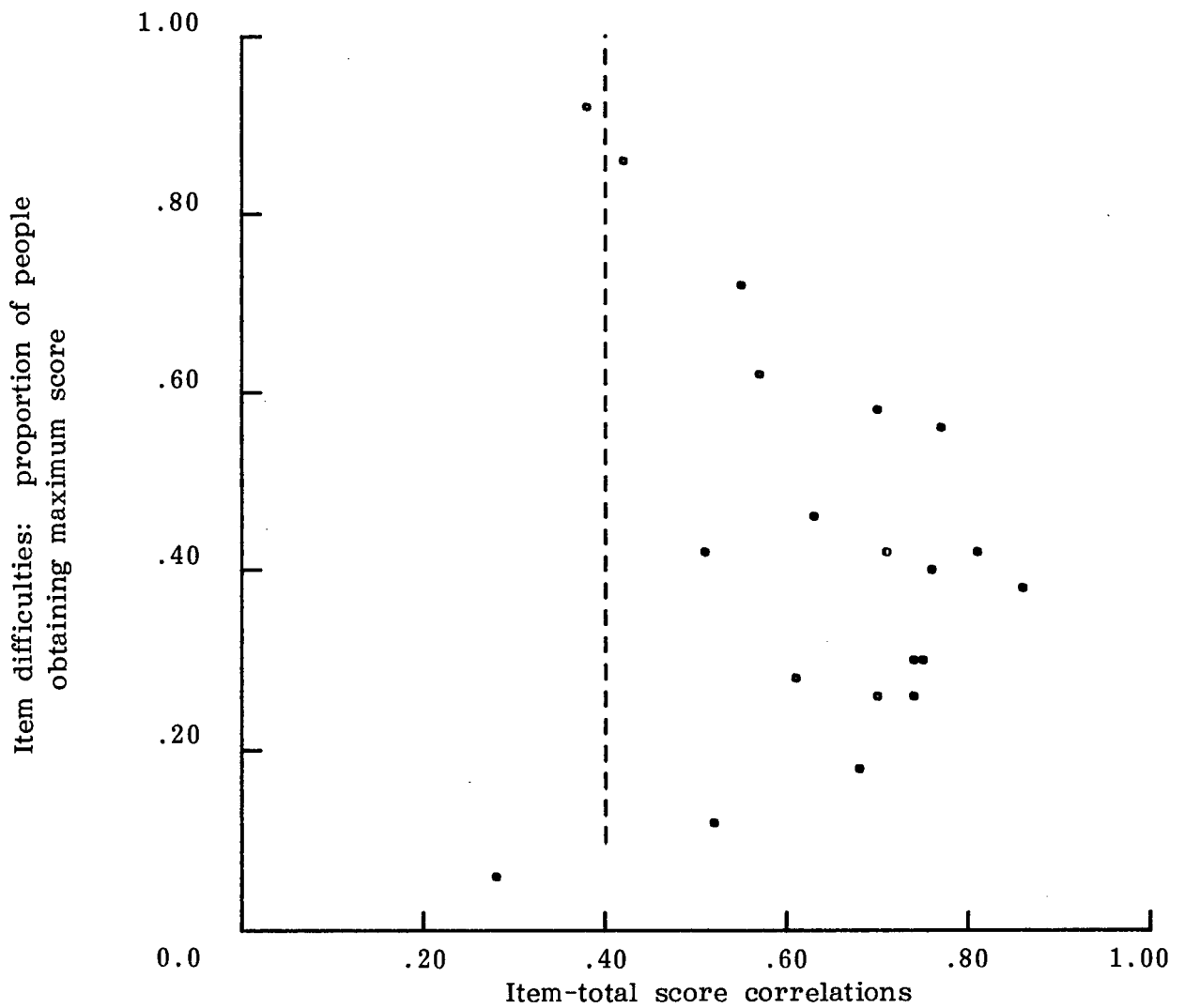


FIGURE 6

Item analysis information for the
Time subtest of the SET

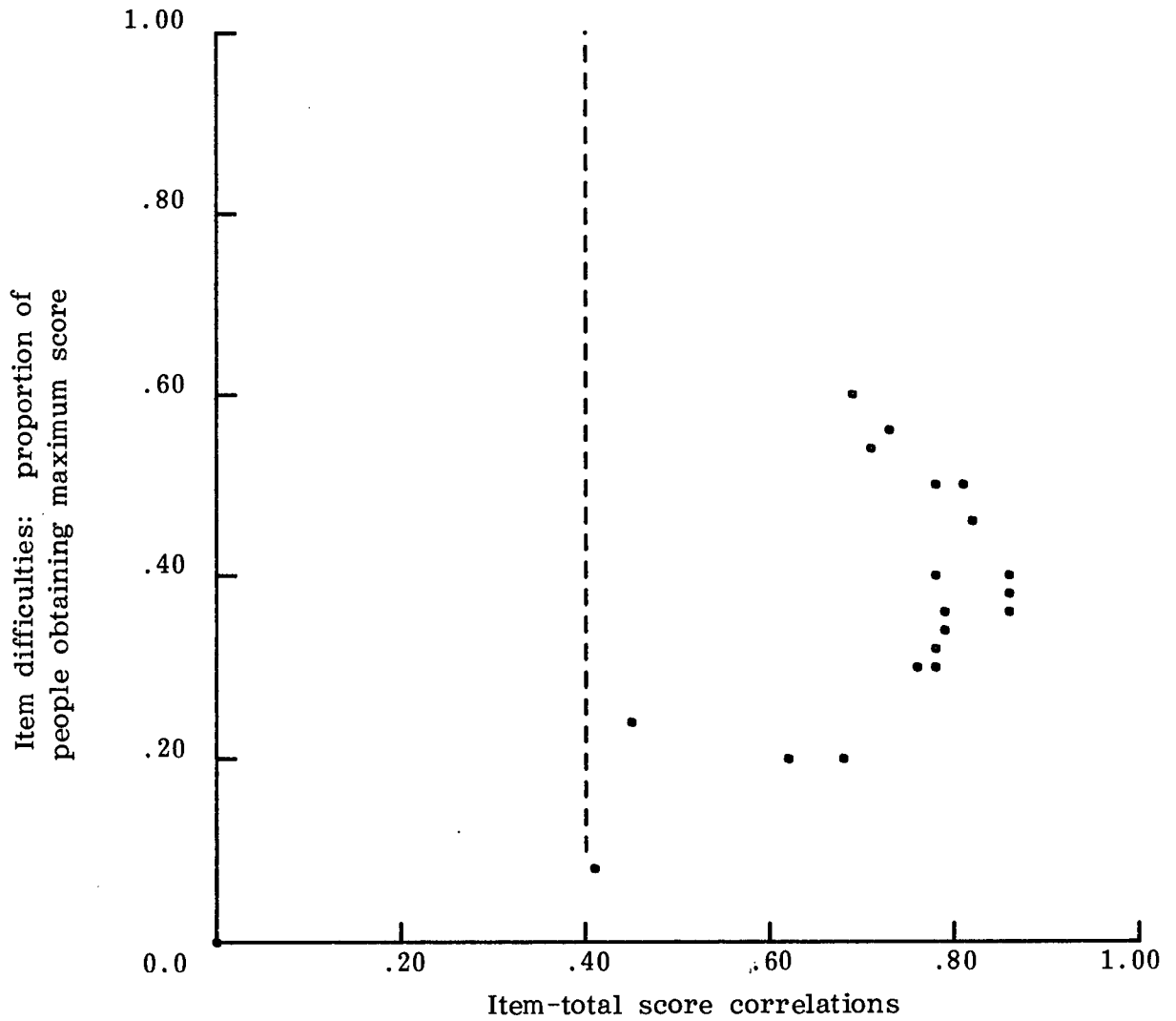


FIGURE 7

Item analysis information for the
Money subtest of the SET

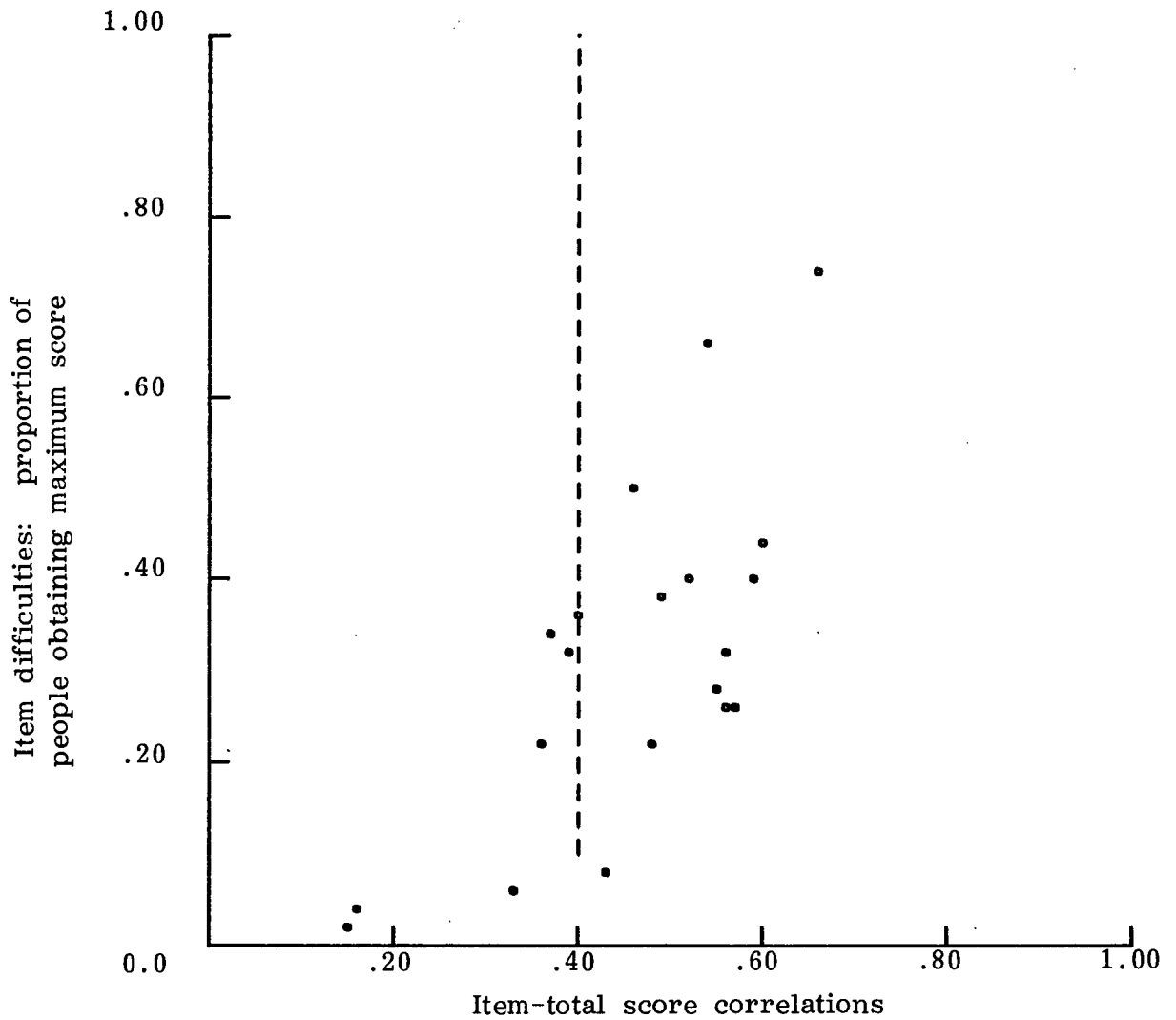


FIGURE 8

Item analysis information for the
Community Awareness subtest of
the SET

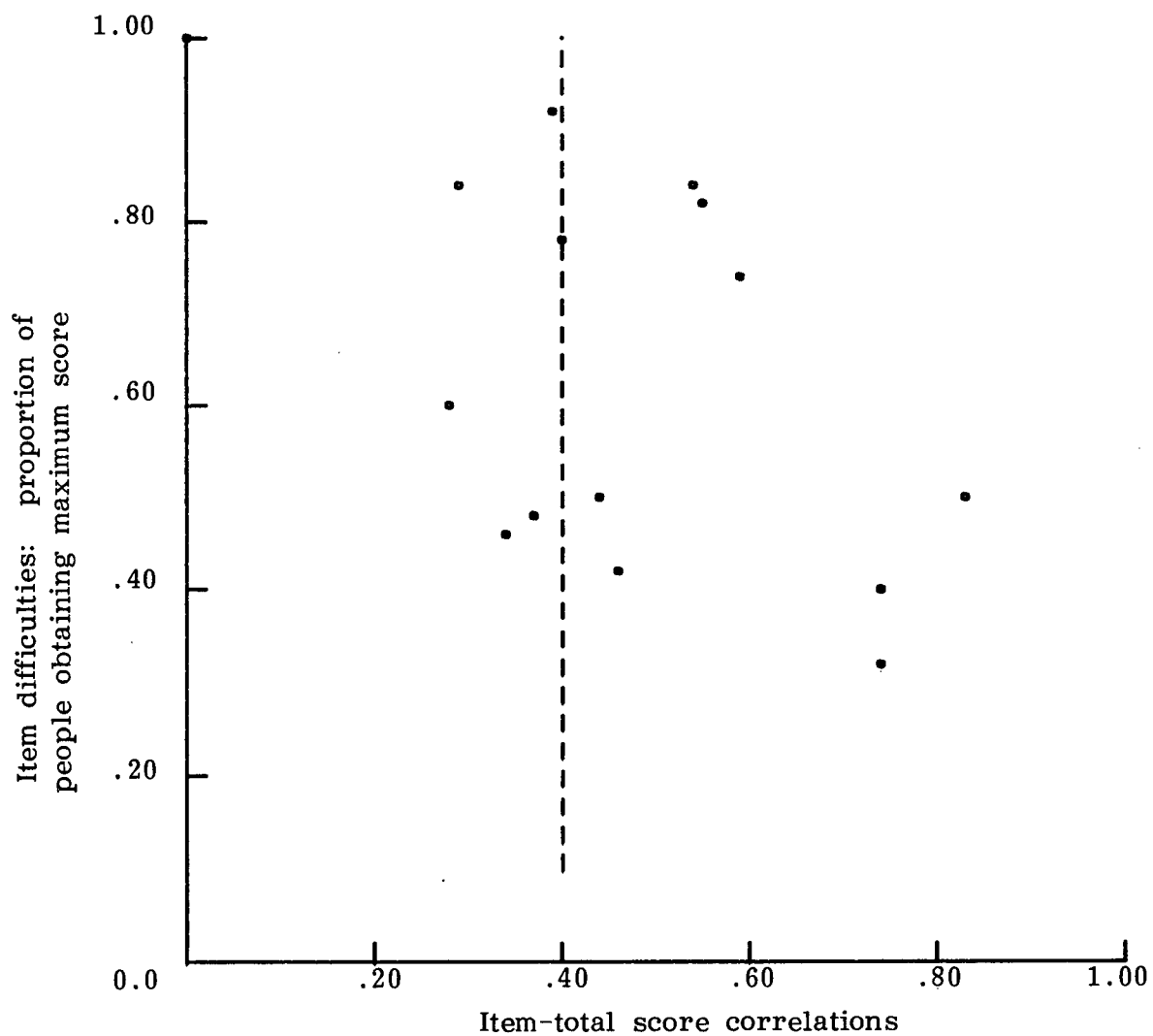
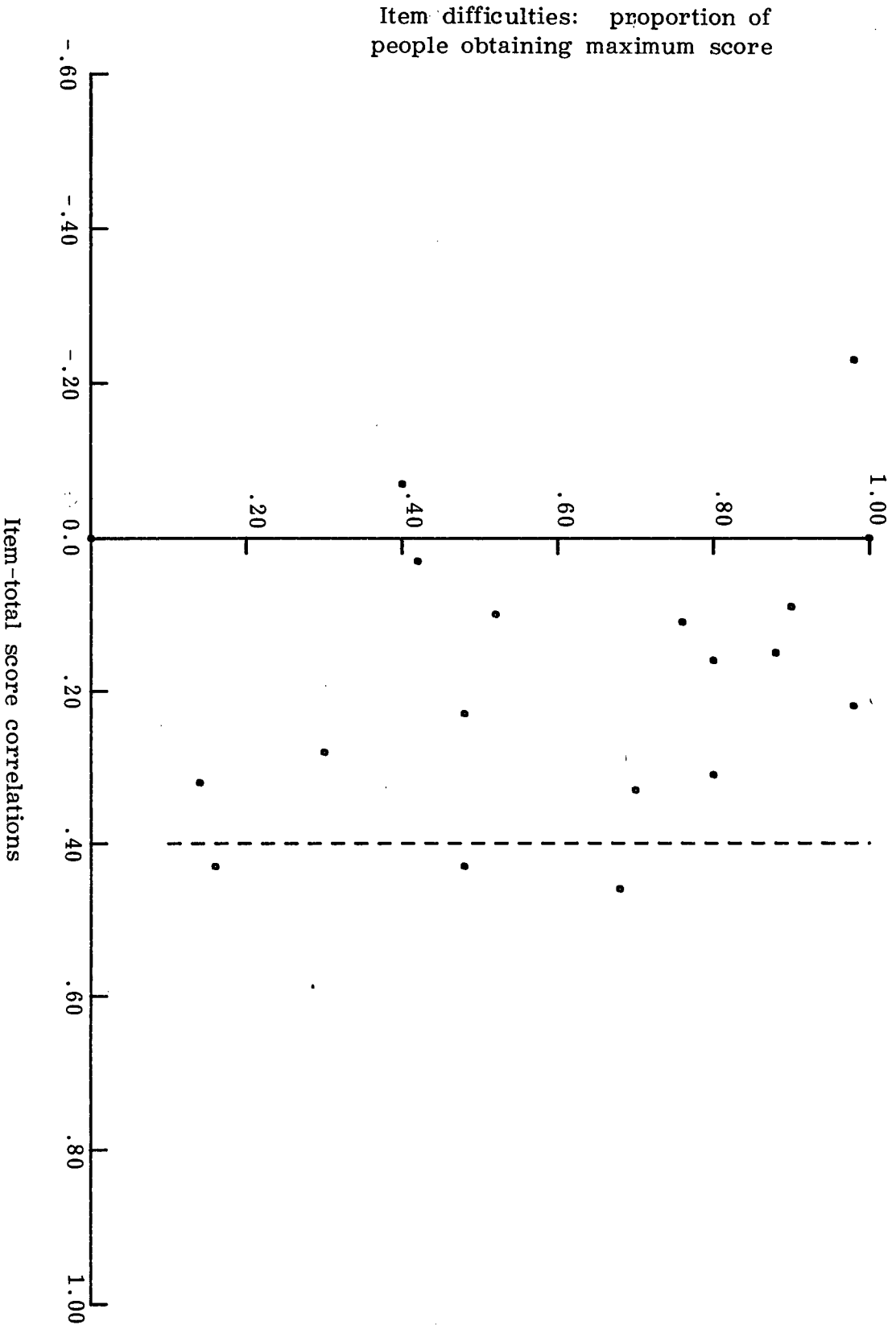


FIGURE 9

Item analysis information for the
Communication subtest of the SET

FIGURE 10

Item analysis information for the Motor Movements subtest of the SET



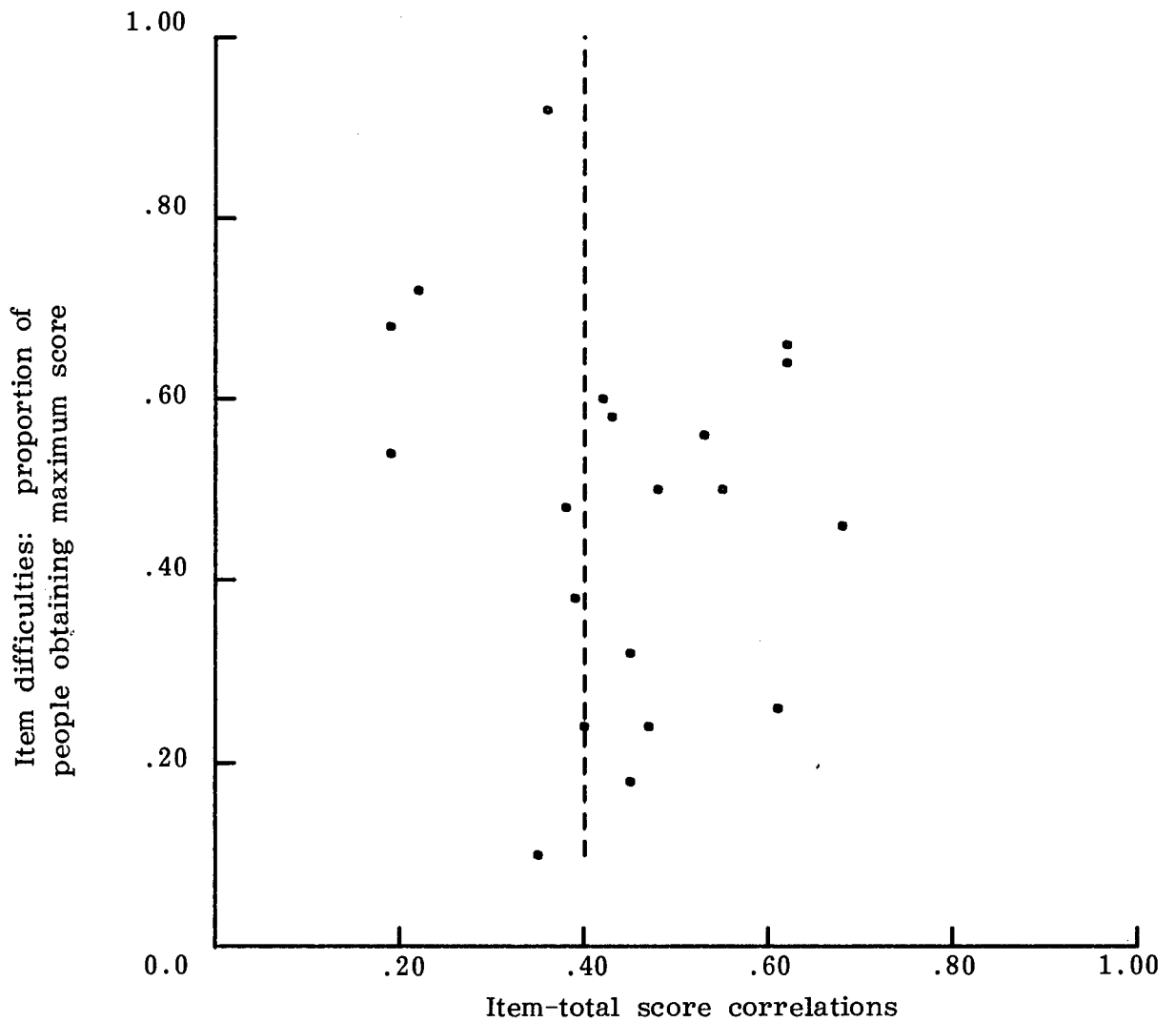


FIGURE 11

Item analysis information for the
Concept Attainment subtest of the
SET

Only one item in the subtest, Money Handling correlates less than .40 with total test scores. No subjects were able to complete a cheque without error.

The results for the subtest, Community Awareness indicate that six items do not discriminate among members of this sample. One is a General Information question, "What place of interest would you take a new friend to see?" Two are Community Cost questions, "How much does it cost to ride a city bus?" and "How much does a loaf of bread cost?" Three are Work Related Information questions, "What is the minimum wage per hour?"; "What is a labour union?"; and, "What is insurance?". These latter three are the most difficult items for this sample ($p = .04, .02, \text{ and } .06$ respectively).

One-half of the items in the Communication subtest correlate less than .40 with total test scores. Of the six items in the Conversation section, the first three are very easy for this sample ($p = 1.00, 1.00, .84$); the two last items of this section, that do not discriminate among members of this sample were of moderate difficulty ($p = .46, .60$). For the Speech section the first two items are very easy for this sample ($p = 1.00, .92$). The last item, which does not discriminate, is of moderate difficulty ($p = .48$). The first item in the Sentence Development section and the first item in the Grammatical Development section are very easy for this sample ($p = 1.00, 1.00$).

Only three items from the subtest Motor Movements, correlate .40 or greater with total test scores. The first measures the second possible

level of performance in the Gross Manual Dexterity task. The next, measures the lowest level of performance of the Fine Dexterity task. And, the last measures the second highest level of performance of the Manipulative Dexterity task. All the items measuring Gross Hand/Eye Coordination, two items measuring balance, and the item reflecting the lowest performance on the Manipulative Dexterity task are very easy for this sample ($p \geq .80$). The items reflecting the highest levels of performance on the Gross Manual Dexterity and Manipulative tasks are very difficult ($p = 0.0$).

For the Concept Attainment subtest, seven items do not appear to be discriminating among members of this sample. These items measure color discrimination, composition of objects, the first two levels of performance on the discrimination task, the first amount concept item, position concepts, and the last conservation item. The first task is very easy ($p = .92$) and the second is difficult ($p = .10$) for this sample.

Stability. Estimates of the generalizability of SET subtest and total test scores, over the universe of occasions, within a three month interval, are presented in Table 37. All coefficients are greater than or equal to .92 except for the subtest Motor Movements where $p = .82$. As can be seen from Table 38 the universe score variance ($\sigma^2(p) + (\sigma^2(p_i))$) is small, relative to those for the first six subtests, whereas the estimates for the persons-by-occasions and the persons-by-occasions-by-items components are of the same magnitude as those for these subtests.

TABLE 37

Generalizability coefficients for subtest and total test
score of the set

Subtest	ρ^2
Reading	.96
Writing	.98
Numbers	.97
Time	.98
Money	.97
Community	.92
Communication	.94
Motor Movements	.82
Concept Attainment	.94
Social Education Test	.99

TABLE 38

Estimates of variance components¹ contributing to σ^2 for the subtest and total test scores of the set

Subtest	σ^2_p	$(1/n_i) \sigma^2_{pi}$	$(1/n_o^1) \sigma^2_{po}$	$(1/n_i^1/n_o^1) \sigma^2_{poi}$
Reading	.10	.002	.002	.001
Writing	.15	.002	.002	.001
Numbers	.09	.004	.002	.002
Time	.10	.003	.001	.001
Money	.12	.002	.002	.001
Community	.04	.003	.002	.002
Communication	.04	.003	.002	.001
Motor Movements	.02	.002	.002	.002
Concept Attainment	.13	.006	.005	.004
Total Test	.15	.002	.001	.001

Results of Analysis - VCL

Test Analysis. Tables 39 and 40 show the mean, range, standard deviation, standard error of measurement, and the Hoyt estimates of internal consistency for each subtest of the VCL, for each analysis sample. Table 41 summarizes these results for the total test.

Within the domain, Basic Work Habits, a floor effect was observed for the subtest, Making Decisions, for both analysis samples, and for the subtests, Independence and Taking Direction for the first analysis sample. Within the domain Work Skills, a floor effect was observed for all three subtests for both samples. In addition, a floor effect was obtained for this domain for the first analysis sample.

A ceiling effect was obtained for the subtests, Independence and Use and Care of Equipment, within Basic Work Habits, for both samples. This effect was also observed for the subtest, Taking Direction, for the first analysis sample. For the domain Work Skills, a ceiling effect was obtained for the subtest, Following Instructions for both samples. A ceiling effect was observed for the domain Acceptance Skills, and for all four subtests, for both samples. These results indicate that more difficult items are needed in these subtests in order to adequately discriminate among high scorers.

The estimates of internal consistency for Basic Work Habits, Work Skills, and Acceptance Skills are .91, .91, and .92, respectively, for analysis sample two. The subtest coefficients range from .65 for Making Decisions to .90 for Task Analysis for the first sample; and, .63 for Taking Direction

TABLE 39

Test analysis information for the subtests of the VCL 1

Subtest	Mean	Range	St. Dev.	Hoyt	SEM
Basic Work Habits	19.92	5 - 35	8.24	.91	2.42
Independence	5.43	0 - 10	2.76	.83	1.01
Making Decisions	4.29	0 - 9	2.29	.65	1.21
Use, Care of Equipment	5.76	1 - 10	2.46	.80	.98
Taking Direction	4.50	0 - 10	2.11	.67	1.08
Work Skills	19.98	0 - 72	14.94	.91	4.52
Speed	2.83	0 - 9	2.30	.79	.95
Following Instructions	5.69	0 - 10	2.55	.83	.93
Task Analysis	11.45	0 - 61	12.86	.90	3.94
Acceptance Skills	22.40	0 - 40	8.61	.92	2.31
Appearance	6.29	0 - 10	2.56	.85	.90
Punctuality	5.45	0 - 10	2.48	.79	1.02
Talking	5.19	0 - 10	2.78	.87	.90
Relationships	5.48	0 - 10	2.53	.81	1.00

TABLE 40

Test analysis information for the subtests of the VCL 2

Subtest	Mean	Range	St. Dev.	Hoyt	SEM
Basic Work Habits	18.62	7 - 34	7.60	.89	2.51
Independence	5.12	1 - 10	2.68	.82	1.02
Making Decisions	3.57	0 - 8	2.11	.64	1.13
Use, Care of Equipment	5.50	1 - 10	2.86	.85	.99
Taking Direction	4.43	1 - 9	1.98	.63	1.07
Work Skills	18.24	4 - 51	12.69	.87	4.46
Speed	2.45	0 - 6	1.60	.49	1.02
Following Instructions	5.60	0 - 10	2.57	.81	.99
Task Analysis	10.19	0 - 39	11.40	.88	3.86
Acceptance Skills	21.69	0 - 40	8.28	.91	2.40
Appearance	5.95	0 - 10	2.40	.81	.95
Punctuality	4.88	0 - 10	2.51	.80	1.01
Talking	5.24	0 - 10	2.95	.87	.95
Relationships	5.62	0 - 10	2.54	.82	.98

TABLE 41

Test analysis summary for the Vocational Checklist

Sample	Mean	Range	St. Dev.	α_{strat}	SEM
VCL 1	62.36	20 - 119	26.81	.96	5.99
VCL 2	58.55	23 - 109	24.01	.95	5.95

to .88 for Task Analysis for the second sample. The internal consistency of the total test is .96 and .95 for samples one and two respectively.

Correlations among subtest and total test scores for both samples are presented in Tables 42 and 43. Basic Work Habits, Work Skills, and Acceptance Skills correlate .85, .76, and .78 respectively, with total test scores for sample one, and, .80, .69, and .80, respectively, for sample two. For both samples the subtest, Task Analysis, correlates lowest with total test scores.

Item Analysis. The results of the item analysis for the subtests of the VCL are presented in Figures 12 to 17. The subtest, Task Analysis, was not included because subjects performed different tasks and different numbers of tasks within each program as well as across programs. For the first analysis sample forty-one of the fifty items correlate .40 or greater with total test scores. For the second sample thirty-six items reach this criterion.

