5-1-1974

The Use of Multidimensional Perceptions as Predictors of Successful Adjustment for Mentally Retarded Adults and Psychiatric Patients

David L. Hanson

THE USE OF MULTIDIMENSIONAL PERCEPTIONS AS PREDICTORS OF SUCCESSFUL ADJUSTMENT FOR MENTALLY RETARDED ADULTS AND PSYCHIATRIC PATIENTS

by

David L. Hanson

Bachelor of Arts, Concordia College, 1967
Master of Arts, University of North Dakota, 1971

A Dissertation

Submitted to the Graduate Faculty

of the

University of North Dakota

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Grand Forks, North Dakota

May
1974

This dissertation submitted by David L. Hanson in partial fulfillment of the requirements for the Degree of Doctor of Philosophy from the University of North Dakota is hereby approved by the Faculty Advisory Committee under whom the work has been done.

(Chairman)

Alice T. Clark

Alfred M. Miranda

John O. Hall

William Johnson
Dean of the Graduate School

Permission

Title THE USE OF MULTIDIMENSIONAL PERCEPTIONS AS PREDICTORS OF SUCCESSFUL ADJUSTMENT FOR MENTALLY RETARDED ADULTS AND PSYCHIATRIC PATIENTS

Department Psychology

Degree Doctor of Philosophy

In presenting this dissertation in partial fulfillment of the requirements for a graduate degree from the University of North Dakota, I agree that the Library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by the professor who supervised my dissertation work or, in his absence, by the Chairman of the Department or the Dean of the Graduate School. It is understood that any copying or publication or other use of this dissertation or part thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of North Dakota in any scholarly use which may be made of any material in my dissertation.

Signature _____

Date _____

ACKNOWLEDGMENTS

I wish to thank Dr. LeRoy A. Stone, a MESA developer, for his extensive help and encouragement in the preparation of this dissertation. His recommendations regarding this methodology were very valuable.

I wish to thank the following for their help in an advisory capacity:

Dr. Alice T. Clark
Dr. John O. Noll
Dr. John D. Tyler
Dr. Alfred M. Mirande

The technical advisors and participants in this study were very helpful and cooperative. I would like to express a special thanks to the following:

Col. Roy Kimbrell
Mr. Allen C. Lindem
Dr. Richard Listiak
Mr. Kenneth Ose
Mr. Ronald Sanden
Mr. Paul Witucki

Finally, I grant a special thanks to my wife, Sharon, for her encouragement and typing assistance.

TABLE OF CONTENTS

LIST OF TABLES

ABSTRACT

Six investigations were conducted to determine the efficacy of employing modifications of the Multidimensional Evaluation Structure Analysis (MESA) methodology to the predictive task of assessing potential vocational outcome for mentally retarded adults and psychiatric patients. Five of the investigations dealt with mentally retarded adults who were seeking or had obtained vocational placement positions. A sixth investigation dealt with psychiatric patients (inpatients and outpatients) and their probable community adjustment success.

Each of the six subject groups (referred to as stimulus-persons) were evaluated by their respective judge group from four different agencies. The SS-1 and SS-2 stimulus-person groups (N = 10 and 11, respectively) were institutionalized at a State School for the Mentally Retarded. The OTC-1 and OTC-2 stimulus-person groups (N = 15 and 11, respectively) were clients at an Opportunity Training Center facility designed to train mentally retarded persons in occupational and self-help skills. The SWS stimulus-person group (N = 8) were clients of a Sheltered Workshop designed to train mentally retarded persons in skills leading to competitive employment. The PU stimulus-person group (N = 14) was selected from a psychiatric unit functioning to assess and treat persons with psychiatric problems. The number of judges used for each investigation ranged from three to nine.

Predictor values were derived from factor loadings extracted by the MESA methodology. MESA is a "content" scaling methodology

that requires a judge to estimate the perceived degree of similarity (using values from zero to 100) between two stimuli. In these investigations, the stimuli were persons (called stimulus-persons) in the groups described above. The judge groups were asked to evaluate their respective stimulus-person group. Also, the judges were asked to indicate on a five-point scale how familiar they were with each stimulus-person.

For the SS-2, OTC-2, SWS, and PU groups, the judges were also asked to estimate the probable degree of success (using a value from zero to 100) that each stimulus-person might experience when placed in the community. The probability values and MESA predictor values were later correlated with follow-up data collected more than one year after the judgments were made.

Two MESA computational routines were employed including the Group Composition Structure Analysis (GCSA) and the Observer Factor Judgment Analysis (OFJA). The GCSA found underlying judgmental criteria used by the judges where the OFJA found clusters of judges demonstrating a commonness of judgmental strategy. MESA predictor values were formulated from GCSA factor loadings and were modified with regard to how much variance each factor (judgment-dimension) accounted for and by how familiar each judge was with the stimulus-persons. The predictor values were correlated with follow-up data. These relationships reflected the efficacy of using the MESA methodology in assessing clinical judgment and staff consensus as predictors of success outcomes for individuals in groups.

Three studies demonstrated that clinical judgment (using MESA and similarity estimations) can adequately predict future success with regard to the samples investigated. The three studies showing little or no predictive success had judges that used different judgmental strategies, as shown by the OFJA analysis.

The six investigations demonstrated that clinical judgment allows for adequate prediction provided the clinical judgments are handled in an objective and systematic way. Using MESA as a predictive format also allows the investigator to determine when prediction will probably fail based on the OFJA analysis. The study demonstrated a methodology that may be used to systematically assess staff consensus and individual clinical judgment in predicting outcome for a variety of groups.

CHAPTER I

INTRODUCTION

There are a variety of diagnostic terms and developmental concepts implying the existence of a human population whose major characteristic appears to be the inability to adequately maintain themselves outside a sheltered environment. The diversity of views depicting this population range from the more traditional treatment of mental retardation as a disease (Robinson and Robinson, 1965) to the views of Braginsky and Braginsky (1971) who challenge the "illness" concept and argue that mental retardation is, in many ways, a myth. In addition, Tarjan, Wright, Eyman and Keenan (1973) claim that " . . . about two-thirds of the individuals diagnosed as retarded lose this label during late adolescence or early adulthood" (p. 372). Whatever the academic outcome of the diagnostic issue, the fact remains that there are individuals who behave in ways that make it difficult for them to gain competitive employment. For the purpose of this study, the term "mentally retarded person" will be used since it is more widely known and has attained a rather high degree of acceptability.

A number of studies (Cobb, 1968; Eagle, 1967; Lui, 1963; Magnifico and Doll, 1962; Nixon, 1970; Rosen, Kivitz, Clark and Floor, 1970; Skaarbrevik, 1971; Windle, 1962; and others) have clearly shown that mentally retarded persons are able to do productive and financially rewarding work. Some of these investigators maintain, however,

that specialized evaluation and training are necessary in order to maximize the probability of their employment opportunities.

Rosen et al. (1970) stated that progress in rehabilitation has changed attitudes to such a large degree that the treatment of mentally retarded people, emphasizing employability, has also changed drastically. Apparently this population is being viewed as more of an economic asset as opposed to a rather costly institutional liability. For example, it appears to be profitable to train handicapped people (including the mentally retarded) to become productive and independent workers where failure to do so increases the prospect of their life-long dependence on tax funds. Two studies (Conley, 1971; Mars, 1967) have shown that financial investment in habilitative programs can produce a profit. This can occur when the money used is later returned in the form of taxes collected from habilitated workers and from the absence of expense for their institutional dependence. Moral issues aside, the monetary benefits of habilitating mentally retarded people (among others) appear to be substantial.

The availability of jobs for mentally retarded people is an important issue since employment options are more restricted for this population due to their limited abilities. The view that our technological advances have decreased the availability of unskilled jobs for the mentally retarded person has been supported by Skaarbrevik (1971) and Komisar (1966) but refuted by Nixon (1970). Komisar (1966) reported that mentally retarded people are not absorbed into technological agricultural communities in the United States as is common in the less advanced rural European countries. However, in

a rather thorough review of the research in this area, Nixon (1970) concluded that:

> Regardless of gaps and limits in our research and data, existing evidence does not justify the conclusion that technological change is changing worker requirements in such a way as to reduce significantly the opportunities for rehabilitation of the mentally retarded. The changes in our technology may actually increase such job opportunities (p. 155).

If this is true, the time may come when mentally retarded people regularly fill a void in the job market for the variety of occupations for which they are well suited. The need for such occupations appears to be increasing since Tarjan et al. (1973) reported that community based vocational programs for mentally retarded people appear to be the wave of the future, acknowledging that institutional release efforts have at least tripled in the past two decades. Vocational placement, then, is and will continue to be an area of importance and worthy of continued study.

The purpose of this study was to demonstrate a method where subjective clinical judgments could be objectively handled to predict future vocational success of mentally retarded persons. It is doubtful that such a method will replace the subjective decision making process primarily used by professionals at the present time. However, the method described in this paper may later be used as a supplemental method in that subjective process. This method was a product of past research efforts and employed a recently proposed multidimensional scaling methodology.

CHAPTER II

REVIEW OF THE LITERATURE

Part I

Prediction of Vocational Habilitation

The literature reporting the community adjustment problems experienced by mentally retarded adults has increased in recent years. The habilitation of this population, focusing on vocational employment, has seemingly been part of the general "non-institutional" effort that has recently characterized the treatment of psychiatric patients.

In Windle's (1962) review of research in vocational habilitation of mentally retarded persons, the prevalent attitude that this group could not be trained was questioned and refuted. As a result, he stressed the need for studying employability factors and suggested that the best method was that which employed statistical significance for decision making. Also, he emphasized the use of generalization and synthesis where diverse findings were found. For example, Windle cited a number of studies where specific and oftentimes conflicting results were obtained. One of his general conclusions implied that statistically predictive studies may not completely replace the ongoing use of clinical judgment but that "basic data" is needed which can be of use in clinical judgment applications. It was suggested that such basic data is most readily derived from experimental

and correlational studies of variables that affect the successful vocational adjustment of mentally retarded adults. Many of the investigators to be mentioned have apparently followed this suggestion.

Eagle (1967) reviewed a number of studies and cited conflicting findings regarding the importance of the age at the time of admission, the effect of institutionalization, the type of preadmission home, the value of previous work experience, the effect of family factors, sex, intellectual level, and personality. Though many of the research findings were conflicting, he summarized by noting that some generalizations could be made though even these were of little value in attempting to predict the probable vocational success for the individual mentally retarded person. Also, Eagle specifically studied the factors involved in vocational adjustment failure and argued for more studies which could lead to a decrease in the institutional re-admission rate, which in one study (where data from various studies was evaluated) appeared to be about 52 percent (Eagle, 1967, p. 241).

A comprehensive review of predictive studies regarding the vocational success of mentally retarded adults was written by Cobb (1968). His summary of research findings concluded that: (1) there is no possibility, to date, of using a simple formula, based on psychometric data, that lends itself to accurate prediction, (2) there is a need to conduct follow-up studies on those studies already conducted rather than continually generate new ones, (3) there is a need to include personality factors in predicting successful vocational adjustment, and (4) there is clear evidence that adult adaptation and vocational adjustment is multidimensional in nature. Other, and more specific,

findings were also summarized in his review. In essence, Cobb's work acknowledged the confusing status of predictor variables as they relate to the vocational adjustment of mentally retarded adults.

A more recent review, though not as complete, was conducted by Hanson (1971). The review grouped vocational adjustment factors into community, family background, institutional, job readiness, and personal factors. He also concluded that the prediction of successful community adjustment was extremely complex and that there were no specific factors (objectively or subjectively assessed) known that could accurately predict the adjustment outcome of mentally retarded adults when placed in the community.