For the domain, Basic Work Habits, all items of the subtests, Independence and Use and Care of Equipment, correlate .40 or greater with total test scores for both analysis samples. The item, "Knows when he has a problem and seeks advice", from the subtest, Making Decisions, does not discriminate among members of either sample. In addition, the item "Can tell the quality of his own work" correlated less than .40 with total scores for the second sample. This item is difficult ($p = .07$). For the subtest Taking Directions, the items, "Treats supervisor as a boss...", and "Seeks advice appropriately" do not discriminate among members of the

TABLE 42

Correlations among subtest and total test scores for the VCL 1¹

Subtest	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Basic Work Habits	1.00														
Independence	.93	1.00													
Making Decisions	.84	.69	1.00												
Use, Care of Equipment	.86	.69	.72	1.00											
Taking Direction	.78	.74	.45	.48	1.00										
Work Skills	.73	.68	.70	.62	.47	1.00									
Speed	.69	.61	.71	.63	.39	.60	1.00								
Following Instructions	.79	.72	.80	.63	.54	.74	.81	1.00							
Task Analysis	.37	.36	.38	.28	.21	.65	.19	.26	1.00						
Acceptance Skills	.68	.60	.67	.54	.50	.53	.63	.75	.001	1.00					
Appearance	.59	.58	.57	.49	.36	.53	.65	.65	.06	.83	1.00				
Punctuality	.63	.58	.59	.52	.43	.59	.51	.69	.23	.83	.76	1.00			
Talking	.50	.38	.55	.36	.41	.33	.41	.58	-.09	.80	.45	.45	1.00		
Relations	.54	.45	.52	.43	.47	.33	.53	.59	-.19	.87	.59	.57	.73	1.00	
Total Test	.85	.76	.80	.78	.55	.76	.74	.77	.33	.78	.71	.66	.57	.64	1.00

¹N=42; $r_{\text{critical}} = .307$ ($p \leq .05$)

TABLE 43

Correlations among subtest and total test scores of the VCL 2¹

Subtest	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Basic Work Habits	1.00														
Independence	.87	1.00													
Making Decisions	.73	.45	1.00												
Use, Care of Equipment	.86	.70	.54	1.00											
Taking Direction	.65	.49	.33	.34	1.00										
Work Skills	.70	.73	.43	.61	.35	1.00									
Speed	.46	.47	.29	.32	.35	.59	1.00								
Following Instructions	.74	.74	.47	.67	.37	.79	.58	1.00							
Task Analysis	.24	.28	.14	.23	.06	.62	-.10	.10	1.00						
Acceptance Skills	.55	.44	.44	.45	.40	.43	.32	.52	.08	1.00					
Appearance	.46	.43	.40	.45	.12	.37	.38	.41	.05	.73	1.00				
Punctuality	.52	.46	.32	.46	.39	.43	.24	.45	.18	.80	.62	1.00			
Talking	.34	.17	.43	.20	.35	.18	.07	.25	.06	.79	.34	.47	1.00		
Relations	.46	.41	.31	.32	.43	.43	.22	.62	-.03	.83	.45	.53	.66	1.00	
Total Test	.80	.73	.57	.71	.46	.69	.46	.69	.27	.80	.75	.64	.51	.69	1.00

¹N=42; $r_{\text{critical}} = .307$ ($p \leq .05$)

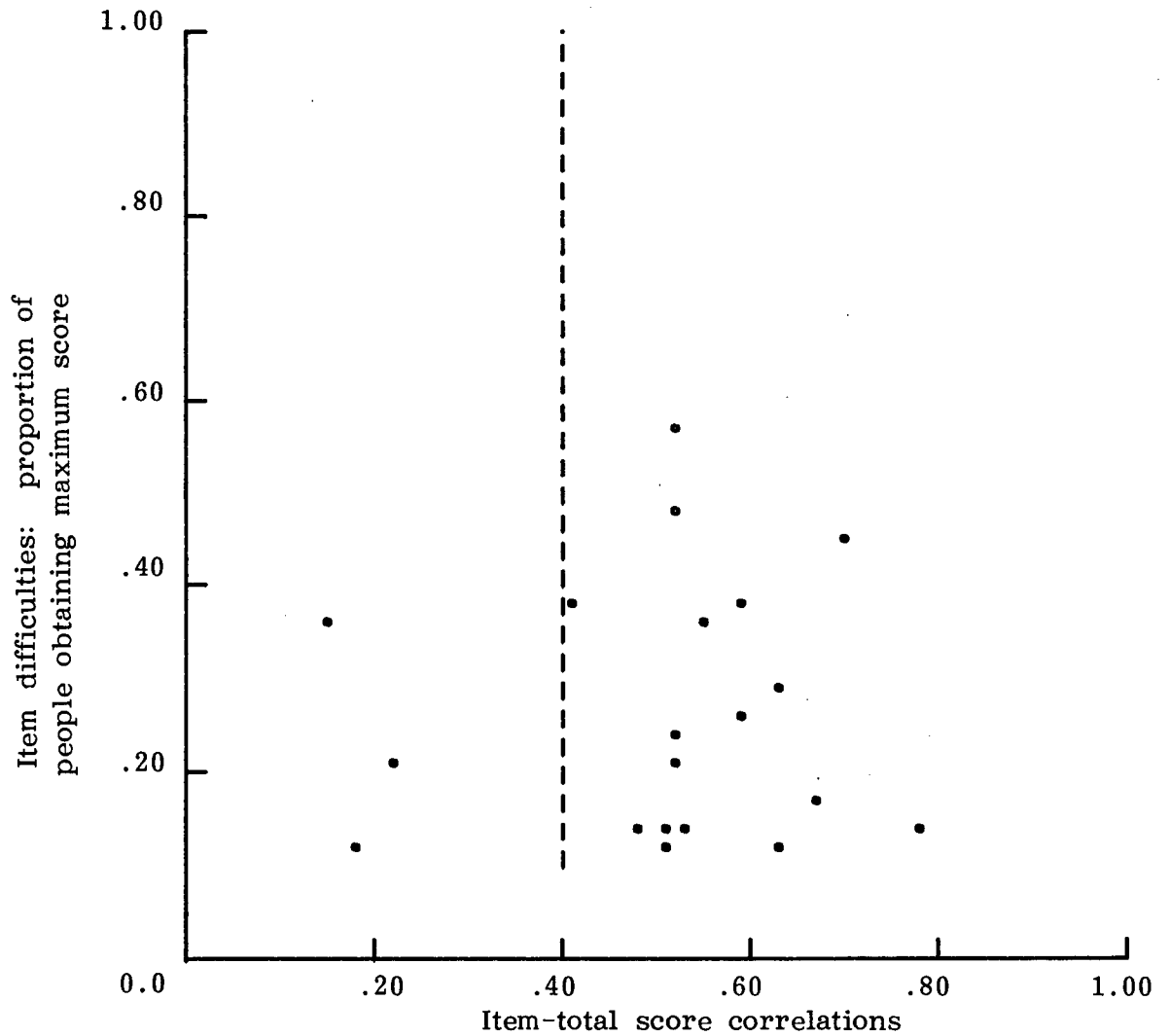


FIGURE 12

Item analysis information for the
subtest Basic Work Habits of the
VCL (1)

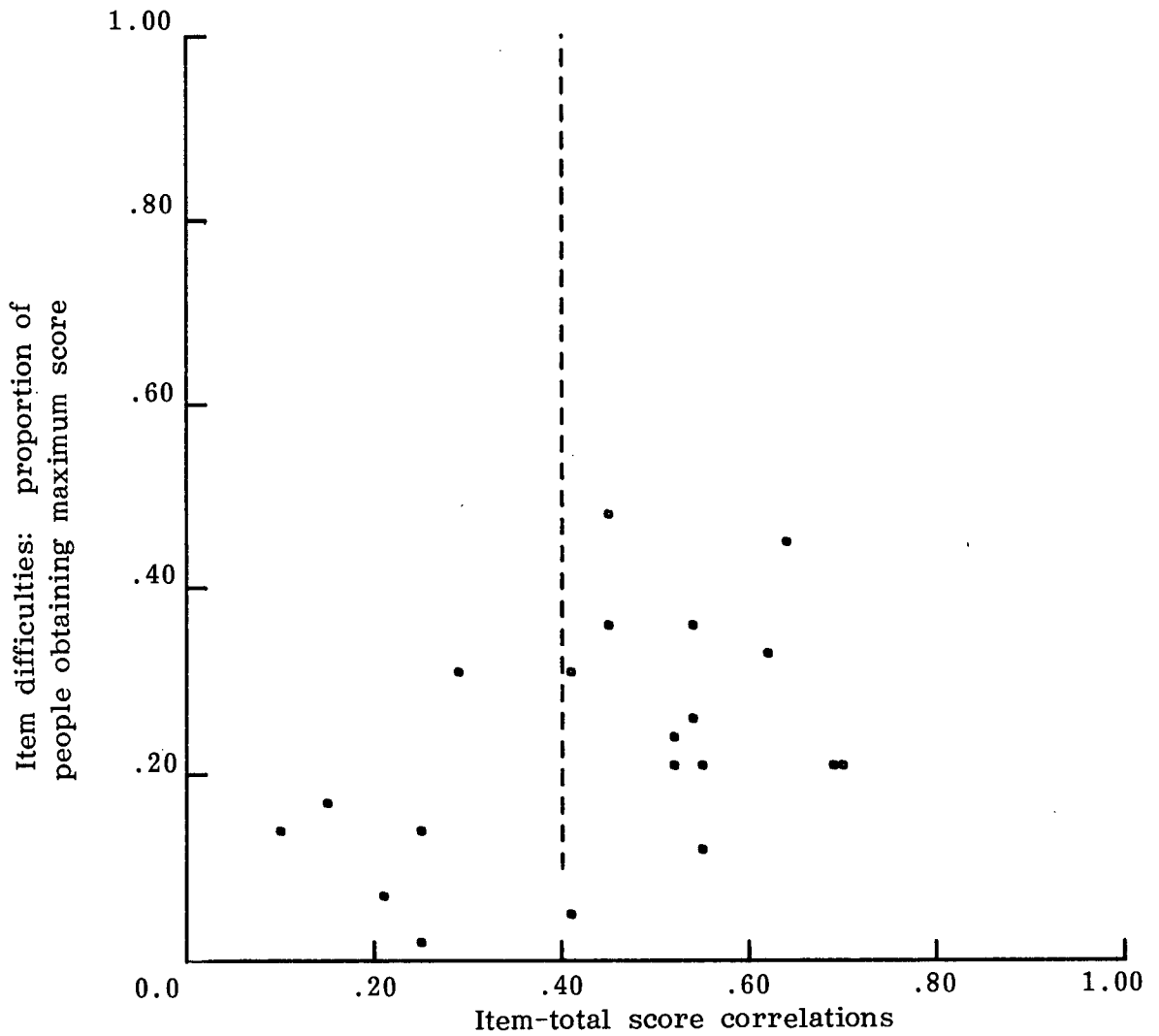


FIGURE 13

Item analysis information for the
subtest Basic Work Habits of the
VCL (2)

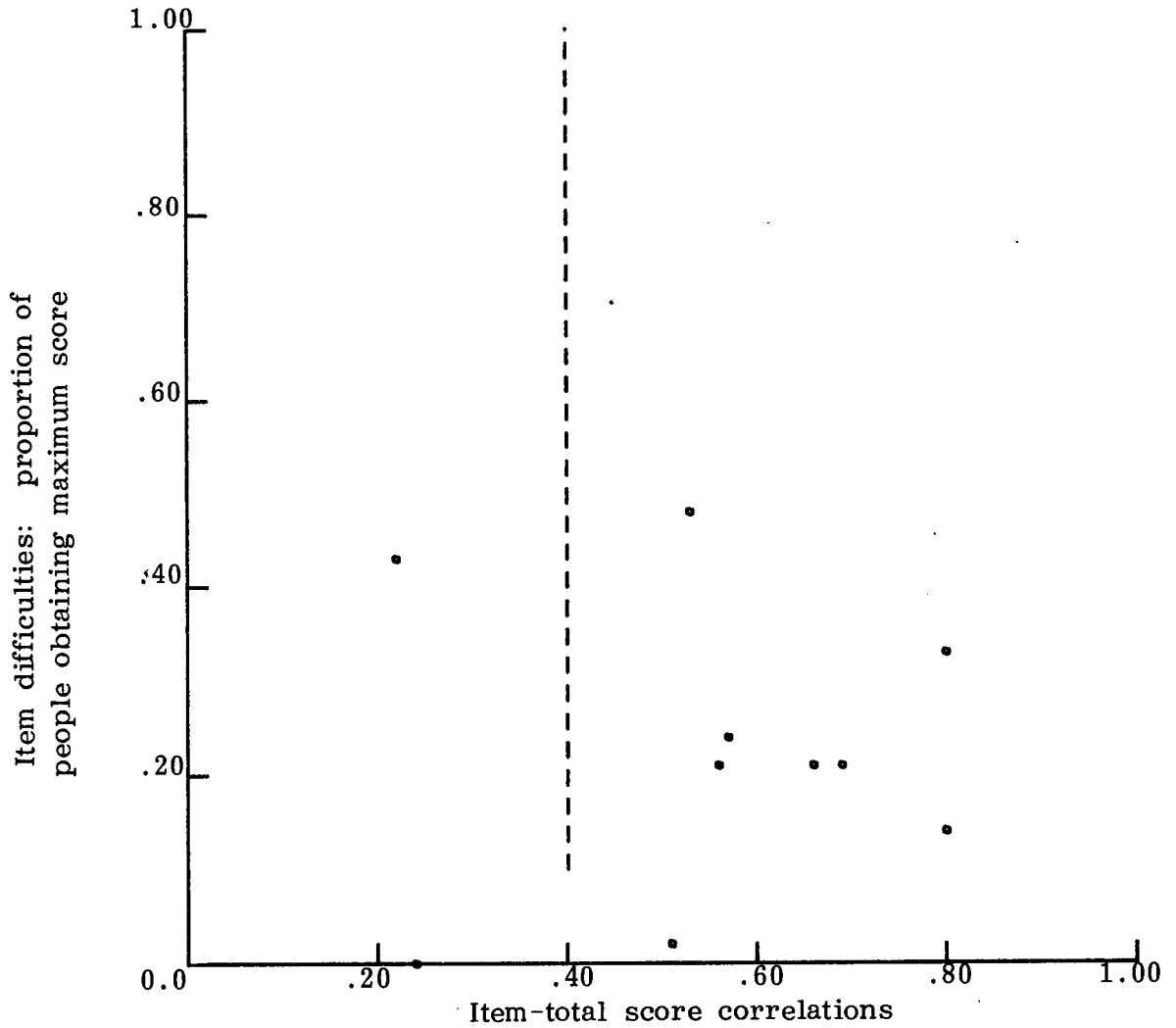


FIGURE 14

Item analysis information for the subtest
Work Skills of the VCL (1)

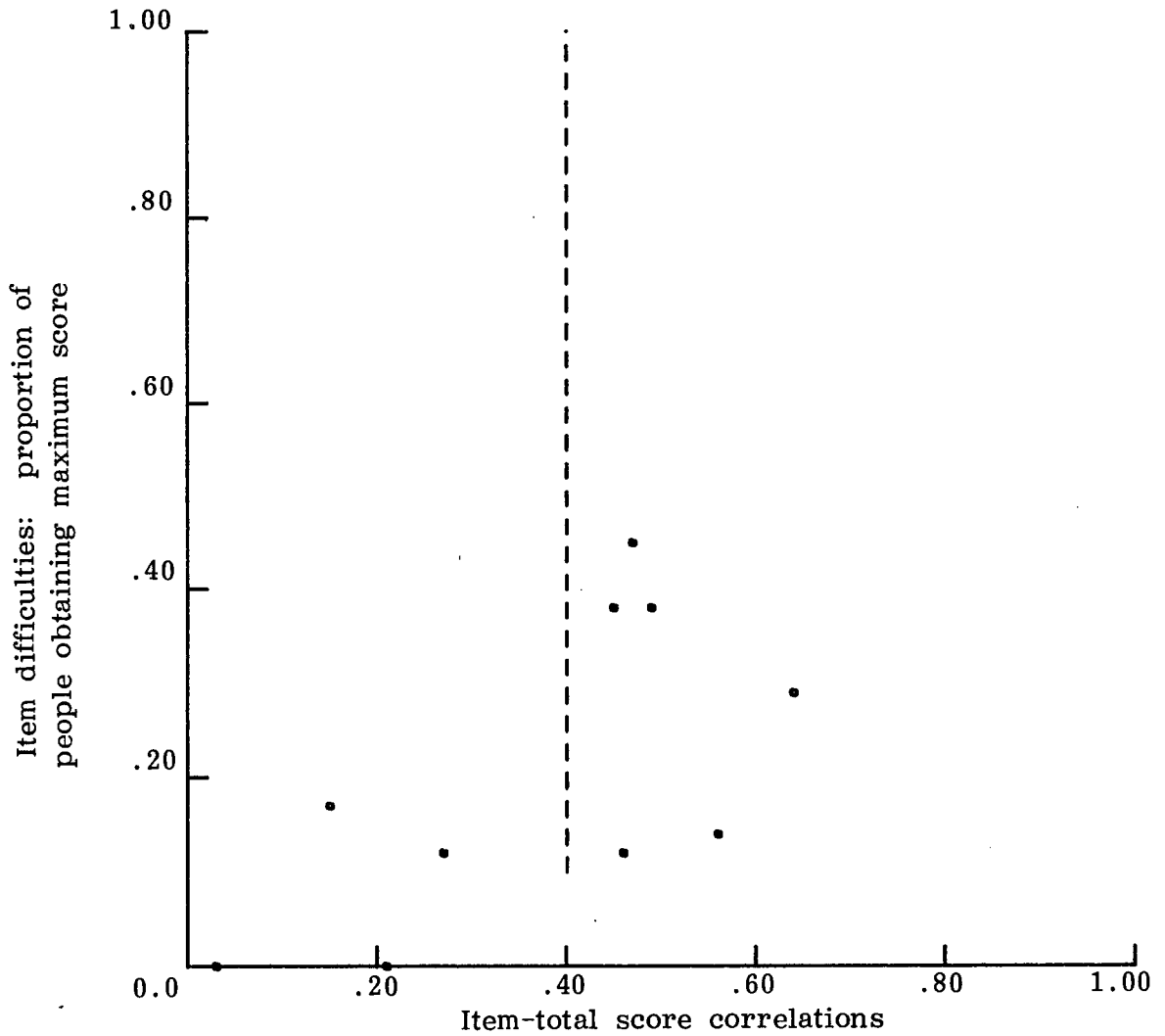


FIGURE 15

Item analysis information for the
subtest Work Skills of the VCL (2)

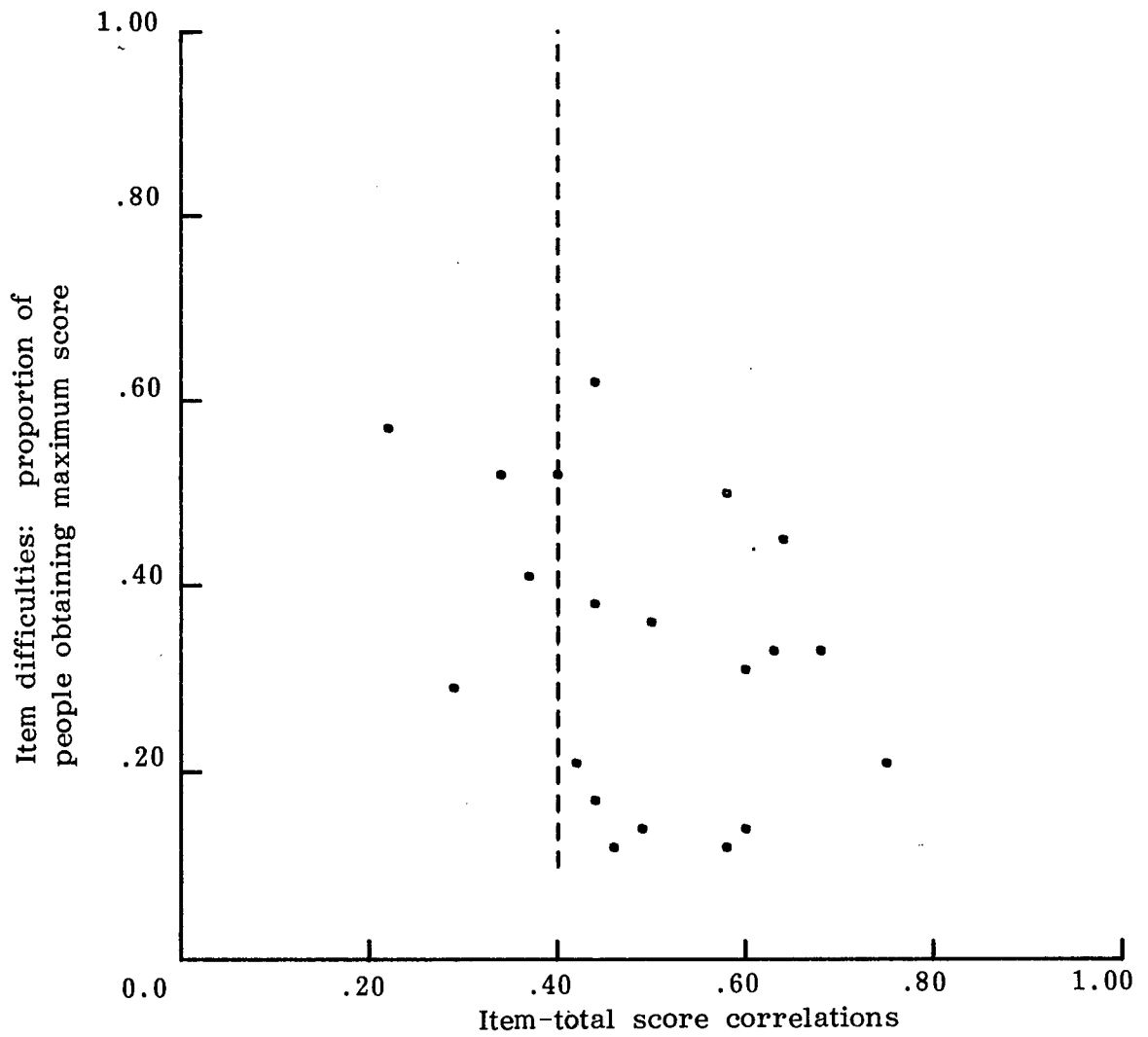


FIGURE 16

Item analysis information for the subtest
Acceptance Skills of the VCL (1)

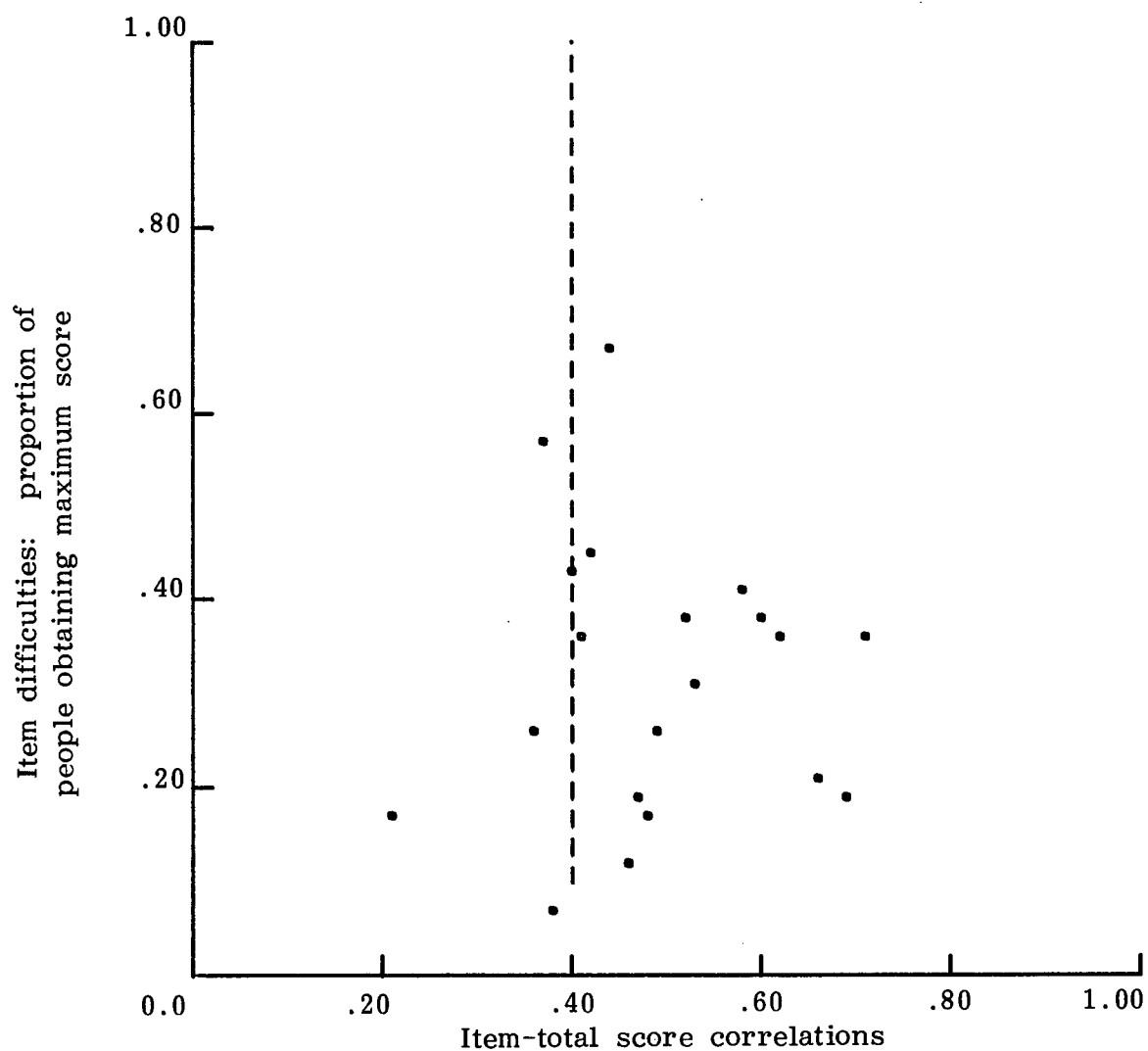


FIGURE 17

Item analysis information for the subtest
Acceptance Skills of the VCL (2)

first sample. For the second sample, only one item of this subtest correlates .40 or greater, "Treats supervisor as a boss...". The remaining four items are more difficult for this sample.

For the domain Work Skills, a number of items correlate less than .40 with total test scores. The item, "Works above 80% industrial standard" from the Speed subtest, does not discriminate among members of either sample. It is very difficult ($p = 0.0$ for both samples). Only one item, "Works at 40-60% industrial standard", discriminates among members of the first sample. For the subtest, Following Instructions, the item, "Learns when he is shown the task in small simple units" correlates less than .40 with total test scores for analysis sample one. The remaining items in this subtest appear to discriminate among members of both samples.

For the domain, Acceptance Skills, all items of the subtest, Appearance, discriminate among members of both samples. The item, "On time after coffee/lunch", from the subtest, Punctuality, correlates less than .40 with total test scores for both samples. In addition, the items "On time to start work" and "Looks after his own appointments" do not discriminate among members of the first and second samples, respectively. The latter item is difficult for the second sample ($p = .07$). The remaining items in this subtest correlate .40 or greater with total test scores. For the subtest, Self-Expression Talking, the item, "Has short talks about things he knows about" does not discriminate among members of the first sample. The items measuring basic social conversation and "Whole conversations about a number of things" correlate less than .40 with total test scores for the

first sample. The former is the easiest item for this sample ($p = .57$). The items measuring other conversation skills are useful in discriminating among members of both samples. For the last subtest, Relations with Co-Workers, the item "Responds to friendly advances...", does not discriminate among members of the first sample. This is the easiest item for this sample ($p = .57$). The remaining items correlate .40 or greater with total test scores.

Generalizability of Ratings. Estimates of the generalizability of ratings over the universe of raters for the VCL are presented in Table 44. Table 45 provides estimates for each component that was used in calculating these coefficients.

In examining the results for each subtest by rater pairs, it can be seen that eleven coefficients are equal to zero: the universe score variance equals zero. Six of the eleven subtests (excluding Basic Work Habits, Work Skills, and Acceptance Skills) have at least one coefficient equal to zero. The subtest, Speed, has the largest number of coefficients (three) that are zero. The first and fourth rater pairs have no ratings where generalizability is zero. Rater pair 3 has four coefficients that are zero. Thus rater bias appears to affect about one half of the subtests. All rater pairs reveal this bias.

Of the remaining seventy-three coefficients, nine are greater than or equal to .80. For the domain Basic Work Habits, one coefficient for total domain scores is equal to .80. No subtest coefficients reach this criterion. For the domain, Work Skills, two coefficients are greater than .80 for the

TABLE 44

Generalizability coefficients by rater pair and across raters for
the subtest and total test scores of the VCL

Subtest	Rater Pair						AR
	1	2	3	4	5	6	
Basic Work Habits	.80	.02	.37	.67	.06	.05	.18
Independence	.74	0.0	.53	.48	.19	.44	.21
Making Decisions	.55	0.0	.27	.10	.48	.12	.16
Use, Care of Equipment	.69	.15	0.0	.09	0.0	.39	.25
Taking Direction	.38	.70	.37	.65	.30	.01	.38
Work Skills	.88	.64	.39	.96	.62	.75	.83
Speed	.45	0.0	0.0	.56	.02	0.0	0.0
Following Instructions	.66	0.0	.35	.60	.80	0.0	.27
Task Analysis	.70	.87	0.0	.96	0.0	.78	.85
Acceptance Skills	.67	.58	.32	.63	.57	.61	.77
Appearance	.64	.57	.39	.84	.34	.80	.75
Punctuality	.68	.17	.70	.60	.11	.48	.61
Talking	.55	.14	.69	.36	.89	.53	.54
Relations	.21	.75	.07	.54	.39	.08	.64
Total Test	.93	.07	.29	.89	.37	.56	.74

TABLE 45

Variance components contributing to universe score variance ($\sigma^2_{\mu p}$) and error variance (σ^2_{δ}) by rater pair and across raters for subtest and total test scores of the VCL

Subtest		Rater Pair							
		1	2	3	4	5	6	Across Raters	
		A B	A B	A B	A B	A B	A B	A B	A B
Basic Work Habits	$\sigma^2_{\mu p}$.05+.002	0 +.001	.01+.01	.09 +.03	.05+.001	0 +.003	.01+.002	
	σ^2_{δ}	.01+.004	.06+.01	.02+.002	.003+.01	.08+.01	.04+.01	.05+.01	
Independence	$\sigma^2_{\mu p}$.07+.01	0 + 0	.07+.01	.11 + 0	.02+.01	0 +.03	.02+.01	
	σ^2_{δ}	.01+.02	.10+.02	.06+.01	.10 +.02	.15+.01	.02+.01	.09+.01	
Making Decisions	$\sigma^2_{\mu p}$.05+ 0	0 + 0	.01+ 0	0 +.01	.07+ 0	0 +.02	.03+.02	
	σ^2_{δ}	.02+.02	.17+.03	.02+.01	.10 +.03	.05+.02	.14+.04	.08+.02	
Use, Care of Equipment	$\sigma^2_{\mu p}$.09+.01	0 +.03	0 + 0	.01 +.01	0 + 0	.02+.02	0 +.01	
	σ^2_{δ}	.03+.02	.12+.02	.07+.01	.13 +.02	.18+.03	.05+.02	.10+.02	
Taking Directions	$\sigma^2_{\mu p}$.01+.004	.03+.09	.01+.003	.06 +.01	.04+.01	.002+0	.04+.01	
	σ^2_{δ}	.01+.01	.02+.005	.01+.01	.02 +.02	.11+.02	.11 +.07	.05+.02	
Work Skills	$\sigma^2_{\mu p}$.05+.002	.02+.005	.001+0	.11 +.01	.01+.001	.04+.001	.04+.01	
	σ^2_{δ}	.003+.003	.01+.002	.001+.001	.004+.002	.005+.001	.01+.01	.01+.003	
Speed	$\sigma^2_{\mu p}$.03+ 0	0 + 0	0 + 0	.06 +.004	0 +.001	0 + 0	0 + 0	
	σ^2_{δ}	.01+.01	.04+.01	.01+.01	.04 +.02	.08 +.02	.08+.01	.05+.02	
Following Instructions	$\sigma^2_{\mu p}$.05+ 0	0 + 0	.01+.003	.05 +.01	.02+.01	0 + 0	.02+.001	
	σ^2_{δ}	.01+.01	.12+.01	.01+.01	.03 +.01	.04+.02	.06+.01	.05+.01	
Task Analysis	$\sigma^2_{\mu p}$.04+.01	.04+.01	0 + 0	.18 +.03	0 + 0	.11+.02	.06+.01	
	σ^2_{δ}	.02+.01	.01+.003	0 +.001	.01 +.003	0 + 0	.02+.02	.01+.01	
Acceptance Skills	$\sigma^2_{\mu p}$.02+.002	.12+.002	.005+.002	.22 +.002	.02+.003	.03+.08	.07+.002	
	σ^2_{δ}	.01+.003	.01+.05	.01+.003	.13 +.004	.01+.005	.01+.005	.02+.004	
Appearance	$\sigma^2_{\mu p}$.04+.01	.07+ 0	.02+ 0	.38 + 0	.004+.02	.04+.01	.12+.01	
	σ^2_{δ}	.02+.01	.03+.02	.03+.01	.06 +.01	.02+.02	.05+.01	.03+.01	
Punctuality	$\sigma^2_{\mu p}$.03+.01	0 +.02	.02+.01	.19 +.01	0 +.01	.02+ 0	.06+.01	
	σ^2_{δ}	.02+.01	.06+.01	.01+.01	.13 +.01	.06+.02	.01+.02	.03+.01	
Talking	$\sigma^2_{\mu p}$.04+ 0	0 +.01	.09+ 0	.11 +.01	.15+.01	.19+.01	.09+.002	
	σ^2_{δ}	.02+.01	.06+.004	.03+.01	.17 +.03	.01+.01	.16+.02	.06+.02	
Relations	$\sigma^2_{\mu p}$.002+.01	.15+.001	0 +.003	.13 +.01	.05+ 0	0 +.001	.09+.003	
	σ^2_{δ}	.03+.01	.03+.02	.02+.01	.10 +.02	.06+.02	.04+.02	.03+.02	
Total Test	$\sigma^2_{\mu p}$.04+.001	0 +.001	.002+.0002	.12 +.002	.12+.007	.01+.002	.03+.001	
	σ^2_{δ}	.001+.001	.02+.001	.001+.001	.01 +.002	.02+.007	.08+.003	.01+.003	

domain scores; and three subtest coefficients (two for Task Analysis and one for Following Instructions) are greater than or equal to .80. For the last domain, Acceptance Skills, no coefficients for domain scores are greater than or equal to .80; two coefficients for the subtest Appearance, and one coefficient for Self-Expression, are greater than or equal to .80. Only rater pair 3 has no coefficients greater than or equal to .80. Rater pair 4 has four, the largest number for all raters.

In examining the sixty-four coefficients that are greater than zero but less than .80, the persons-by-judges variance component contributes most to the observed score variance of twenty-eight coefficients. All domains and subtests, except Task Analysis, have at least one coefficient where the pj component is the largest or equal to the largest variance component contributing to the ratio. For the subtests, Independence, Taking Directions, and Relations with Co-Workers, four of the six coefficients have the pj component contributing the most to the observed score variance. In addition, all rater pairs have at least two coefficients where the pj component is the largest. For rater pair 3, the pj component in eight of the fourteen coefficients contributes most to observed score variance.

The generalizability coefficients across all raters for Basic Work Habits, Work Skills, and Acceptance Skills are .18, .83, and .77, respectively. The subtest coefficients range from 0.0 for the subtest, Speed, to .85 for Task Analysis. The generalizability of total scores is .74. The pj component contributes most to the observed score variance for the domain, Basic Work Habits, and its four subtests. Within the Work Skills domain, the subtests,

Speed and Following Instructions also reveal this bias. No subtests within Acceptance Skills, nor the total test coefficient reveal this bias.

Results of Analysis - Difficulty Section of the VCL

Test analysis. Tables 46 and 47 show the mean, range, standard deviation, standard error of measurement, and the Hoyt estimate of internal consistency for the subtests and total tests, for both analysis samples.

For both analysis samples, a floor effect was obtained for all subtests except Basic Work Habits, indicating that not all subjects displayed these difficulties. In addition, a ceiling effect was observed for the subtests, Independence, Use and Care of Equipment, Following Instructions, and Punctuality for analysis sample one. This effect was observed for the subtests Independence and Following Instructions, for the second sample.

The Hoyt coefficients for Basic Work Habits, Work Skills, and Acceptance Skills are .62, .72, and .35, respectively, for the first sample; and .77, .83, and .60, respectively for the second sample. The internal consistency for the total test is .90 for sample one and .95 for sample two.

Correlations among subtest and total test scores are presented in Tables 48 and 49. For the first analysis sample, difficulty scores for Basic Work Habits correlate highest with total difficulty scores. For the second sample, difficulties in Basic Work Habits and Work Skills correlate highest with total test scores.

Item Analysis. The results of the item analyses for the Difficulty subtests are presented in Figures 18 and 19. For the first sample fourteen of the thirty-one items correlate .40 or greater with total test scores. For

TABLE 46

Test analysis information for the Difficulty subtests of the VCL 1

Subtest	Mean	Range	St. Dev.	Hoyt	SEM
Basic Work Habits	6.29	0 - 17	3.85	.77	1.76
Independence	1.83	0 - 6	1.56	.77	.62
Making Decisions	1.83	0 - 5	1.08	.12	.83
Use, Care of Equipment	1.10	0 - 6	1.34	.61	.68
Taking Direction	1.52	0 - 4	1.13	.33	.75
Work Skills	2.98	0 - 10	2.85	.83	1.07
Speed	1.55	0 - 6	1.63	.75	.71
Following Instructions	1.43	0 - 6	1.58	.81	.56
Acceptance Skills	4.31	0 - 11	3.01	.60	1.81
Appearance	1.38	0 - 5	1.36	.42	.84
Punctuality	.90	0 - 6	1.19	.60	.62
Talking	1.10	0 - 3	.98	.07	.78
Relations	.93	0 - 3	1.00	.09	.78
Total Test	27.14	4 - 68	17.58	.95	4.04

TABLE 47

Test analysis information for the Difficulty subtests of the VCL 2

Subtest	Mean	Range	St. Dev.	Hoyt	SEM
Basic Work Habits	5.98	1 - 16	3.06	.62	1.81
Independence	1.50	0 - 6	1.11	.46	.67
Making Decisions	2.12	0 - 5	1.23	.44	.75
Use, Care of Equipment	.98	0 - 5	1.32	.59	.69
Taking Direction	1.38	0 - 5	1.17	.32	.78
Work Skills	2.07	0 - 10	2.05	.72	1.00
Speed	.88	0 - 5	1.21	.66	.62
Following Instructions	1.19	0 - 6	1.19	.60	.62
Acceptance Skills	4.26	0 - 10	2.47	.35	1.91
Appearance	1.19	0 - 5	1.35	.42	.83
Punctuality	.90	0 - 4	1.12	.45	.68
Talking	1.21	0 - 3	.92	0	.88
Relations	.95	0 - 3	.96	.04	.77
Total Test	24.62	4 - 66	13.11	.90	4.10

TABLE 48

Correlations among Difficulty subtest and total test scores for VCL 1¹

Subtest	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Basic Work Habits	1.00													
Independence	.60	1.00												
Making Decisions	.53	.33	1.00											
Use, Care of Equipment	.67	-.01	.03	1.00										
Taking Direction	.74	.28	-.02	.61	1.00									
Work Skills	.73	.51	.47	.45	.43	1.00								
Speed	.63	.53	.40	.27	.41	.86	1.00							
Following Instructions	.61	.33	.40	.50	.31	.85	.45	1.00						
Acceptance Skills	.64	.35	.14	.44	.71	.43	.39	.35	1.00					
Appearance	.56	.11	.13	.53	.62	.40	.39	.30	.62	1.00				
Punctuality	.50	.35	.11	.31	.49	.34	.33	.25	.63	.22	1.00			
Talking	.11	.15	-.02	-.04	.19	.17	.09	.21	.40	-.07	-.10	1.00		
Relations	.19	.18	.07	.06	.19	-.01	-.01	-.01	.59	-.01	.22	.29	1.00	
Total Test	.94	.57	.45	.62	.75	.82	.71	.68	.81	.62	.57	.25	.31	1.00

¹N=42; $r_{\text{critical}} = .307$ ($p \leq .05$)

TABLE 49

Correlations among Difficulty subtest and total test scores for VCL 2¹

Subtest	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Basic Work Habits	1.00													
Independence	.79	1.00												
Making Decisions	.74	.56	1.00											
Use, Care of Equipment	.69	.28	.28	1.00										
Taking Direction	.79	.45	.45	.53	1.00									
Work Skills	.75	.61	.47	.54	.62	1.00								
Speed	.65	.59	.43	.33	.60	.89	1.00							
Following Instructions	.68	.50	.42	.63	.50	.89	.58	1.00						
Acceptance Skills	.67	.54	.43	.44	.62	.77	.71	.66	1.00					
Appearance	.57	.49	.34	.34	.52	.52	.50	.42	.76	1.00				
Punctuality	.70	.48	.44	.61	.58	.61	.51	.58	.67	.42	1.00			
Talking	.20	.23	.25	-.04	.17	.54	.53	.43	.64	.23	.24	1.00		
Relations	.23	.15	.06	.19	.29	.35	.34	.30	.55	.20	.04	.36	1.00	
Total Test	.91	.73	.62	.76	.76	.92	.82	.81	.87	.68	.74	.48	.40	1.00

¹N=42; $r_{\text{critical}} = .307$ ($p \leq .05$)

FIGURE 18

Item analysis information for the Difficulty subtests of the VCL (1)

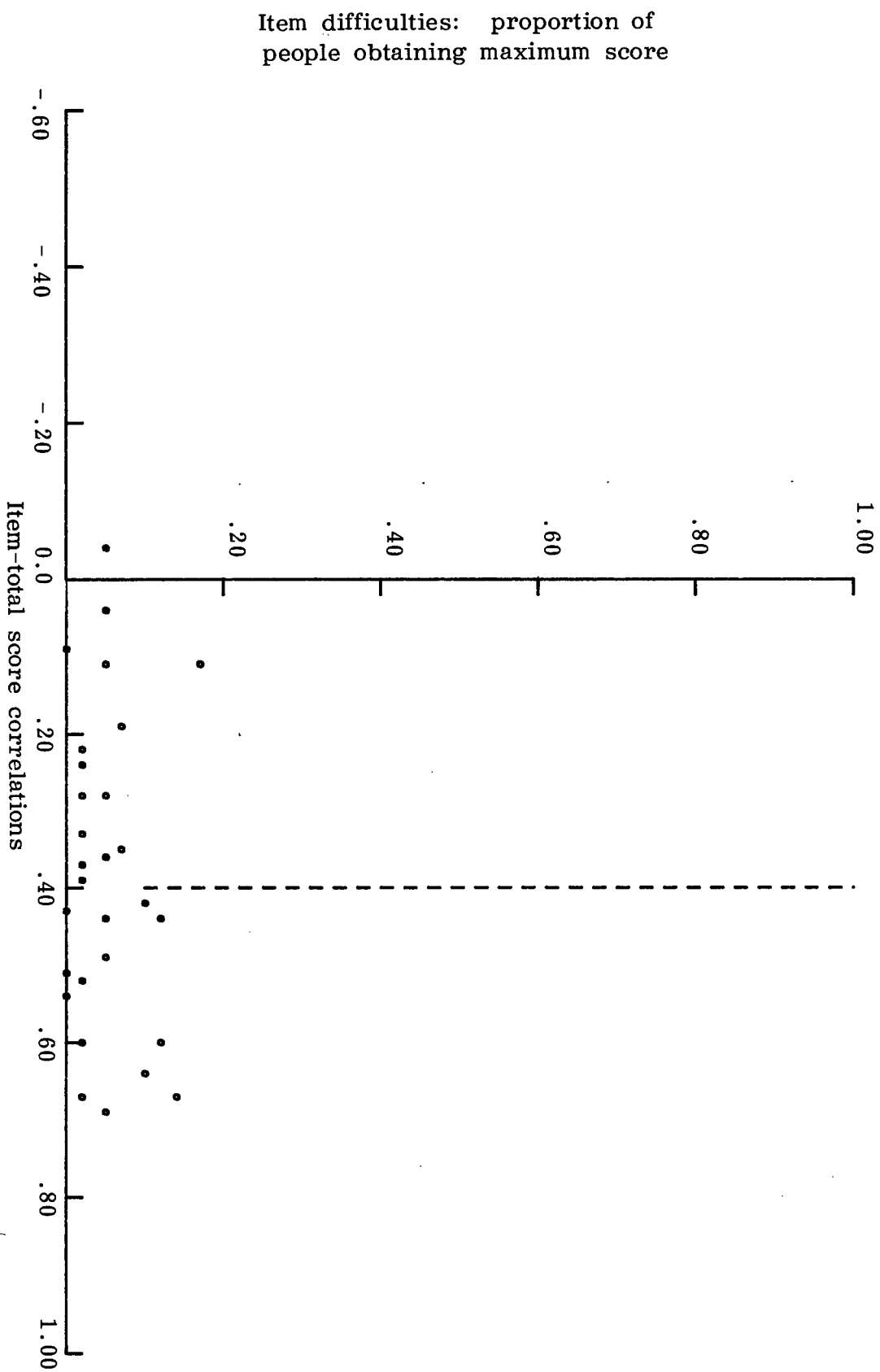
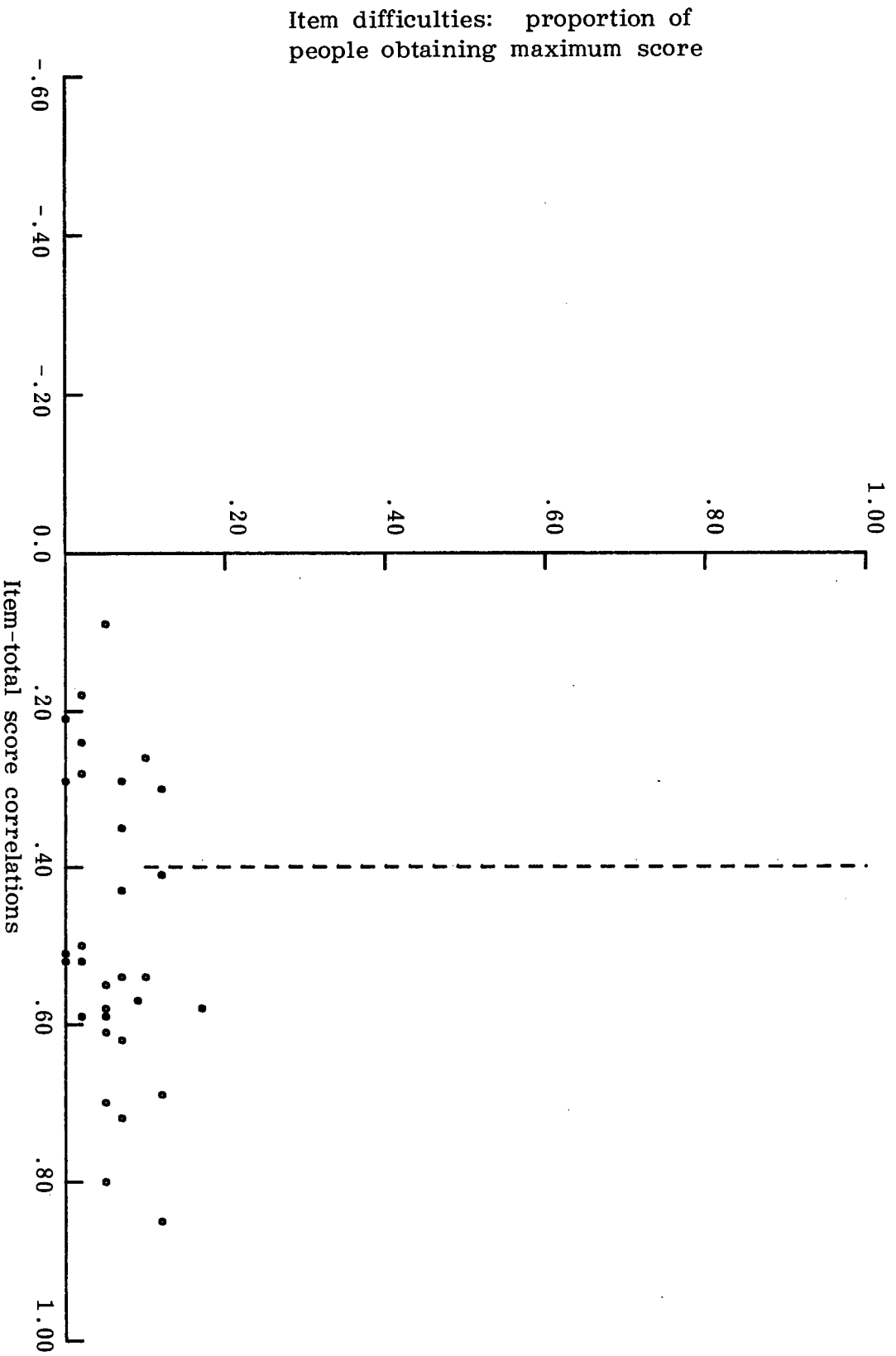


FIGURE 19

Item analysis information for the Difficulty subtests of the VCL (2)



the second sample, twenty-one items reach or exceed this criterion.

For the subtest, Basic Work Habits, five items discriminate among members of the first sample and eight items discriminate among members of the second sample. The items that correlate less than .40 with the total scores of both samples are:

1. concerned about trying new things
2. borrows things without asking
3. sees supervisor as family or friend

The following items do not discriminate among members of the first sample:

1. stops work when tired of the job
2. fools around when he has the chance
3. waits for supervisor to tell him what to do
4. would rather not work around machinery

The following item does not discriminate among members of the second sample:

1. "always" asking what to do next or if he's done well

For the second subtest, Work Skills, only one item, "Can't apply the words used in work" does not discriminate among members of the first sample.

For the subtest, Acceptance Skills, three items correlate .40 or greater with total test scores for the first sample, and six correlate .40 or greater with total test scores for the second sample. The following items do not discriminate among members of both samples:

1. doesn't care about how he looks
2. hard to understand what he says
3. doesn't speak very much
4. talks "all the time"
5. order co-workers around, bullies them, or makes fun of them
6. "always" touching people or telling them how much he likes them

The following items do not discriminate among members of the first sample:

1. either comes too early or stays long after he should be going home

2. stays away or come late on purpose
3. is shy

Generalizability of ratings. Estimates of the generalizability of ratings over the universe of raters for the Difficulty subtests are presented in Table 50. Table 51 provides estimates for each variance component used in obtaining these coefficients.

Of the seventy-eight coefficients for subtests by rater pairs, eleven are equal to zero. For the domain, Basic Work Habits, one coefficient is zero for domain scores; one coefficient is zero for the subtest, Making Decisions; three coefficients are zero for the subtest Use and Care of Equipment; and one coefficient is zero for Taking Direction. For the domain, Work Skills, one coefficient for domain scores, and one for each of the subtests equals zero. For the domain, Acceptance Skills, the subtests Punctuality and Relations with Co-Workers have two and one coefficients equal to zero, respectively. Rater pairs 1 and 2 have no coefficients that are equal to zero. Rater pair 6 has six subtest ratings where generalizability is zero.

Four of the remaining sixty-seven coefficients are greater than or equal to .80. These are for ratings on the subtests, Use and Care of Equipment, Taking Direction, Work Skills, and Relations with Co-Workers. Rater pairs 2,3 and 5 have 2, 1 and 1 subtest ratings that are greater than or equal to .80.

Of the remaining sixty-three coefficients that are greater than zero but less than .80, twenty-four reveal persons-by-judges bias. Ratings

TABLE 50

Generalizability coefficients by rater pair and across raters
for Difficulty subtest and total test scores of the VCL

Subtest	Rater Pair						AR
	1	2	3	4	5	6	
Basic Work Habits	.70	.21	.48	0.0	.22	.15	.45
Independence	.78	.27	.78	.08	.24	.46	.45
Making Decision	.33	.31	.68	0.0	.04	.44	.07
Use, Care of Equipment	.69	.75	0.0	0.0	.90	0.0	.48
Taking Direction	.63	.10	.85	0.0	.44	.13	.28
Work Skills	.76	.80	.64	.52	.02	0.0	.66
Speed	.59	.67	.43	.18	.36	0.0	.54
Following Instructions	.68	.68	.66	.49	.06	0.0	.65
Acceptance Skills	.47	.77	.65	.24	.47	.39	.67
Appearance	.34	.78	.54	.08	.58	.37	.60
Punctuality	.61	.27	0.0	.28	.49	0.0	.47
Talking	.47	.52	.73	.65	.78	.47	.68
Relations	.50	.89	0.0	.25	.18	.50	.42
Total Test	.86	.63	.80	.02	.09	.22	.72

TABLE 51

Variance components contributing to the universe score variance (σ^2_{up}) and error variance (σ^2_{ϵ}) for the Difficulty subtest and total test scores of the VCL

Subtest		Rater Pair													
		1		2		3		4		5		6		Across Raters	
		A	B	A	B	A	B	A	B	A	B	A	B	A	B
Basic Work Habits	$\sigma^2_{\mu\rho}$.02+.01		.003+.01		.04+.003		0 + 0		.01+.01		0 + .003		.02+.003	
	σ^2_{δ}	.01+.01		.04 + .01		.004+.004		.04+.03		.06+.01		.01+.01		.02+.01	
Independence	$\sigma^2_{\mu\rho}$.09+.02		.04 + .04		.08+.02		.01+ 0		.04+ 0		.03+.02		.06+.02	
	σ^2_{δ}	.03+.02		.16 + .04		.01+.02		.11+.02		.15+.05		.04+.02		.08+.02	
Making Decisions	$\sigma^2_{\mu\rho}$.01+.03		.04 + 0		.11+ 0		0 + 0		0 + .01		0 + .06		.01+ 0	
	σ^2_{δ}	.03+.04		.04 + .04		.02+.03		.22+.10		.12+.04		.05+.02		.07+.05	
Use, Care of Equipment	$\sigma^2_{\mu\rho}$.05+.06		.11 + .13		0 + 0		0 + 0		.43+.02		0 + 0		.03+.03	
	σ^2_{δ}	.02+.03		.01 + .07		.02+.01		.13+.05		.03+.62		.02+.02		.03+.03	
Taking Direction	$\sigma^2_{\mu\rho}$.08+ 0		0 + .01		.02+.04		0 + 0		.09+ 0		0 + .01		.02+ 0	
	σ^2_{δ}	.01+.04		.06 + .02		.01+.01		.05+.03		.06+.05		.05+.04		.03+.03	
Work Skills	$\sigma^2_{\mu\rho}$.06+.001		.10 + 0		.05+.01		.04+.001		0 + .003		0 + 0		.05+.002	
	σ^2_{δ}	.04+.01		.09 + .02		.02+.01		.03+.01		.13+.02		.03+.002		.02+.01	
Speed	$\sigma^2_{\mu\rho}$.06+.003		.14 + 0		.01+ 0		.01+.001		.07+ 0		0 + 0		.06+ 0	
	σ^2_{δ}	.02+.02		.02 + .04		.003+.01		.06+.01		.08+.04		.03+.01		.03+.02	
Following Instructions	$\sigma^2_{\mu\rho}$.09+ 0		.07 + 0		.19+.01		.07+.01		0 + .02		0 + 0		.08+.01	
	σ^2_{δ}	.02+.03		.01 + .03		.08+.03		.05+.03		.24+.03		.05+.01		.02+.02	
Acceptance Skills	$\sigma^2_{\mu\rho}$.03+.004		.03 + .01		.01+.01		.01+.001		.01+.01		0 + .01		.02+.01	
	σ^2_{δ}	.003+.01		.01 + .004		.003+.01		.02+.01		.02+.02		.01+.01		.01+.01	
Appearance	$\sigma^2_{\mu\rho}$.03+.01		.11 + .11		.03+.05		.02+ 0		.07+.05		.06+.02		.08+.03	
	σ^2_{δ}	.04+.03		.06 + .01		.04+.04		.20+.05		.05+.04		.02+.03		.04+.03	
Punctuality	$\sigma^2_{\mu\rho}$.07+.01		0 + .01		0 + 0		0 + .02		.16+ 0		0 + 0		.04+.01	
	σ^2_{δ}	.02+.03		.03 + .01		.01+.01		.03+.02		.12+.04		.01+.01		.03+.02	
Talking	$\sigma^2_{\mu\rho}$.03+.02		.02 + .02		.06+.05		.06+.004		.10+.12		0 + .10		.03+.07	
	σ^2_{δ}	.02+.03		.01 + .03		.01+.03		.01+.03		.02+.04		.08+.03		.02+.02	
Relations	$\sigma^2_{\mu\rho}$.02+.02		.10 + .07		0 + 0		.02+ 0		0 + .02		0 + .15		.01+.04	
	σ^2_{δ}	.01+.04		0 + .02		.01+.03		.02+.03		.08+.02		.11+.04		.03+.03	
Total Test	$\sigma^2_{\mu\rho}$.03+.002		.03 + .003		.03+.002		.001+0		.001+.004		0 + .002		.03+.002	
	σ^2_{δ}	.002+.002		.02 + .003		.01+.002		.02+.004		.04+.004		.01+.002		.01+.003	

for all subtests, except Use and Care of Equipment and Self-Expression - Talking, have at least one coefficient to which the p_j component contributes most to the observed score variance. The subtest Independence, has four, the largest number for any subtest. The third pair of raters has no subtest ratings where the p_j component is the largest or equal to the largest variance component. The other rater pairs have from three (for rater pairs 2 and 6) to eight (for rater 5) coefficients to which the p_j component contributes most to observed score variance.

The generalizability coefficients for total test scores by rater pair range from .02 for rater pair 4 to .86 for rater pair 1. For rater pairs 4, 5 and 6, the p_j component contributes most to observed score variance.

The coefficients of generalizability for the subtests across all raters range from .07 for Making Decisions to .68 for Relations with Co-Workers. For all subtests within the domain, Basic Work Habits, the p_j component contributes most to observed score variance. The generalizability coefficient for total test scores is .72.

Results of Analysis - RCL

Test Analysis. The mean, range, standard deviation, standard, standard error of measurement, and Hoyt coefficient of internal consistency for each subtest of the RCL for both analysis samples are presented in Tables 52 and 53. Table 54 summarizes these results for the total test.

For the first analysis sample, a floor effect was obtained for the subtest, Cooking and Home Management. For this sample a ceiling effect was observed for seven subtests: Cleanliness, Appearance and Eating,

TABLE 52

Test analysis information for the subtests of the RCL 1

Subtest	Mean	Range	St. Dev.	Hoyt	SEM
Personal Routines	61.98	27 - 87	14.38	.93	3.74
Cleanliness	14.46	6 - 20	4.52	.92	1.22
Appearance, Eating	15.83	8 - 20	3.41	.84	1.28
Room Management	11.73	6 - 20	4.08	.86	1.45
Time Management	11.83	3 - 20	4.19	.84	1.60
Health	8.12	3 - 18	3.57	.79	1.55
Community Awareness	39.49	12 - 69	14.64	.93	3.91
Transportation	9.17	1 - 20	4.82	.90	1.45
Shopping	10.56	1 - 20	4.92	.92	1.34
Leisure	8.39	3 - 15	3.29	.66	1.81
Budgeting	5.59	2 - 15	3.12	.75	1.47
Cooking, Home Management	5.78	0 - 12	3.19	.74	1.56
Social Maturity	59.32	26 - 80	13.13	.91	3.99
Communication	10.44	2 - 18	4.27	.85	1.59
Consideration	13.59	4 - 19	3.43	.71	1.75
Getting Friends	14.61	7 - 20	3.39	.78	1.50
Keeping Friends	10.07	4 - 15	3.01	.63	1.73
Handling Problems	10.61	3 - 19	3.58	.81	1.48

TABLE 53

Test analysis information for the subtests of the RCL 2

Subtest	Mean	Range	St. Dev.	Hoyt	SEM
Personal Routines	59.42	25 - 90	15.12	.93	3.85
Cleanliness	14.32	6 - 20	4.60	.92	1.21
Appearance, Eating	15.10	8 - 20	3.69	.84	1.39
Room Management	11.57	3 - 20	5.07	.91	1.45
Time Management	11.27	4 - 18	3.84	.81	1.59
Health	7.15	2 - 18	3.70	.80	1.58
Community Awareness	37.72	13 - 80	17.41	.95	3.81
Transportation	8.63	2 - 20	4.68	.86	1.63
Shopping	9.95	0 - 19	4.79	.88	1.55
Leisure	8.52	2 - 20	4.19	.81	1.75
Budgeting	5.22	0 - 13	3.37	.80	1.42
Cooking, Home Management	5.40	1 - 14	3.48	.77	1.58
Social Maturity	58.97	23 - 92	14.54	.92	4.01
Communication	9.95	0 - 18	4.77	.87	1.61
Consideration	14.45	5 - 20	4.14	.85	1.54
Getting Friends	13.95	7 - 20	3.23	.73	1.60
Keeping Friends	10.25	5 - 18	3.39	.73	1.68
Handling Problems	10.38	2 - 17	3.57	.78	1.60

TABLE 54

Test analysis summary for Residential Checklist

Sample	Mean	Range	St. Dev.	α_{strat}	SEM
RCL 1	160.78	72 - 224	36.75	.96	6.93
RCL 2	156.13	61 - 262	41.42	.97	7.04

Room Management, Time Management, Transportation, Shopping, and Getting Friends. For the second sample, a floor effect was obtained for three subtests: Shopping, Budgeting, and Communication. A ceiling effect was observed for six subtests: Cleanliness, Appearance and Eating, Room Management, Transportation, Leisure, Consideration, and Getting Friends. These results indicate that these subtests do not contain sufficient easy and/or hard items for these samples.

The Hoyt coefficients for Personal Routines, Community Awareness, and Social Maturity are .93, .93, and .91, respectively, for sample one; and, .93, .95, and .92, respectively for sample two. For sample one, the subtest coefficients range from .63 for Keeping Friends to .92 for Cleanliness and Shopping. For the second sample these values range from .73 for Getting Friends and Keeping Friends to .92 for Cleanliness. The estimates of internal consistency for the total test are .96 and .97 for samples one and two, respectively.

Correlations among subtest and total test scores for both samples are presented in Tables 55 and 56. For the first analysis sample, Personal Routines, Community Awareness, and Social Maturity correlate .88, .88, and .86 with total scores. These values are .84, .89, and .91 for the second sample.

Item Analysis. The results of the item analyses for the subtests of the RCL are presented in Figures 20 to 25. Eighty-four of the one hundred and fifty items of the RCL correlate .40 or greater with total test scores for the first sample. Ninety-five items reach or exceed this criterion for the second sample.