Predictor and Criterion Variable Studies

Several investigators (Cobb, 1968; Eagle, 1967; Windle, 1962) have reviewed studies relating measures of ability or demographic data to various measures of vocational "success-failure" as the criterion variable. Several important and more recent studies will be mentioned which specifically attempted to relate predictor and criterion variables using multivariate analysis and a variety of tests and rating scales.

Studies Using Multivariate Analysis

A predictive study was conducted by Song and Song (1969) and " . . . clearly demonstrated that in order to obtain a higher level of prediction, one should not consider several variables simultaneously" (p. 570). For example, the investigators determined that vocational placement success was associated with a higher age, a longer period

of institutionalization, the white race, better work habits and lower scores on the Information, Comprehension, and Similarities subtests of the WAIS. In addition, they argued for the value of the Intelligence Quotient (IQ) score as a "good" predictor, stating that intelligence appears to be the " . . . most important predictor when other relevant factors are taken into consideration" (p. 570). Madison (1964) suggested that all available data be "pooled" to find those variables with the most predictive power where Song and Song (1969) suggested that the mass of prediction variables used in studies like their own be factor analyzed to reduce the number of predictor variables.

A comprehensive multivariate study that reduced predictor variables was conducted by Stephens (1964) and is described and critically evaluated by Cobb (1968). She used 141 criterion variables (80 continuous and 61 dichotomous) descriptive of successful adjustment and 85 predictor variables drawn from 17 different test instruments. Using factor analysis, she reduced the number of predictor variables. The resultant analysis pointed out, among other things, that "success" was a very ambiguous concept in that there were a variety of personal criterion conditions that were termed "equally" successful and yet vastly different. Also, as Cobb (1968) pointed out, her suggestion that successful adjustment could be a profile representation rather than a "point" representation seemed to be a good one. However, Cobb was critical of the "hope" that Stephens (1964) had of using her adjustment battery of tests for predicting success in that no cross-validation study was attempted

and, indeed, would be difficult given the complex nature of the Stephen's data.

Another factor analytic reduction of predictor variables was made by Rosen et al. (1970). They acknowledged the variety of inconsistent findings found in the literature and argued that such inconsistency may have resulted from the absence of good criterion variables that are used to develop suitable prediction variables. Their effort attempted to operationally define criterion variables so that the investigator would be able to further study the problem with standardized measures of "success adjustment." The study employed 29 demographic, psychometric, and behavioral measurements as predictor variables and 22 criterion variables derived from structured interviews and adjustment measures. Factor analysis of both predictors and criterion measures yielded five Predictor Factors and six Criterion Factors (and measurement items for each) which suggested " . . . the possibility that accurate decisions regarding selection from discharge can be made from psychometric scores and assessment of work performance within the institution" (p. 733). They suggested caution, however, until proposed cross-validation studies were completed.

A multiple regression study was conducted by Kraus (1972) using nine objective predictor variables and eight objective criterion variables. He found that living in the community under supervision was the best possible situation needed to develop social adjustment. The study apparently confirmed and refuted a number of previous findings dealing with relationships between pre-placement characteristics and post-placement "success-failure."

A predictive study was conducted by Erickson (1966) and employed a number of predictors including objective test data and "teacher prognosis" ratings, among many others. "Success" was assessed by judgments made by four judges employed in the program. The ratings were then related to the varying combinations of predictor variables. The Adjective Checklist, the Wechsler Adult Intelligence Scale (WAIS) Picture Arrangement subtest, the WAIS Performance IQ, the Stanford Arithmetic Achievement Test, and the Steadiness Test were found to be the best five variables, of the 15 originally used, that related to the favorable ratings of success ($\underline{R}$ = .68). However, no cross-validation of the beta weights derived from the regression analysis was undertaken which can be considered one weakness of the Kraus and Erickson studies.

A study using multiple regression that included a cross-validation procedure was that completed by Comstock (1971). Using the Peabody Picture Vocabulary Test, the Ravens Progressive Matrices, the Vineland Social Maturity Scale and the subject's reaction time, Comstock found a high multiple correlation coefficient ($\underline{R}$ = .75) with criterion values derived from an employability model. The multiple correlation coefficient from the cross-validation was lower ($\underline{R}$ = .62) based on 20 subjects who were randomly selected from the original 81 subjects. As one would expect, the multiple correlation coefficient of the cross-validation sample was lower than the coefficient derived from the first sample. However, it could still be considered moderately high. A great deal of the variation is left unexplained with such a coefficient and the resultant beta weights may prove to lack the desired accuracy as predictors of success of an individual case.

A rather lengthy longitudinal study (from four to 10 years) was conducted by Hoffman (1969) on 569 "mentally handicapped" adults who were discharged into the community. The criterion variables were the absence or presence of police contact, the self-support ability, the wage level, the stability of interpersonal ties, and the marital status. A number of predictor variables (126) were used from data collected prior to discharge and some predictive relationships were found, using multiple regression. Hoffman concluded that on the " . . . basis of longitudinal independent variables available at the time of discharge, it would seem that a fairly effective degree of prediction is possible for the several types of post-discharge behavior recorded some 4 to 10 years after discharge" (p. 45).

Several statements can be made regarding the studies discussed above. The need for cross-validation studies of proposed predictor variables is evident. The effort by Comstock (1971) appeared to be a good beginning and when more predictor variables can be shown to demonstrate cross-validated predictive value, the quest for a general prediction scheme will be advanced.

The multivariate regression analyses have the advantage of objectively combining a number of variables to derive an optimal set of beta weights for prediction. However, it is doubtful that any general linear equation will be able to predict vocational success-failure unless individual differences (such as personality variables), differential training experiences, and the different vocational skills required for jobs are considered. For example, the best linear equation derived from the most recent evidence regarding predictor variables may suggest that Subject A has a high probability of succeeding

on vocational placement. However, the individual's fondness for the outdoors may increase his success as a construction worker while he may be a failure as a dishwasher. In this case, the need for matching the adult to a vocation for which he is well suited would be important. This argument has been stressed by some investigators (Bae, 1968; Engel, 1952; Jackson and Butler, 1963). However, no study reviewed here using regression analysis has attempted to investigate the importance of compatibility between the job placement and the mentally retarded adult.

The Use of Tests and Rating Scales

Tests and rating scales have been employed in an attempt to isolate factors that will predict future vocational success for mentally retarded adults. Also, a variety of tests and rating scales that attempt to measure the degree of successful community adjustment for use as criterion measures have been employed by some investigators.

Assessment and Potential Success. Tests and rating scales have been used or devised to assess the potential for success before actual job placement is decided by staff personnel. Many studies have been and will be discussed that use measures originally devised for purposes such as measures in intelligence, social competence and other attributes. Still other tests and rating scales have been devised specifically for measurement of potential job success of mentally retarded adults.

A study was conducted by Engelhardt (1970) using the Detroit Tests of Learning Aptitude (comprised of 10 intellectual functioning

measures). Included in the investigation was a cross-validation procedure and as a result of his analysis, Engelhardt stated that it was possible " . . . to identify the potential rehabilitant or nonrehabilitant, with 75% accuracy, at time of placement in a special education program for socially maladjusted boys" (p. 548). Although the study used a population of "socially maladjusted" boys, it appeared that at least some of the studies using "mentally retarded" persons dealt with those that could have been described as socially maladjusted.

The Elkin (1968) study used a number of predictive variables and correlated them with actual work performance, where no personality or demographic variables were used. She found that the Social Knowledge Scale was significantly related to male, but not female, performance on the job. Also, Elkin found that " . . . the O'Conner Finger Dexterity Test was significantly related [$r = .58$, $p < .01$] to performance on the domestic jobs" (p. 537).

A variety of other tests designed for specific purposes, yet applied to the predictive task of selecting probable successful mentally retarded adults for job placement, have been used in a number of studies. Prominent in the array of such tests has been the Wechsler Adult Intelligence Scale (Clark and Foster, 1970; Fry, 1956; Jackson and Butler, 1963; Johnson, 1970; Kaufman, 1970; Larson, 1964; Rosen et al., 1970; Stephens, Peck, and Veldman, 1968) and other IQ measures (Barrett, Relos, and Eisele, 1965; Bernstein, 1969; Hartzler, 1953; Kraus, 1972; Madison, 1964; Tarjan, Dingman, Eyman and Brown, 1960). In addition, the Illinois Test of Psycholinguistic abilities (Clark and Foster, 1970), the Peabody Picture

Vocabulary Test (Comstock, 1971), the Ravens Progressive Matrices (Comstock, 1971), the Vineland Social Maturity Scale (Comstock, 1971; Johnson, 1970) and the General Aptitude Test Battery (Hendle, 1971) have been used as measures for prediction with varying degrees of success. Other, less popular tests, have also been used for this purpose.

The Laradon Occupational Success Predictive Battery is comprised of 17 tests developed from 101 tests of aptitude and personality variables. A cross-validation of its predictive ability was completed by Mars (1966) on 60 mentally retarded adults with a resultant multiple correlation of $\underline{R} = .54$. The investigator suggested that the battery was " . . . somewhat effective" (p. 48) but expressed disappointment regarding such a low multiple correlation.

Brolin (1972) conducted a study of 193 former clients of an evaluation center to determine if rehabilitation services improved client vocational success and what variables contributed to such success. He found three instruments that were appropriate for use as evaluators of vocational potential for mentally retarded adults. The three were the Vocational Capacity Scale (VCS), the Prevocational Laboratory Summary (PVLS), and the Client Evaluation Report (CER). Brolin reported that " . . . since many of the demographic client variables were not generally found significant to the client's rehabilitation, more emphasis should be directed on external factors like the family and community--and particularly the client's job environment" (p. 650). A cross-validation of the three measures was suggested in order to validate the predictive value, particularly for the VCS since it was found to be the measure most related to vocational success.

A rating scale technique was employed by Warren (1961) and the factor structure of the scale was found to be Personality and Social Adjustment, Work Habits and Efficiency, and a General Factor. Though no cross-validation study was attempted, Warren (1961) stated that " . . . the rating scale is a useful predictor of potential employment" (p. 633).

Assessment of Actual Success. Daniels and Stewart (1971) devised a scale called the Vocational Adjustment Rating Scale (VARS) that consists of 62 items to be used as a checklist. It has four to five response categories for each item. The VARS instrument has two parts that can be used independently where one part assesses factors indirectly related to work success and the other part assesses factors directly related to work success. The authors claimed that the instrument is reliable (having high inter-rater reliability) and has acceptable content validity.

Another checklist rating scale of criterion variables is the San Francisco Vocational Competency Scale (Levine and Elzey, 1968). This rating scale can be used by persons familiar with the ability and work habits of the mentally retarded adults on job placement and consists of 30 items. The scale is relatively simple to score and Downie (1969) reported that the scale may be of some value to the vocational counselor interested in the assessment of actual job performance.

There are only a few rating scales designed to assess vocational placement success of mentally retarded adults. Most likely, the demand for such scales is small since criterion measures of income, work stability, and others are considered adequate measures of actual

success, thus minimizing the need for scores which tend to give abstract measures of success.

The use of tests and rating scales for objectively assessing the actual and potential vocational placement success of mentally retarded adults has been extensive. The rationale for such attempts is varied and oftentimes such measures show empirical relationships where no logical relationship or face validity is evident. Criticism of the use of tests and rating scales is included with an overall criticism of the general predictive strategy oftentimes directed toward this body of research.

The Predictive Strategy: Problems and Criticism

The use of objective tests and rating scales as predictor variables for criteria such as successful community adjustment is a common methodology in the studies reviewed thus far. There are, however, a number of objections to the use of such predictor variables, even in cases where predictive success was demonstrated. In Cobb's (1968) comprehensive review of the literature, he stated clearly that no simple formula existed to date that could accurately predict differential success-failure outcomes for mentally retarded adults after being placed in the community. In an earlier article, Cobb (1967) stressed that it was not " . . . that [test scores] are without value, but [that] alone they do not give us proof of future outcome" (p. 23).