TABLE 55
Correlations among subtest and total test scores
of the RCL 1¹

Subtest	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Personal Routines	1.00																		
Clean.	.82	1.00																	
App./Eat.	.77	.70	1.00																
Rm. Man.	.70	.60	.38	1.00															
Tm. Man.	.75	.34	.54	.36	1.00														
Heal.	.58	.31	.23	.13	.49	1.00													
Community Awareness	.65	.36	.40	.28	.73	.59	1.00												
Trans.	.51	.20	.33	.11	.57	.71	.80	1.00											
Shop.	.69	.47	.51	.34	.68	.51	.91	.71	1.00										
Lei.	.14	.04	.16	-.09	.41	-.01	.46	.12	.32	1.00									
Bud.	.49	.22	.18	.26	.56	.55	.75	.51	.65	.11	1.00								
Ck. Hm. Man.	.52	.37	.22	.43	.45	.40	.76	.46	.62	.29	.57	1.00							
Social Maturity	.64	.35	.56	.33	.79	.31	.64	.37	.64	.44	.49	.46	1.00						
Commun.	.58	.30	.31	.14	.73	.62	.74	.52	.68	.47	.59	.51	.73	1.00					
Cons.	.35	.27	.33	.34	.35	-.08	.12	-.07	.16	.17	.17	.08	.65	.19	1.00				
Get. Fr.	.56	.36	.49	.31	.53	.32	.66	.45	.62	.41	.26	.62	.72	.56	.23	1.00			
Kp. Fr.	.37	.13	.36	.27	.58	-.03	.29	.09	.28	.24	.30	.22	.81	.37	.66	.37	1.00		
Hand. Prob.	.49	.22	.59	.17	.68	.13	.48	.29	.55	.28	.41	.23	.82	.46	.42	.49	.71	1.00	
Total Test	.88	.59	.66	.50	.86	.57	.88	.65	.86	.39	.66	.67	.86	.78	.41	.73	.55	.67	1.00

¹N=41; $r_{\text{critical}} = .304$ ($p \leq .05$)

TABLE 56
Correlations among subtest and total test scores for RCL 2¹

Subtest	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Personal Routines	1.00																		
Clean.	.76	1.00																	
App./ Eat.	.82	.76	1.00																
Rm. Man.	.84	.69	.69	1.00															
Tm. Man.	.72	.25	.34	.42	1.00														
Heal.	.43	-.12	.12	.08	.66	1.00													
Community Awareness	.58	.03	.24	.29	.80	.85	1.00												
Trans.	.37	-.12	-.02	.11	.72	.73	.89	1.00											
Shop.	.52	.01	.23	.19	.75	.81	.90	.84	1.00										
Lei.	.45	.11	.24	.32	.48	.50	.76	.56	.50	1.00									
Bud.	.49	-.02	.19	.17	.75	.80	.86	.72	.74	.54	1.00								
Ck. Hm. Man.	.66	.18	.31	.47	.70	.78	.82	.56	.68	.64	.68	1.00							
Social Maturity	.67	.22	.50	.37	.70	.73	.75	.66	.76	.48	.67	.59	1.00						
Commun.	.38	-.07	.10	.03	.64	.81	.79	.74	.83	.45	.71	.62	.73	1.00					
Cons.	.60	.30	.50	.43	.57	.40	.41	.33	.44	.22	.42	.32	.83	.38	1.00				
Get. Fr.	.55	.19	.49	.38	.39	.58	.59	.48	.58	.54	.34	.54	.73	.64	.49	1.00			
Kp. Fr.	.44	.20	.44	.28	.29	.41	.38	.30	.37	.30	.38	.27	.76	.24	.69	.41	1.00		
Hand. Prob.	.60	.25	.40	.36	.71	.49	.62	.60	.64	.33	.62	.45	.76	.40	.62	.25	.65	1.00	
Total Test	.84	.37	.57	.56	.84	.77	.89	.74	.84	.65	.77	.79	.91	.73	.68	.70	.59	.75	1.00

¹
N = .40; r_{critical} = .308 (p ≤ .05)

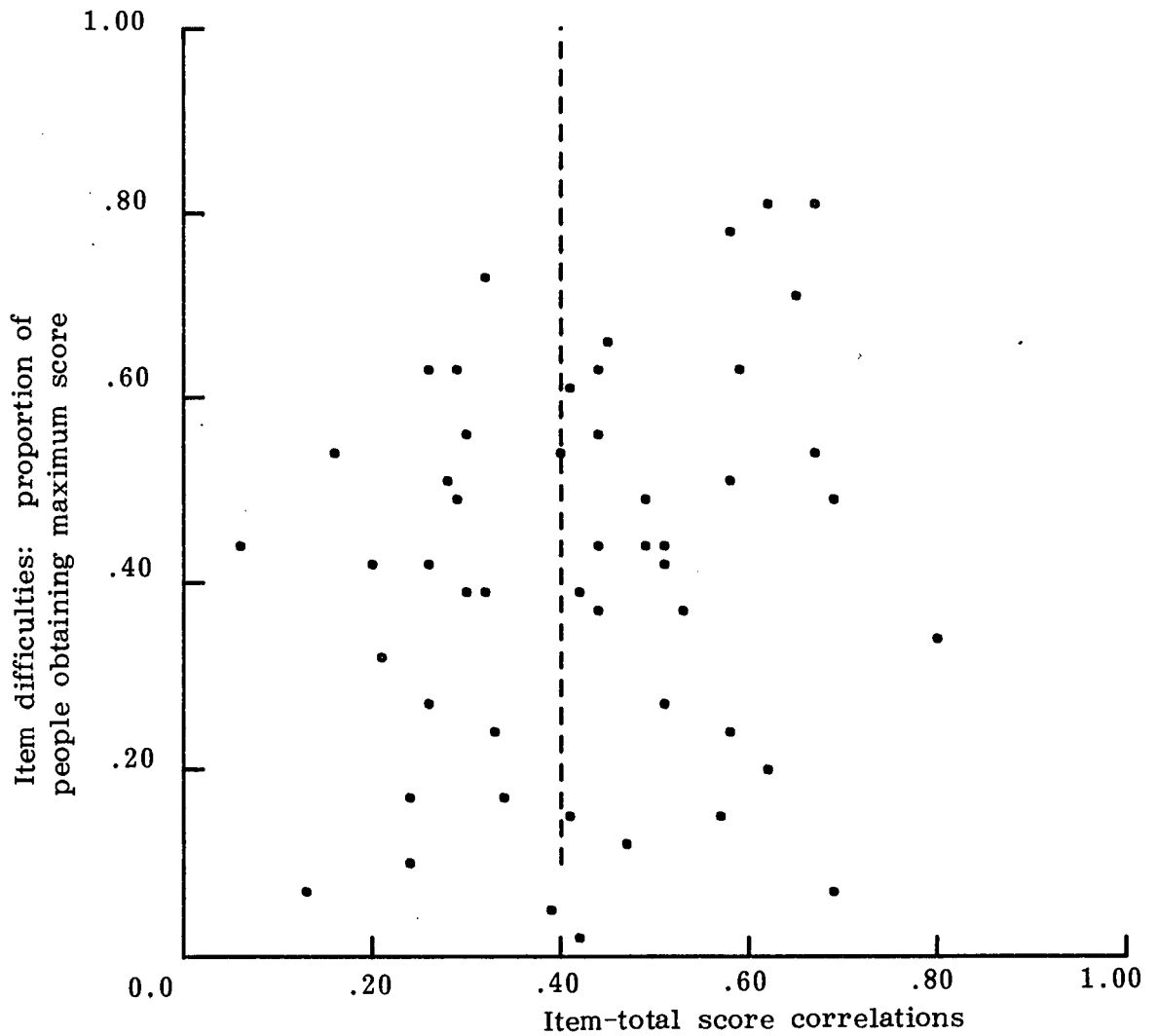


FIGURE 20

Item analysis information for the subtest
Personal Routines of the RCL (1)

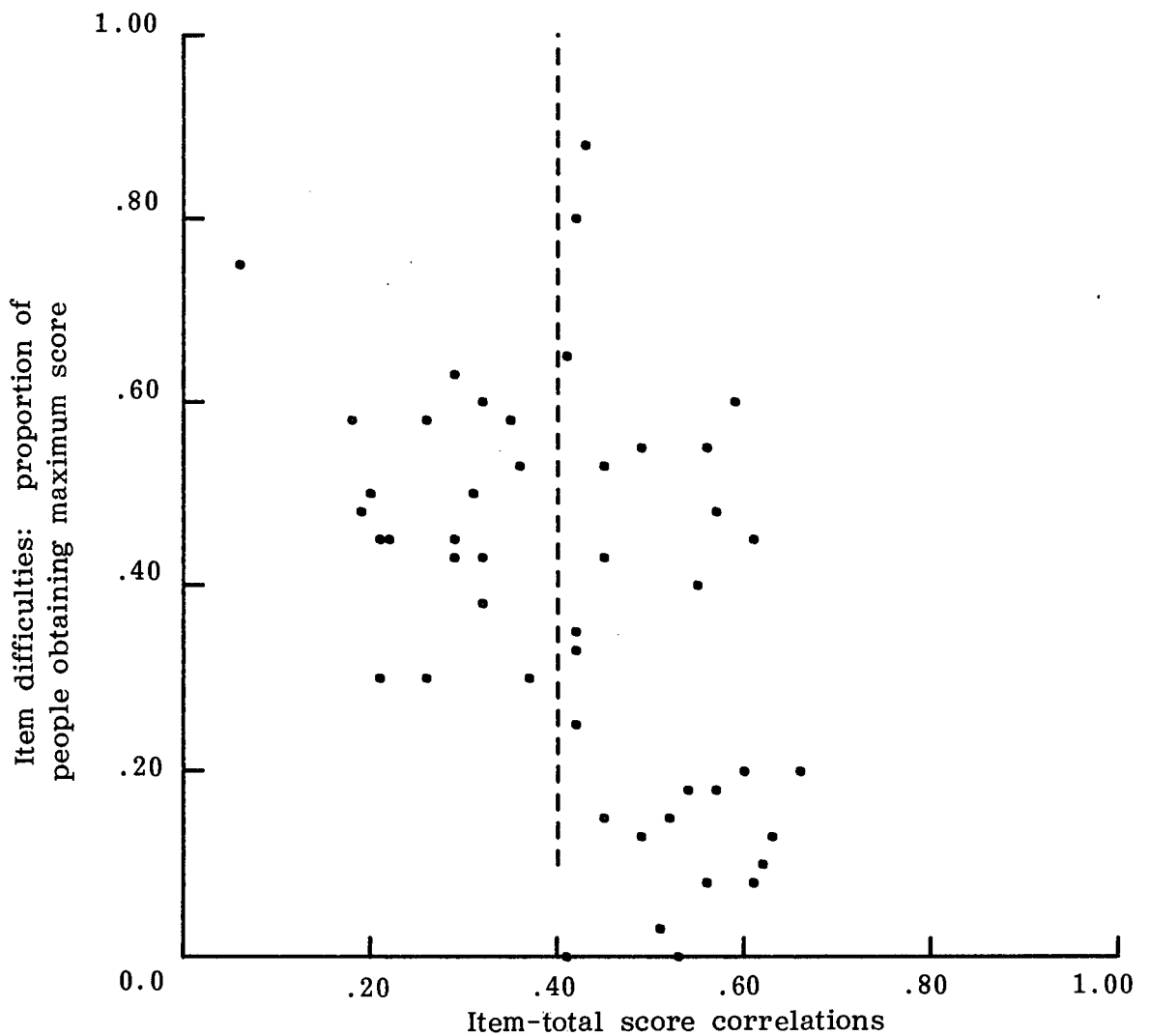


FIGURE 21

Item analysis information for the
subtest Personal Routines of the
RCL(2)

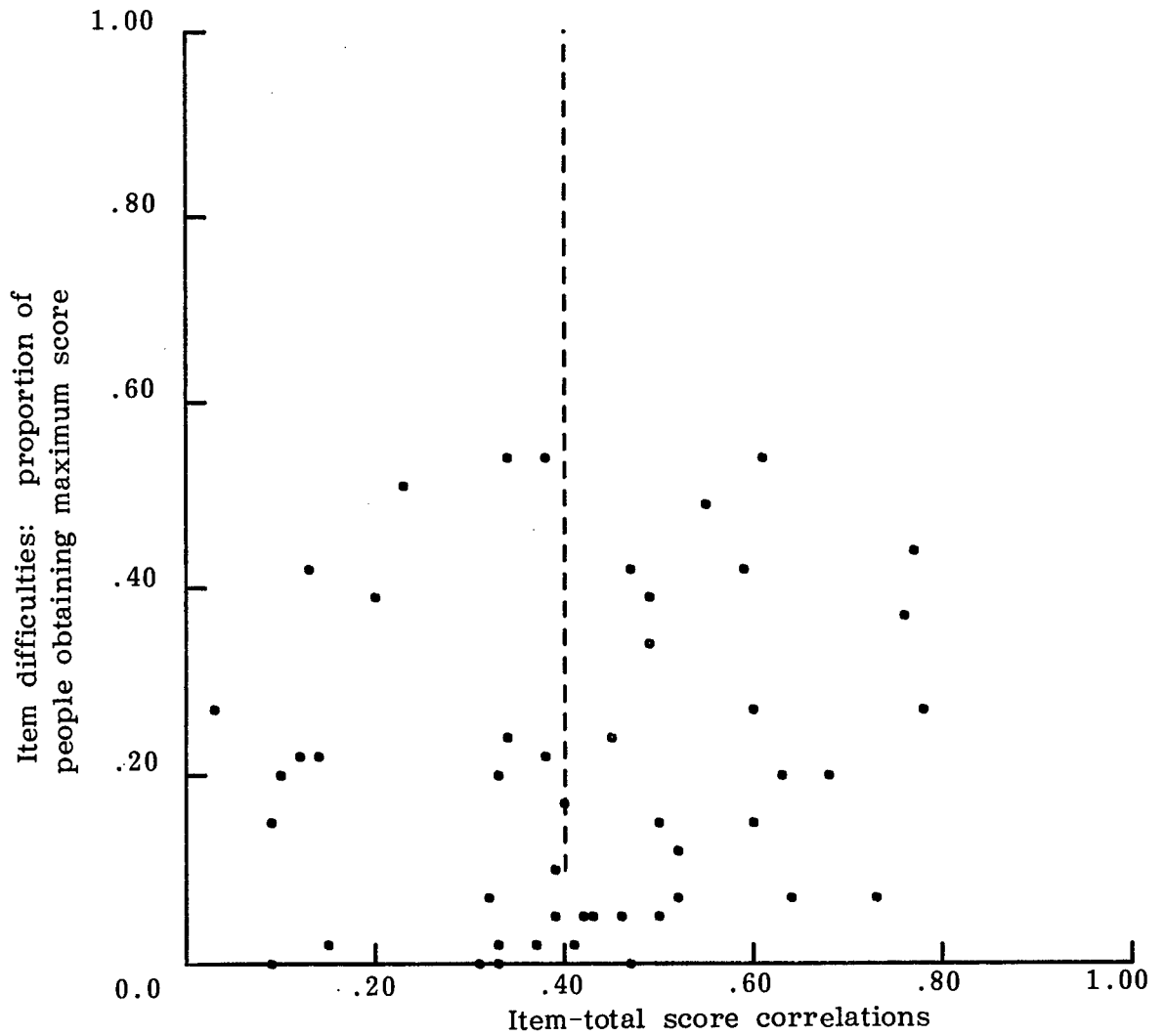


FIGURE 22

Item analysis information for the
subtest Community Awareness of
the RCL (1)

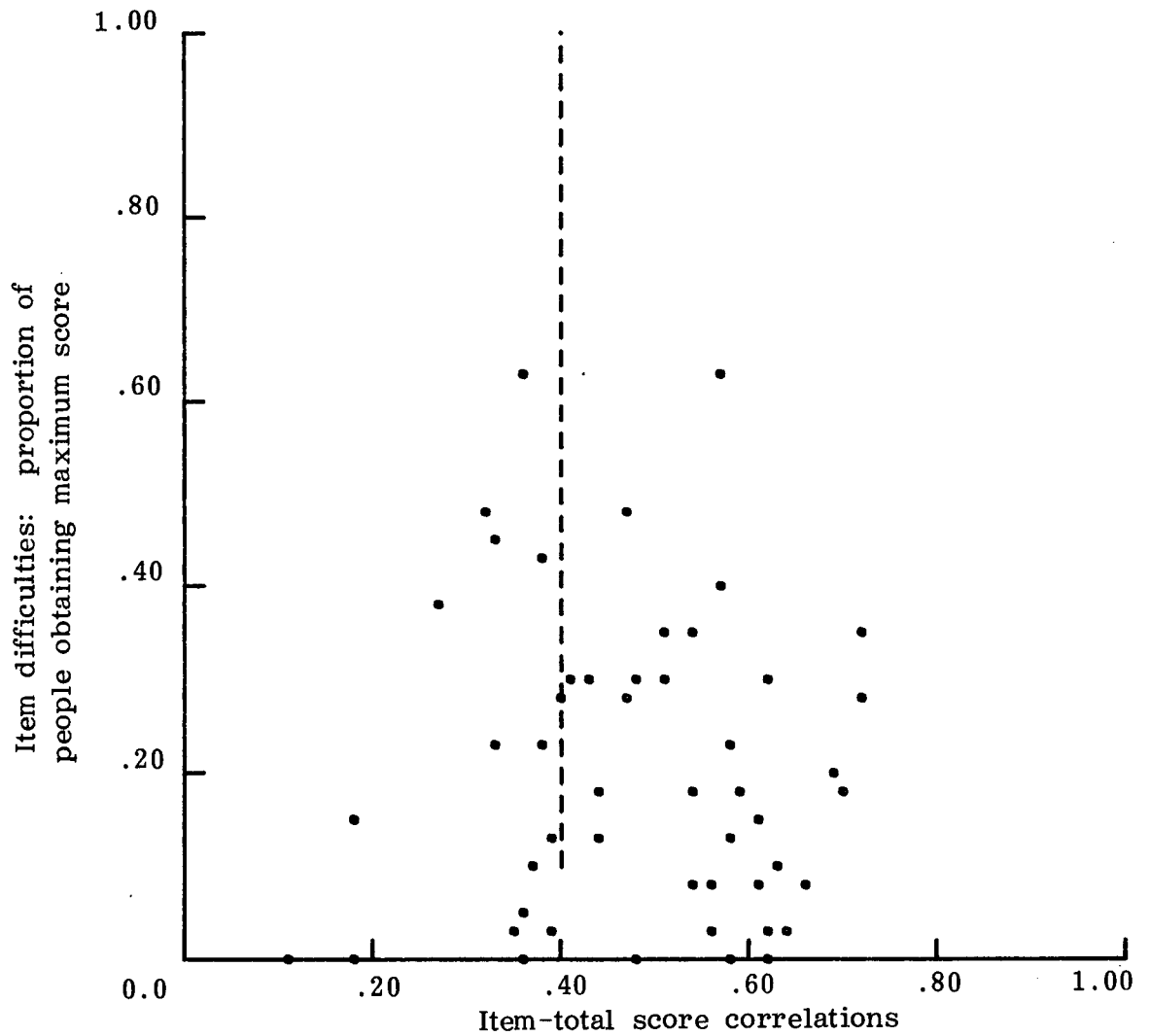


FIGURE 23

Item analysis information for the
subtest Community Awareness of
the RCL (2)

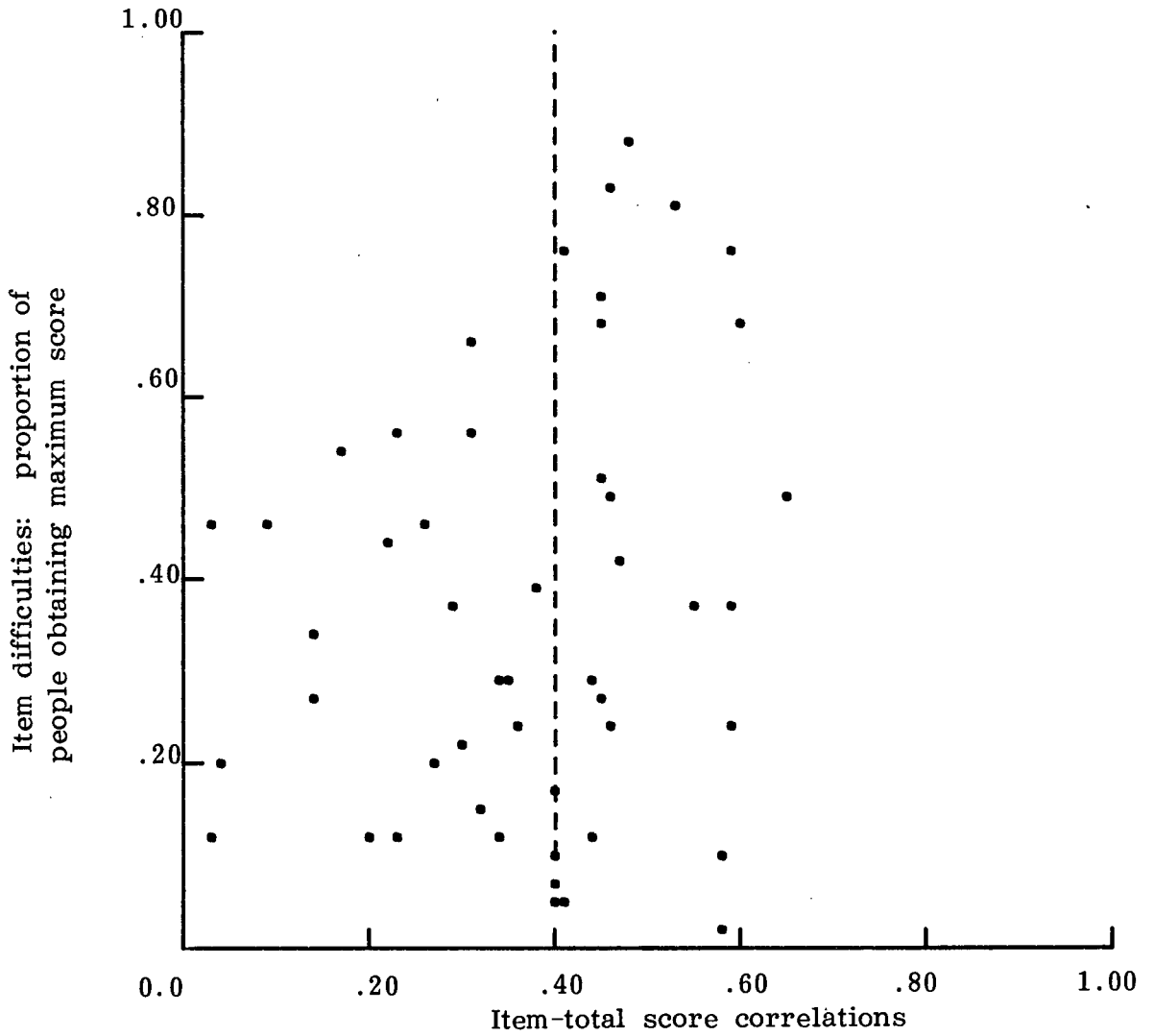


FIGURE 24

Item analysis information for the
subtest Social Maturity of the RCL (1)

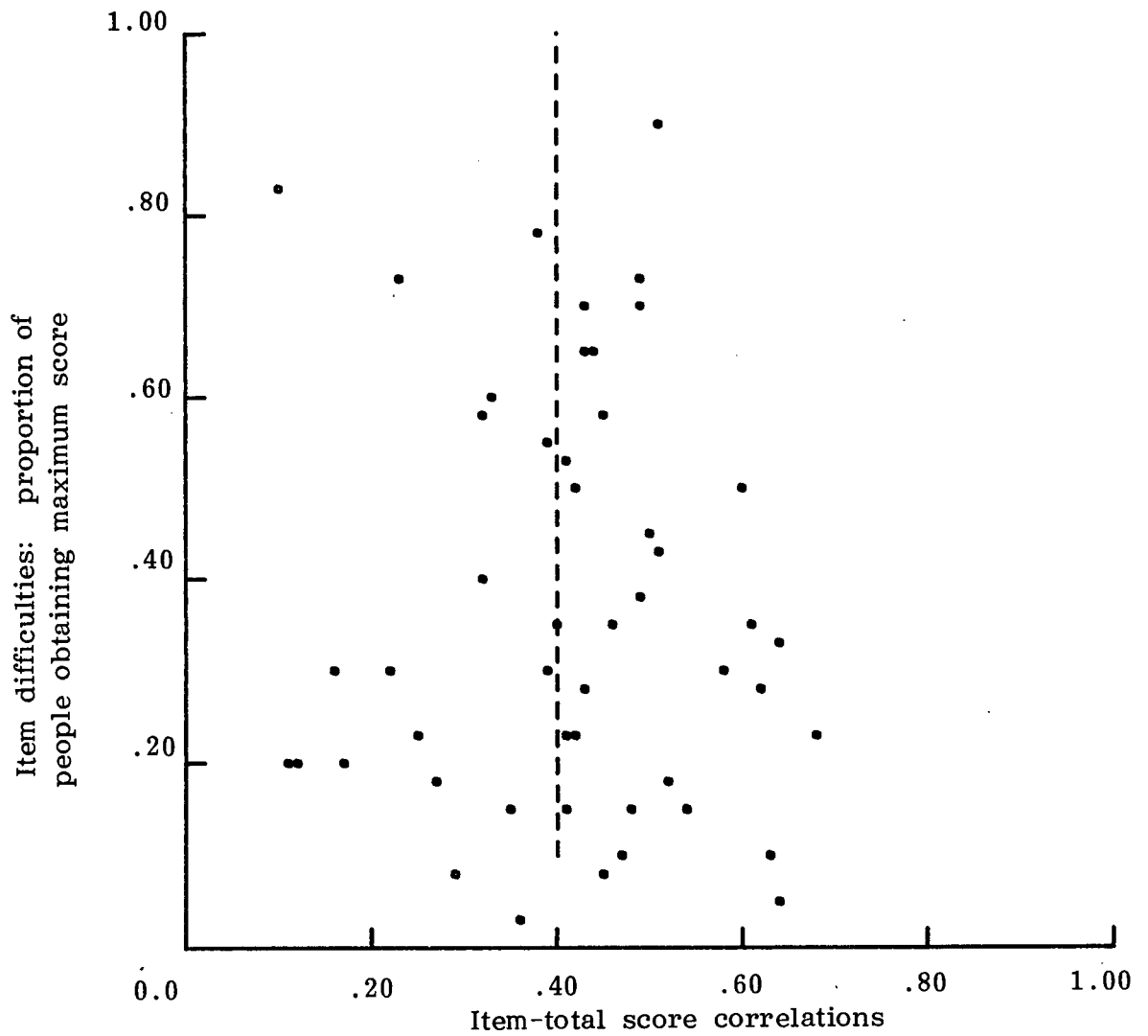


FIGURE 25

Item analysis information for the
subtest Social Maturity of the RCL (2)

For the subtest Cleanliness seven items appear to discriminate among members of the first sample. Three items correlate less than .40 with the total scores. These are:

1. washes all over when bathing or showering
2. bathes or showers when needed
3. brushes teeth properly and regularly

For the second sample, only two items of this subtest appear to be discriminating. These are:

1. cleans and clips nails; keeps ears clean
2. covers mouth when coughing or sneezing and uses a handkerchief

The remaining eight items correlate .40 or greater with total test scores.

For both samples, six items from the subtest Appearance and Eating appear to be discriminating. The following items correlate less than .40 with total test scores for both samples:

1. dresses himself neatly
2. changes underwear and clothing regularly
3. clothes "go together"

In addition the items

1. wears the right kind of clothes for the occasion and the weather
2. chews food with mouth closed, and does not speak with mouth full or make loud noises

do not discriminate among members of the first and second samples, respectively.

For the subtest Room Management, only four items discriminate among members of the first sample and five among members of the second sample.

The following items correlate less than .40 with total test scores for both samples:

- 1, when undressing separates dirty from clean clothes
- 2, keeps room tidy/or accepts sharing of responsibility of room cleaning

3. cleans wash basin and bath tub after use
4. clean regularly

In addition, the items

1. makes his bed neatly
2. can find his clothes in his drawers

do not discriminate among members of the first and second samples respectively.

For the subtest, Time Management the item

1. has and sets own alarm

correlates less than .40 with total test scores for sample one. The remaining nine items discriminate among members of this sample. The items

1. gets up on his own
2. meets routine time deadlines on his own

do not discriminate among members of the second analysis sample. The remaining eight items correlate .40 or greater with total test scores.

For the subtest, Health, the following items do not appear to be discriminating among members of either sample:

1. eats a balanced diet
2. aware of differences between men and women

These are the easiest items for both samples. In addition, five other items do not discriminate among members of the first analysis sample

1. aware of medical problems and needed precautions
2. handles medications reliably on his own
3. understands his own sexual development
4. knows how babies are conceived and born
5. aware of birth control process, venereal disease symptoms, prevention and cure

Thus, eight items from this subtest discriminate among members of the first sample, whereas only three correlate .40 or greater with total test scores for the second sample.

For the subtest, Transportation, two items correlate less than .40 with total scores for sample one. These are

1. travels with someone on the bus
2. behaves appropriately when riding on the bus

The remaining eight items correlate .40 or greater with total test scores.

Two additional items do not appear to discriminate among members of the second sample:

1. uses bus for short trips
2. transfers to another bus line when necessary

Eight items are useful in discriminating among members of this sample. All the items in this subtest are difficult for these samples. The p values range from 0.0 to .34 for sample one and 0.0 to .30 for sample two. All four items however, approach the criterion of .40 ($r = .38, .33, .38, \text{ and } .39$).

All items of the subtest, Shopping, correlate .40 or greater with total test scores for both samples.

A number of items from the subtest, Leisure, do not appear to be discriminating among members of either sample. These are

1. does casual things that don't require much planning with a friend
2. does things by himself
3. satisfied with his use of leisure, not continuously asking what to do or complaining about lack of something to do
4. has an active sport that he does
5. has a creative hobby that he does
6. follows something (spectator) which requires special interest or knowledge

In addition, three more items correlate less than .40 with the total scores of sample one. These are

1. takes part in planned/supervised leisure

2. plans outings occasionally with friends
3. belongs to a group where he can meet people

Thus, only one item discriminates among members of the first sample whereas four items discriminate among members of the second sample.

For the subtest, Budgeting, two items do not discriminate among members of either sample:

1. has bank account of his own
2. understands dangers of credit buying

No subjects received the highest scale value for the first item. An additional item correlates less than .40 with the total scores of sample two:

1. reads and understands billing statements

The following item correlates less than .40 with the total scores of sample one

1. does not usually run out of money before next allowance/cheque

For both samples seven items measuring budgeting skills correlate .40 or greater with total test scores.

For both samples, the following items from the subtest, Cooking and Home Management, correlate less than .40 with total scores

1. does major cleaning
2. entertains for an evening
3. takes charge of home/apartment when necessary for extended periods of time

In addition, four other items do not appear to discriminate among members of sample one:

1. prepares his own breakfast or lunch
2. helps with food preparation
3. prepares a variety of simple nutritious meals
4. takes an interest in how the house/room looks

Two additional items correlate less than .40 with the total scores of sample two

1. uses a simple cookbook
2. knows the instructions on household labels

The items from this subtest are difficult for both samples. The p values range from 0.0 to .24 for sample one and 0.0 to .30 for sample two. Only three items discriminate among members of sample one and five among members of sample two.

All items from the subtest, Communication, correlate .40 or greater with total test scores for both samples, except for the item:

1. knows how and whom to contact in emergency

which does not discriminate among members of the second analysis sample.

For the subtest, Consideration, only one item discriminates among members of the first sample:

1. is polite - uses please, thank you, etc.

Nine items correlate less than .40 with total test scores. For the second sample, only two items fail to discriminate:

1. looks at people when talking to them
2. doesn't bully others

For the subtest, Getting Friends, the following items fail to discriminate among the members of either sample:

1. takes his turn in group duties (this item approaches the criterion of .40)
2. makes his own friends, doesn't rely on staff, volunteers, family, etc.
3. finds a way to introduce himself to someone he wants to know

In addition, the item:

1. offers help when someone is sick, upset or having trouble doing something

correlates less than .40 with the total scores of sample one; and, the items

1. smiles at and greets people he recognizes
2. sits with someone at coffee/meals, or on the bus
3. friendly with people of both sexes

correlate less than .40 with the total scores of sample two. Thus six items discriminate among members of the first sample and four among members of the second sample.

For the subtest, Keeping Friends, five items discriminate among members of the second sample; only two discriminate among members of the first sample. The following items correlate less than .40 with total test scores for both samples:

1. doesn't take things too seriously or overreact
2. works out a compromise with a friend; doesn't always expect to have his own way
3. can take a hint when someone wants to leave or wants him to leave
4. shows affection without embarrassing others
5. doesn't expect more of friends than they can give

In addition, the following items do not discriminate among members of the first sample:

1. does things for himself
2. when he has offended someone, apologizes and tries not to do it again
3. if necessary breaks relationships without becoming enemies

For the last subtest, Handling Problems, seven items discriminate among members of the first sample; six discriminate among members of the second sample. The following items correlate less than .40 with total test scores for

both samples.

1. doesn't give up easily
2. accepts responsibilities for his decisions

In addition the item

1. learns from his errors; doesn't keep making the same mistake
correlates less than .40 with total scores for the first sample; and, the
items

1. tries hard to do well
2. works out his own solutions to problems

do not discriminate among members of the second analysis sample.

Generalizability of Ratings. Estimates of the generalizability of ratings over the universe of raters for the subtests of the RCL are presented in Table 57. Table 58 provides estimates for each component used in calculating these coefficients.

In examining the results for subtests by rater pair, it can be seen that eleven coefficients are estimated to be zero. For the domain Personal Routines, one coefficient for the domain scores is zero; two coefficients, each, for the subtests, Cleanliness, Appearance and Eating, and Health are zero; and, one coefficient for Time Management is zero. For the domain, Community Awareness, two coefficients for the subtest Budgeting are zero; and, one coefficient for Cooking and Home Management is zero. There are no coefficients equal to zero for the domain scores of Social Maturity, nor for the subtest scores. Rater pairs 1 and 5 have no subtest ratings where generalizability is zero. Rater pair 2 has five coefficients that are equal to zero,

TABLE 57

Generalizability coefficients by rater pair and across raters for subtests and total test scores of the RCL

Subtest	Rater Pair						AR
	1	2	3	4	5	6	
Personal Routines	.89	.01	.54	0.0	.94	.31	.22
Cleanliness	.73	0.0	.32	0.0	.69	.60	.12
Appearance, Eating	.86	0.0	0.0	.02	.70	.16	.12
Room Management	.56	.95	.85	.37	.76	.47	.27
Time Management	.81	.01	0.0	.26	.73	.82	.50
Health	.68	0.0	.02	0.0	.56	.04	.43
Community Awareness	.82	.04	.02	.49	.70	.67	.62
Transportation	.81	.69	.58	.80	.78	.78	.81
Shopping	.81	.82	.54	.89	.60	.05	.68
Leisure	.47	.30	.11	.42	.54	.75	.27
Budgeting	.80	0.0	.20	.48	.51	0.0	.48
Cooking, Home Management	.12	0.0	.03	.10	.84	.13	.05
Social Maturity	.72	.03	.47	.61	.76	.70	.71
Communication	.84	.88	.06	.88	.92	.86	.81
Consideration	.52	.03	.77	.62	.13	.69	.62
Getting Friends	.59	.47	.21	.64	.65	.70	.17
Keeping Friends	.36	.28	.82	.31	.42	.45	.68
Handling Problems	.77	.04	.20	.87	.08	.57	.71
Total Test	.68	.24	.36	.43	.62	.48	.45

TABLE 58

Variance components contributing to universe score variance ($\sigma^2_{\mu\rho}$) and error variance (σ^2_{δ}) by rater pair and across raters for subtest and total score of the RCL

Subtest		Rater Pair						Across Raters
		1	2	3	4	5	6	
Personal Routines	$\sigma^2_{\mu\rho}$ σ^2_{δ}	.03 + .002 .002+.002	0 + .0001 .01+.004	.05+ 0 .03+.002	0 + 0 .10+.002	.05+.002 .001+.002	.01+.001 .01+.001	.01+.001 .04+.002
Cleanliness	$\sigma^2_{\mu\rho}$ σ^2_{δ}	.09 + .01 .03 + .01	0 + 0 .08+.01	.04+.001 .09+.01	0 + 0 .03+.01	.07 + .001 .02 + .01	.04+.003 .02+.004	.01+.002 .09+.01
Appearance, Eating	$\sigma^2_{\mu\rho}$ σ^2_{δ}	.05 + .01 .005+.005	0 + 0 .01+.02	0 + 0 .06+.01	0 + .003 .17+.003	.07++ 0 .02 + .01	0 + .01 .04+.01	.01+.004 .06+.01
Room Management	$\sigma^2_{\mu\rho}$ σ^2_{δ}	.03 + .01 .02 + .01	.18+.02 .01+.01	.17+.01 .02+.01	.03+ 0 .03+.01	.10 + .002 .02 + .01	.03+.01 .03+.01	.02+.004 .08+.01
Time Management	$\sigma^2_{\mu\rho}$ σ^2_{δ}	.09 + .0003 .01 + .01	0 + .003 .18+.02	0 + 0 .08+.01	.02+ 0 .05+.01	.07 + 0 .02 + .01	.11+.003 .02+.01	.05+ 0 .04+.01
Health	$\sigma^2_{\mu\rho}$ σ^2_{δ}	.06 + 0 .01 + .01	0 + 0 .10+.01	0 + .001 .05+.01	0 + 0 .03+.01	.02 + .002 .01 + .01	0 + .002 .03+.01	.03+ 0 .03+.01
Community Awareness	$\sigma^2_{\mu\rho}$ σ^2_{δ}	.04 + .001 .01 + .002	0 + .001 .02+.003	0 + .001 .05+.001	.02+.001 .02+.002	.03 + .001 .01 + .002	.02+.001 .01+.002	.03+.001 .02+.003
Transportation	$\sigma^2_{\mu\rho}$ σ^2_{δ}	.12 + 0 .02 + .01	.02+.03 .02+.01	.06+.01 .04+.01	.05+.01 .01+.01	.09 + .01 .02 + .01	.18+ 0 .02+.01	.13+.004 .02+.01
Shopping	$\sigma^2_{\mu\rho}$ σ^2_{δ}	.09 + .002 .01 + .01	.09+.001 .01+.01	.06+ 0 .04+.01	.09+.004 .04+.01	.08 + .003 .04 + .01	0 + .004 .08+.004	.09+.004 .04+.01
Leisure	$\sigma^2_{\mu\rho}$ σ^2_{δ}	.03 + .002 .02 + .02	.03+ 0 .01+.03	0 + .01 .08+.01	.02+.01 .03+.01	.07 + .004 .05 + .01	.01+.01 .01+.01	.01+.01 .04+.01
Budgeting	$\sigma^2_{\mu\rho}$ σ^2_{δ}	.08 + 0 .01 + .01	0 + 0 .06+.02	.01+.01 .05+.01	.01+ 0 .002+.01	.02 + .002 .01 + .01	0 + 0 .02+.01	.03+ 0 .02+.01
Cooking, Home Manage.	$\sigma^2_{\mu\rho}$ σ^2_{δ}	0 + .003 .01 + .01	0 + 0 .03+.01	0 + .002 .07+.01	0 + .004 .03+.01	.05 + .01 .003+.01	.01+.0002 .02+.01	0 + .002 .03+.01
Social Maturity	$\sigma^2_{\mu\rho}$ σ^2_{δ}	.04 + .001 .01 + .002	.0 + .001 .03+.004	.01+.001 .01+.001	.05+ 0 .03+.003	.02 + .002 .004+.002	.03+.001 .01+.001	.04+.001 .01+.002
Communication	$\sigma^2_{\mu\rho}$ σ^2_{δ}	.07 + .01 .01 + .01	.21+ 0 .001+.03	0 + .003 .04+.01	.16+.0004 .01+.01	.21 + .01 .01 + .01	.08+.01 .01+.01	.14+.01 .02+.01
Consideration	$\sigma^2_{\mu\rho}$ σ^2_{δ}	.05 + .01 .04 + .01	0 + .01 .32+.01	.13+.002 .03+.01	.06+.001 .03+.01	0 + .003 .01 + .01	.04+.01 .02+.01	.06+.003 .03+.01
Getting Friends	$\sigma^2_{\mu\rho}$ σ^2_{δ}	.04 + .01 .02 + .01	.03+.003 .02+.01	0 + .01 .01+.01	.04+.002 .01+.01	.02 + .01 .02 + .01	.02+.003 .02+.01	.004+.01 .04+.01
Keeping Friends	$\sigma^2_{\mu\rho}$ σ^2_{δ}	.02 + .001 .02 + .01	0 + .01 .01+.01	.11+.01 .02+.01	.01+.004 .03+.01	.01 + .01 .02 + .01	.03+.001 .03+.01	.05+.001 .01+.01
Handling Problems	$\sigma^2_{\mu\rho}$ σ^2_{δ}	.10 + .003 .02 + .01	0 + .004 .10+.01	.002+.002 .02+.01	.10+.01 .01+.01	0 + .002 .01 + .01	.08+.004 .06+.01	.07+.002 .02+.01
Total Test	$\sigma^2_{\mu\rho}$ σ^2_{δ}	.07 + .001 .02 + .01	.02+.001 .05+.01	.02+.001 .02+.01	.02+.01 .03+.01	.06 + .003 .03 + .01	.03+.003 .02+.02	.03+.001 .03+.01

Of the remaining ninety-seven coefficients, twenty-two are greater than or equal to .80. The domain, Personal Routines, accounts for seven of these. Two coefficients are greater than .80 for each of the following: Personal Routines (domain scores), Room Management, and Time Management. The subtest, Appearance and Eating has one coefficient greater than .80. For the domain Community Awareness, the subtests Budgeting, and Cooking and Home Management, and the domain scores each have one coefficient greater than .80. The subtests, Transportation and Shopping have two and three coefficients, respectively, that are greater than or equal to .80. Thus, this domain accounts for an additional eight coefficients. Finally, the domain, Social Maturity, accounts for the remaining seven coefficients that are greater than or equal to .80. Five of these are for the subtest Communication, the largest number for any subtest; Keeping Friends and Handling Problems account for the remaining two. All rater pairs have at least two coefficients that are greater than or equal to .80. For rater pair 1, eight ratings have generalizabilities that exceed this criterion.

There are seventy-five coefficients that are greater than zero, but less than .80. In examining the variance components that contribute to these coefficients, it can be seen that for thirty-eight, the persons-by-judges component contributes most to observed score variance. All domains and subtests, except Transportation, have at least one coefficient where the pj component is the largest or equal to the largest variance component. For the subtest, Keeping Friends, the pj component contributes most to observed score variance for ten coefficients.

The generalizability of total scores by rater pair ranges from .04 for rater pair 2 to .87 for rater pair 4. The p_j component contributes most to the coefficients for rater pairs 2, 3, and 5.

The generalizability coefficients for subtests, across all raters, range from .05 for Cooking and Home Management to .81 for Communication and Transportation. For eight subtests, the p_j component contributes most to the observed score variance. Five of these occur for the domain Personal Routines and four of its five subtests; two are for the subtests Leisure, and Cooking and Home Management, in the second domain; and the last is for the subtest, Getting Friends.

Results of Analysis - Validity

Practical Academics, Language Development, and Motor Development.

The correlation matrices for these traits are presented in Tables 59 and 60. The results suggest evidence for the convergent validity of at least three of these traits. The values in the validity diagonals are all significant except for the trait, Motor Movements for the first analysis sample.

Evidence for the discriminant validity of Language Development is indicated. The validity coefficient for this trait is greater than the correlations among traits where neither trait nor method is common (e.g. $r_{A_2B_1}$; $r_{C_2A_1}$, etc.). However, this value is not greater than all correlations among traits, measured by the same test (e.g. $r_{A_1B_1}$; $r_{B_1C_1}$; $r_{A_2B_2}$, etc.).

The pattern of correlations for different traits measured by different tests, and different traits measured by the same test is the same for the first analysis sample, except for those correlations among traits as measured by

TABLE 59

Intercorrelations of Language Development, Practical Academic Skills, and Motor Movements as measured by the ABS Part I (1) and the set and correlations of these traits with MA

Trait		AFI				ABS (1)				MA
		A ₁	B ₁	C ₁	D ₁	A ₂	B ₂	C ₂	D ₂	
AFI										
Lang. Dev. (Sp. + Wr.)	A ₁	(.99)								.89
Number, Time Con.	B ₁	.92	(.99)							.90
Money Handling	C ₁	.89	.95	(.97)						.83
Motor Movements	D ₁	.46	.44	.43	(.82)					.52
ABS										
Lang. Dev. (Sp. + Wr.)	A ₂	[.70]	.57	.56	.28	(.76)				.58
Number, Time Con.	B ₂	.49	[.49]	.48	.19	.76	(.69)			.53
Money Handling	C ₂	.54	.47	[.49]	.20	.62	.70	(.47)		.53
Motor Movements	D ₂	-.11	-.15	-.18	[.31]	.24	.17	-.08	(.70)	-.15

N=48; $r_{\text{critical}} = .282$ ($p \leq .05$)

TABLE 60

Intercorrelations of Language Development, Practical Academic Skills, and Motor Movements measured by the ABS Part I (2) and the set and the correlations of these traits with MA

Trait		AFI				ABS (2)				MA
		A ₁	B ₁	C ₁	D ₁	A ₂	B ₂	C ₂	D ₂	
AFI										
Language Dev.	A ₁	(.99)								.89
Number, Time Con.	B ₁	.92	(.99)							.90
Money Handling	C ₁	.88	.92	(.97)						.83
Motor Movements	D ₁	.50	.49	.47	(.82)					.52
ABS										
Language Dev.	A ₂	[.68]	.62	.57	.42	(.76)				.57
Number, Time Con.	B ₂	.62	[.53]	.48	.43	.74	(.69)			.50
Money Handling	C ₂	.54	.44	[.46]	.52	.66	.56	(.47)		.31
Motor Movements	D ₂	-.13	-.14	-.09	[.26]	.24	.12	.18	(.70)	.25

¹N=48; r_{critical} = .282 (p ≤ .05)

the AFI. For the second sample, the same pattern occurs among correlations for traits measured by the two tests.

Except for the same pattern of correlations, there is no additional evidence for the discriminant validity of the remaining three traits. The validity coefficients are not greater than correlations among traits measured by the different tests, nor among traits measured by the same test.

Despite the evidence for convergent validity, and support for the discriminant validity of Language Development, an examination of the correlations of these traits with Mental Age (MA) reveals that all correlations are significant except for Motor Movements, as measured by the ABS; and, these coefficients are greater than those in the validity diagonals for all traits measured by the AFI. For the ABS, the correlations between Number and Time Concepts and MA and Money Handling and MA are greater than the validities of these traits, for analysis sample one.

Self-Help Skills. The correlation matrices for these traits are presented in Tables 61 and 62. The results indicate convergent validity of all traits except Cleanliness for the first analysis sample and Leisure for both analysis samples.

There is some evidence for the discriminant validity of the traits, Appearance and Eating, Room Management, Transportation, and Budgeting. The validity coefficients for the first two traits are greater than the correlations among traits where neither trait nor method is common. For the latter two traits, the diagonal values are greater than all but one coefficient, each. However, no validity coefficient is greater than all

TABLE 61

Intercorrelations of Self-Help Skills as measured by the RCL (1) and the ABS Part I (1)
and correlations of these traits with Mental Age

Trait		AFI (1)										ABS (1)										MA
		A ₁	B ₁	C ₁	D ₁	E ₁	F ₁	G ₁	H ₁	I ₁	A ₂	B ₂	C ₂	D ₂	E ₂	F ₂	G ₂	H ₂	I ₂			
AFI																						
Cleanliness	A ₁	(.12)																			-.24	
Appearance, Eat.	B ₁	.65 (.12)																			-.22	
Room Management	C ₁	.61 .40 (.27)																			-.31	
Transportation	D ₁	.15 .33 .09 (.81)																			.44	
Shopping	E ₁	.40 .46 .32 .71 (.68)																			.50	
Leisure	F ₁	.06 .11 -.07 .19 .32 (.27)																			.05	
Budgeting	G ₁	.22 .16 .28 .53 .65 .10 (.48)																			.43	
Cooking, Home Man.	H ₁	.30 .15 .41 .46 .60 .28 .57 (.05)																			.15	
Social Maturity	I ₁	.28 .49 .32 .42 .64 .33 .46 .45 (.71)																			.30	
ABS																						
Cleanliness	A ₂	[.07] .54 .02 .19 .17 .01 .01 -.05 .37 (.56)																			.39	
Appearance, Eat.	B ₂	.07 [.55] -.04 .11 .03 -.03 -.12 -.19 .24 .90 (.45)																			.14	
Room Management	C ₂	.01 .23 [.37] .13 .25 -.04 .25 .30 .27 .46 .37 (.44)																			-.002	
Transportation	D ₂	.23 .48 -.01 [.49] .55 .01 .32 .10 .45 .47 .48 .28 (.68)																			.53	
Shopping	E ₂	-.19 .11 -.19 .31 [.37] .08 .08 .09 .27 .27 .27 .16 .58 (.68)																			.51	
Leisure	F ₂	.24 .51 .07 .25 .25 [.22] .16 .07 .28 .59 .63 .37 .51 .15 (.66)																			-.11	
Budgeting	G ₂	.11 .32 .01 .30 .50 .05 [.39] .14 .43 .43 .36 .39 .68 .53 .41 (.47)																			.53	
Cooking, Home Man.	H ₂	.003 .15 .36 .28 .48 .16 .18 [.50] .52 .30 .18 .68 .28 .43 .20 .35 (.33)																			.18	
Social Maturity	I ₂	-.11 .23 -.13 .19 .19 .22 .05 -.01 [.37] .40 .42 .16 .36 .36 .49 .51 .35 (.50)																			.25	

¹N=41; r_{critical} = .304 (p ≤ .05)

TABLE 62

Intercorrelations of Self-Help as measured by the RCL (2) and the ABS Part I (2)
and correlations of these traits with Mental Age

		AFI										ABS (2)										MA										
		A ₁	B ₁	C ₁	D ₁	E ₁	F ₁	G ₁	H ₁	I ₁	A ₂	B ₂	C ₂	D ₂	E ₂	F ₂	G ₂	H ₂	I ₂													
AFI																																
Cleanliness	A ₁	(.12)																				.25										
Appearance, Eating	B ₁	.77 (.12)																				.25										
Room Management	C ₁	.68	.67	(.27)																		.10										
Transportation	D ₁	-.10	.12	.09	(.81)																	.65										
Shopping	E ₁	.76	.83	.83	.45	(.68)															.86											
Leisure	F ₁	-.09	.08	.13	.75	.41	(.27)														.39											
Budgeting	G ₁	.09	.31	.23	.77	.58	.82	(.48)													.66											
Cooking, Home Man.	H ₁	.15	.27	.33	.49	.47	.56	.53	(.05)													.67										
Social Maturity	I ₁	.24	.51	.38	.38	.68	.67	.75	.48	(.71)													.21									
ABS																																
Cleanliness	A ₂	[.32]	.41	.22	.12	.35	.11	.28	.19	.18	(.56)													-.007								
Appearance, Eating	B ₂	.44	[.47]	.23	-.02	.35	-.03	.14	.04	.15	.91 (.45)													.02								
Room Management	C ₂	.41	.23	[.64]	.11	.49	.07	.18	.06	.20	.30	.29	(.44)													.02						
Transportation	D ₂	.10	.19	.04	[.32]	.28	.52	.52	.16	.34	.36	.41	.08	(.68)													.38					
Shopping	E ₂	.22	.13	.28	.30	[.37]	.46	.40	.31	.31	.10	.14	.23	.60 (.68)													.31					
Leisure	F ₂	.09	.08	.05	-.13	.01	[.02]	.04	-.09	.03	.34	.47	.28	.52	.27 (.66)													.32				
Budgeting	G ₂	.11	.13	.09	.32	.30	.58	[.55]	.21	.33	.24	.24	.07	.75	.74	.35 (.47)													.31			
Cooking, Home Man.	H ₂	.30	.15	.45	.05	.36	.22	.22	[.24]	.12	.22	.23	.53	.28	.62	.16	.41	(.33)													.02	
Social Maturity	I ₂	.02	.11	.05	.19	.20	.39	.50	.33	[.35]	.34	.25	-.02	.35	.26	-.02	.36	.34	(.50)													.17

¹ N=40; $r_{\text{critical}} = .308$ ($p \leq .05$)

correlations among traits as measured by the same tests. In addition, there is no consistent pattern of correlations for different traits measured by different tests, and different traits measured by the same test.

In examining the relationship of these traits to MA, it can be seen that, for both analysis samples, and, as measured by both the AFI and the ABS, the traits, Transportation, Shopping, and Budgeting correlate significantly with MA. In addition, for sample one, Social Maturity as measured by the AFI, and Cleanliness, as measured by the ABS, also correlate significantly with MA. For the second sample, the traits Leisure, as measured by both tests, and Cooking and Home Management as measured by the AFI, correlate significantly with MA. The correlations between Shopping and MA, and Budgeting and MA exceed the validities of these traits as measured by the AFI for both samples and the ABS for sample one. This also occurs for the traits, Transportation as measured by the AFI for sample two, and, the ABS for both samples; Leisure as measured by both tests for sample two; Cooking and Home Management, (AFI) for sample two; and, Cleanliness (ABS) for sample one.

Vocational Skills. The correlation matrices for these traits are presented in Tables 63 and 64. There is evidence for the convergent validity of Independence, for both samples; Self-Expression, for sample one, and, Basic Work Habits and Appearance for the second sample.

Some support is evident for the discriminant validity of the traits, Independence and Self-Expression for the first analysis sample. The validity coefficients for these traits are greater than correlations where neither trait

TABLE 63

Intercorrelations of Vocational Skills as measured by the VCL (1) and the ABS Part I (1)
and correlations of these traits with Mental Age

		AFI						ABS (1)						MA
		A ₁	B ₁	C ₁	D ₁	E ₁	F ₁	A ₂	B ₂	C ₂	D ₂	E ₂	F ₂	
AFI														
Basic Work Habits	A ₁	(.82)												.31
Speed	B ₁	.56	(.54)											.03
Independence	C ₁	.75	.71	(.14)										.27
Self-Expression	D ₁	.26	.41	.51	(.68)									.40
Relations ē Co. W.	E ₁	.31	.54	.53	.73	(.42)								.20
Appearance	F ₁	.50	.65	.62	.46	.61	(.60)							.19
ABS														
Basic Work Habits	A ₂	[.23]	.11	.32	.03	.00	.22	(.33)						.004
Speed	B ₂	.01	[.12]	.09	.08	.15	.06	.17	(0.0)					-.06
Independence	C ₂	.27	.31	[.40]	.18	.21	.36	.83	.27	(.56)				.02
Self-Expression	D ₂	-.27	-.19	-.06	[.39]	.12	.07	.22	.19	.26	(.57)			.42
Relations ē Co. W.	E ₂	.12	.24	.24	.11	[.26]	.33	.66	.23	.88	.25	(.50)		.25
Appearance	F ₂	.23	.36	.37	.06	-.04	[.22]	.52	.11	.57	.22	.47	(.45)	.11

¹
n=42; $r_{\text{critical}} = .307$ ($p \leq .05$)

TABLE 64

Intercorrelations of Vocational Skills as measured by the VCL (2) and the
ABS Part I (2) and correlations of these traits with Mental Age

		AFI (2)						ABS (2)						MA
		A ₁	B ₁	C ₁	D ₁	E ₁	F ₁	A ₂	B ₂	C ₂	D ₂	E ₂	F ₂	
AFI														
Basic Work Habits	A ₁	(.82)												.45
Speed	B ₁	.21	(.54)											.21
Independence	C ₁	.67	.46	(.14)										.14
Self-Expression	D ₁	.33	.06	.33	(.68)									.55
Relations with Coworkers	E ₁	.37	.33	.43	.66	(.42)								.22
Appearance	F ₁	.51	.38	.49	.34	.46	(.60)							.19
ABS														
Basic Work Habits	A ₂	[.49]	.06	.43	.10	.17	.28	(.33)						.01
Speed	B ₂	.37	[.11]	.32	.20	.08	.32	.49	(0.0)					.28
Independence	C ₂	.69	.23	[.53]	.04	.27	.20	.67	.40	(.56)				.13
Self-Expression	D ₂	.37	.28	.29	[.28]	.29	.24	.46	.21	.41	(.57)			.05
Relations with Coworkers	E ₂	.58	.29	.39	-.08	[.05]	.17	.59	.37	.83	.44	(.50)		.17
Appearance	F ₂	.33	.19	.33	-.09	.12	[.33]	.54	.38	.55	.15	.36	(.45)	-.06

¹
N = 42; r_{critical} = .307 (p ≤ .05)

nor method is common. However, neither coefficient is greater than all the correlations among traits measured by the same test. In addition no consistent pattern of correlations among different traits measured by the same test and different traits measured by different tests is apparent.

An examination of the relationship of these traits with MA reveals that Basic Work Habits and Self-Expression, as measured by the AFI, for both samples, and Self-Expression, as measured by ABS for sample one, correlate significantly with MA. In addition these coefficients for Self-Expression are of the same order or greater than the validities of this trait. For the first analysis sample, the correlation between MA and Basic Work Habits, as measured by the AFI, is greater than the validity of this trait.

Because the subtests of the SET show very high positive correlations with MA, a closer examination of the relationship of this test to measured intelligence was made. In Table 65 the correlations of the subtests and total scores of the SET with MA are corrected for attenuation. Estimates of the internal consistency (.93) and the stability (.86) of the Stanford Binet were obtained from Silverstein (1969) and Himmelstein (1968), respectively.

ABS Part II. An attempt was made to assess the validity of the ABS Part II, using, as a criterion, reported incidents of maladaptive behaviors. Almost all reported incidents fell into only five of the thirteen categories of the ABS Part II. The mean scores for the three groups (NIN, LIN, and HIN) on reported incidents are compared with mean scores on the five categories of the ABS Part II in Figures 26 to 30. The ABS means, for four of the five behavior domains, follow the pattern $\bar{x}_H > \bar{x}_L > \bar{x}_N$. For the domain,

TABLE 65

Intercorrelations of the subtest and total test scores of the set with RSB Mental Age corrected for attenuation due to error in estimates of the internal consistencies and stabilities of these tests

Subtest	Correlations with Mental Age		
	Uncorrected	Corrected for Internal Consistency Error	Corrected for Stability Error
Reading	.87	.93	.95
Writing	.83	.87	.90
Numbers	.87	.94	.96
Time	.95	1.00	1.00
Money	.83	.88	.92
Community	.83	.92	.94
Communication	.75	.83	.84
Motor Movements	.52	.67	.63
Concept Attainment	.87	.98	.97
Total Test	.92	.96	1.00

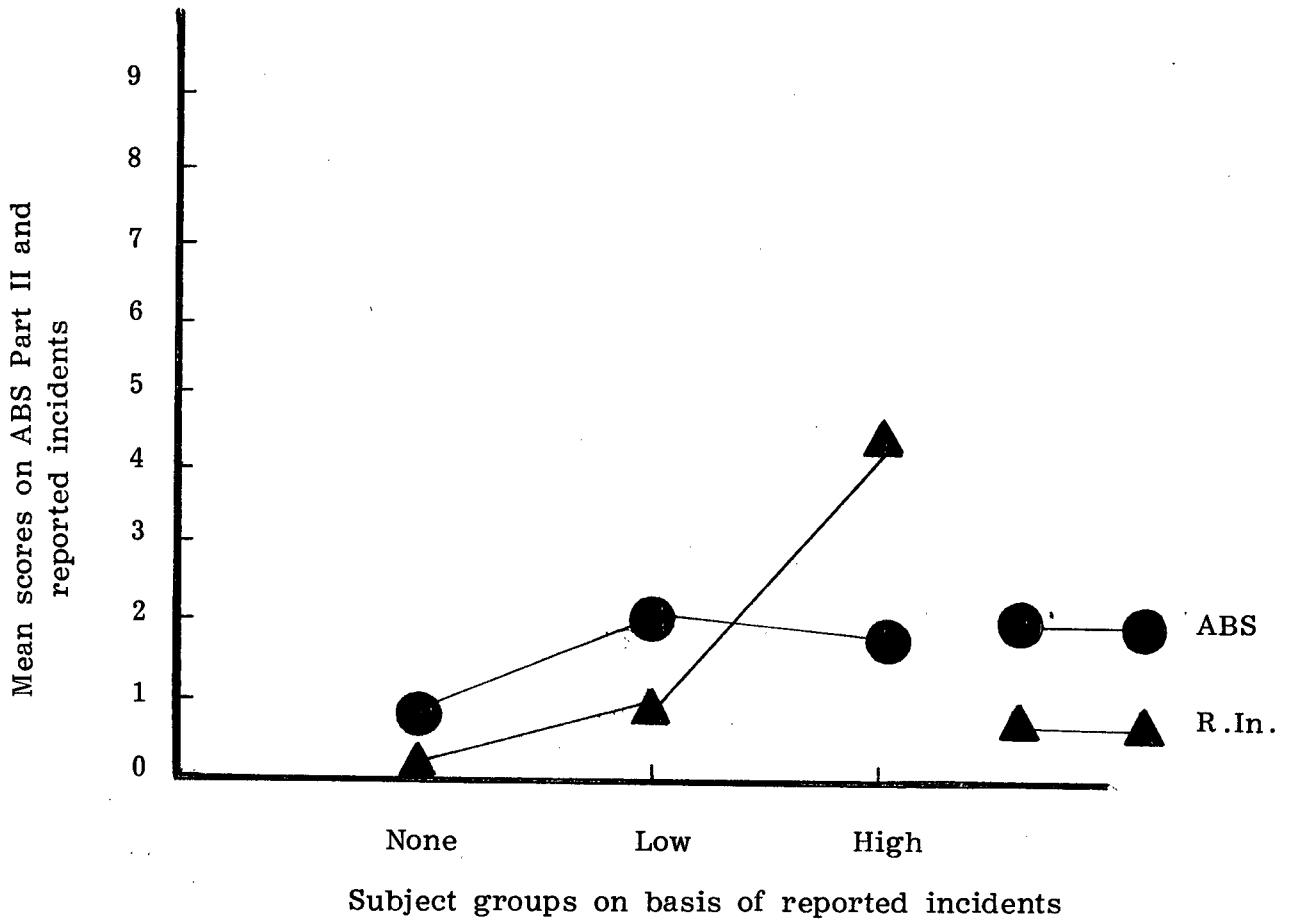


FIGURE 26

Comparison of mean scores on the ABS Part II and on reported incidents of maladaptive behavior for the domain Violent and Destructive Behavior

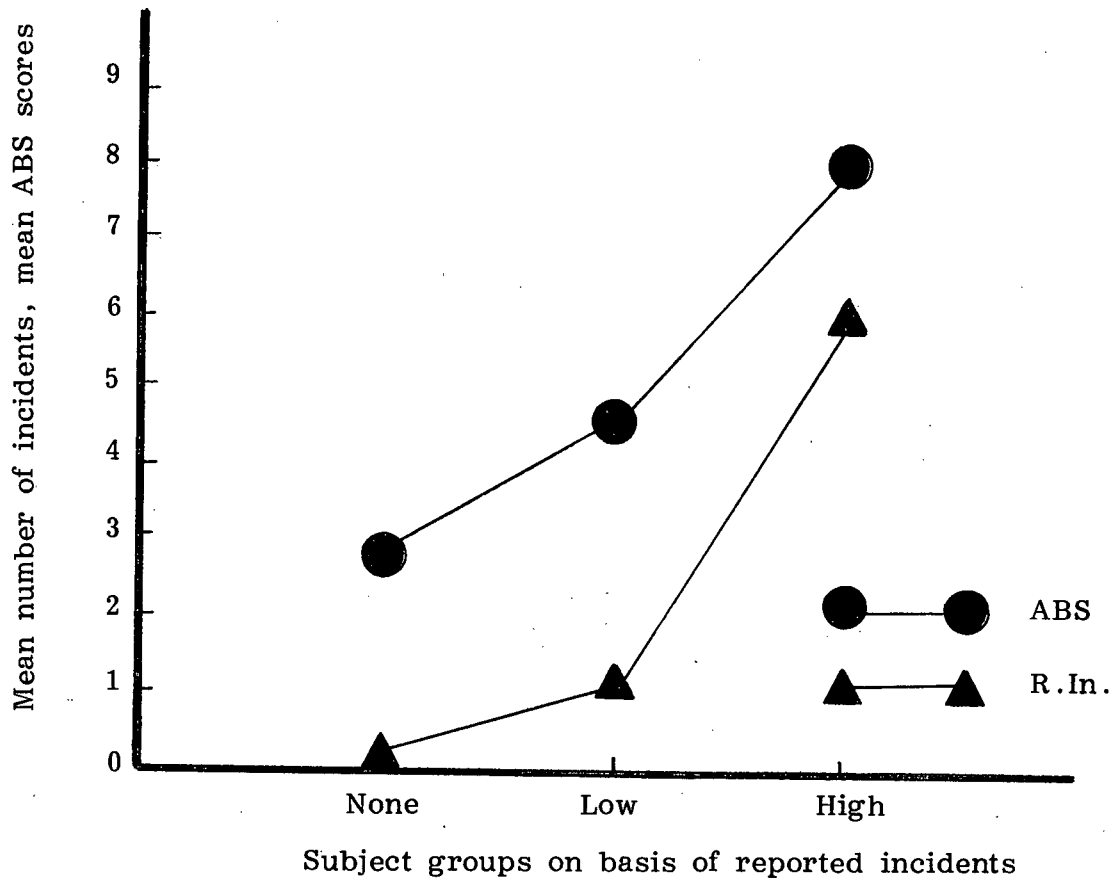


FIGURE 27

Comparison of mean scores on the ABS Part II and on reported incidents of maladaptive behavior for the domain Antisocial Behavior

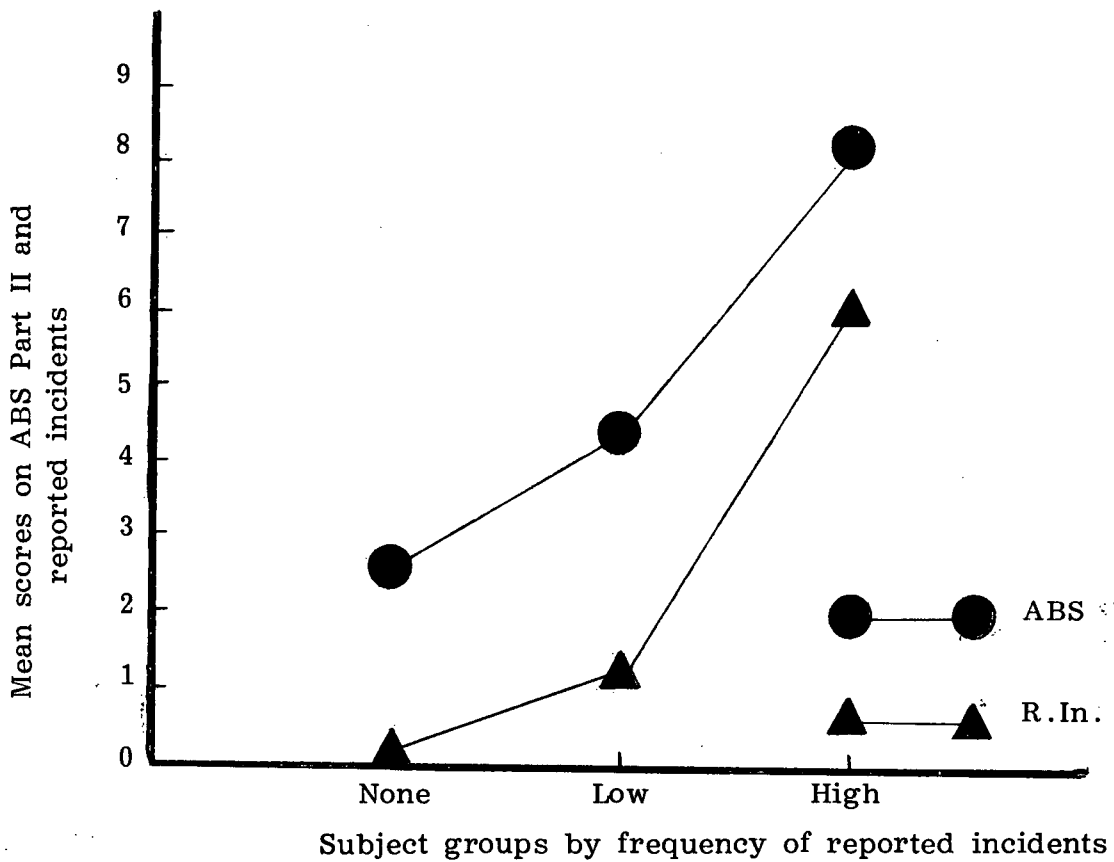


FIGURE 28

Comparison of mean scores on ABS Part II and on reported incidents of maladaptive behavior for the domain Rebellious Behavior