Predictor Variable Problems

Cobb and Epir (1966) reported that under a few broad classifications including formal test measures, background history factors, and observational factors, it has been found that " . . . literally

hundreds of variables have been recorded and a very large proportion of them show some sort of relationship with some criterion variable" (p. 11). The question of which variables are the most important is difficult since all predictor variables have drawbacks restricting their ability to predict criterion outcomes.

Gardner and Giampa (1970) stated that "unidimensional" measures were inadequate as predictor variables. After supporting the hypotheses that " . . . inappropriate social and emotional behavior is independent of behavioral competence," they maintained that the " . . . unidimensional measures" (e.g. [intelligence quotient] and [social quotient], currently in use are inadequate to the task of individual therapeutic programming" (p. 169). Their argument assumed that success in training programs was a function of behavioral control and competence, which they claim to have demonstrated. Cobb (1968) also argued that community adaptation is multidimensional in nature and agreed that unidimensional measures are inadequate measures of adaptation, even when used in combination.

From another view, Masland, Sarason, and Gladwin (1958) have criticized the use of test measures (such as IQ tests) that were developed to "diagnose" or "define" mentally retarded people. In short, they argued that such conventional tests used to predict social and vocational success-failure have been inadequate since they were not designed to measure such outcomes and cannot be expected to produce valid predictions. Though it could be argued that such conventional test measures tap "factors" common to adaptation and vocational adjustment, the Masland et al. comment noted that the only practical

"predictive tests" would be ones that were specifically designed to quantify those personal factors that seem to be related to actual vocational success. For example, Bolton, Butler and Wright (1968) formulated four factors that seemed to relate to rehabilitation success. They were greater family responsibility, better education and vocational preparation, younger age at acceptance into the community, and better personal and social adjustment. However, if one accepts these factors as truly representative of qualities indicative of the successfully placed mentally retarded adult, which may be erroneous in view of contradictory research findings, the fact remains that objective tests of factors such as these do not exist and hence complicate the measurement problem.

Perhaps the most widely investigated predictor variable studied has been the measured IQ. In Eagle's (1967) review, he cited five studies that related lower IQ scores to vocational success, three studies that related higher IQ scores to vocational success and 20 studies that demonstrated IQ scores to have no practical predictive value. To date, no study has clearly demonstrated whether or not IQ scores are valuable predictor variables and this may be due to the fact that IQ tests were not designed as measures of vocational adjustment.

Jastak (1967) also expressed criticism of IQ scores and argued that IQ tests and related measures do not assess the "mental competence" required for successful vocational adjustment. A similar argument was advanced by Barrett et al. (1965) who reported that the IQ score does reflect the true nature of total mental ability of the retarded" (p. 106).

The general use of test scores as predictor variables was criticized by Hendel (1971) and he suggested a possible alternative research approach. He argued that the " . . . most promising approach to ability measurement of retarded individuals may be to concentrate on what happens within the testing process itself rather than on modification of the peripheral conditions such as those reported in this paper" (p. 616). He used the General Aptitude Test Battery (GATB) and found little hope for predicting vocational success in the individual case with the use of the GATB scores. His suggestion, however, reflected the finding that test taking behavior is quite different (even with identical resultant scores) and such behavior may be of use in assessing skills related to vocational success.

There are additional difficulties in establishing precisely which "mental ability" scores in a regression equation are indicative of adjustment success and which are indicative of adjustment failure. For example, Song and Song (1969) reported that scores on the Comprehension subtest of the WAIS were inversely related to good job efficiency, for mentally retarded adults. In contrast, Kaufman (1970) reported that Comprehension subtest scores were the most valuable predictor variables in his study where a high score was indicative of successful vocational employment. He stated that " . . . arithmetic functioning in addition to comprehension is a most important factor to consider in preparing mentally retarded young adults for the employment market" (p. 779). Such contradictory findings illustrate one of the difficulties in determining viable predictor variables.

In summary, the predictive strategy is usually based on predictor variables that have questionable validity, reliability, and

predictability. In addition, the fact that temporal influences may obscure the relationship between predictor variables and the current functioning of any mentally retarded person indicates that the predictive strategy requires more research effort to isolate valid predictive variables. To complicate matters further, it appears as though criterion variables (that being predicted) are not precisely measured either.

Criterion Variable Problems

The question of what "successful" vocational-community adjustment is may not be elementary. The importance of job skills, personality, and assessed behavioral-mental abilities have relevance to this question.

The need for mentally retarded adults to obtain job skills before being placed in a vocation is seldom debated. However, the precise description of what those job skills might be is lacking, as has been pointed out by a number of investigators (Cobb, 1968; Cohen, 1960; DiMichael, 1960; Engel, 1952; Jackson and Butler, 1963; Rosen et al., 1970; and others). The problem is complex since tenuous predictor variables have been used to predict somewhat intangible and poorly defined criterion outcomes. The problem is even more complex if personality variables are important ones to consider as suggested by a number of investigators (Bae, 1968; Cobb, 1968; Cohen, 1960; Engel, 1952; Huber and Soforenko, 1963; Jackson and Butler, 1963; Michal-Smith, 1950; Nihira, 1969; Song and Song, 1969; Stephens et al., 1968; and Windle, 1962). The measurement and predictive

application of personality variables would indeed increase the complexity of an already complex problem.

The description and measurement of 141 criterion variables by Stephens (1964) provides the investigator with a minimum number of criterion variables from which to choose. Most of these variables could probably fall under categories listed as social adjustment, community adjustment, vocational adjustment, personal adjustment, and others. The determination of "success" has thus been difficult since all of these criterion categories and variables of " . . . positive adaptation . . . have frequently been obscured by the criterion of merely remaining outside of the institution . . ." (Cobb, 1968, p. 139). There are other less specific problems inherent in the predictive strategy.

Other General Problems

Windle's (1962) review of the prognosis for mentally retarded people summarized the research effort in one way by stating that the findings had limited reliability and generalizability. Also, Eagle (1967) summarized his review in part by stating that "Published data on more than forty release characteristics for their prognostic value for success or non-success in community placement of previously institutionalized retardates showed wide disagreement as to their utility" (p. 241). More recently, Cobb (1968) stated that with " . . . few exceptions [the studies reviewed] constitute isolated investigations, with little attempt to develop a coherent body of knowledge, to replicate or cross-validate, to expand the basis of generalization from small local population samples to wider populations, to systematize

and standardize the measurement of independent and dependent variables, or to fit predictive investigation to coherent developmental theory" (p. 138).

It seems clear and well established that predictive studies of future vocational success and community adjustment for mentally retarded adults are still needed to isolate variables and/or procedures by which the selection process can be improved. The practical problem then was succinctly stated by Cobb (1967). He argued that "Our job in personnel selection . . . is to pick the people who will fill those slots with the maximum probability of success, and the minimum probability of failure" (p. 11). He went even further to say that:

> . . . we do not have to wait for a nice recipe book of test procedures in order to make good judgments about clients. We can use a clinical judgment, if you like, which is more qualitative than quantitative, which coordinates and puts together the objective information and general opinions estimate as to the probability of success (p. 23).

The impact and validity of this latter statement shall be considered next.

Statistical and Clinical Prediction

Cobb's (1967) stated cited above is important since the debate of "clinical versus statistical prediction" has been receiving increased attention in recent years. By "clinical judgment," most investigators would probably agree with Thorne (1960) who stated that:

> Clinical judgment is operationally defined as involving the ability to make good (sound) decisions after gathering and evaluating all pertinent evidence, weighing possible alternatives in terms of past experience or normative probabilities, and arriving at problem solutions which reflect basic science orientations (the cultural value system against which scientists operate) (p. 128).

Also, writing on the nature of clinical judgment, Bieri, Atkins, Briar, Leaman, Miller, and Tripodi (1966) stated:

> It is apparent today that whatever else it may be, the cognitive behavior which we call judgment, even in the controlled setting of the psychophysical experiment, is an intricate amalgam of cognitive processes, which may include perception, concept formation, thinking, memory, discrimination, decision making, and creative imagination (p. 5).

Though many other descriptions of clinical judgment could be made, these two statements probably satisfy the need for a definition without removing the obvious ambiguity of the term.

The statistical decision making process is quite different. Perhaps Sawyer's (1970) use of the term "mechanical" describes the process clearly by inferring that subjectivity is removed entirely and data is derived by strictly objective methods. Sawyer reported, however, that many studies attempting some form of prediction use a combination of statistical-clinical decision making process to denote decisions that were made clinically but based on psychometric data.

Sawyer (1970) examined the statistical-clinical combinations used in a variety of studies and defined eight categories ranging from "pure clinical" to "pure statistical." In addition, Sawyer made a distinction between what normally occurs in making clinical or statistical predictions and whether the data was collected (measured) in a statistical or clinical fashion. Thus, attributes could be clinically or statistically measured and combined to be used later in a clinical or statistical prediction process. To complicate matters even further, some judges are better than others (Bolton et al., 1968). Also, some statistical models for prediction are better than others, given identical input data.

Support for the use of clinical judgment as the source for the decision making process is meager. Though Cobb (1967; 1969) and Cobb and Epir (1966) supported the use of clinical judgment in the prediction of vocational success for mentally retarded adults, others find such judgment inadequate, as will be discussed shortly. Cobb (1969) stresses, however, that reliance on the vocational counselor's " . . . exercise of his best 'clinical' judgment" is due to the absence of measures for prediction possibilities (p. 147). Justman (1968) agreed and stated:

> The best that the counselor can do is gather as much relevant evidence as he can, relate whatever data he has collected about the present status of the individual on the one hand and the job requirement on the other, add a dash of intuition, and make an estimate of success (p. 153).

In addition, the efficacy of clinical judgment was partially promoted when Sawyer (1970) stated that:

> There is good reason to think that incorporating a clinician into data collection might improve prediction. The psychometric test may fail to provide an appropriate response alternative, or may fail to ask the most crucial question in the first place. The clinician, though subject to similar errors, is favored by feedback; in effect collecting and analyzing data at the same time, he can adjust as he goes (p. 78).

Clinical judgment, however, is seldom investigated in the majority of applied settings. The "feedback" that Sawyer (1970) mentions or the "evaluating" that Thorne (1960) mentions are normally encountered in research projects and rarely encountered in the applied vocational placement setting. When the use of clinical judgment has been adequately studied, many investigators find the process less than adequate to the task of prediction.

Critics of clinical judgment methods used in prediction generally point to errors in clinical prediction as related to the diagnosis or prognosis judgments in the clinical setting. In the area of vocational rehabilitation for mentally retarded adults, three studies (Rosen et al., 1970; Fields and Gibson, 1972; Gibson and Fields, 1970) reflect the errors in judgment that are often made.

Rosen et al. (1970) studied 29 demographic, psychometric, and behavioral measures as they related to 22 criterion variables of community functioning for mentally retarded adults. They found that " . . . while institutional raters were positively influenced by the subject's verbal intelligence, verbal ability was not an important correlate of later community judgment" (p. 731). Generally, they found that non-verbal abilities were better predictors of successful vocational-community adjustment. The finding reveals that clinical judgment may be based on "intuitive" relationships that may have no research basis. A good example could include the "intuitive" notion that higher IQ scores would be indicative of success, which has not been established in the literature as discussed previously. In addition, many clinicians "intuitively" know that social intelligence or judgment can be assessed by the WAIS Picture Arrangement subtest. However, Wechsler (1958) stated that " . . . social intelligence is just general intelligence applied to social situations" and Wechsler has questionable data for even that assertion (p. 75).