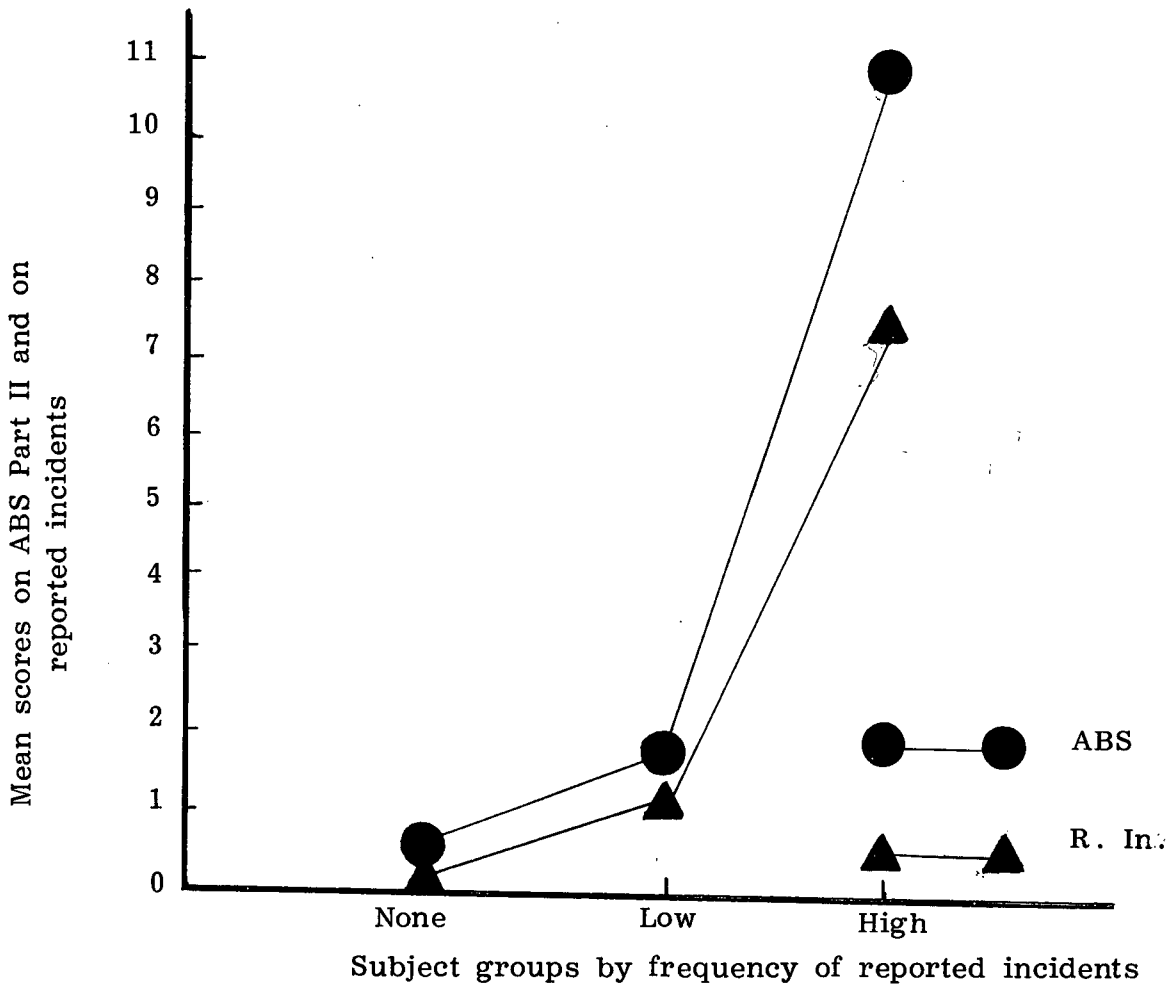


FIGURE 29

Comparison of mean scores on the ABS Part II and on reported incidents of maladaptive behavior for the domain Untrustworthy Behavior

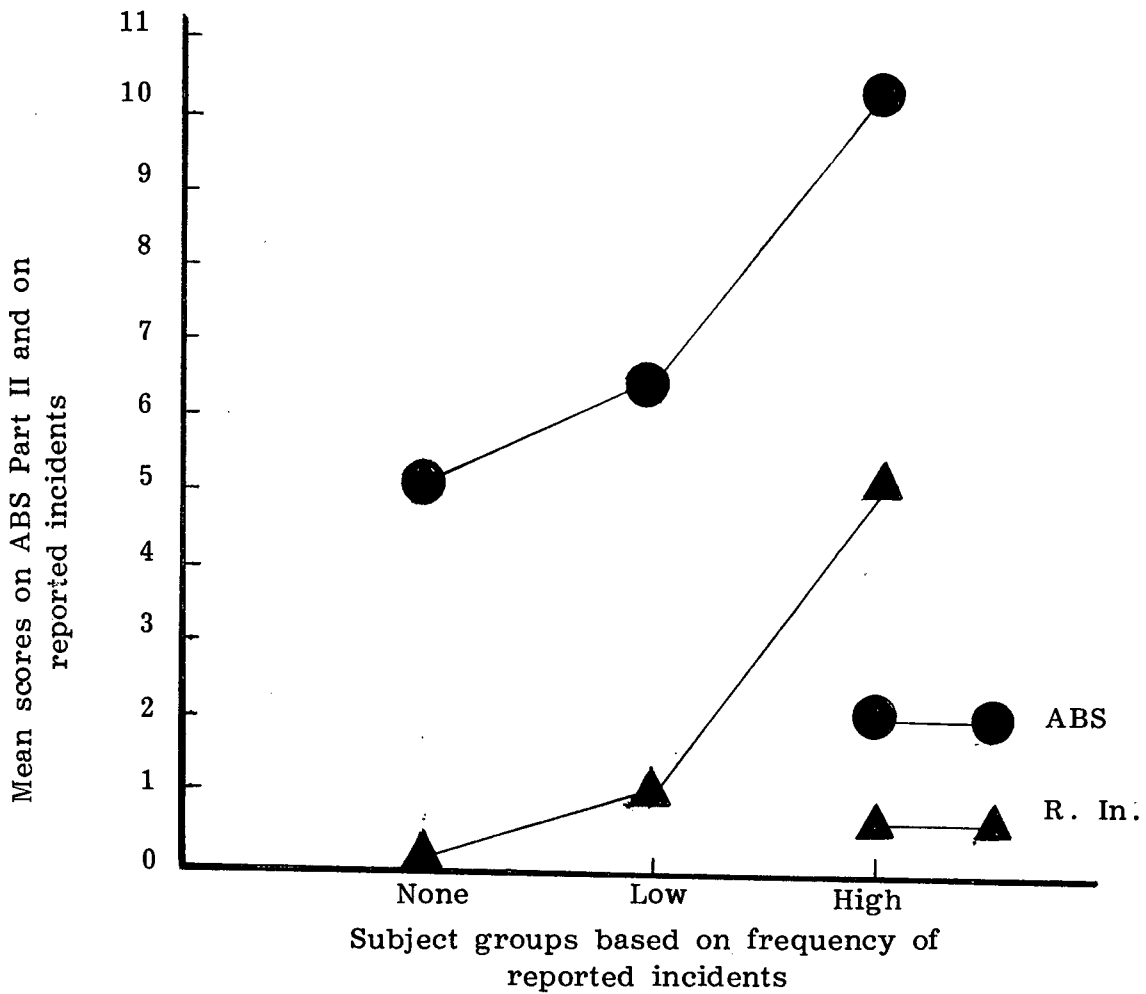


FIGURE 30

Comparison of mean scores on the ABS Part II and on reported incidents of maladaptive behavior for the domain Psychological Disturbances

Violent and Destructive Behavior, the mean scores for the low and high incidents groups are approximately equal.

Chapter V

Discussion

The results of this study, presented in the previous chapter, will be discussed in terms of five purposes of this study: the reliability of the scales; their ability to discriminate among members of the sample; their validity; the relationship of adaptive behavior to measured intelligence; and the usefulness of the scales for program development.

Reliability

Internal Consistency. For both Parts I and II of the ABS and for the tests of the AFI, the internal consistency estimates that were obtained in this study are sufficiently high and comparable to those for standardized tests such as the Stanford-Binet Intelligence Scale. However, for the rating scales in particular (ABS, VCL, and RCL), the Hoyt coefficients for some subtests are low ($r_{xx} \leq .85$). This is, in part, due to the length of these subtests. For the ABS Part I the number of items ranges from two for Responsibilities to twenty-two for Independent Functioning. With the exception of the subtests Economic Activity for both samples and Language Development for the first sample, the subtests with reliabilities less than .80 contain seven or fewer items. Similarly, the subtests of the ABS Part II contain from one to seven items. The highest coefficient of internal consistency for both analysis samples is for the longest subtest, Psychological Disturbances. For the VCL the coefficients for most of the five item subtests are less than .85. However, the internal consistencies of domain scores (Basic Work Habits, Work Skills, and Acceptance Skills) are sufficiently high ($r_{xx} \geq .87$). Similarly, the coefficients for the domain scores of the RCL (Personal Routines, Community Awareness, and Social Maturity) are greater than .90 for both samples, although those for some of the ten item subtests are less than .85.

The observed differences among the two samples are probably due to two factors

1. although the majority of subjects in both samples were the same there were some differences that might account for the obtained

results. These differences arose when only one set of ratings was available for a given group of subjects. Thus each sample included only one half of the subjects in these groups.

2. Differences among the obtained results for both samples may be due to rater bias. That is, raters may have tended to assign different scores to the same subjects.

Stability. The generalizability of SET scores over occasions (within a three month interval) is high for all subtests except Motor Movements. The traits measured by this subtest appear to be less stable than those measured by other subtests. Changes in persons' overall scores over occasions (po component) and changes in persons' scores on specific items over occasions (poi component) seem to contribute equally to the observed score variance.

Interrater Reliability. Estimates of the generalizability of ratings indicate great variation in the consistency of judgments for different rater pairs as well as for different subtests. Some inconsistencies were found to be the result of a particular type of rater bias: there was a tendency for different raters to assign different scores to the same people. As a result, the persons-by-judges interaction (pj component) contributed most to observed score variance. The presence of this bias is probably due to characteristics of the scales as well as a result of differences in the opportunities available to raters for observing particular behaviors.

At least three sources of ambiguity appear in the items of the ABS, the VCL, and the RCL. First, some items do not specify the context in

which behaviors are to be observed. The items from the subtest Responsibilities from the ABS Part I suffer from this deficiency. Secondly, some items do not indicate the procedures for measuring levels of achievement. The subtests of the domains Basic Work Habits and Work Skills from the VCL, Personal Routines from the RCL, and the subtest Number and Time Concepts from the ABS Part I reveal this shortcoming. Finally, standards for evaluating performance are not provided for all items. Items from the subtests Occupation Domestic and Occupation General of the ABS Part I do not provide specific standards of performance to be used in assigning scores.

The opportunities for raters to observe the behaviors measured by these scales may have differed. This difference may be due to the nature of certain behaviors. For example, the subtests from the domains Social Maturity of the RCL and Acceptance Skills of the VCL, and the subtest Social Maturity from the ABS Part I measure behaviors that subjects may display in some situations but not others and in the presence of some raters but not others. Differences in opportunities for observation may also be due to the structure of the facility. This is especially likely for programs where only one staff member normally worked with the subjects. It may also account for the inconsistencies in the ratings on the VCL and the occupational subtests of the ABS Part I for the Activity groups. At the time of this study, the majority of the subjects in these groups were involved in work activities only on a weekly basis. The staff members who worked with them were not equally involved in these work programs.

Finally, raters appear to have differed in the extent to which they used the opportunities for observation that were available to them. The inconsistencies in ratings suggest that many raters relied on what they remembered about a given individual's behavior. Objective evidence of this is found in the ratings for the Work Skills domain of the VCL. The differences in rater judgments suggest that standards for assessing performance such as the supervisor's rate of task completion, was not systematically determined and then applied.

The rater bias (the persons-by-judges interaction or pj component) that resulted from ambiguities in the test items and from differences in opportunities for observation affected all scales. The results for the ABS Part I indicate that the pj component contributed most to the observed score variance for at least one subtest for each pair of raters. The results over rater pairs indicate that this bias occurred in all subtests for at least one pair of raters. For the ABS Part II this effect was observed for all rater pairs and for eleven of the fourteen subtests. For the VCL, this occurred for twelve of the fourteen subtests and for five of the six rater pairs. And, for the RCL, this occurred for sixteen of the eighteen subtests and for all rater pairs. In terms of the proportion of coefficients reflecting this bias, the pj component accounts for the largest portion of observed score variance in one third of the coefficients for the VCL and the RCL; whereas, for the ABS Parts I and II, this effect was observed in only one quarter of the coefficients. The greater specificity of the VCL and the RCL items may account for this difference.

Although only one type of rater bias can be demonstrated quantitatively in this study, incidental evidence of other types of bias is available. The experimenter's review of ratings suggests that different raters probably do attend to different behaviors. This selective perception can be demonstrated by the case of one male subject who was seen engaged in self-stimulating behavior (hand waving, rocking) on several occasions. One rater indicated that this man never displayed these behaviors. The second rater indicated that these behaviors occurred frequently (on ABS Part II).

In addition to coefficients of generalizability that were found to be low as a result of rater bias, there were also coefficients for some ratings that were equal to zero. The lack of universe score variance in these ratings may reflect valid homogeneity in the samples, particularly since program samples are probably based on ability and disability groupings. This explanation appears feasible for subtests measuring skills that are easy for all subjects, for those that are difficult for all subjects, and for those measuring maladaptive behaviors (ABS Part II and the VCL Difficulty subtests), where the majority of subjects were rated as not displaying those behaviors. However, for other subtests such as Responsibilities from the ABS Part I, the zero coefficients may be the result of a tendency for a given rater to assign similar scores to all subjects, hence restricting universe score variance.

The results of the generalizability analysis are not unique to this study since the conditions described that produced low coefficients of inter-rater reliability are typical of many settings where these scales might be

used. In fact the reliability coefficients obtained in this study are comparable to those cited by the authors of the ABS. Their analysis, however, did not isolate the particular types of error that affect the measurement of adaptive behavior. The problem of improving the reliability of ratings for future users of these scales appears to depend on the elimination of ambiguities from the scales themselves and on systematic training and monitoring of raters.

Discrimination

Rating Scales. The results of the item analyses provide some evidence regarding the ability of these tests to discriminate among members of this sample. For the ABS Part I, fifty-seven and fifty-nine percent of the items discriminate among members of the first and second analysis samples respectively. Eighty-two percent of the VCL items discriminate among members of sample one; seventy-two percent, among members of sample two. For the RCL, fifty-six and sixty-three percent of the items discriminate among members of the first and second samples, respectively.

At least three types of inefficient items appear to be present in the ABS Part I, the VCL, and the RCL:

1. those representing minimal skills or capacities displayed by the majority of subjects;
2. those representing skills or capacities that are influenced by institutional regulations and/or other environmental restrictions;
3. ambiguous items which are either non-specific with respect to context, or which measure behaviors that are difficult to assess, and which therefore may solicit raters' opinions.

Included in the first category from the ABS Part I are those non-discriminating items from the subtests Independent Functioning, Physical

Development, Language Development, and the item Table Clearing from the subtest Occupation Domestic. These items are judged by raters to be very easy for this sample. The items from the first subtest represent basic self-care skills that may be viewed as minimal criteria for admission to a facility such as Bevan Lodge, which is not equipped to provide intensive personal care. Although the non-discriminating items in the subtests Physical Development and Language Development also seem to represent basic capacities displayed by all subjects, this is in part due to the restriction of range in sampling. Subjects with profound physical or sensory handicaps were excluded from this study. Finally the item Table Clearing seems to represent a skill that all subjects have acquired and displayed.

No items from the VCL appear to belong in this category. Only two items from the RCL appear to be very easy for this sample. These are:

- a. smiles and greets people he knows
- b. aware of the differences between men and women.

The second category of non-discriminating items, those measuring behaviors that are affected by institutional regulations and/or other environmental restrictions, includes those from the subtests Occupation Domestic and Occupation General from the ABS Part I. Since the tasks measured by the former are carried out by members of the Laundry and Housekeeping program, the majority of subjects have no opportunity to display these skills. Similarly, the item Work Fitness from the latter subtest measures skills that few subjects have the opportunity to display in the types of

work programs in which they are engaged. Although the items in this category tend to be difficult for this sample, several subjects appear to have received partial marks for these items. This is probably because raters were influenced by the direction:

If you have never actually seen the person perform the act...
but are convinced that he could perform it if given the opportunity
answer that he can perform the act.

(Nihira et al., 1969)

Several items from the VCL are included in the second category of non-discriminating items. The item measuring Quality Control Skills from the subtest Making Decisions is judged to be very difficult for this sample. In fact very few subjects have the opportunity to do quality control work since it is done on a regular and systematic basis by the staff in most work programs. The non-discriminating items from the subtest Speed are also judged to be difficult for this sample. However, these items may be inefficient because raters do not appear to have systematically determined and then applied the suggested standard of performance (the supervisor's rate of task completion). The non-discriminating items from the subtest Punctuality clearly belong in this second category of inefficient items. All subjects took coffee and lunch breaks in their work areas and thus had little opportunity to display independence in this area. In addition, many subjects were not informed of appointments until they were required for one.

From the RCL the following items appear to be affected by institutional regulations or other environmental restrictions: those from the subtests Room Management, Time Management, Health (items regarding the handling

of medications), Transportation, Budgeting, and Cooking and Home Management. These subtests measure the ability to perform tasks that are carried out by the staff (Health and Budgeting) or a program group (Room Management and Cooking and Home Management), or which residents have no opportunity to acquire or display in the relatively isolated environment of Bevan Lodge (Transportation and Time Management).

Several items from the ABS Part I are included in the third category of inefficient items. At least one item Perseverance from the subtest Self-Direction is ambiguous because it does not specify the context in which the behaviors are to be observed. The item Interaction with Others from the subtest Socialization seems inappropriate for the target population since it is couched in terms of children's behavior. The remaining non-discriminating items from these subtests and from the subtest Responsibilities appear ambiguous because they seem to be measuring behaviors that are more difficult to assess in an objective way than the skill-oriented items of other subtests. They also measure behaviors that subjects may display in the presence of one rater but not another, and in some situations but not others.

From the VCL, the non-discriminating items from the subtests Taking Direction, Following Instructions, Self-Expression, and Relations with Co-workers are included in the third category of inefficient items. The ambiguities in these items may reflect differences in behaviors that raters attend to as well as differences in the behaviors displayed in the presence of different raters.

Finally, the non-discriminating items from the subtests Cleanliness, Appearance and Eating, Health (those measuring the extent of sex education),

Transportation (those referring to behavior rather than ability to use public transport), Leisure, Communication, Consideration, Getting Friends, Keeping Friends, and Handling Problems from the RCL are included in the third category. Although many of these items are skill-oriented, they do not lend themselves to easy assessment. This is partly due to the high order of specificity of information required as well as variations in the availability of regular opportunities for observation.

In summary, the ABS Part I contains a number of behavioral domains that measure minimal skills. Although these are not useful in discriminating among members of this sample, they have been found to be important dimensions in the assessment of the more severely handicapped (Nihira et al., 1969). The AFI, however, contains many domains that are useful for discriminating among individuals with higher levels of functioning. Subtests such as Budgeting and Time Management are difficult for this sample, but they are appropriate for individuals in more independent situations (Nihira et al., 1969)

Social Education Test. Seventy-three percent of the items in this test discriminate among members of this sample. The majority of non-discriminating items from this test appear to be too easy or too difficult for this sample. These include the items from the subtests Reading, Numbers, Time, Money, and Communication.

The non-discriminating items from the subtest Community Awareness appear to measure knowledge that subjects in a relatively isolated environment would have little opportunity to acquire.

Very few items from the subtest Motor Movements discriminate among members of this sample. Some are very easy whereas others are very difficult for this sample. This is partly due to the restriction in range of abilities displayed by the subjects as well as the nature of the abilities measured. Although the information provided by this subtest is important it may be best regarded as supplementary.

The subtest Concept Attainment appears to be very useful for assessing basic concepts especially for those subjects who are unable to complete any items from the academic subtests. The Discrimination and Directional Concept sections appear to be very useful. However, performance on the task of sorting screws is difficult to interpret. The dimensions of color, length, and shape are confounded in the test materials. In addition, many subjects appeared to have difficulty understanding the instruction "Put the screws that are the same together". These individuals were unable to sort any of the screws correctly. However, after the first demonstration in the training session, they were able to complete the task without error. Similarly, the instructions for the first Amount Concept item appeared to confuse many examinees. This may account for the ambiguity in this item. Finally, the last Conservation item may be non-discriminating because of the tendency to select the last option as the correct one. In this case the last one is correct.

Measures of Maladaptive Behavior. For the ABS Part II, fifty-nine and fifty percent of the items discriminate among members of analysis samples one and two respectively. For most of the non-discriminating items, the majority of subjects were rated as not displaying these behaviors. For

the few subjects rated as displaying these problems, low scores were assigned. In addition, the difference between the categories of occasionally and frequently occurring behaviors may not be meaningful. The numerical estimates provided for these two classes of behavior overlapped for more than one half of the items. This indicates that different raters use different criteria in assigning behaviors to the occasional and frequent categories of occurrence.

For the Difficulty subtests of the VCL, forty-five and sixty-eight percent of the items appear to discriminate among members of the first and second analysis samples, respectively. An examination of the p (difficulty) values for the non-discriminating items indicates that very few subjects are considered to display these difficulties to a severe extent. This does not necessarily mean that subjects have mastered the skill behaviors in these domains. It is possible that they display other difficulties not assessed by these tests. For example, an individual may not be able to work above 80% industrial standard (Speed subtest) but his coordination may be adequate, he may be strong enough, he may be quite able to work with small objects and he may have work rhythm, in terms of a good sequence for completing a task (Difficulties). His difficulty may be a tendency to repeat a quality control step at the end of a task. It is also conceivable for a person to have mastered the skill behaviors and yet display difficulties. For example, he may have mastered the goals cited in the subtest Independence, and still tend to stop work every hour when he is tired.

Validity

The results presented in the multi-trait multi-method framework indicate support for the convergent validity of sixteen traits. However, evidence for discriminant validity emerged for only seven of these. Although these results suggest a predominance of method variance, this conclusion must be viewed cautiously and assessed within the following constraints.

First, the traits within each matrix are correlated. As a results, the correlations among different traits measured by different tests are elevated. In addition, the validity coefficients are attenuated because of the low interrater reliabilities of the rating scales. The combination of these two factors exaggerates the effect of method variance. A third consideration is the apparent restriction in range for minimal skills. Since this restriction does not occur for all traits, nor for minimal skills as measured by all tests, the results cannot be interpreted confidently. The fourth constraint is that the two methods of assessing the Self-Help skills and Vocational skills are correlated since the same raters were involved. Thus the results reflect this dependence in method. Finally, there may be incomplete matching of traits as measured by both tests. This is of particular concern for the Vocational skills matrix since many items from the ABS were not specific with respect to the work context although they measured relevant behaviors.

Some evidence was presented for the practical validity of five domains of the ABS Part II: Violent and Destructive Behavior, Antisocial Behavior, Rebellious Behavior, Untrustworthy Behavior, and Psychological Disturbances.

These domains were the most frequently used by raters in their assessments of the residents. In addition, approximately ninety-two percent of all reported incidents of maladaptive behavior were categorized in these domains. These types of behavior may occur in response to the frustrations inherent in an institutional environment. They would not necessarily occur in a more normal situation. Thus this dimension of maladaptive behavior may not be an appropriate one for diagnosis and classification of the mentally retarded.

The fact that these maladaptive behaviors do occur is evident. However, a more appropriate approach to their assessment might be one based on operant techniques where antecedent events as well as contingencies of reinforcement are identified.

Adaptive Behavior and Measured Intelligence

The relationship between adaptive behavior and measured intelligence was also examined in this study. The correlations of SET subtest and total test scores with RSB Mental Age are the most striking. These relationships are stronger than the ones reported by other studies. This may be due to a restriction in the range of abilities displayed by subjects in those studies. Other aspects of adaptive behavior as measured by both scales also show significant correlations with Mental Age. In addition, the non-significant correlations may be attenuated because of the low (interrater) reliabilities of the scales as well as by the restriction in range that occurred for some domains.

In general, it would appear that both the measures of adaptive behavior examined in this study and intelligence tests measure the same abilities.

Although assessment is presumably more reliable for intelligence tests, the information provided by them is not useful for programming purposes. From the practical point of view, then, improved measures of adaptive behavior are necessary. From the theoretical point of view, however, adaptive behavior and intelligence as measured by IQ tests do not appear to be distinguishable constructs in this sample of retarded adults.

Program Development

In terms of format and content, the ABS appears to be a more appropriate measure for individuals in institutional environments; whereas, the AFI is more suitable for community based facilities. First, the ABS includes all skill areas in one test, making assessment by one rater possible. The AFI, however, is divided into Vocational, Residential, and Academic tests, for situations where these programs are separate. Secondly, the ABS contains a number of subtests measuring minimal self-care skills that have been found to be important dimensions in the assessment of profoundly handicapped individuals, who are usually found in large institutional settings. The AFI, on the other hand, has large sections devoted to the measurement of behaviors such as vocational abilities and skills demonstrating community awareness that are characteristic of individuals with mild handicaps living in more independent situations.

The AFI provides a more reliable assessment of practical academic skills. The SET contains more items than the equivalent domains of the ABS. Assessment using the SET is more objective with standards for administration and criteria for scoring correct and incorrect responses.

Although the academic subtests of the ABS (e.g. Economic Activity, Number and Time Concepts) appear to discriminate among members of the sample, this may be the result of rater bias rather than valid individual differences. The profile of SET scores also reveals a given individual's weak and strong areas, and therefore suggest appropriate areas for programming. However, it does not preclude the administration of a more detailed task sequence for a particular academic area (e.g. Reading) to determine the individual's present level of success in that area.

The ABS Part I poses several difficulties for program development for self-help, vocational, and academic skills. First, the subtests are of different lengths. As a result, the profile reveals very little about the relative strengths and weaknesses of a given individual. Even on an item basis difficulties are apparent. A given item may contain several sections. A score, unless it is the highest or lowest possible may represent many different combinations of abilities and/or disabilities which contribute to that item. In addition items are on different scales. Thus performance on one item cannot always be directly compared to performance on another.

In contrast, all subtests of the SET, the VCL, and the RCL contain equal numbers of items. Thus within a given area, academic, vocational or residential, it is possible to determine relative weakness and strength on the basis of subtest scores. Within a given test, all items use the same scale. For the VCL and the RCL, one category of response is available for individuals who are learning a skill or behavior but who have not yet mastered it, a category not provided by the ABS.

Programming for maladaptive behaviors on the basis of the ABS Part II poses several problems. The subtests of this scale are of unequal length; and items use different scales. For example, individuals can obtain scores of 0-10 for the item "is withdrawn or shy". For the item "seems to feel persecuted", scores range from 0-14. Thus this pinpointing of target behaviors is difficult.

All but one of the VCL Difficulty subtests are of equal length. In addition, all items use the same scale. However, the domains consist of different numbers of subtests so that profiles of domain scores are not useful for determining relative strengths and weaknesses. The shortness of these tests also prohibits assessment of all but the most common behavior problems displayed in a vocational setting.

Both scales measuring maladaptive behaviors, while providing useful guidelines in assessing these behaviors, may be more appropriately replaced by operant techniques that permit definition of specific target behaviors as well as their antecedent and reinforcing events.

In summary, the structure and format of these scales suggest that the ABS is more appropriate for program planning in institutional settings, whereas the AFI is useful for community based facilities. Several shortcomings of the ABS were described which do not occur in the AFI rating scales. However, the usefulness of the ABS and the VCL and RCL is limited by their low (interrater) reliabilities. Standardized procedures for observation and specific criteria for assessment may improve the quality of results obtained for these scales.

Limitations

The major limitation of this study is the result of the sampling. Since subjects and raters could not be sampled randomly, the generalizability of the results of this study is limited. The relatively large number of raters that were required because of the structure of Bevan Lodge meant that the samples for the generalizability analyses were very small. Comparison of individual rater pairs across tests and of raters for one test were not possible because of changes in raters and the lack of control for rater variables. It was not possible, within the scope of this study, to assess the stability of the rating scales, norms for different traits, or the predictive validity of these tests.

Summary

In terms of the purposes of this study the following results were obtained.

1. Reliability

- A. Internal Consistency. Estimates of the internal consistency of the ABS and the AFI were sufficiently high and comparable to those of standardized tests such as the Stanford-Binet.
- B. Stability. The stability of the SET was also adequate although the subtest Motor Movements appears to measure less stable traits than the other subtests.
- C. Interrater Reliability. The results indicate great variation among rater pairs and across raters in the extent to which they agree in their judgments. Some of these observed inconsistencies appear to be the result of rater bias (persons-by-judges interaction). The presence of this bias is probably due to ambiguities in the scales as well as differences in the opportunities available to raters for observing particular behaviors. The problem of improving the reliability of ratings appears to depend on the elimination of ambiguities from the scales themselves and on systematic training and monitoring of raters.

2. Discrimination

Although many items from the ABS and the AFI were found to be useful in discriminating among members of this sample,

both tests contained ambiguous items that require improvement as well as items that are too easy or too difficult for this sample. A further examination of the results also indicated that the ABS would be more useful with severely handicapped individuals, whereas the AFI seems more appropriate for individuals functioning at a higher level and in more independent situations.

3. Validity

Evidence for the convergent validity of sixteen of the nineteen traits was provided. However, evidence of discriminant validity emerged for only seven of these. The practical validity of the ABS Part II was also assessed. The evidence supports the validity of each of five domains examined.

4. Adaptive Behavior and Measured Intelligence

Many aspects of adaptive behavior as measured by the ABS and the AFI were found to be significantly correlated with RSB Mental Age. The relationship of the SET and measured intelligence was most striking. Adaptive behavior scales are necessary because of the information they provide for programming. However the hypothesis that adaptive behavior is distinct from intellectual functioning was not supported by the results of this study.

5. Usefulness for Program Development

The ABS appears to be appropriate for program development in institutional settings, whereas the AFI is useful for community based facilities. Several shortcomings of the ABS were described which do not occur in the AFI rating scales. However, the usefulness of the ABS, the VCL, and the RCL is limited by their low interrater reliabilities.

References

- Adams, J. Adaptive behavior and measured intelligence in the classification of mental retardation. American Journal of Mental Deficiency, 1973, 78 (1), 77-81.
- Balthazar, E.E. and G.E. English. A system for the social classification of the more severely mentally retarded. American Journal of Mental Deficiency, 1969, 74 (3), 361-366.
- Beerman, L., Bellamy, T., Di Rocco, P., Friedland, M., Foss, G., and E. Steinbock. Civil rights of the mentally retarded. Working paper No. 69, Eugene: University of Oregon, Rehabilitation Research and Training Centre in Mental Retardation, May 1973.
- Benton, A. Psychological evaluation and differential diagnosis in mental retardation. In Mental Retardation, Stevens, H. and R. Heber (ed.). Chicago: University of Chicago Press, 1964.
- Campbell, D.T. and D.W. Fiske. Convergent and discriminant validation by multitrait-multimethod matrix. Psychological Bulletin, 1959, 56 (2), 81-105.
- Cornfield, J. and J. Tukey. Average values of mean squares in factorials. Annals of Mathematical Statistics, 1956, 27, 907-949.
- Cronbach, L.J., Rajaratnam, N. and G.C. Gleser. Theory of generalizability: A liberalization of reliability theory. The British Journal of Statistical Psychology, 1963, 16 (2), 137-162.
- Cronbach, L.J., Schönemann, P. and D. McKie. Alpha coefficients for stratified parallel tests. Educational and Psychological Measurement, 1965, 25 (2), 291-312.
- Cronbach, L.J., Gleser, G.C., Nanda, H., and N. Rajaratnam. The Dependability of Behavioral Measurements: Theory of Generalizability for Scores and Profiles. New York: John Wiley & Sons, Inc., 1972.
- Gardner, J.M. and F.L. Giampa. Utility of three behavioral indices for studying severely and profoundly retarded children. American Journal of Mental Deficiency, 1971, 76 (3), 352-356.
- Gleser, G.C., Cronbach, L.J. and N. Rajaratnam. Generalizability of scores influenced by multiple sources of variance. Psychometrika, 1965, 30 (4), 395-418.

- Grossman, H.J. Manual on Terminology and Classification in Mental Retardation. American Association on Mental Deficiency, Special Publication Series #2, 1973 (revised).
- Heber, R. A Manual on Terminology and Classification in Mental Retardation. (Monograph Supplement), American Journal of Mental Deficiency, 1959.
- Heber, R. A Manual on Terminology and Classification in Mental Retardation. (Monograph Supplement), American Journal of Mental Deficiency, 1961 (2nd edition).
- Himmelstein, P. Use of the Stanford-Binet, form LM, with retardates: A review of recent research. American Journal of Mental Deficiency, 1968, 72 (2), 69-76.
- Hurst, J.G. The meaning and use of different scores obtained between the performance on the Stanford-Binet Intelligence Scale and the Vineland Social Maturity Scale. Journal of Clinical Psychology, 1962, 18, 153-160.
- Leland, H., Shelhaas, M., Nehera, K., and R. Foster. Adaptive behavior: A new dimension in the classification of the mentally retarded. Mental Retardation Abstracts, 1967, 4 (3), 359-387.
- Leland, H. The relationship between intelligence and mental retardation. American Journal of Mental Deficiency, 1969, 73 (1), 533-536.
- Marlett, N.J. Adaptive Functioning Index. Calgary: Vocational Rehabilitation and Research Institute, 1973.
- Marlett, N.J. Adaptive Functioning Index: Manual #2, Training. Calgary: Vocational Rehabilitation and Research Institute, 1974.
- Nunnally, J.C. Psychometric Theory. New York: McGraw Hill Book Co., 1967.
- Rajaratnam, N., Cronbach, L.J. and G.C. Gleser. Generalizability of stratified parallel tests. Psychometrika, 1965, 30 (1), 39-56.
- Robinson, H.M. and N.M. Robinson. The Mentally Retarded Child. New York: McGraw Hill Book Co., 1965.
- Terman, L.M. and M.A. Merrell. Stanford-Binet Intelligence Scale. Boston: Houghton Mifflin Co., 1960.
- Wolfensberger, W. The Principle of Normalization in Human Services. Toronto: National Institute on Mental Retardation through Leonard Crainford, 1972.

Appendix A

Adaptive Behavior is a term that refers to the effectiveness of the individual in coping with the natural and social demands of his environment. Since 1959, the American Association on Mental Deficiency has recommended that Adaptive Functioning be included with intelligence as a major dimension in identifying and classifying the mentally retarded.

The purpose of my study is to investigate the usefulness of two new measures of Adaptive Behavior: The AAMD Adaptive Behavior Scales and the VRRI Adaptive Functioning Index.

The AAMD Adaptive Behavior Scale consists of two parts. The first is designed to assess the individual's skills and habits in ten behavior domains considered important to the maintenance of personal independence in daily living. The second part is designed to provide measures of maladaptive behavior related to personality and behavior disorders.

The VRRI Adaptive Functioning Index consists of three parts. The first, The Social Education Test is a test of practical academic skills and related concept development. The second and third parts are behavior check lists designed to assess vocational performance and residential living skills.

I will need your assistance to complete this study. You will be asked to rate selected individuals from your group on the Adaptive Behavior Scale and on the two behavior check lists of the Adaptive Functioning Index.

I hope that this study will not only provide useful information about the scales, but will also help you focus on the needs of individuals in your group.

TO:

FROM: S. Mongrain

Re: Adaptive Behavior Scales

Practice: _____

Residents to be assessed:

Instructions

1. Please complete the practice assessment and return it to me before beginning the other assessments.
- * 2. On the answer sheet cover, please complete the following items:
 - a) Name of the Person being rated
 - b) Sex of the Person being rated
 - c) Rater's name
 - d) Date of A.B. Assessment
 - e) How long have you known this person?
3. There are three types of items:
 - a) For the first type, select the one statement that applies to the individual you are rating. The number of this statement is the individual's score on that item.
 - b) For this second type of item, indicate whether or not a behavior occurs. Select all statements that apply to the individual. Space is provided for additional examples in the "Other" category. The individual's score is equal to the number of statements selected. e.g., If a, c, and f, are selected, the individual's score on that item is 3. Enter this on the line in the upper right corner of each item.
 - c) This type of item is similar to the second type, except that you must also estimate the frequency of the behavior. Check "Occasionally" for behaviors that occur once in a while or now

and then. Check "Frequently" for behaviors that occur quite often or habitually. Give 1 point for each "occasionally". Give 2 points for each "frequently".

- * In addition, please estimate the frequency of each statement you select numerically. e.g. 1/wk; 2/da; 5/mo etc. enter these estimates in the margin next to each statement you select.
- 4. If you are not sure how to rate an individual on a particular item, try to observe that person in the appropriate situation.
- 5. If you are not sure how to rate an individual on a particular item and if you do not have the opportunity to observe that behavior, but are convinced that if given the opportunity he/she could perform if answer that he/she can perform the act.
- *6. DO NOT LEAVE BLANKS
- *7. Do not consult with any other staff member when you are completing your assessments. These must be done independently or the results will not be valid.

Special Instructions

- Part I
- Item 8 If (f) is circles, give 7 points for Item 8.
 - Item 11 Circle (5) in case of male.
 - Item 30 If zero is circles in Item 29, give 5 points for Item 30.
 - Item 34 If "YES" is checked, give 6 points for this item.
 - Item 35 If Item 34 is "NO" circle zero in Item 35.
 - Item 36 If Item 34 is "NO" circle zero for Item 36
 - Item 37 If Item 34 is "NO" circle zero for Item 37.
 - Item 52 If zero is circled in Item 51 give 5 points for Item 52.
 - Item 53 If zero is circled in Item 51 give 6 points for Item 53.

THANK YOU VERY MUCH FOR YOUR COOPERATION!

* Instructions marked with * are specific to this study.

TO:

FROM: S. Mongrain

Re: Adaptive Functioning Index

	<u>Vocational</u>	<u>Residential</u>
Residents to be assessed:		

Instructions

1. Please complete the practice assessment and return it to me before beginning the other assessments.
2. Please do not consult with any other staff member when you are completing your assessments. These must be done independently or the results will not be valid.

Vocational Check List

1. On the cover page of this check list, please write the resident's name, your name (supervisor), and on the line marked "FILE" write the date of the assessment.
2. For the items on pages 2, 3, 4, 6, 7,:
 - a) Record your assessment in Column 1.
 - b) Difficulties: (in the box) Check (✓) each item which describes a problem being experiences. If the problem is severe, use two checks (✓✓). Total the number of checks for each difficulty section and record on the line marked difficulty.
 - c) Goal Behaviors: (in order of increasing difficulty) Read each statement and assign a mark as follows:
 - 0 - if the resident does not show or has not had the chance to show the behavior.
 - 1 - if the resident is making a concerted effort but has not quite mastered the behavior.

- 2 - if the resident has mastered the skill or behavior and does it on his own.

d) Record the totals for each section in the box at the end of column 1.

3. For the Task Analysis Sheet, page 5, the regular tasks the resident performs are listed according to their skill level. Rate the resident's competence in each of these tasks using the following number ratings:
- 1 - Work needs to be checked several times a day.
 - 2 - Work o.k. if checked once a day.
 - 3 - Good but needs checking 1 - 2 times per week.
 - 4 - Quality good and stays good.

* DO NOT COMPUTE the CXS total as this is unnecessary.

*4. Do not complete the scoring sheet on the cover page.

Residential Check List

1. On the cover page of this check list, please write the resident's name, your name (supervisor), and on the line marked location write the date of the assessment.
2. Read each statement (many statements are followed by examples) and assign marks as follows in Column 1:
 - 0 - if the resident does not show or does not have the chance to show the behavior.
 - 1 - if the resident is making a concerted effort but has not quite mastered the behavior, or requires assistance or prompting.
 - 2 - if the trainee has mastered the skill and does it on his own.
3. Record the totals for each section in the box at the end of column 1.
- *4. Do not complete the cover page scoring sheet.

THANK YOU

* Instructions marked with * are specific to this study.

INFORMATION ABOUT RATER

NAME _____

AGE _____

SEX _____

EDUCATION (check appropriate category)

GRADUATE SCHOOL _____

OTHER SPECIAL TRAINING _____

FOUR YEARS COLLEGE _____

(please stipulate above:) _____

TWO YEARS COLLEGE _____

HIGH SCHOOL _____

ELEMENTARY LEVEL _____

JOB CLASSIFICATION _____

LENGTH OF TIME WORKING AT BEVAN LODGE (to June 1, 1974) _____

Supplementary Instructions for the Administration of the SET

Subtest: Writing

Application Form

Read questions if necessary; do not explain questions.

Subtest: Numbers

Item 8

Do not indicate with motion of the hand cutting the lone in half.

Subtest: Motor Movements

Items 7 - 9

Demonstrate exercise if necessary.

Subtest: Concept Attainment

Conservation

Do not change wording of items to "...Are there more, less, or the same..." since this changes the nature of the task.

In general, read all questions as they are presented in the protocol. This is especially important for

Money Handling: Item 20.

Supplementary Scoring Criteria for the SET

Subtest: Reading

Oral Reading

If the subject makes the same error more than once (e.g. Upton, Peru, floods), count as one error.

Subtest: Writing

Item 1: Please print the letters 's', 'b', 'd'.
Upper or lower case letters are correct.

Item 2: Please write your name here.
Appropriate capitalization and spelling defines 'good form'.

Application Form

Date of Birth: day, month, and year.

Marital Status:

- a) S or M
- b) single or married crossed out on form
- c) single or married written or printed in space

Present Address

- a) Bevan Lodge
- b) 'Home' address
- c) must include
 - 1. apartment number if applicable
 - 2. street number
 - 3. street name
 - 4. city name
 - 5. province; if city is in B.C., province optional since mail would arrive without this specified.

Whom to Notify: name, not position nor relationship to client

Employer: a) name of person and/or company

When did you leave your last job: correct within one year

In general, for Application Form, if a score is an odd number of points round it up to the nearest even number.

Subtest: NumbersItem 3: How many blocks are there?

- a) If incorrect on both trials score zero
- b) If incorrect on first trial but correct on second trial, score one.
- c) If correct on first trial score two.

Subtest: Community AwarenessItem 5: "The Green Sheet", "The Free Press", "The Vancouver Sun", "The Province", other newspapers that are confirmed by staff.Item 9: Limited to places of interest on Vancouver Island or Lower Mainland.Item 10: 8¢Item 11: \$1.50 - \$3.25Item 12: 15¢ - 40¢Item 13: 33¢ - 65¢Item 14: \$2.50/hr.Subtest: Concept AttainmentItems 1, 2, 3, 4, 8, 9, 10, 11, 14, 15, 16

- a) Score zero if subject responds incorrectly to
 - 1. all sections of these items
 - 2. two sections of three part items.
- b) Score one if subject responds correctly to
 - 1. one section of a two part item
 - 2. two sections of a three part item.
- c) Score two if subject responds correctly to all parts of the item.

Discrimination (Sorting) Task

- a) If subject correctly sorts 9 - 12 triplets on first trial, score one point each for 5, 6, and 7.
- b) If subject improves on second trial to criterion of 9 - 12 triplets, score one point for 5 and 6.
- c) If subject improves by three triplets on second trial but does not reach criterion of 9 - 12 triplets, score one point for 5.

- d) If subject achieves criterion of 9 - 12 triplets on second trial but has not increased his score by 3 triplets, score one point for 6.

Conservation

- a) If subject responds correctly to both items 18 and 19 on the first trial, score one for each item.
- b) If subject requires two trials for item 18, but only one trial for item 19, score one for item 18.
- c) If subject requires two trials for both items score one for item 18.
- d) If subject does not respond correctly to one item but does for the other after two trials, score one point for 18.
- e) If subject does not respond correctly to either item after two trials each score zero for 18 and 19.

Appendix B

TABLE B1

Correlations of subtest and total test scores of the ABS Part I (1) with biodemographic variables and mental age¹

Subtest	Age	Sex	Length at Bevan Lodge	Length of Previous Placement	Total Time in Institutions	RSB Mental Age
Ind. Func.	.002	-.27	-.13	.06	.04	.43
Phys. Dev.	-.21	-.08	.08	.06	.04	-.14
Econ. Act.	.21	-.17	-.02	.10	.11	.59
Lang. Dev.	.04	-.11	-.13	-.05	-.05	.58
Num. & T. Con.	.05	-.14	-.07	.12	.09	.53
Occ. Dom.	.01	.37	.03	.004	.001	.09
Occ. Gen.	.09	.15	.06	.23	.22	.004
Self-Dir.	.02	.01	.04	.16	.17	.05
Respon.	.20	.10	.06	.17	.14	.02
Socializ.	-.02	-.14	.06	-.06	-.01	.25
ABS Pt. I	.05	-.11	-.05	.08	.08	.45

¹N=48; $r_{\text{critical}} = .282$ ($p \leq .05$)

TABLE B2

Correlations of subtest and total test scores of the ABS Part I (2) with biodemographic variables and mental age¹

Subtest	Age	Sex	Length at Bevan Lodge	Length of Previous Placement	Total Time in Institutions	RSB Mental Age
Ind. Funct.	.11	.01	-.07	.06	.03	.35
Phys. Dev.	-.15	-.04	-.06	-.18	-.20	-.23
Econ. Act.	.03	-.15	-.18	-.05	-.08	.51
Lang. Dev.	-.03	-.05	-.22	-.08	-.10	.57
NTC	-.10	-.09	-.35	-.21	-.26	.50
Occ. Dom.	.05	.37	.01	.05	.01	-.02
Occ. Gen.	.14	.07	.03	.02	.004	.01
Self-Dir.	.13	-.01	.01	.08	.05	.12
Respon.	.19	-.07	-.17	-.05	-.09	.13
Socializ.	.05	.08	-.25	-.21	-.24	.17
ABS Part I (2)	.07	.02	-.17	-.05	-.09	.39

¹N=48; $r_{\text{critical}} = .282$ ($p \leq .05$)

TABLE B3

Item analysis information for the ABS Part I (1 and 2): Independent Functioning

Item	Mean		Standard Deviation		r_{subtest}		r_{total}		Difficulty	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
1	6.79	6.69	.54	.59	.57	.48	.52	.41	.85	.73
2	3.15	3.04	.88	.92	.64	.47	.70	.60	.40	.35
3	3.85	3.96	.36	.20	-.05	.29	-.23	.33	.85	.96
4	9.46	9.19	1.07	1.61	.53	.31	.48	.26	.77	.75
5	4.94	4.85	.25	.46	.10	.17	.20	.32	.94	.90
6	7.96	7.96	.29	.20	.39	.14	.12	-.19	.98	.96
7	5.81	5.88	.49	.39	.55	.36	.33	.23	.85	.90
8	7.90	7.92	.59	.40	.43	.41	.36	.27	.98	.91
9	5.50	5.40	1.05	1.11	.31	.50	.24	.37	.65	.71
10	5.15	5.02	.97	1.30	.45	.53	.54	.58	.35	.38
11	5.85	5.90	.36	.31	.56	.21	.31	.15	.85	.90
12	9.23	9.25	1.10	.98	.41	.45	.35	.42	.58	.52
13	8.13	8.23	1.53	1.33	.63	.42	.52	.32	.63	.65
14	5.63	5.67	.82	.56	.71	.54	.62	.38	.77	.71
15	6.48	6.08	1.07	1.40	.60	.50	.57	.35	.75	.63
16	6.96	6.98	.20	.14	.25	.34	-.01	.21	.96	.98
17	7.00	7.00	.01	.003	.002	.002	.01	.01	1.00	1.00
18	6.02	6.00	.14	.003	.06	.002	.06	.01	1.00	1.00
19	3.58	3.58	.74	.65	.55	.54	.61	.58	.73	.67
20	4.50	4.63	1.70	.56	.59	.60	.59	.42	.48	.56
21	5.02	5.02	1.84	1.86	.55	.37	.64	.58	.29	.31
22	5.71	5.60	1.49	1.33	.50	.52	.66	.75	.04	.02

TABLE B4

Item analysis information for the ABS Part I (1 and 2): Physical Development

Item.	Mean		Standard Deviation		r _{subtest}		r _{total}		Difficulty	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
1	3.81	3.90	.39	.31	.22	-.02	-.07	-.29	.81	.90
2	3.92	3.98	.28	.14	-.06	-.09	.42	.37	.92	.98
3	5.79	5.44	.99	1.29	.39	.50	.20	.41	.27	.27
4	6.88	6.79	.39	.50	.43	.65	.10	.19	.90	.83
5	5.79	5.89	.46	.39	.26	.24	.33	.22	.81	.90
6	3.83	3.81	.52	.45	.54	.56	.07	.11	.90	.83

TABLE B5

Item analysis information for the ABS Part I (1 and 2): Economic Activity

Item	Mean		Standard Deviation		r_{subtest}		r_{total}		Difficulty	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
1	3.73	3.52	1.43	1.35	.65	.66	.69	.74	.06	.33
2	3.81	3.73	1.58	1.55	.54	.82	.57	.67	.15	.13
3	3.73	3.67	1.32	1.31	.66	.71	.65	.75	.38	.33
4	3.60	3.65	1.16	1.25	.62	.77	.60	.69	.04	.04

TABLE B6

Item analysis information for the ABS Part I (1 and 2): Language Development

Item	Mean		Standard Deviation		r_{subtest}		r_{total}		Difficulty	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
1	3.31	3.38	1.88	1.67	.78	.64	.60	.54	.25	.17
2	7.00	6.88	.003	.87	.003	-.13	.01	-.10	1.00	1.00
3	5.63	5.65	.70	.79	.42	.51	.31	.40	.73	.79
4	3.33	3.13	.75	.80	.68	.64	.55	.43	.48	.50
5	4.04	3.98	1.09	1.12	.56	.47	.58	.48	.46	.44
6	3.50	3.19	1.94	1.97	.77	.69	.68	.58	.29	.25
7	4.46	4.35	.87	1.04	.43	.44	.54	.49	.67	.65
8	4.81	4.71	.57	.71	.31	.02	.35	.17	.88	.83
9	4.63	4.65	1.67	1.66	.64	.57	.51	.55	.04	.06

TABLE B7

Item analysis information for the ABS Part I (1 and 2): Number and Time Concepts

Item	Mean		Standard Deviation		r_{subtest}		r_{total}		Difficulty	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
1	4.75	4.65	1.50	1.48	.65	.48	.74	.57	.44	.38
2	5.40	5.52	1.99	1.97	.73	.39	.74	.44	.48	.54
3	3.42	3.48	.87	.85	.57	.62	.45	.47	.63	.69

TABLE B8

Item analysis information for the ABS Part I (1 and 2): Occupation Domestic

Item	Mean		Standard Deviation		r_{subtest}		r_{total}		Difficulty	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
1	2.44	2.33	.58	.66	.50	.47	.13	.24	.48	.44
2	3.96	3.88	1.88	1.84	.55	.57	.26	.38	.35	.38
3	3.25	2.96	.93	.99	.57	.60	.53	.47	.48	.33
4	2.02	2.06	.91	1.02	.55	.68	.54	.61	.08	.13
5	2.83	2.71	.56	.71	.44	.31	.37	.26	.92	.85
6	4.02	3.77	1.04	1.26	.62	.61	.38	.62	.44	.42

TABLE B9

Item analysis information for the ABS Part I (1 and 2): Occupation General

Item	Mean		Standard Deviation		r_{subtest}		r_{total}		Difficulty	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
1	2.33	2.38	.60	.49	.01	.25	.30	.53	.29	.38
2	5.44	5.46	1.05	.92	.54	.65	.47	.60	.71	.65
3	6.17	6.02	1.33	1.47	.51	.63	.40	.51	.65	.60

TABLE B10

Item analysis information for the ABS Part I (1 and 2): Self-Direction

Item	Mean		Standard Deviation		r_{subtest}		r_{total}		Difficulty	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
1	4.69	4.48	.66	.90	.36	.46	.19	.52	.79	.69
2	3.27	3.25	.84	.76	.77	.65	.67	.53	.50	.44
3	6.42	6.31	1.16	1.06	.74	.77	.50	.58	.69	.58
4	4.46	4.52	.92	.83	.81	.53	.59	.39	.67	.69
5	6.21	6.00	1.30	1.37	.76	.76	.51	.58	.65	.52
6	4.54	4.54	.85	.96	.46	.44	.53	.49	.75	.75
7	4.27	3.98	1.82	1.97	.69	.68	.69	.68	.13	.10

TABLE B11

Item analysis information for the ABS Part I (1 and 2): Responsibilities

Item	Mean		Standard Deviation		r_{subtest}		r_{total}		Difficulty	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
1	3.44	3.35	.62	.76	.30	.34	.29	.52	.50	.50
2	4.90	2.79	.59	.62	.30	.34	.58	.41	.08	.06

TABLE B12

Item analysis information for the ABS Part I (1 and 2): Socialization

Item	Mean		Standard Deviation		r_{subtest}		r_{total}		Difficulty	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
1	3.21	3.58	1.58	1.49	.24	.62	.17	.49	0.0	0.0
2	2.92	2.88	1.35	1.30	.47	.69	.49	.48	0.0	0.0
3	5.04	5.13	1.03	.94	.34	.34	.67	.38	.42	.44
4	3.25	3.10	.86	.88	.44	.44	.29	.25	.48	.38
5	2.48	2.50	.68	.65	.60	.48	.56	.52	.04	.08
6	6.21	6.15	.94	.95	.44	.27	.37	.23	.44	.38
7	7.15	7.15	.85	.80	.14	.14	.27	.50	.42	.38

TABLE B13

Generalizability coefficients by rater pair for the subdomains of
Independent Functioning

Subdomain	Rater Pair							
	1	2	3	4	5	6	7	8
Eating Skills	.78	.54	.46	.10	.34	.49	0.0	.25
Toilet Use	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Cleanliness	.51	.63	0.0	.42	.70	.58	.15	0.0
Appearance	.77	.15	0.0	0.0	.66	.74	.44	.29
Care of Cloth.	.86	0.0	0.0	0.0	0.0	.47	0.0	0.0
Dressing, Unor.	0.0	0.0	0.0	1.00	0.0	0.0	0.0	0.0
Locomotion	.82	.44	.80	.50	.23	.97	.33	0.0
Gen. Ind. F.	.75	.47	0.0	.04	.75	.82	.18	.03

TABLE B14

Generalizability coefficients by rater pair for the subdomain of
Physical Development

Subdomain	Rater Pair							
	1	2	3	4	5	6	7	8
Sensory Dev.	1.00	0.0	0.0	0.0	.60	0.0	0.0	0.0
Motor Dev.	.82	.51	0.0	.44	0.0	.15	.93	0.0

TABLE B15

Generalizability coefficients by rater pair for the subdomains of
Economic Activity

Subdomain	Rater Pair							
	1	2	3	4	5	6	7	8
Money Handling & Budgeting	.46	.72	.78	0.0	.81	.90	.83	0.0
Shopping Skills	.37	.63	0.0	.54	.97	.88	.61	0.0

TABLE B16

Generalizability coefficients by rater for the subdomains of
Language Development

Subdomain	Rater Pair							
	1	2	3	4	5	6	7	8
Speaking and Writing	.70	.58	.94	0.0	.94	.92	.86	.44
Comprehension	.82	.42	.96	.37	.79	.82	.91	0.0
General Language Development	.65	.13	.48	.39	.80	.63	.29	.70

TABLE B17

Generalizability coefficients by rater pair for the subdomain of
Occupation Domestic

Subdomain	Rater Pair							
	1	2	3	4	5	6	7	8
Cleaning	.31	.59	0.0	.51	.22	0.0	0.0	.91
Kitchen Duties	.06	.29	0.0	.06	.07	.48	.81	.43
General Occupation Domestic	.37	.33	0.0	.37	.74	0.0	.36	0.0

TABLE B18

Generalizability coefficients by rater pair for the subdomains of
Self-Direction

Subdomain	Rater Pair							
	1	2	3	4	5	6	7	8
Sluggishness	.59	.20	0.0	0.0	0.0	0.0	0.0	0.0
Initiative	.80	.44	.54	.31	.33	.23	.87	0.0
Persistence	.80	.29	.64	.28	.50	.15	.74	0.0
Planning & Org.	.70	0.0	0.0	.83	0.0	.73	0.0	0.0
Self-Direction General	.77	.56	0.0	0.0	0.0	1.00	.68	.60

TABLE B19

Correlations of subtest and total test scores of the ABS Part II (1) with biodemographic variables and mental age¹

Subtest	Age	Sex	Length at Bevan Lodge	Length of Previous Placement	Total Time in Institutions	RSB Mental Age
Viol. & Des. Beh.	-.14	.15	-.08	-.09	-.10	-.13
Antisocial Beh.	-.20	.20	-.09	-.23	-.25	-.03
Rebellious Beh.	-.17	.27	-.04	-.17	-.19	-.08
Untrust. Beh.	-.004	.10	-.08	-.04	-.08	.14
Withdrawal	.13	.08	.13	.15	.17	-.28
Ster. Beh. & Odd. M.	-.04	-.05	.29	.14	.19	.12
Inapp. Inter. Man.	-.08	.22	.05	.03	.04	-.10
Unacc. Voc. Hab.	-.15	.18	-.12	-.22	-.20	.24
Unacc./Ecc. Hab.	-.06	.27	.27	.08	.05	-.001
Self-Abusive Beh.	-.04	-.08	.06	-.03	-.01	-.22
Hyper. Tend.	-.20	.08	.08	-.16	-.26	-.05
Sex. Aber. Beh.	-.10	.26	.18	-.05	-.12	-.18
Psycho. Disturb.	-.03	.28	-.01	-.09	-.16	-.06
Use of Med.	.07	.21	.05	.20	.18	.03
ABS Part II (1)	-.13	.29	-.003	-.12	-.18	-.07

¹ N=48; $r_{\text{critical}} = .282$ ($p \leq .05$)

TABLE B20

Correlations of subtest and total test scores of the ABS Part II (2) with biodemographic variables and mental age¹

Subtest	Age	Sex	Length at Bevan Lodge	Length of Previous Placement	Total Time in Institutions	RSB Mental Age
Viol. & Des. Beh.	-.15	.27	.13	.06	.09	-.04
Antisocial Beh.	.04	-.02	.07	.18	.19	.01
Rebellious Beh.	-.15	.17	.04	.02	.02	-.17
Untrust. Beh.	-.04	.08	0	-.01	-.004	-.06
Withdrawal	-.02	-.003	-.07	.09	.08	-.08
St. Beh., Odd Man..	-.02	-.11	.11	-.02	.02	.04
Inapp. Inter. Man.	.05	.12	.13	.10	.12	.07
Unacc. Voc. Hab.	.06	-.05	.11	-.02	.02	.27
Unacc./Ecc. Hab.	-.05	-.04	.06	.01	.02	.02
Self-Abusive Beh.	-.21	.06	-.07	-.08	-.07	.04
Hyperactive Tend.	.001	.03	.16	-.004	.04	.21
Sex. Aber. Beh.	-.19	.12	.11	-.01	.03	-.27
Psych. Disturb.	.05	.16	.15	.07	.09	-.10
Use of Med.	-.03	.31	.08	.04	.03	.06
ABS Pt. II (2)	-.04	.14	.14	.09	.11	-.03

¹N=48; $r_{\text{critical}} = .282$ ($p \leq .05$)

TABLE B21

Intercorrelations of subtest scores of the ABS Parts I and II (1)

		VDB	AB	RB	UB	W	SBOM	IIM	UVH	UEH	SAB	HT	SAB	PD	UM
1	Ind. F.	-.25	-.33	-.45	-.27	-.34	-.10	-.31	-.08	-.12	-.03	-.13	-.35	-.36	-.09
2	Phys. D.	.05	.04	.07	.02	-.15	.12	.02	.18	.08	-.31	.17	.02	-.01	-.18
3	Econ. A.	-.36	-.23	-.35	-.16	-.26	.15	-.20	.04	-.17	-.06	-.26	-.30	-.35	-.03
4	Lang. D.	-.33	-.16	-.27	-.12	-.38	.29	-.19	.13	-.11	.04	-.12	-.23	-.32	.08
5	NTC	-.50	-.29	-.38	-.09	-.33	.17	-.13	-.11	.05	-.05	-.10	-.15	-.32	.17
6	Occ. D.	.10	.16	.03	-.04	-.08	.10	.05	.09	-.02	-.17	-.04	.07	.07	-.16
7	Occ. G.	-.22	-.42	-.48	-.45	-.35	.12	-.18	-.12	-.22	-.12	-.16	-.18	-.41	-.03
8	Self-D.	-.16	-.42	-.56	-.44	-.28	.09	-.36	-.07	-.29	-.05	-.17	-.26	-.46	-.14
9	Respon.	-.33	-.37	-.46	-.44	-.22	-.10	-.23	-.27	-.09	-.15	-.05	.006	-.28	.05
10	Soc.	-.50	-.37	-.47	-.42	-.24	.16	-.19	.06	-.44	-.14	-.16	-.22	-.57	-.18

N=48 $r_{\text{critical}} = .282$ ($p \leq .05$)

TABLE B22

Intercorrelations of subtest scores of the ABS Parts I and II (2)

		VDB	AB	RB	UB	W	SBOM	IIM	UVH	UEH	SAB	HT	SB	PD	M
1	Ind. Funct.	-.17	-.25	-.49	-.17	-.37	-.28	-.16	-.34	-.55	-.35	-.07	-.41	-.24	-.13
2	Phys. Dev.	-.25	-.52	-.02	-.05	-.26	-.34	-.01	-.37	-.10	-.20	-.20	.04	-.41	-.43
3	Econ. Act.	.07	-.16	-.24	-.05	-.18	-.07	-.08	-.03	-.16	-.24	-.07	-.14	-.22	-.25
4	Lang. Dev.	.01	-.22	-.17	-.04	-.24	.08	-.02	.08	.02	-.17	-.01	-.12	-.18	-.14
5	NTC	-.03	-.06	-.15	.003	-.16	.05	.01	.06	.10	-.18	.03	.05	-.01	-.16
6	Occ. Dom.	.15	-.22	-.08	.17	-.13	-.38	-.17	-.22	-.18	-.19	-.30	.004	-.12	-.03
7	Occ. Gen.	-.08	-.30	-.50	-.12	-.37	-.28	-.28	-.30	-.49	-.27	-.27	-.32	-.06	-.16
8	Self-Dir.	-.01	-.28	-.58	-.25	-.51	-.42	-.27	-.38	-.49	-.32	-.26	-.39	-.31	-.12
9	Respon.	.04	.08	-.14	-.03	-.12	-.15	-.11	-.17	-.18	-.08	-.06	-.05	.03	-.07
10	Socializ.	-.14	-.33	-.33	-.22	-.40	-.25	-.10	-.15	-.24	-.04	-.30	-.07	-.32	-.26

N=48 $r_{\text{critical}} = .282$ ($p \leq .05$)

Appendix C

TABLE C1

Correlations of subtest and total test scores of the SET (1) with biodemographic variables¹

		Chronological Age	Sex	Length at Bevan	Length of Prev. Inst.	Total Inst.	Mental Age
A	Reading	0.134	-.069	-.114	-.014	-.037	.866
B	Writing	.127	.020	-.092	-.069	-.080	.825
C	Numbers	.065	-.028	-.116	.086	.049	.873
D	Time	.118	-.108	-.062	.071	.047	.950
E	Money	.238	-.110	-.045	.120	.094	.832
F	Community	.156	-.073	-.111	-.024	-.045	.830
G	Communication	.041	.023	-.151	-.026	-.056	.748
H	Motor Movements	-.295	-.015	-.232	-.086	-.126	.522
I	Concept Attainments	-.059	-.162	-.247	-.055	-.103	.867
J	Total Score	.099	-.066	-.131	.010	-.020	.920

¹N=50 r_{critical} = .276 (p ≤ .05)

TABLE C2

Correlations of subtest and total test scores of the SET (2) with biodemographic variables¹

		Chronological Age	Sex	Length at Bevan	Length of Prev. Inst.	Total Inst.	Mental Age
A	Reading	-.160	-.067	-.225	-.027	-.074	.935
B	Writing	-.119	-.008	-.233	-.092	-.130	.864
C	Numbers	-.088	.012	-.171	.051	.003	.919
D	Time	-.112	-.063	-.148	.135	.079	.856
E	Money	-.075	-.053	-.195	.090	.030	.912
F	Community	-.003	-.133	-.266	.049	-.019	.863
G	Communication	-.285	-.012	-.346	-.074	-.140	.765
H	Motor Movements	-.380	-.147	-.328	-.039	-.107	.473
I	Concept Attainment	-.230	-.082	-.340	-.011	-.087	.887
J	Total Score	-.159	-.058	.013	.013	-.048	.944

¹N=30 $r_{\text{critical}} = .355$ ($p \leq .05$)

TABLE C3

Item analysis information for the Reading subtest of the SET

Item	r_{total}	Difficulty
1	.67	.72
2	.81	.58
3	.69	.62
4	.80	.58
5	.85	.50
6	.86	.46
7	.35	.08
8	.67	.28
9	.86	.48
10	.75	.32
11	.67	.26
12	.55	.14
13	.86	.42
14	.57	.14
15	0.0	0.0
16	.68	.26
17	.41	.16
18	.59	.26
19	.36	.12
20	.59	.14

TABLE C4

Item analysis information for the Writing subtest of the SET

Item	r_{total}	Difficulty
1	.73	.62
2	.76	.48
3	.76	.46
4	.85	.38
5	.71	.68
6	.82	.40
7	.72	.22
8	.77	.58
9	.86	.44
10	.87	.42
11	.85	.38
12	.83	.34
13	.81	.28
14	.70	.20
15	.61	.14
16	.84	.36
17	.84	.36
18	.84	.38
19	.61	.26
20	.51	.16

TABLE C5

Item analysis information for the Numbers subtest of the SET

Item	r_{total}	Difficulty
1	.46	.84
2	.44	.40
3	.63	.56
4	.62	.74
5	.75	.30
6	.71	.30
7	.47	.40
8	.67	.34
9	.09	.18
10	.72	.44
11	.84	.42
12	.84	.36
13	.77	.36
14	.75	.28
15	.66	.18
16	.77	.28
17	.64	.18
18	.49	.10
19	.67	.18
20	.37	.06

TABLE C6

Item analysis information for the Time subtest of the SET

Item	r_{total}	Difficulty
1	.38	.92
2	.42	.86
3	.57	.62
4	.51	.42
5	.70	.58
6	.71	.42
7	.55	.72
8	.77	.56
9	.81	.42
10	.76	.40
11	.63	.46
12	.61	.28
13	.75	.30
14	.70	.26
15	.86	.38
16	.74	.30
17	.74	.26
18	.52	.12
19	.68	.18
20	.28	.06