Two other studies (Fields and Gibson, 1972; Gibson and Fields, 1970) specifically stated the kinds of errors that occur when clinical judgment is employed. Gibson and Fields (1970) found that " . . . expert judges [failed] to employ [in their predictions] many of the

cues they identified as important . . ." and that there was a " . . . disparity between the cues identified by the experts and those evidently used by the receiver group, e.g., employer, landlady, job mates, and other significant members of the community environment" (p. 562). In another study, Fields and Gibson (1972) found that the initial clinical prediction was for success and that a clinical prediction of failure was based on more data. In addition, they found that the clinician simply does not use all the data available, as would be employed by a multiple regression approach using all quantifiable variables.

The problem of such human judgment, in explaining clinical judgment errors, was summarized by Bolton et al. (1968). They argued that " . . . it just is not reasonable to expect the human brain to be able to analyze and correlate a set of data as efficiently as a statistical method which is designed for maximal predictive efficiency" (p. 17).

In the best of direct comparisons, the statistical method of prediction appears to be superior to the clinical judgment approach. For example, many studies (Bernstein, 1969; Bolton et al., 1968; Brolin, 1972; Cobb, 1969; Comstock, 1971; Eagle, 1967; Engelhardt, 1970; Erickson, 1966; Fry, 1956; Fields and Gibson, 1972; Gibson and Fields, 1970; Hoffman, 1969; Johnson, 1970; Kraus, 1972; Rosen et al., 1970; Sawyer, 1970; Shafter, 1957; Song and Song, 1969; Warren, 1961; Windle, 1962; and others) have implied in some way that the relationships between predictor and criterion variables could serve as objective information leading to statistical prediction of vocational success. Very few

studies, if any, maintain that subjective clinical prediction can do the same. Sawyer (1970) reviewed 45 studies which convincingly found statistical prediction superior to clinical prediction. Bolton et al. (1968) reported that statistical prediction was better than the "average" clinical prediction though some individual clinicians were quite accurate. It appears that the majority of the studies comparing statistical and clinical prediction appear to support the superiority of the statistical prediction process over the clinical prediction process.

In summary, the clinical versus statistical (and combinations thereof) debate is, as yet, not completely settled regarding vocational placement since no method or combination of methods appears to be capable of accurately predicting the vocational outcome for mentally retarded adults. As a result of Goldstein's (1964) investigation, the implication was made that neither descriptive nor psychometric data was sufficient for predicting extrainstitutional adjustment. Perhaps the most important reason for this implication is the extreme complexity and multidimensional nature of the variables, as stressed by Cobb (1968) and others.

The Madison (1964) study classified each mentally retarded patient as a success or failure following one year of temporary discharge from an institution where employment success-failure was the criterion variable. One variable, of many, found significantly related to success was the "professional staff agreement" regarding work suitability on vocational placement. The finding is important, particularly in light of a later review by Eagle (1967) where the importance of staff consensus was stressed again. Eagle stated that:

> The importance of staff consensus has not received appropriate attention or evaluation in the past, particularly since each placement failure is a failure in the collective professional judgment of all those participating in the decision to remove the specific retardate from the institution (p. 237).

The present study exploited the multidimensional nature of the predictive problem by using a multidimensional approach in the evaluation of unique clinical judgments (similarity estimations) that, in sum, represented a staff consensus like that often used in an applied clinical setting.

Part II

Multidimensional Evaluation Structure Analysis (MESA)

Multidimensional Evaluation Structure Analysis (MESA) was the multidimensional scaling methodology employed in the current investigation. MESA is a recently proposed multidimensional scaling methodology (Coles and Stone, 1972; Stone and Coles, 1970; Stone, Coles and Lindem, 1970) and Stone (1972) cited 18 studies (many unpublished) that employed the MESA methodology with excellent success. In addition, Hanson and Stone (1974) found in two of those 18 studies that MESA meaningfully determined the underlying judgment-decisions used by staff who judged a number of mentally retarded adults regarding their actual and potential job placement success.

The MESA methodology uses similarity estimations as input where zero indicates "no similarity" and 100 indicates "identity" between two stimuli. The perceived similarity is a sole function of the stimulus "content" a judge perceives though each judge may be directed to make similarity estimations regarding a specific question (e.g., "How similar are Person A and Person B with respect to intelligence?").

Obviously, other judgmental directives could be made. Even with such directives, judges use several judgmental criteria since their perception of similarity is generally considered multidimensional, rather than unidimensional, in nature. Likewise, stimuli are not unidimensional as argued by Stevens (1966) who stated that: "All acts of judgment seem to demand some degree of selective abstraction, for no stimulus is one-dimensional" (p. 391).

With MESA, each stimulus is paired with every other stimulus in the group of stimuli that is being judged (evaluated). If $m(m - 1)$ similarity estimations are desired (where m equals the number of stimuli in the group), then each stimulus-pair will be evaluated twice. For example, if 10 (m) stimuli comprise a group, then 90 similarity estimations [10 (10 - 1)] will be completed by each judge, and each stimulus-pair will be evaluated twice. In this case, the order of the two identical stimulus-pairings is reversed (S_{ij} to S_{ji}) so that all stimulus-pairs will have an S_{ij} stimulus-pair counterpart. For example, the first stimulus-person pairing may be "J. Jones and T. Doe." However, the second time this particular stimulus-person pairing appears, it would read "T. Doe and J. Jones." This also allows the investigator to estimate the reliability of the judges by correlating S_{ij} and S_{ji} similarity estimations for each pair of stimuli evaluated. If $[m(m - 1)/2]$ similarity estimations are desired, each stimulus-pair will be evaluated only once by each judge and the judges will have half the judgmental task as required in the previously described case. For example, if 10 (m) stimuli comprise a group then 45 similarity estimations [10 (10 - 1)/2] will be completed by each judge, and each stimulus-pair will be evaluated

only once. In this case, the reliability estimation can be obtained by using the "parallel-judge" method as described by Coles (1970).

Normally, the stimulus-person pairs are presented to the judges on computer cards and the similarity estimations for each stimulus-pair are written directly on the cards. A deck of computer cards, then, is required for each judge and each deck is shuffled a number of times before judgments are made to eliminate any systematic presentation of the stimulus-pairs. Thus each judge obtains a deck of computer cards where the stimulus-pairs are presented in a randomized order. After the judges have made their similarity estimations, the completed decks are collected and the similarity estimations are key-punched to expedite the data gathering and analytical processes.

The MESA methodology has two major analytical routines from which factor structures are obtained. One or both may be of interest to the investigator.

Group Composition Structure Analysis (GCSA) proceeds by averaging (arithmetic mean statistic) the similarity estimations (across judges) for each stimulus-pair evaluated. The resultant mean similarity estimation matrix (across judges) will have unity values (1.00) in the diagonal since stimulus-persons paired with themselves would have received similarity estimations of "100," (i.e., "identity") though these judgments are not made by the judges. All similarity estimations are given two decimal points for the purpose of computation, thus the value of 1.00 will be in the diagonal instead of 100.

The column vectors of the mean similarity estimation matrix are then inter-correlated (each column with every other column using the product-moment model). The resultant similarity correlation

matrix will have unity values (1.00) since any column will correlate perfectly (r = 1.00) with itself. Also, a matrix with unity diagonal values lends itself readily to principal components factor analysis. The similarity correlation matrix is further analyzed by factor analysis (principal components) and rotated using the varimax method (Kaiser, 1958). Consequently, the resultant factors are orthogonal (independent) and when the stimuli are persons, " . . . should be regarded as idealized types of persons, and the loadings of actual people specify to what extent they are mixtures of various types" (Stone et al., 1970, p. 8). In addition, such factors can be considered judgment-dimensions and shall be referred to as such throughout this paper.

The second major MESA analytical routine is Observer Factor Judgment Analysis (OFJA) and requires that the similarity estimations made for each stimulus-pair be inter-correlated across judges. The resultant interjudge similarity correlation matrix has unity values (1.00) in the diagonal and is further analyzed by factor analysis (principal components method). The resulting factors (considered judge-dimensions throughout this paper) should be viewed as "clusters" of judges or " . . . groups based on the 'commonness' of stimulus evaluation patterns (Stone et al., 1970, p. 12). In other words, the OFJA judge-dimensions places participating judges in groups based on common similarity estimation patterns as they each perceived the stimulus-pairs. In this study, the stimulus-pairs were mentally retarded persons and psychiatric patients who are referred to as stimulus-persons throughout this paper.

Statement of Problem

The problems encountered by investigators studying the predictive assessment of vocational success for mentally retarded adults have been characterized as complex due to the multidimensional nature of the predictor and criterion variables. In addition, the lack of follow-up studies (even short term ones) and lack of cross-validation investigations has been described as a drawback in the discovery of a method where accurate predictive assessment can be made.

The purpose of this study was to devise and cross-validate a predictive assessment procedure using clinical judgment as the primary predictor variable and MESA as the primary instrument for clinical judgment measurement. The attempt is limited in that the rank success within specific groups was predicted and no claim is made that the procedure can predict the absolute probability of success for any subject in any population. The procedure employs MESA (described earlier) where the "clinical judgments" are made in terms of similarity estimations and where the directive question required a comparison between mentally retarded subjects (stimulus-persons) regarding their actual or potential vocational placement success. Also, a group of psychiatric patients were evaluated where successful community adjustment was predicted.

CHAPTER III

METHOD

Introduction

In an earlier investigation (Hanson and Stone, 1974), the Multidimensional Evaluation Structure Analysis (MESA) methodology was used to determine the underlying judgmental criteria used by groups of professionals who judged several groups of mentally retarded adults. The directive question related to actual and potential placement success of these adults. Each judgment-dimension appeared to be polar (the judgment-dimensions were generally bipolar) representing positive attributes indicative of vocational success and negative attributes indicative of vocational failure. The judgment-dimensions were interpreted by the judges.

It was hypothesized that suitable combinations of the judgment-dimensions for the two previous studies would be systematically related to future success variables, such as income, length of placement, and others. Relationships found between judgment-dimensions and follow-up data, however, could not be considered truly predictive unless cross-validation studies were successfully completed. Consequently, four additional studies were conducted to determine the predictive value of the MESA methodology.

A complete description of subjects, judges, materials, and procedure of the previous two studies is found in the Hanson (1971)

study. A report of important factors in those studies is found in Hanson and Stone (1974) and a brief report will be described here.

Two Previous Investigations

The SS-1 and OTC-1 Groups

Hanson and Stone (1974) used 10 mentally retarded adults (ages 16 to 57 years) from the Grafton State School for the Mentally Retarded in Grafton, North Dakota. These adults (the SS-1 stimulus-persons) were characterized as "potential" vocational placements since all were institutionalized and few had even been placed in the community. A summary of biographical and examination data for the SS-1 group is found in Table 1.

TABLE 1

SUMMARY OF BIOGRAPHICAL AND EXAMINATION DATA FOR THE SS-1 AND OTC-1 STIMULUS-PERSONS

Group	Median Age in Years and Months	Age Range in Years and Months	Mean Years and Months with Agency	Median Age in Years and Months When Admitted to Agency	Mean Full Scale WAIS IQ
SS-1	23-11	16-8 to 57-9	16-1.7	12-1.5	63.1[a]
OTC-1	22-7	18-1 to 54-7	2-1.3	20-5	73.0[b]

[a]Includes two Wechsler Intelligence Scale for Children (WISC) exams.

[b]Includes one WISC exam.

Four judges were used to evaluate (using similarity estimations) the SS-1 group and all four were familiar with the 10 stimulus-persons. The MESA analysis of these similarity estimations extracted three

judgment-dimensions accounting for 84 percent of the judgmental variance. On the first judgment-dimension (Socialization), stimulus-persons with high positive loadings were considered more socially mature. On the second judgment-dimension (Independent Functioning), stimulus-persons with high negative loadings were seen as being able to function with less supervision. The third judgment-dimension (Institutional Background) was a monopolar factor where those with high negative loadings were considered to have greater placement potential. In each case, those who were more socially mature, able to function more independently, and had less of an institutional background were thought of as having more potential for vocational placeeent success. It should be noted that "positive-negative" loadings on judgment-dimensions do not reflect "positive-negative" success attributes according to algebraic signs. That is, a high negative loading may, in fact, reflect a positive success attribute for that stimulus-person.