TABLE C7

Item analysis information for the Money subtest of the SET

Item	r_{total}	Difficulty
1	.69	.60
2	.82	.46
3	.73	.56
4	.78	.50
5	.71	.54
6	.78	.40
7	.78	.32
8	.78	.30
9	.86	.40
10	.79	.34
11	.79	.36
12	.81	.50
13	.86	.36
14	.86	.38
15	.68	.20
16	.76	.30
17	.41	.08
18	.62	.20
19	0.0	0.0
20	.45	.24

TABLE C8

Item analysis information for the Community subtest of the SET

Item	r_{total}	Difficulty
1	.40	.36
2	.54	.66
3	.48	.22
4	.49	.38
5	.66	.74
6	.60	.44
7	.46	.50
8	.57	.26
9	.37	.34
10	.59	.40
11	.52	.40
12	.39	.32
13	.36	.22
14	.55	.28
15	.16	.04
16	.56	.32
17	.56	.26
18	.15	.02
19	.33	.06
20	.43	.08

TABLE C9

Item analysis information for the Communication subtest of the SET

Item	r_{total}	Difficulty
1	0	1.00
2	0	1.00
3	.29	.84
4	.40	.78
5	.28	.60
6	.34	.46
7	0	1.00
8	.39	.92
9	.59	.74
10	.46	.42
11	.44	.50
12	.37	.48
13	.59	.74
14	.74	.32
15	0	1.00
16	.54	.84
17	.83	.50
18	0	1.00
19	.55	.82
20	.74	.40

TABLE C10

Item analysis information for the Motor Movements subtest of the SET

Item	r_{total}	Difficulty
1	.10	.52
2	.03	.42
3	-.07	.40
4	0	1.00
5	.09	.90
6	.16	.80
7	.22	.98
8	.11	.76
9	.23	.48
10	-.23	.98
11	.15	.88
12	.33	.70
13	.43	.16
14	0	0
15	.46	.68
16	.28	.30
17	0	0
18	.31	.80
19	.43	.48
20	.32	.14

TABLE C11

Item analysis information for the Concept Attainment subtest of the SET

Item	r_{total}	Difficulty
1	-.36	.92
2	.43	.58
3	-.35	.10
4	.40	.24
5	-.19	.68
6	-.38	.48
7	.45	.18
8	.62	.64
9	.55	.50
10	.48	.50
11	.45	.32
12	-.22	.72
13	.53	.56
14	.68	.46
15	-.39	.38
16	.42	.60
17	.61	.26
18	.62	.66
19	.47	.24
20	-.19	.54

Appendix D

TABLE D1

Correlations of subtest and total test scores of the VCL (1) and biodemographic variables and mental age¹

	Age	Sex	Length at Bevan Lodge	Length of Previous Placement	Total Time in Institutions	RSB Mental Age
Basic Work Habits	.11	-.04	-.11	.01	-.03	.22
Independence	.11	-.15	-.17	.02	-.03	.24
Making Decisions	.09	.03	.06	.07	.04	.25
Use, Care of Equipment	.05	.18	-.04	-.05	-.09	.05
Taking Directions	.14	-.20	-.23	.01	-.03	.20
Work Skills	.03	.03	.02	-.13	-.21	.09
Speed	.13	-.15	-.03	-.03	.06	.03
Following Instructions	.11	-.17	.05	.04	.02	.32
Task Analysis	-.02	.09	.03	-.15	-.23	.04
Acceptance Skills	.19	.04	0.0	.09	.08	.38
Appearance	.13	.02	.01	.17	.13	.19
Punctuality	.24	.07	.05	.22	.21	.47
Talking	.19	-.08	0.0	-.01	0.0	.40
Relationships	.06	.13	-.01	-.08	-.07	.20
Total Test	.11	.02	-.03	-.04	-.10	.24

¹N=42; $r_{\text{critical}} = .307$ ($p \leq .05$)

TABLE D2

Correlations of subtest and total test scores of the VCL (2) with biodemographic variables and mental age¹

	Age	Sex	Length at Bevan Lodge	Length of Previous Placement	Total Time in Institutions	RSB Mental Age
Basic Work Habits	.21	-.07	-.17	-.01	-.04	.19
Independence	.23	-.15	-.09	-.07	-.08	.03
Making Decisions	.14	.03	-.02	.18	.16	.23
Use, Care of Equipment	.17	.03	-.22	.01	-.04	.17
Taking Directions	.12	-.13	-.21	-.13	-.15	.20
Work Skills	.28	0.0	-.12	.04	.01	.13
Speed	-.23	-.27	-.28	.11	-.16	.21
Following Instructions	-.06	-.02	-.28	-.21	-.24	.07
Task Analysis	.36	.05	-.03	.11	.09	.10
Acceptance Skills	.16	.14	-.14	.04	0.0	.45
Appearance	.09	.10	-.14	.06	.02	.19
Punctuality	.13	0.0	-.14	.01	-.02	.44
Talking	.18	.17	-.06	.13	.10	.55
Relationships	.10	.16	-.11	.10	-.11	.23
Total Test	.27	.03	-.17	.03	-.01	.29

¹N=.42 $r_{\text{critical}}=.307$ ($p \leq .05$)

TABLE D3

Item analysis information for the Basic Work Habits subtest of the VCL (1 & 2)

Item	Mean		Standard Deviation		r_{subtest}		r_{total}		Difficulty	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
1	1.45	1.36	.71	.69	.64	.62	.52	.45	.57	.48
2	1.26	1.14	.67	.61	.70	.63	.59	.54	.38	.26
3	.83	1.00	.70	.66	.62	.59	.67	.55	.17	.21
4	1.02	.83	.75	.76	.71	.69	.63	.70	.29	.21
5	.86	.79	.75	.78	.50	.55	.52	.52	.21	.21
6	1.02	.93	.72	.75	.52	.45	.59	.52	.26	.24
7	1.21	.91	.68	.62	-.03	.16	.15	.10	.36	.14
8	.76	.62	.69	.62	.41	.30	.53	.21	.14	.07
9	.71	.62	.71	.70	.69	.56	.78	.55	.14	.12
10	.57	.50	.74	.60	.53	.53	.51	.41	.14	.05
11	1.43	1.38	.59	.62	.45	.63	.52	.64	.48	.45
12	1.24	1.12	.66	.71	.70	.63	.55	.41	.36	.31
13	1.10	1.19	.62	.71	.63	.64	.52	.45	.24	.36
14	1.29	1.10	.74	.76	.75	.75	.70	.62	.45	.33
15	.71	.71	.67	.81	.43	.66	.63	.69	.12	.21
16	1.10	1.26	.58	.63	.43	.55	.22	.54	.21	.36
17	1.29	1.21	.64	.61	.61	.45	.41	.29	.38	.31
18	.74	.83	.67	.70	.36	.33	.18	.15	.12	.17
19	.91	.83	.62	.66	.57	.42	.48	.25	.14	.14
20	.48	.29	.71	.51	.22	.19	.51	.25	.12	.02

TABLE D4

Item analysis information for the subtest Work Skills of the VCL (1)

Item	Mean		Standard Deviations		r_{subtest}		r_{total}		Difficulty	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
1	.93	.83	.71	.70	.65	.36	.69	.15	.21	.17
2	.95	.76	.70	.66	.66	.51	.66	.27	.21	.12
3	.64	.67	.82	.69	.63	.16	.56	.46	.21	.12
4	.26	.17	.50	.38	.65	.22	.51	.21	.02	0.0
5	.05	.02	.22	.15	.43	.16	.24	.03	0.0	0.0
6	1.36	1.36	.62	.53	.47	.54	.22	.45	.43	.38
7	1.41	1.36	.63	.66	.72	.63	.53	.47	.48	.45
8	1.12	1.05	.59	.73	.78	.72	.57	.64	.24	.29
9	1.14	1.21	.72	.72	.61	.52	.80	.49	.33	.38
10	.67	.62	.72	.73	.64	.63	.80	.56	.14	.14

TABLE D5

Item analysis information for the subtest Acceptance Skills of the VCL (1)

Item	Mean		Standard Deviation		r_{subtest}		r_{total}		Difficulty	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
1	1.45	1.33	.59	.53	.62	.57	.58	.71	.50	.36
2	1.26	1.24	.63	.58	.78	.66	.50	.53	.36	.31
3	1.29	1.31	.64	.56	.68	.67	.44	.41	.38	.36
4	1.21	1.14	.65	.78	.68	.63	.68	.52	.33	.38
5	1.07	.93	.75	.71	.54	.49	.60	.66	.31	.21
6	1.24	1.17	.62	.73	.61	.66	.63	.62	.33	.36
7	1.33	1.10	.61	.66	.63	.57	.37	.49	.41	.26
8	1.21	1.07	.57	.51	.52	.48	.29	.21	.29	.17
9	1.19	1.14	.83	.81	.62	.62	.64	.58	.45	.41
10	1.48	.41	.71	.63	.50	.60	.46	.38	.12	.07
11	1.55	1.50	.63	.63	.56	.52	.44	.37	.62	.57
12	1.41	1.26	.70	.74	.64	.84	.34	.40	.52	.43
13	.86	.91	.68	.79	.77	.73	.44	.36	.17	.26
14	.79	.81	.68	.71	.77	.73	.49	.48	.14	.17
15	.60	.76	.73	.76	.73	.69	.60	.47	.14	.19
16	1.50	1.41	.63	.59	.43	.59	.22	.42	.57	.45
17	1.48	1.64	.59	.53	.64	.55	.40	.44	.52	.67
18	.91	1.19	.73	.74	.64	.72	.42	.60	.21	.38
19	.91	.76	.73	.76	.66	.56	.75	.69	.21	.19
20	.69	.62	.68	.70	.59	.64	.58	.46	.12	.12

Appendix E

TABLE E1

Correlations of subtest and total test scores of the RCL (1) with Biodemographic variables and mental age¹

Subtest	Age	Sex	Length at Bevan Lodge	Length of Previous Placement	Total Time in Institutions	RSB Mental Age
Personal Routines	.23	-.01	.04	.30	.26	-.04
Cleanliness	.01	-.02	0.0	.19	.17	-.24
Appearance, Eating	.09	-.01	-.04	.20	.17	-.22
Room Management	.25	.20	.14	.24	.24	-.31
Time Management	.30	-.16	-.02	.23	.19	.37
Health	.17	-.10	.01	.20	.18	.39
Community Awareness	.18	-.17	-.03	.12	.10	.38
Transportation	.18	-.28	-.16	.02	.02	.44
Shopping	.20	-.20	-.11	.10	.06	.50
Leisure	.03	-.04	-.01	-.02	-.02	.05
Budgeting	.26	-.26	.11	.27	.25	.43
Cooking, Home Management	.11	-.12	-.13	.19	.19	.15
Social Maturity	.26	-.28	-.15	.20	.14	.30
Communication	.11	.24	-.13	.17	.12	.58
Consideration	.35	-.18	-.14	.17	.12	.15
Getting Friends	-.02	-.03	-.20	.04	-.01	.02
Keeping Friends	.29	-.28	-.08	.12	.08	-.03
Handling Problems	.27	-.34	-.03	.23	.19	.26
Total Test	.10	-.10	-.02	-.01	-.03	.39

¹N=41; $r_{\text{critical}} = .304$ ($p \leq .05$)

TABLE E2

Correlations of subtest and total test scores of the RCL (2) with biodemographic variables and mental age¹

Subtest	Age	Sex	Length at Bevan Lodge	Length of Previous Placement	Total Time in Institutions	RSB Mental Age
Personal Routines	-.05	.01	-.04	.03	-.02	.37
Cleanliness	-.12	.05	-.05	.08	.03	.25
Appearance, Eating	-.05	-.16	-.15	-.02	-.08	.25
Room Management	-.19	.35	.05	.02	-.03	.10
Time Management	.11	-.13	.07	.12	.12	.34
Health	.06	-.14	-.09	-.14	-.12	.42
Community Awareness	.24	-.16	-.02	-.04	-.04	.42
Transportation	.23	-.29	-.03	-.04	-.01	.41
Shopping	.20	-.13	-.06	-.02	-.03	.44
Leisure	-.11	-.14	-.18	.21	-.21	.11
Budgeting	.40	-.10	.12	-.11	.13	.27
Cooking, Home Management	.16	.17	.11	-.06	-.04	.24
Social Maturity	.08	-.11	-.01	.01	-.02	.21
Communication	.11	-.12	.02	-.08	.05	.34
Consideration	.04	.01	.01	.09	.04	-.21
Getting Friends	-.10	.04	-.03	-.24	-.22	.35
Keeping Friends	.16	-.10	.11	.19	.17	-.01
Handling Problems	.07	-.22	-.11	.10	.04	.24
Total Test	.25	-.17	-.05	.23	.18	.25

¹ N=40 $r_{\text{critical}} = .308$ ($p \leq .05$)

TABLE E3

Item analysis information for the Cleanliness subtest of the RCL (1 and 2)

Item	Mean		Standard Deviation		r_{subtest}		r_{total}		Difficulty	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
1	1.63	1.56	.49	.50	.72	.53	.44	.18	.63	.58
2	1.73	1.73	.45	.51	.71	.62	.32	.06	.73	.75
3	1.61	1.58	.54	.59	.65	.65	.29	.29	.63	.63
4	1.37	1.35	.62	.62	.80	.80	.51	.32	.44	.43
5	1.56	1.53	.59	.60	.69	.77	.41	.26	.61	.58
6	1.46	1.43	.56	.55	.84	.89	.49	.29	.49	.45
7	1.32	1.45	.65	.60	.75	.78	.26	.31	.42	.50
8	1.24	1.18	.70	.71	.80	.80	.41	.42	.39	.35
9	1.10	1.10	.77	.63	.52	.53	.80	.42	.34	.25
10	1.44	1.43	.50	.64	.61	.74	.49	.20	.44	.50

TABLE E4

Item analysis information for the Appearance and Eating subtest of the
RCL (1 and 2)

Item	Mean		Standard Deviation		r_{subtest}		r_{total}		Difficulty	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
1	1.39	1.15	.59	.70	.42	.44	.44	.42	.44	.33
2	1.56	1.55	.50	.60	.56	.66	.30	.32	.56	.60
3	1.42	1.45	.63	.55	.33	.37	.29	.19	.49	.48
4	1.54	1.35	.51	.62	.45	.58	.16	.45	.54	.43
5	1.37	1.30	.54	.69	.48	.59	.30	.29	.39	.43
6	1.81	1.80	.40	.41	.84	.66	.67	.42	.81	.80
7	1.61	1.48	.50	.72	.57	.62	.45	.32	.66	.60
8	1.78	1.80	.42	.41	.79	.64	.58	.42	.78	.80
9	1.78	1.88	.48	.34	.79	.62	.62	.43	.81	.88
10	1.59	1.35	.59	.58	.47	.45	.59	.55	.63	.40

TABLE E5

Item analysis information for the Room Management subtest of the RCL (1 and 2)

Item	Mean		Standard Deviation		r_{subtest}		r_{total}		Difficulty	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
1	1.34	1.10	.58	.71	.43	.64	.32	.26	.39	.30
2	1.51	1.40	.55	.67	.52	.64	.40	.31	.54	.50
3	1.39	1.40	.54	.71	.67	.74	.20	.45	.42	.53
4	1.42	1.48	.55	.60	.69	.72	.06	.36	.44	.53
5	1.15	1.15	.62	.66	.68	.76	.26	.21	.27	.30
6	1.37	1.45	.58	.68	.66	.72	.51	.56	.42	.55
7	1.32	1.33	.57	.69	.73	.79	.44	.61	.37	.45
8	1.17	1.18	.67	.75	.62	.77	.21	.32	.32	.38
9	.49	.38	.71	.63	.44	.32	.47	.56	.12	.08
10	.59	.73	.74	.72	.37	.64	.41	.52	.15	.15

TABLE E6

Item analysis information for the Time Management subtest of the RCL (1 and 2)

Item	Mean		Standard Deviation		r_{subtest}		r_{total}		Difficulty	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
1	1.68	1.55	.52	.60	.65	.53	.65	.59	.71	.60
2	.66	.48	.86	.68	.32	.53	.33	.62	.24	.10
3	1.00	1.15	.74	.66	.48	.36	.51	.37	.27	.30
4	1.49	1.63	.64	.54	.47	.48	.44	.41	.56	.65
5	1.44	1.40	.67	.59	.68	.28	.67	.21	.54	.45
6	1.46	1.50	.55	.60	.68	.37	.69	.49	.49	.55
7	1.46	1.38	.60	.67	.55	.57	.58	.57	.51	.48
8	.98	.90	.72	.67	.56	.55	.58	.57	.24	.18
9	.71	.55	.60	.64	.70	.59	.69	.61	.07	.08
10	.95	.75	.59	.67	.39	.61	.57	.63	.15	.13

TABLE E7

Item analysis information for the Health subtest of the RCL (1 and 2)

Item	Mean		Standard Deviation		r_{subtest}		r_{total}		Difficulty	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
1	1.29	1.13	.60	.52	.47	.38	.53	.60	.37	.20
2	.90	.68	.70	.69	.52	.47	.62	.49	.20	.13
3	.46	.38	.67	.54	.57	.69	.24	.51	.10	.03
4	.32	.40	.61	.74	.59	.59	.13	.45	.07	.15
5	.12	.10	.40	.30	.69	.50	.42	.41	.02	0.0
6	1.44	1.25	.63	.77	.08	.21	.28	.22	.51	.45
7	.90	.90	.66	.67	.48	.51	.34	.54	.17	.18
8	1.63	1.53	.49	.60	.37	.45	.26	.35	.63	.58
9	.83	.65	.70	.80	.51	.60	.24	.66	.17	.20
10	.22	.15	.53	.36	.46	.61	.39	.53	.05	0.0

TABLE E8

Item analysis information for the Transportation subtest of the RCL (1 and 2)

Item	Mean		Standard Deviation		r_{subtest}		r_{total}		Difficulty	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
1	1.07	1.10	.65	.67	.64	.58	.45	.47	.24	.28
2	1.54	1.43	.51	.59	.58	.34	.61	.47	.54	.48
3	1.42	1.30	.63	.69	.64	.50	.55	.38	.49	.43
4	.88	.70	.81	.76	.83	.76	.60	.70	.27	.18
5	.71	.73	.81	.82	.62	.73	.38	.58	.22	.23
6	.39	.35	.59	.62	.69	.65	.39	.50	.05	.08
7	.44	.50	.59	.72	.67	.48	.46	.44	.05	.13
8	1.24	1.23	.73	.80	.59	.54	.47	.33	.42	.45
9	1.24	1.05	.70	.75	.69	.63	.49	.51	.39	.30
10	.24	.25	.54	.49	.61	.61	.43	.64	.05	.03

TABLE E9

Item analysis information for the Shopping subtest of the RCL (1 and 2)

Item	Mean		Standard Deviation		r_{subtest}		r_{total}		Difficulty	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
1	1.29	1.18	.72	.71	.72	.66	.77	.72	.44	.35
2	1.27	1.23	.71	.66	.67	.47	.59	.54	.42	.35
3	1.02	1.13	.61	.76	.77	.77	.63	.51	.20	.35
4	.88	.78	.64	.73	.74	.76	.60	.54	.15	.18
5	1.20	1.15	.72	.80	.80	.66	.76	.57	.37	.40
6	.95	1.03	.67	.73	.78	.72	.68	.72	.20	.28
7	1.54	1.60	.51	.55	.34	.29	.38	.57	.54	.63
8	1.07	.90	.69	.67	.80	.62	.78	.59	.27	.18
9	.68	.43	.61	.50	.80	.56	.73	.48	.07	0.0
10	.66	.55	.62	.68	.48	.63	.52	.63	.07	.10

TABLE E10

Item analysis information for the Leisure subtest of the RCL (1 and 2)

Item	Mean		Standard Deviation		r_{subtest}		r_{total}		Difficulty	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
1	1.24	1.15	.70	.66	.32	.53	.20	.41	.39	.30
2	1.29	1.20	.68	.72	.33	.55	.13	.27	.42	.38
3	.73	.55	.59	.75	.26	.60	.32	.61	.07	.15
4	1.46	.140	.64	.63	.29	.55	.34	.32	.54	.48
5	1.42	1.63	.67	.49	.29	.39	.23	.36	.51	.63
6	.66	.63	.79	.84	.29	.46	.10	.38	.20	.23
7	.42	.48	.74	.75	.40	.23	.09	.18	.15	.15
8	.76	.80	.80	.79	.40	.35	.12	.33	.22	.23
9	.07	.28	.35	.60	.32	.65	.15	.54	.02	.08
10	.34	.43	.53	.64	.43	.65	.41	.61	.02	.08

TABLE E11

Item analysis information for the Budgeting subtest of the RCL (1 and 2)

Item	Mean		Standard Deviation		r_{subtest}		r_{total}		Difficulty	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
1	1.27	1.20	.59	.61	.43	.52	.49	.62	.34	.30
2	1.17	1.18	.59	.64	-.03	.45	.03	.43	.27	.30
3	.15	.18	.36	.39	.59	.54	.47	.58	0.0	0.0
4	.32	.30	.47	.46	.03	.12	.09	.11	0.0	0.0
5	.49	.33	.71	.62	.65	.76	.52	.66	.12	.08
6	.95	.95	.63	.78	.54	.35	.40	.39	.17	.28
7	.39	.25	.59	.49	.49	.62	.50	.62	.05	.03
8	.42	.25	.59	.44	.59	.72	.43	.62	.05	0.0
9	.17	.30	.50	.52	.54	.46	.42	.35	.05	.03
10	.27	.30	.50	.56	.46	.44	.33	.36	.02	.05

TABLE E12

Item analysis information for the Cooking and Home Management subtest
of the RCL (1 and 2)

Item	Mean		Standard Deviation		r_{subtest}		r_{total}		Difficulty	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
1	.66	.68	.66	.80	.62	.65	.39	.69	.10	.20
2	1.00	.98	.71	.80	.50	.44	.34	.48	.24	.30
3	.27	.25	.45	.49	.40	.75	.31	.56	0.0	.03
4	.78	.68	.69	.69	.61	.46	.50	.58	.15	.13
5	.20	.13	.46	.34	.31	.50	.41	.36	.02	0.0
6	.61	.55	.63	.68	.46	.20	.64	.37	.07	.10
7	.95	.95	.71	.64	.02	.36	.14	.44	.22	.18
8	.98	.85	.65	.62	.50	.46	.33	.39	.20	.13
9	.29	.20	.51	.46	.42	.50	.37	.39	.02	.03
10	.05	.15	.22	.36	.17	.24	.33	.18	0.0	0.0

TABLE E13

Item analysis information for the Communication subtest of the RCL (1 and 2)

Item	Mean		Standard Deviation		r_{subtest}		r_{total}		Difficulty	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
1	1.24	1.23	.66	.66	.67	.61	.59	.40	.37	.35
2	1.68	1.65	.61	.58	.44	.65	.59	.49	.76	.70
3	1.63	1.60	.58	.71	.73	.67	.60	.49	.68	.73
4	1.20	1.13	.72	.72	.64	.58	.55	.64	.37	.33
5	1.39	1.38	.67	.71	.57	.69	.46	.60	.49	.50
6	.90	.73	.77	.82	.64	.72	.59	.68	.24	.23
7	1.10	1.00	.70	.75	.45	.51	.44	.43	.29	.28
8	.27	.30	.55	.61	.52	.56	.40	.45	.05	.08
9	.54	.48	.71	.75	.40	.54	.44	.48	.12	.15
10	.49	.48	.60	.64	.42	.45	.41	.29	.05	.08

TABLE E14

Item analysis information for the Consideration subtest of the RCL (1 and 2)

Item	Mean		Standard Deviation		r_{subtest}		r_{total}		Difficulty	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
1	1.71	1.70	.56	.46	.42	.34	.41	.43	.76	.70
2	1.54	1.65	.55	.58	.51	.53	.31	.43	.56	.70
3	1.46	1.53	.64	.60	.63	.66	.17	.45	.54	.58
4	1.37	1.53	.62	.55	.33	.58	.22	.39	.44	.55
5	1.56	1.50	.67	.75	.59	.66	.31	.43	.66	.65
6	1.29	1.38	.75	.63	.40	.66	.26	.50	.46	.45
7	1.37	1.48	.66	.68	.40	.60	.09	.32	.46	.58
8	1.42	1.45	.59	.60	.42	.67	.03	.42	.46	.50
9	.90	1.20	.70	.72	.05	.44	.04	.49	.20	.38
10	.98	1.05	.72	.75	.11	.37	.36	.58	.24	.30

TABLE E15

Item analysis information for the Getting Friends subtest of the RCL (1 and 2)

Item	Mean		Standard Deviation		r_{subtest}		r_{total}		Difficulty	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
1	1.83	1.83	.38	.39	.64	.34	.47	.10	.83	.83
2	1.88	1.90	.33	.30	.41	.36	.48	.51	.88	.90
3	1.81	1.75	.40	.49	.52	.51	.53	.38	.81	.78
4	1.20	1.23	.78	.77	.39	.51	.47	.51	.42	.43
5	1.32	1.23	.61	.58	.49	.26	.38	.39	.39	.30
6	1.39	1.20	.67	.69	.51	.64	.65	.61	.49	.35
7	1.12	.78	.68	.70	.50	.52	.34	.54	.29	.15
8	1.27	1.33	.63	.62	.39	.44	.29	.32	.37	.40
9	1.66	1.70	.53	.52	.44	.26	.45	.23	.68	.73
10	1.15	1.03	.65	.77	.44	.16	.35	.22	.29	.30

TABLE E16

Item analysis information for the Keeping Friends subtest of the RCL (1 and 2)

Item	Mean		Standard Deviation		r_{subtest}		r_{total}		Difficulty	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
1	1.51	1.48	.60	.60	.23	.25	.23	.41	.56	.53
2	1.27	1.28	.59	.60	.18	.62	.14	.46	.34	.35
3	.85	.93	.62	.69	.22	.25	.20	.11	.12	.20
4	1.02	1.13	.47	.52	.09	.35	.03	.12	.12	.20
5	1.00	1.08	.71	.62	.40	.20	.46	.41	.24	.23
6	.83	.88	.67	.65	.39	.30	.32	.35	.15	.15
7	1.15	1.00	.62	.64	.15	.45	.14	.17	.27	.20
8	.66	.78	.62	.70	.32	.47	.40	.41	.07	.15
9	.98	.90	.69	.67	.47	.44	.30	.27	.22	.18
10	.81	.83	.64	.59	.52	.62	.34	.47	.12	.10

TABLE E17

Item analysis information for the Handling Problems subtest of the RCL (1 and 2)

Item	Mean		Standard Deviation		r_{subtest}		r_{total}		Difficulty	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
1	1.71	1.65	.46	.48	.35	.62	.45	.44	.71	.65
2	1.07	.93	.69	.66	.59	.59	.45	.52	.27	.18
3	.73	.80	.50	.61	.73	.70	.58	.63	.02	.10
4	.93	1.00	.65	.75	.57	.50	.40	.62	.17	.28
5	.78	.65	.61	.58	.40	.46	.40	.64	.10	.05
6	1.05	1.08	.59	.62	.45	.37	.27	.42	.20	.23
7	1.51	1.60	.51	.50	.43	.41	.45	.33	.51	.60
8	1.17	1.05	.63	.64	.41	.29	.34	.25	.29	.23
9	.88	1.05	.60	.75	.41	.22	.23	.16	.12	.30
10	.78	.58	.61	.55	.57	.42	.58	.36	.10	.03

Appendix F

Examples of Reported Incidents of Maladaptive Behaviors
and the Domain (ABS Part II) in which they were Classed

Violent and Destructive Behavior

1. M.E. slapped P.S. across the face.
2. J.C. broke 'another' ashtray in the dining room.
3. D.L. tore the buttons off the chesterfield in the lounge.
4. L.S. punched a resident for using the bathroom (L.S.) had just cleaned.

Antisocial Behavior

1. J.C. noisy until 1 a.m.
2. J.C. playing radio loudly at bedtime and disturbing everyone.
3. L.S. upset, used very abusive language.

Rebellious Behavior

1. J.C. made unauthorized long distance call 'again'.
2. K.B. refused to eat diet desert.

Untrustworthy Behavior

1. P.W. using other people's radios without permission.
2. J.L. took C.E.'s ring.
3. A.L. caught going through K.L.'s drawers.
4. P.W. took S.M.'s rings and sweater.

Self-Abusive Behavior

1. J.S. scratched hands till bleeding.

Hyperactive Tendencies

1. J.C. hyperactive all evening.
2. J.C. would not stop talking all evening.

Sexually Aberrrrant Behavior

1. J.L. caught in lounge with his hands where they shouldn't be on a female resident.

Psychological Disturbances

1. M.M. went to R.C.M.P. in Courtenay to get himself committed; agitated.
2. L.S. upset, crying; said everyone hates her.
3. S.F. faked a seizure.

TABLE F1

Reported incidents of Maladaptive Behavior, classified according
to the domains of the ABS Part II

Domain	Mean No. of Incidents	Total No. of Incidents	No. of Person in NIN	No. of Person in LIN	No. of Person in HIN
Violent and Destructive Beh.	2.37	45	32	12	7
Antisocial Beh.	3.25	39	39	8	4
Rebellious Beh.	3.00	33	40	7	4
Untrustworthy Beh.	2.17	13	45	5	1
Withdrawal	0	0	51	0	0
Stereotyped Beh. & Odd Man.	0	0	51	0	0
Inappropriate Inter- Personal Manners	0	0	51	0	0
Unacceptable Vocal Habits	0	0	51	0	0
Unacceptable or Eccentric Habits	0	0	51	0	0
Self-Abusive Beh.	4.00	4	50	0	1
Hyperactive Tend.	3.50	7	49	1	1
Sexually Aber. Beh.	2.00	2	50	0	1
Psychol. Disturb.	2.23	29	38	10	3

TABLE F2

Mean scores for no, low and high incidents groups on subtests of the ABS Part II and on reported incidents of Maladaptive Behavior

Subtest	ABS (1)			ABS (2)			Critical Incidents		
	NIN	LIN	HIN	NIN	LIN	HIN	NIN	LIN	HIN
Violent and Destructive Beh.	.83	2.73	1.86	.90	1.58	2.17	0.0	1.08	4.57
Antisocial Beh.	2.32	3.63	13.67	3.78	3.75	3.00	0.0	.73	7.00
Rebellious Beh.	2.89	3.00	11.67	2.45	5.86	4.67	0.0	1.29	6.00
Untrustworthy Beh.	.31	3.00	16.00	.50	1.20	6.00	0.0	1.20	7.00
Psychological Dist.	3.94	5.22	16.67	6.22	7.30	4.00	0.0	1.20	5.67

TABLE F3

Combined mean scores for no, low and high incidents groups
on subtests of the ABS Part II for samples one and two.

Subtest	ABS (1) and (2) Combined		
	NIN	LIN	HIN
Violent and Destructive Behavior	.87	2.13	2.00
Antisocial Behavior	3.05	3.69	8.34
Rebellious Behavior	2.67	4.43	8.17
Untrustworthy Behavior	.41	2.10	11.00
Psychological Disturbance	5.06	6.31	10.34

Correlations of subtests of the AFI and ABS Part I that were combined in the multitrait-multimethod matrices

TABLE F4

AFI	Lang. Dev. ABS (1)	Lang. Dev. ABS (2)
Reading	.64	.64
Writing	.66	.65
Communication	.59	.66

TABLE F5

AFI	NTC ABS (1)	NTC ABS (2)
Numbers	.49	.51
Time	.55	.46

TABLE F6

AFI	App. + Eat. AFI 3(1)	App. + Eat. AFI 3(2)
Appearance	.48	.42
Eating Skills	.54	.41

TABLE F7

ABS	RM. MAN. AFI 3(1)	RM. MAN. AFI 3(2)
Cleaning	.42	.64
Care of Clothing	.15	.56

TABLE F8

ABS	CK. HM. MAN. AFI 3(1)	CK. HM. MAN. AFI 3(2)
Kitchen Duties	.47	.27
Occ. Dom. Gen.	.38	.26

TABLE F9

ABS	IND. AFI 1	MAK. DEC. AFI 1	IND. AFI 2	MAK. DEC. AFI 2
Initiative	.35	.42	.51	.32
Persistence	.42	.40	.63	.39
Gen. Self. Dir.	.25	.18	.48	.16

TABLE F10

AFI	SOC. ABS 1	SOC. ABS 2
Communication	.15	.35
Consideration	.09	.31
Getting Friends	.42	.45
Keeping Friends	.32	.20
Handling Prob.	.37	.26

TABLE F11

AFI	OCC. GEN. 1	OCC. GEN. 2
Use, Care Equipment	.18	.56
Task Analysis	.17	.24
Punctuality	.22	.27