The second Hanson and Stone (1974) study used 15 mentally retarded adults who were termed "successfully employed" because they had been working from periods ranging between two and 52 weeks. They were being followed at the Opportunity Training Center located in Grand Forks, North Dakota. These "actual" working stimulus-persons (OTC-1 group) ranged from 18 to 54 years of age and a summary of their biographical and examination data is found in Table 1.

Four judges were used to evaluate (using similarity estimations) the OTC-1 stimulus-person and all four judges were familiar with them in varying degrees. The MESA analysis of the fifteen OTC-1 stimulus-

persons extracted three judgment-dimensions that accounted for 75 percent of the judgmental variance. High negative loadings on the Employment Stability, Inappropriate Intellectual Independence, and Unexpected Vocational Outcome judgment dimensions were interpreted as predictive of vocational success.

In comparison with the SS-1 group, the OTC-1 group was a younger, more intellectually capable group. Also, only three of the OTC-1 stimulus-persons had ever been institutionalized where the SS-1 group consisted of stimulus- persons who were institutionalized at the time of the study.

The SS-1 and OTC-1 studies were used as the basis for the development of a prediction scheme based on modifications of MESA. Predictor and success values had to be determined in the course of this development.

Predictor Values

The judgment-dimensions extracted in the SS-1 and OTC-1 studies were modified and combined to reflect predictor values indicative of the vocational success experienced within the specific group evaluated. The predictor values were ranked and considered measurements of predicted outcome. The predictor values were derived from a combination of the MESA judgment-dimensions with two modifications. The "amount of variance accounted for" by each judgment-dimension and the degree to which judges were familiar with the stimulus-persons formed the basis of the MESA modifications. These modifications were hypothesized to improve the predictive efficiency of the judgment-dimensions in both the SS-1 and OTC-1 studies.

The "Weighted" Judges Factor. One problem encountered in using the loadings on the judgment-dimensions as predictor values is that some judges were more familiar with the stimulus-persons than others. This was documented from the familiarity ratings (based on a five-point scale) made by each judge where a five indicated "know very well" and one indicated "do not know very well." The mean familiarity rating that each judge assigned to the stimulus-persons is found in Table 2.

TABLE 2

MEAN JUDGE FAMILIARITY RATINGS AND JUDGE "WEIGHT" VALUE FOR SS-1 AND OTC-1 GROUPS

		Judges			
Study		A	B	C	D
	Raw Mean[a]	4.40	4.70	3.10	3.00
SS-1					
	Weighted[b] Value	2.00	3.00	1.00	1.00
	Raw Mean[a]	3.47	4.07	5.00	2.13
OTC-1					
	Weighted[b] Value	2.00	3.00	4.00	1.00

[a]Where five indicated "know very well" and one indicated "do not know very well."

[b]Where the lowest "raw mean" value in each judge group was set at 1.00 and the corresponding values for other judges were lowered proportionately and rounded to the nearest whole number.

In assigning "weights" to the judges, the undocumented assumption was made that a judge who was quite familiar with the stimulus-

persons would be more qualified to evaluate the stimulus-persons on the judgmental directive (e.g., potential vocational success) than a judge who was not as familiar with the stimulus-persons. To this end, the lowest raw mean familiarity value for a group of judges was given a value of 1.00, and the other raw mean values for each judge were adjusted accordingly. In addition, the adjusted values were rounded off to the nearest whole number. These "weighted" values served to give all four judges differential "importance" based on how familiar they were with the stimulus-persons.

This weighting scheme was implemented by duplicating the deck of judgmental computer cards depending on the judges' "weighted" value and proceeding with the normal MESA computations. For example, the original MESA analysis of the SS-1 group was handled as a four judge problem, since there were four judges. The "weighted" MESA analysis was handled as a seven judge problem since the "weighted" values for Judges A, B, C, and D of the SS-1 group were 2.0, 3.0, 1.0, and 1.0 (where 2.0 + 3.0 + 1.0 = 7.0), respectively. The "weighted" MESA analysis resulted in a "weighted" principal components factor analysis with resultant "weighted" judgment-dimensions. The use of these "weighted" judgment-dimensions for prediction, then, accounted for the fact that some judges were more familiar with the stimulus-persons than others. This was hypothesized as a factor that was important in the prediction of vocational success.

The "Variance Accounted For" Factor. To derive the predictor values from the "weighted" judgment-dimensions described above, the undocumented assumption was made that the degree to which any

judgment-dimension accounted for portions of the total accountable variance, the more "important" is that particular judgment-dimension. For example, the three judgment-dimensions of the SS-1 group (see Table 12) accounted for 53, 18, and 13 percent of the total judgmental variance. Using the judgment-dimension accounting for the least amount of variance (judgment-dimension III, in this case equals 13 percent) as the base of 1.00 (where [13%/13%] = 1.00), judgment-dimension I was valued as 4.08 (53%/13%) and judgment-dimension II was valued as 1.38 (18%/13%). Notice the divisor is 13 percent in each case since that is the smallest amount of variance accounted for in this study. Consequently, judgment-dimension I can be viewed as being 4.08 times more "important" than judgment-dimension III and judgment-dimension II can be viewed as 1.38 times more "important" than judgment-dimension III.

The use of these percentages as indicators of "importance" was desired since it did not appear logical to formulate predictor values from judgment-dimensions that would be treated equally. The judgment-dimension that accounted for the most variance did so because the judges were primarily considering that dimension relative to the directive question. The scheme used to determine "importance" was arbitrary but logical in that a simple formulation was employed, that of comparative "variances accounted for" between the judgment-dimensions.

Ranking. Predictor values for each stimulus-person were based on factor loadings from unweighted and "weighted" judgment-dimensions. For both the SS-1 and OTC-1 studies, the poles of the judgment-dimensions were arranged by reflecting the algebraic sign (plus to minus and minus to plus) so that the positive pole represented

positive attributes indicative of vocational success. This was done since a positive loading on one judgment-dimension and a negative loading on another, within the same factor structure, could have both been indicative of success attributes. It should be noted that changing the sign of factor loadings in factor analysis does not alter the potency of that loading. However, algebraic signs are required to give meaning to factor structures.

The predictor values were determined by the generalized formula that follows:

$$[(X_1V_1+X_2V_2+\ldots+X_{\underline{i}}V_{\underline{i}})/\Sigma V_{\underline{i}}]$$

Where: $X_{\underline{i}}$ = the stimulus-person's factor loading on the $\underline{i}$th judgment-dimension.

$V_{\underline{i}}$ = the relative "importance" of the $\underline{i}$th judgment-dimension.

$\Sigma V_{\underline{i}}$ = the sum total of all $V_{\underline{i}}$ values.

For example, Stimulus-person-1 in the SS-1 study had unweighted factor loadings (after algebraic signs were changed so that positive loadings were indicative of positive attributes) of +.96, -.03, and +.12 of Judgment-dimension I, II, and III, respectively. The relative "importance" of judgment-dimensions I, II, and III were found to be 4.08, 1.38, and 1.00, respectively. The sum total of $V_{\underline{i}}$ values was 6.46 (where 4.08 + 1.38 + 1.00 = 6.46). Placing these values into the generalized formula we have as follows:

$$[(X_1V_1+X_2V_2+\ldots+X_{\underline{i}}V_{\underline{i}})/\Sigma V_{\underline{i}}] = \text{Predictor Value}$$

$$[[[(.96)(4.08)]+[(-.03)(1.38)]+[(.12)(1.00)]]/6.46] = \text{Predictor Value}$$

$$[[(3.92)+(-.04)+(.12)/6.46] = \text{Predictor Value}$$

$$[4.00/6.46] = .62$$

It should be noted that the predictor value of .62 for Stimulus-person-1 of the SS-1 study is not a probability value. It is a value that has meaning only as it compares to the other stimulus-persons within the SS-1 group, where higher values represent higher predicted potential of vocational placement success. Notice also that if the "amount of variance accounted for" is not considered (i.e., all three judgment-dimensions considered equally "important"), the average factor loading for Stimulus-1 would be +.35, where:

$$[[[(.96)+(-.03)+(.12)]/3.00] = .35].$$

In this example, the "variance accounted for" predictor value for Stimulus-1 is higher (.62 versus .35) when the "importance" of the judgment-dimension is considered. This occurred because he had a high positive loading on Judgment-dimension I that accounted for the most variance.

Given the predictor values derived from combinations of these factors, the stimulus-persons were ranked from highest to lowest based on the predictor values. These ranks, in turn, were then correlated with rank success values derived from follow-up information.

Success Values

In order to determine the validity of the predictor values based on the modification of MESA judgment-dimensions, follow-up information was gathered approximately 15 months after the original similarity estimations were made by the SS-1 and OTC-1 judges.

These success values were not intended to define "successful community-vocational adjustment" for the stimulus-persons but were designed to serve as indicators of such success.

Information regarding income and number of weeks on job placement were collected for most of the stimulus-persons. This information was ranked and then correlated with the rank predictor values to determine the relationship between MESA predictor values and criterion values.

Prediction

It was hypothesized that the relationships found between predictor and criterion values of the SS-1 and OTC-1 studies could be applied to future studies, in an attempt to predict future success of mentally retarded adults before they are placed on the job. Consequently, two additional studies were conducted to determine the empirical predictive value of MESA (and the MESA modification procedures for prediction). This cross-validation attempt was conducted at the same two agencies as were the SS-1 and OTC-1 studies.

Cross-Validation Studies

The SS-2 and OTC-2 Groups

Two cross-validation studies were conducted to determine the predictive value of MESA judgment-dimensions regarding the vocational success of two selected groups of mentally retarded adults. Both groups were derived from the two agencies used in the two previous studies (SS-1 and OTC-1) and most of the judges used in the present studies had also been employed in the previous studies.

SS-2 Group Subjects

One group of stimulus-persons was selected from Grafton State School for the Mentally Retarded in Grafton, North Dakota. This agency maintained a Vocational Placement Department that determined and coordinated the placement opportunities given to qualified mentally retarded adults in residence there. On June 1, 1972, 14 male residents were to be considered as "placement possibilities" within the next year. To further define this group, they were all male, presently institutionalized, had no previous placement, and ranged in age from 21 to 60 years of age. A more detailed summary of biographical and examination data for the SS-2 group is found in Table 3.

SS-2 Group Judges

Three judges were used to evaluate (using similarity estimations) the SS-2 stimulus-persons. All three judges were employees of the State School for the Mentally Retarded and were familiar with the stimulus-persons in varying degrees. The SS-2 judges participated, in various ways, in placement decisions made regarding the stimulus-persons. The judges were selected on this basis and were regarded as "expert" judges for this study. Two of the judges used in the SS-2 study had also been used as judges in the previous SS-1 study and a detailed description of SS-2 judges is found in Table 4.

OTC-2 Group Subjects

A second group of stimulus-persons was selected from the Opportunity Training Center in Grand Forks, North Dakota. This agency accepted mentally retarded adults from a variety of referral sources

TABLE 3

DESCRIPTION OF STIMULUS-PERSONS FOR THE SS-2 GROUP

Stimulus-Persons	Age in Years and Months[a]	Age When Admitted to State School in Years and Months	Years and Months Institutionalized	Measured IQ[b]
1	23-4	13-7	9-9	35[c]
2	27-0	14-1	12-11	?
3	51-9	14-5	37-4	50[d]
4	53-0	12-1	40-9	46[d]
5	28-10	8-8	20-2	41[e]
6	32-10	22-3	10-7	?
7	60-0	29-5	30-5	51[c]
8	21-8	19-2	2-6	50[d]
9	26-3	15-8	10-7	?
10	26-10	15-2	9-8	44[c]
11	36-3	11-8	24-5	43[c]
12	29-2	11-10	17-4	47[c]
13	26-9	12-5	13-7	58[d]
14	32-0	19-4	12-8	71[d]

[a]As of June 1, 1972.

[b]Examinations administered in July, 1959, or more recently.

[c]Stanford-Binet Intelligence for Children examination.

[d]Wechsler Adult Intelligence Scale (WAIS) examination.

[e]Goodenough Draw-a-Person.

TABLE 4

DESCRIPTION OF JUDGES FOR THE SS-2 GROUP

Judge	Sex	Occupational Title	Occupational Duties	Years and Months Performing the Duties Described
A[a]	Male	Caseworker	Arrange job, resthome, and foster home placements. Coordinate evaluation program for outpatients.	2-9
B[a]	Male	Program Director	In charge of all non-medical resident programs. Some psychological evaluations and counseling.	11-0
C	Male	Director, Outpatient Services	Adjustment and vocational counseling. Program Evaluation.	0-9

[a]These two judges were also used to judge the earlier SS-1 group of stimulus-persons.

and attempted to train each person in various aspects of community adjustment and occupational skills. As of May 15, 1972, 11 male adults (OTC-2 group) had been accepted into the training program between September, 1971 and March, 1972. They ranged in age from 16 to 34 years of age. A more detailed description of the OTC-2 stimulus-persons is found in Table 5.

OTC-2 Group Judges

Four judges were used to evaluate the OTC-2 stimulus-persons and all were familiar with them in varying degrees. Three of the judges were employees of the Opportunity Training Center and one was a part-time psychology consultant (having a Ph.D.) who had less individual contact with the stimulus-persons. All four judges participated, in various ways, in vocational placement decisions made regarding the stimulus-persons. The judges were selected on this basis and were considered expert judges for this study. A more detailed description of the OTC-2 judges is found in Table 6.

Materials

The judges for the SS-2 and OTC-2 groups were given the materials necessary to complete their similarity estimations for their respective groups. When possible, they were seen individually and given the materials with a brief explanation.

The three judges of the SS-2 group were given a four page instruction form (Appendix A) and a deck of judgmental computer cards. The first two pages of the instruction form ("Instructions To Judge") explained the purpose of the study and specifically explained the use of the similarity estimation as the expression

TABLE 5

DESCRIPTION OF STIMULUS-PERSONS FOR THE OTC-2 GROUP

Stimulus-Persons	Age in Years and Months[a]	Years and Months With Agency[a]	Age in Years and Months When Admitted to Agency	Years and Months Institutionalized (If any)	Full Scale WAIS IQ[b]	Highest Academic Grade Attended
1	21-2	0-7	20-7	-	66	12th
2	29-6	0-4	29-2	-	65	7th
3	28-3	0-6	27-9	-	60	8th
4	34-6	0-4	34-2	-	62	6th
5	28-5	0-3	28-2	-	74	8th
6	28-9	0-8	28-1	-	83	8th
7	17-6	0-6	17-0	1-6[c]	68	7th
8	19-10	0-4	19-6	-	89	12th[d]
9	17-5	0-8	16-9	-	64	8th
10	16-8	0-3	16-5	-	80	11th
11	24-1	0-2	23-11	-	73	8th

[a]As of May 15, 1972.

[b]All examinations were administered in October, 1959, or more recently.

[c]State School for the Mentally Retarded.

[d]High School diploma from a special education program.

TABLE 6

DESCRIPTION OF JUDGES FOR THE OTC-2 GROUP

Judge	Sex	Occupational Title	Occupational Duties	Years and Months Performing the Duties Described
A	Male	Executive Director	Organize and coordinate Training Center Programs.	8-0
B	Male	Placement Director and Social Worker	Help trainees secure employment and help them take care of personal needs.	7-5
C	Female	Psychological Consultant (Ph.D.)	Psychological evaluation of new trainees, research, and teaching.	7-0
D	Male	Training Shop Supervisor	Teach the use of tools, electrical equipment, time and money. Show educational films and supervise construction activities in employment training.	5-8

of the "general degree of similarity" between each stimulus-pair. The SS-2 judge group was asked to direct their similarity estimates toward the "potential success on vocational placement" question. Instructions for the use of the judgmental computer cards were also included here.

The third page of the instruction form ("Judge Information and Rating Sheet") asked for biographical information regarding the judges. In addition, this page contained a section called "Ratings of Stimulus-Persons." Here, all the SS-2 stimulus-persons were listed and the judge was asked to rate (on a five-point scale where one indicated "do not know very well" and five indicated "know very well") his degree of familiarity with all of the stimulus-persons.

The fourth page was entitled the "Probability Estimation Sheet." A simple explanation of the concept of probability was given and each judge was asked to express (in the form of a probability estimate) his or her estimation of the stimulus-persons probable future success. All the stimulus-persons were listed with a convenient place to write this estimation.

The OTC-2 group of judges were given an instruction form similar to that given the SS-2 judge group (see Appendix B). The first two pages ("Instructions To Judge") were the same, asking them to make similarity estimations based on the stimulus-person's "potential success on vocational placement." The third and fourth pages differed only in that the stimulus-persons of the OTC-2 group were presented.

Judges for both groups were given a deck of computer cards with the stimulus-pairs printed at the top. The SS-2 judges were given their respective judgmental computer cards and the OTC-2 judge group were given their stimulus materials. The instructions clearly stated that similarity

estimations were to be written at the lower left-hand corner of the computer cards with respect to the stimulus-pair printed at the top of the computer card.

The SS-2 group of judges were required to make 182 [$m(m - 1)$, where m equalled 14] similarity estimation judgments. The OTC-2 judge group was required to make 110 similarity estimations [11(11 - 1)]. The SS-2 judge group made their judgments on or near June 1, 1972, where the OTC-2 judge group made their judgments on or near May 15, 1972. The instruction forms and computer cards were then collected shortly after they had been delivered to the judges.

Procedure

The similarity estimations made on the judgmental computer cards by the SS-2 and OTC-2 groups of judges were key-punched. The data was then analyzed by MESA. Both Group Composition Structure Analysis (GCSA) and Observer Factor Judgment Analysis (OFJA) routines were computed and extracted judgment-dimensions and judge-dimensions for these groups.

The extracted judgment-dimensions were then listed on separate sheets, one judgment-decision per sheet and the stimulus-persons were reordered on each judgment-dimension from largest positive loadings to largest negative loadings. Thus each judge viewed respective stimulus-persons on a continuum from +1.00 to -1.00, for each judgment-dimension. The judges were interviewed separately to gain their interpretations of the judgment-dimensions and each judgment-dimension was examined separately during the interview.

Based on these interviews, the investigator named the judgment-dimensions by collating the descriptions given by the judges. As a whole, the judges reported that the judgment-dimensions "made sense" and interpretation was not difficult for them. It was, however, sometimes difficult to assign a name to the judgment-dimension that completely represented the judge's interpretations. For this reason, all of the judgment-dimensions were also described.

Prediction

Based on the relationships found between judgment-dimensions (accounting for the judge "weight" value) extracted from the earlier SS-1 and OTC-1 studies and follow-up data, predictions were made regarding the rank success that SS-2 and OTC-2 stimulus-persons would have based on the predictor value. These predictions were then compared to the actual success experienced by the SS-2 and OTC-2 stimulus-persons based on follow-up data collected later.

The probability estimates (derived from the judges' "Probability Estimation Sheet") were averaged across judges for each stimulus-person. In addition, the probability estimates were averaged with respect to the "weight" value of each judge to determine which had been more effective in predicting the actual success of the stimulus-persons.

Generalization Studies

Two additional studies were conducted to determine whether or not MESA (and the described modifications) could be used as a predictive tool at agencies where no previous studies of this kind had been

completed. One study used a group of mentally retarded adults who were being trained in a sheltered workshop setting and another group was selected from a psychiatric unit, consisting of people who were receiving inpatient or outpatient psychiatric care.

The Sheltered Workshop (SWS) Group

One group of stimulus-persons were selected from the Northwestern Area Sheltered Workshop facility (SWS group) located in Thief River Falls, Minnesota. The Sheltered Workshop functioned as a training facility for mentally retarded adults. Due to their limitations, they had been referred for training leading to vocational and self-help skills. The ultimate goal of training was to assist the person in areas of vocational skills and help him become more self-sufficient in the community.

Subjects

Eight stimulus-persons were selected from the Sheltered Workshop and comprised the SWS group. All SWS stimulus-persons were males, between the ages of 18 and 54 years, and had been admitted as clients to the Sheltered Workshop between September, 1971, and December, 1971. The SWS stimulus-persons were selected primarily because they were viewed as trainees who would possibly leave the workshop within one year and obtain competitive employment. As of May 22, 1972, eight trainees were viewed as probable placements within the next year and biographical data describing them is found in Table 7.

TABLE 7

DESCRIPTION OF STIMULUS-PERSONS FOR THE SWS GROUP

Stimulus-Persons	Age in Years and Months	Months With Agency[a]	Age in Years and Months When Admitted to Agency	Years and Months Institutionalized (If any)	Highest Academic Grade Attended
1	23-9	6	23-3	-	?[b]
2	24-8	7	24-1	-	10th
3	54-0	7	53-5	2-0	?
4	24-7	7	24-0	-	9th
5	21-6	5	21-1	-	7th
6	23-8	7	23-1	-	12th
7	18-9	5	18-4	-	11th
8	18-5	8	17-9	-	12th

[a]As of May 22, 1972.

[b]Attended school until 17 years of age, grade unknown.

Judges

Four judges employed by the Northwestern Area Sheltered Workshop were used to evaluate the eight stimulus-persons in the SWS group. All four were familiar with the stimulus-persons and contributed, in varying degrees, to vocational placement decisions. A more complete description of the SWS judges is found in Table 8.

Materials

The four SWS judges were given the materials needed to complete their similarity estimations. They were not seen individually but were given their materials by Judge A.

A four page instruction form (Appendix C) was given to them and the first two pages ("Instructions To Judge") were identical to those given the SS-2 and OTC-2 group judges. These instructions introduced the purpose of the study, explained the use of similarity estimations, and described the use of their judgmental computer cards. The instructions directed the judges to estimate the generalized similarity between each stimulus-pair based on their knowledge of potential vocational placement success.

Pages three and four were, again, identical to those given the SS-2 and OTC-2 groups. Page three asked the judges for biographical information and also asked them to rate their familiarity with the stimulus-persons based on the five-point familiarity scale. Page four briefly described the meaning of probability and asked the judges to estimate the probable success (chances out of 100) they expect each mentally retarded adult to have on future vocational placements. All

TABLE 8

DESCRIPTION OF JUDGES FOR THE SWS GROUP

Judge	Sex	Occupational Title	Occupational Duties	Years and Months Performing the Duties Described
A	Male	Executive Director	Coordinator of programs and placements; counselor.	0-11
B	Male	General Shop Supervisor	Supervision of workshop activities and staff consultant.	1-0
C	Male	Shop Foreman	Teaching of work skills to clients.	0-6
D	Female	Work Adjustment Trainer	Supervision of work and workers; teaching of good work habits and skills.	0-8

the stimulus-persons were listed and space was provided for their probability estimates.

Procedure

All four judges were given their materials and all similarity estimations were made on or near May 22, 1972. Shortly thereafter, the investigator collected the biographical data on the stimulus-persons, the instruction forms, and the decks of computer cards with the similarity estimations written on them.

The similarity estimations were key-punched and then analyzed by the two MESA analyses (GCSA and OFJA). A judgment-dimension and a judge-dimension were extracted and later interpreted by Judges A and B. For clarity, the stimulus-persons were again arranged from highest (+1.00) to lowest (-1.00) on the basis of the judgment-dimension loadings. The two judges reported that their interpretation task was quite simple. The "weighted" judge and "variance accounted for" modifications were also computed and predictor values were obtained.

Prediction

Based on the judgment-dimension produced by the modified MESA analysis, predictions were made regarding the rank success each stimulus-person would have. The probability estimates (also "weighted") were used to predict rank success and compared to the MESA method. Follow-up data regarding the actual success experience by the stimulus-persons was collected approximately one year after judgments were made. The rank of actual success was compared to the rank predicted success rankings to explore the efficacy of employing MESA as a predictive tool.

The Psychiatric Unit (PU) Group

A second group of stimulus-persons were selected from a Psychiatric Unit. Specific identification of this treatment facility has been withheld at the request of the facility. This study was conducted to determine whether or not the use of MESA, as a predictive tool, could predict community adjustment of psychiatric patients who would probably sever contacts within the following year.

Subjects

As of June 12, 1972, there were 14 male psychiatric patients (PU group) who were non-alcoholics, admitted between November 1, 1971 and March 31, 1972, and who were between the ages of 22 and 49 years old. Also, these 14 patients (PU stimulus-persons) had engaged in a minimum of five contacts (where "contact" meant any personal encounter with a staff member on any day) since their admission to the unit. The stimulus-persons of the PU group were selected based on these criteria. A more complete description is found in Table 9.

There were a variety of diagnostic labels placed on these stimulus-persons and in, some cases, one stimulus-person had more than one "diagnosis." None of the stimulus-persons were considered "alcoholic" since this diagnostic group was eliminated from the selection due to widely held beliefs regarding their future adjustment success. It was decided that such a characteristic group would create judgmental "noise" and they were not included in the study.

Judges

Nine judges were selected who were employees of the Psychiatric Unit and all were familiar with the PU stimulus-persons. The

TABLE 9

DESCRIPTION OF STIMULUS-PERSONS FOR THE PU GROUP

Stimulus-Persons	Age in Years and Months[a]	Marital Status	Months Since First Admission[a]	Age in Years and Months When First Admitted[a]	Previous Psychiatric Admissions
1	36-3	Married	8	23-1	-
2	49-1	Married	9	48-4	-
3	23-9	Married	8	23-1	-[b]
4	22-7	Single	3	22-4	2[b]
5	23-5	Married	4	23-1	-
6	45-10	Widowed	3	45-7	-
7	32-1	Married	10	31-3	-
8	40-1	Married	7	39-6	-
9	22-10	Single	7	22-3	-
10	30-5	Married	5	30-0	-
11	37-6	Married	7	36-11	2[c]
12	36-11	Married	4	36-7	-
13	25-11	Single	4	25-7	-[b]
14	22-6	Single	6	22-0	1[b]

[a]As of June 12, 1972.

[b]Psychiatric Unit.

[c]Two other hospitals.

The nine PU judges were all involved in treatment methods and most were involved with discharge decisions. A more complete description of the PU judges is found in Table 10.

Materials

All nine judges were given instruction forms and a deck of judgmental computer cards. The materials were given out at a staff meeting by Judge D and he gave a brief explanation of the purpose of the study and answered any questions.

TABLE 10

DESCRIPTION OF JUDGES FOR THE PU GROUP

Judge	Sex	Occupational Title	Occupational Duties	Years and Months Performing the Duties Described
A	Female	Group Therapy Technician	Individual therapy; group therapy, occupational therapy, recreational therapy.	3-0
B	Male	Recreation Therapist	Group and individual counseling, direct recreation program.	2-8
C	Male	Psychologist	Group and individual therapy, testing.	4-8
D	Male	Clinical Psychologist (Ph.D.)	Group and individual therapy, testing.	1-0
E	Female	Certified Occupational Therapt Assistant	Occupational, recreational, group, individual, and family therapy.	2-8
F	Male	Mental Health Specialist	Coordination, supervision, inservice training, etc.	12-10
G	Female	Registered Nurse	Individual and group therapy.	2-6

TABLE 10--Continued

Judge	Sex	Occupational Title	Occupational Duties	Years and Months Performing the Duties Described
H	Female	Recreational Therapist	Individual, group, recreational, family, and occupational therapy.	2-9
I	Male	Occupational Therapist	Occupational therapy program, work with psychiatric patients.	2-8

The first two pages of the instruction form ("Instructions to Judge") were similar to those given to the SS-2, OTC-2 and SWS groups (Appendix D). These judges were asked to estimate the general degree of similarity between the stimulus-persons regarding their potential community adjustment success. The purpose of the study, the use of similarity estimations, and the use of the computer cards were explained in these instructions.

The third page asked the PU judges to complete a form requiring biographical data. At the bottom of that page, the judges were asked to rate their familiarity with the stimulus-persons based on a five-point scale, where one indicated "do not know very well" and five indicated "know very well."

The fourth page of the instruction form ("Probability Estimation Sheet") gave a brief description of what a probability value was and asked the judges to estimate the probable success that each stimulus-person would have regarding community adjustment. They were allowed to use probability values from zero to 100 for this estimation.

In addition, all nine judges were given a deck of judgmental computer cards with the stimulus-persons printed on the top of the card in varying pair combinations. Since 14 stimulus-persons were evaluated, each judge was required to make 182 similarity estimations [14(14-1)].

Procedure

All nine judges completed their judgmental task on (or after) May 22, 1972. The instruction sheets and computer cards were collected shortly after that date.

The similarity estimations were then analyzed by MESA and both GCSA and OFJA were computed. The resultant judgment-dimensions and judge-dimensions were presented to the same staff members for interpretation. Judges C, D, and F were primarily responsible for interpreting the dimensions and did so by assigning personal attributes to the poles of the bipolar judgment dimensions. This information was recorded and the investigator "named" the judgment-dimensions and judge-dimensions based on those interviews. Again, the "names" assigned to the dimensions, based on the judges' interpretations, were somewhat difficult to create. However, since each dimension was also defined and described, the "name" could be fully explained.

Prediction

Based on the MESA unweighted and "weighted" analyses, the judgment-dimensions were modified to predictor values based on the amount of "variance accounted for" by each judgment-dimension. These values were called predictor values and represented the best possible combination of staff consensus regarding successful community adjustment of all stimulus-persons.

More than one year later, follow-up data was collected regarding how successful the stimulus-persons had actually been in adjusting to community life. This data included known facts about the stimulus-persons and rankings made by staff who were not originally part of the judge group. This follow-up data was not regarded as directly defining successful community adjustment but, rather, served as a crude indicator of such success.

The predictor values and follow-up data values were ranked and the correlational relationship between them was found. The probability

estimations were also related to the follow-up data in both "weighted" and unweighted manners. The way in which modified MESA compared to the probability estimation method of predicting success was also assessed.

Summary

The rather complex analysis of the SS-1, OTC-1, SS-2, SWS, and PU groups provided information regarding the efficacy of using MESA (with modifications) as a predictive tool, given that it has been well established as a descriptive tool. The predictor values are well defined and have been fully described. The success values, however, were difficult to specify. This difficulty was discussed earlier (see Chapter II) and similar problems were encountered in this study. Basic follow-up data was collected which hopefully was indicative of success but, surely, did not totally define successful adjustment.

The purpose of using ordinal measurement (ranks) was due to the fact that no absolute probabilities of success could be determined, given the nature of the data. Consequently, the rank success within a specific group was predicted with no way of determining the absolute probability of success. Information was lost using ordinal (rank) measurement (since both predictor and criterion data were in ratio measurement form) but the simplest solution was desired, given the complex nature of the data and the lack of knowledge regarding sampling distributions of the data.

CHAPTER IV

RESULTS

Introduction

The similarity estimations made by judges for the original State School group (SS-1) and the original Opportunity Training Center group (OTC-1) were collected on June 1, 1971. The follow-up data on both groups was collected on August 1, 1972. Thus, 14 months elapsed between the prediction of success (using similarity estimations) and the determination of actual success for these two groups.

The cross-validation groups were evaluated at different times. The SS-2 group judgments were made on June 1, 1972 and the follow-up data was collected on July 1, 1973. The OTC-2 group judgments were made on May 15, 1972, and the follow-up data was collected on July 1, 1973. Thus, 13 months elapsed between judging and follow-up for the SS-2 group and 13.5 months elapsed between judging and follow-up for the OTC-2 group.

The Sheltered Workshop (SWS) judges made their judgments on May 22, 1972 and the follow-up data was collected on July 1, 1973. The Psychiatric Unit (PU) judges made their judgments on June 12, 1972 and the follow-up data was collected on July 1, 1973. Thus, the SWS stimulus-person group was allowed slightly more than 13 months before the assessment of vocational success was made where the PU stimulus-person group was allowed slightly less than 13

months before the assessment of community adjustment success was made.

In all six studies, the stimulus-persons were totally unaware of the judging and assessment procedures. In addition, the assessment procedures were largely dependent on the data available, since none of the agencies involved made any special effort to collect data specifically for this study. Consequently the relationships to be reported between predicted and actual success left much to be desired methodologically in these exploratory studies.

Two Previous Investigations

The SS-1 Study

The 10 stimulus-persons described in the previous SS-1 study (Hanson and Stone, 1974) were judged by four professionals using the similarity estimations required by the Multidimensional Evaluation Structure Analysis (MESA) methodology. Since $m(m - 1)$ stimulus-pairs were presented to each judge (where m equals the number of stimulus-persons), each stimulus-person pair was presented twice, but in alternate positions. For example, the stimulus-persons were presented first in the S_{ij} position and then in the S_{ji} position on another computer card. Thus a judgmental reliability estimate could be determined by correlating the similarity estimations made for each stimulus-person pair in the S_{ij} position by each stimulus-person pair in the S_{ji} position. For the SS-1 group, this reliability estimate was high ($r = .90$, $df = 43$, $p < .001$).

Prediction of Success

The dimensional analysis for the SS-1 group began by averaging (across judges) the similarity estimations made for each stimulus-person pair. The column vectors for the resultant mean similarity estimation matrix were intercorrelated (Pearson product-moment correlations) to form a similarity correlation matrix. The results of these computations are found in Table 11. This similarity correlational matrix, in turn, was factor analyzed (principal components) and rotated (varimax) to produce the rotated Group Composition Structure Analysis (GCSA) factor matrix found in Table 12. Judgment-dimensions I, II, and III were interpreted by the judges as Socialization, Independent Functioning and Institutional Background, respectively. It was determined that high positive factor loadings on Judgment-dimension I was indicative of personal attributes desired for successful vocational outcome where high negative factor loadings on judgment-dimensions II and III were indicative of success. For clarity, these judgment dimensions were reflected where necessary so that positive factor loadings were indicative of probable vocational success. These dimensions were called Predictor-dimensions and are found in Table 13 for the SS-1 group.

It should be noted that the quantitative value of the factor loadings were not altered by reflecting the algebraic signs of Judgment-dimensions II and III. The predictor-dimensions (Table 13) were then used to compute predicator values to be described below.

In order to account for the differential value of each judge as an evaluator of potential success, another GCSA dimensional analysis was completed. It was obvious that some judges were more familiar

TABLE 11

ARITHMETIC MEAN (ACROSS JUDGES) SIMILARITY ESTIMATIONS, SIMILARITY CORRELATION COEFFICIENTS, MEAN SIMILARITY, AND STANDARD DEVIATIONS OF THE SIMILARITY ESTIMATIONS FOR THE SS-1 GROUP

	Stimulus-Persons									
Stimulus-Persons	1	2	3	4	5	6	7	8	9	10
1	1.00	.67	.46	.66	.59	.79	.43	.93	.83	.69
2	.09	1.00	.60	.53	.64	.58	.66	.71	.68	.64
3	-.79	-.16	1.00	.61	.63	.54	.70	.53	.55	.59
4	.21	-.57	-.22	1.00	.62	.66	.56	.67	.72	.73
5	-.12	-.03	.00	-.06	1.00	.63	.53	.65	.64	.71
6	.75	-.24	-.68	.24	-.04	1.00	.50	.80	.78	.79
7	-.81	.12	.57	-.40	-.31	-.74	1.00	.44	.46	.46
8	.97	.12	-.73	.18	.00	.75	-.85	1.00	.79	.75
9	.83	.02	-.67	.37	-.02	.72	-.83	.79	1.00	.78
10	.50	-.17	-.47	.43	.28	.71	-.80	.57	.67	1.00
Mean Similarity	.70	.67	.62	.67	.66	.70	.57	.73	.72	.71
Standard Deviation	.19	.13	.15	.13	.13	.15	.18	.17	.15	.14

Note:

Similarity estimates are above the major diagonal (unity values) and correlational similarities are below the major diagonal.

TABLE 12

ROTATED GCSA FACTOR MATRIX FOR THE SS-1 GROUP

	Unweighted Judgment Dimensions			
Stimulus-Persons	I	II	III	h^2
1	.96	.03	-.12	.94
2	.11	.90	-.03	.83
3	-.83	-.08	.05	.70
4	.24	-.84	-.04	.77
5	-.03	.05	.97	.94
6	.84	-.26	.02	.77
7	-.86	.23	-.37	.93
8	.95	.06	.01	.91
9	.90	-.14	.03	.82
10	.68	-.36	.43	.77
Amount of variance accounted for	53%	18%	13%	84%

TABLE 13

THE UNWEIGHTED PREDICTOR-DIMENSIONS FOR THE SS-1 GROUP

	Unweighted Predictor-Dimensions		
Stimulus-Persons	I	II	III
1	.96	-.03	.12
2	.11	-.90	.03
3	-.83	.08	-.05
4	.24	.84	.04
5	-.03	-.05	-.97
6	.84	.26	-.02
7	-.86	-.23	.37
8	.95	-.06	-.01
9	.90	.14	-.03
10	.68	.36	-.43
Amount of variance accounted for	53%	18%	13%

with the stimulus-persons than others. The original study by Hanson (1971) reported mean familiarity values that were determined by asking each judge to rate the degree of familiarity he had with each stimulus-person. The results of this evaluation were demonstrated in Chapter III (see Table 2, p. 36).

The procedure for this "weighted" MESA analysis was identical to the original analysis except that each judge was "weighted" in view of how familiar he was with the stimulus-persons. Thus, if one judge rated his familiarity with an average 1.0, then the latter judges' similarity estimations would be represented once where the former judge's similarity estimations would be represented twice.

The analytical result of "weighting" the value of the judges' similarity estimations is found in Table 14. A comparison of Table 14 with Table 11 (unweighted analysis) illustrated that "weighting" the similarity estimations tended to create some variation in both the mean similarity estimation matrix and the similarity correlation coefficient matrix. This "weighted" similarity correlation coefficient matrix was also factor analyzed (principal components) and rotated (varimax) to produce the "weighted" GCSA factor matrix found in Table 15. These "weighted" judgment-dimensions are similar to those of the unweighted judgment-dimensions (see Table 12).

The intercorrelations between unweighted judgment-dimensions and "weighted" judgment-dimensions (Table 16) indicates that the unweighted and weighted analysis produced similar results for the SS-1 group. As a result, the "weighted" judgment-dimensions did not require a revised interpretation by the judges. Since the unweighted judgment-dimensions were interpreted in such a way as to indicate that

TABLE 14

ARITHMETIC MEAN (ACROSS JUDGES) SIMILARITY ESTIMATIONS, SIMILARITY CORRELATION COEFFICIENTS, MEAN SIMILARITY, AND STANDARD DEVIATIONS OF THE SIMILARITY ESTIMATIONS FOR THE WEIGHTED SS-1 GROUP

	Stimulus-Persons									
Stimulus-Persons	1	2	3	4	5	6	7	8	9	10
1	1.00	.63	.45	.71	.65	.78	.44	.93	.84	.71
2	-.26	1.00	.64	.55	.65	.56	.72	.67	.64	.64
3	-.83	.08	1.00	.61	.64	.55	.70	.52	.56	.61
4	.43	-.62	-.34	1.00	.64	.71	.55	.71	.75	.75
5	.13	-.15	-.09	.00	1.00	.69	.53	.69	.68	.73
6	.72	-.56	-.65	.46	.19	1.00	.55	.81	.80	.82
7	-.83	.43	.53	-.56	-.46	-.70	1.00	.44	.46	.46
8	.97	-.24	-.78	.41	.20	.74	-.87	1.00	.83	.76
9	.87	-.36	-.68	.54	.17	.75	-.86	.85	1.00	.76
10	.54	-.36	-.42	.52	.39	.74	-.81	.62	.66	1.00
Weighted Mean Similarity	.71	.67	.63	.70	.69	.73	.58	.74	.73	.72
Standard Deviation	.18	.13	.15	.13	.12	.15	.18	.17	.15	.14

TABLE 15

THE WEIGHTED GCSA FACTOR MATRIX (ROTATED) FOR THE SS-1 GROUP

	Weighted Judgment-Dimensions			
Stimulus-Persons	I	II	III	h^2
1	.96	-.18	.07	.95
2	-.04	.91	-.12	.84
3	-.90	.01	.04	.80
4	.32	-.82	-.06	.77
5	.05	.01	.95	.91
6	.69	-.49	.20	.76
7	-.72	.41	-.48	.92
8	.94	-.17	.17	.94
9	.84	-.36	.16	.86
10	.53	-.46	.51	.74
Amount of variance accounted for	47%	23%	15%	85%

TABLE 16

CORRELATION COEFFICIENTS ($\underline{r}$) BETWEEN UNWEIGHTED AND WEIGHTED GCSA FACTOR MATRICES FOR THE SS-1 GROUP

	Weighted Judgment-Dimensions		
Unweighted Judgment-Dimensions	1	2	3
1	.990	-.255	.077
2	-.484	.955	-.140
3	.327	-.204	.955

positive factor loadings on I and negative factor loadings on II and III were indicative of successful vocational employment, the weighted predictor-dimensions (Table 17) were derived by reflecting the algebraic signs on "weighted" judgment-dimensions II and III (Table 15).

TABLE 17

THE WEIGHTED PREDICTOR-DIMENSIONS FOR THE SS-1 GROUP

	Weighted Predictor-Dimensions		
Stimulus-Persons	I	II	III
1	.96	.18	-.07
2	-.04	-.91	.12
3	-.90	-.01	-.04
4	.32	.82	.06
5	.05	-.01	-.95
6	.69	.49	-.20
7	-.72	-.41	.48
8	.94	.17	-.17
9	.84	.36	-.16
10	.53	.46	-.51
Amount of variance accounted for	47%	23%	15%

The unweighted and "weighted" predictor-dimensions did lend themselves to two methods of computing predictor values. As described in Chapter III, the best predictor-value was hypothesized to be the result of a "weighted" GCSA analysis where the "amount of variance accounted for" by each judgment-dimension (now in the form of a predictor-dimension) would be more "important" if it accounted for more judgmental variation than the other judgment-dimensions. As a consequence, the unweighted and "weighted" predictor-dimensions

were separately computed to derive predictor values where "variance accounted for" was considered in deriving predictor values and where it was not considered by treating each predictor-dimension as equal in "importance." The formula for this procedure is described in Chapter III (see page 39) and all computations were executed by a computer. The results of these computations ("derived" predictor values) are found in Table 18.

Since the "derived" predictor values were computed from predictor-dimensions which were arranged with positive algebraic signs as indicative of successful placement, these "derived" values indicated the predicted degree of success with higher positive values indicative of probable vocational success. These were ranked with the highest rank (1.0) indicating the highest expectation of success for that stimulus-person. All four methods of assessing predictor values are listed in Table 18 with corresponding ranks. Stimulus-person 6 was excluded from the predictor value table since it was found that follow-up data on that stimulus-person was very inadequate.

Follow-up of Actual Success

The follow-up information was gathered by the agency that made the vocational placements for the SS-1 group. A summary of this information is found in Table 19. Some of the figures obtained were "best estimates" on the part of the agency though no bias was involved since the agency judges were totally unaware of the predictor values that had been computed. It should be noted that several stimulus-persons were placed but had not earned any money. This happened when the agency

TABLE 18

PREDICTOR VALUES (DERIVED AND RANK) FROM GCSA MODIFICATIONS FOR THE SS-1 GROUP

	Predictor Values							
	Unweighted GCSA (variance not accounted for)		Unweighted GCSA (variance accounted for)		Weighted GCSA (variance not accounted for)		Weighted GCSA (variance accounted for)	
Stimulus-Persons	Derived	Rank	Derived	Rank	Derived	Rank	Derived	Rank
1	.35	2	.62	1	.36	2	.57	1
2	-.25	7	-.12	6	-.28	7	-.25	7
3	-.27	8	-.51	8	-.32	9	-.51	9
4	.37	1	.34	5	.40	1	.41	4
5	-.35	9	-.18	7	-.30	8	-.14	6
6[a]								
7	-.24	6	-.53	9	-.22	6	-.42	8
8	.29	4	.58	3	.31	4	.54	2
9	.34	3	.59	2	.35	3	.53	3
10	.20	5	.44	4	.16	5	.33	5

[a]Stimulus-person-6 suffered an accident shortly after being placed and thus was excluded from the study.

TABLE 19

EMPLOYMENT SUCCESS FIGURES FOR THE SS-1 GROUP

	Follow-up Information			
Stimulus-Persons	Location	Weeks Placed[a]	Income Earned[a]	Income Earned Per Weeks Placed
1	OTC[b]	37	$1000.00	$27.03
2	Svee[c]	20	0.00	0.00
3	Svee	35	0.00	0.00
4	Svee	41	958.91	23.39
5	OTC	63	500.00	7.94
6[d]				
7	None[e]	0	0.00	0.00
8	Direct[f]	65	2538.05	39.05
9	Svee	57	1057.68	18.56
10	Svee	41	162.54	3.96

[a]From June 1, 1971 to August 1, 1972.

[b]Opportunity Training Center in Grand Forks, North Dakota.

[c]Svee Home (Training Center) in Fargo, North Dakota.

[d]Stimulus-person-6 suffered an accident shortly after being placed and thus was excluded from the study.

[e]Not placed as of August 1, 1972.

[f]Placed directly on the job with no other training.

placed a client at a training center for more training before an actual job placement was made.

The follow-up data was ranked and these values are found in Table 20. The ranks represent the actual success experienced by nine

TABLE 20

RANK EMPLOYMENT SUCCESS FOR THE SS-1 GROUP

	Rank Employment Success		
Stimulus-Persons	Rank Weeks Placed[a]	Rank Income Earned[a]	Rank Earned Income Per Weeks Placed
1	6	3	2
2	8	8	8
3	7	8	8
4	4.5	4	3
5	2	5	5
6[b]			
7	9	8	8
8	1	1	1
9	3	2	4
10	4.5	6	6

[a]From June 1, 1971 to August 1, 1972.

[b]Stimulus-person-6 suffered an accident shortly after being placed and thus was excluded from the study.

stimulus-persons, 14 months after predictions of their success was formulated. These success measures were not the only measures of success but they were the only ones that could be gathered from the agency that appeared to indicate successful adjustment. The success ranks were correlated (Spearman Rank-Order Correlation Coefficient) with the predictor value ranks to determine the viability of using

the described GCSA modifications as a method of predicting success. These correlations are found in Table 21.

TABLE 21

SPEARMAN RANK CORRELATION COEFFICIENTS (r_S)[a] BETWEEN RANK SUCCESS AND PREDICTOR VALUES FOR THE SS-1 GROUP[b]

	Rank Weeks Placed	Rank Income Earned	Rank Earned Income Per Weeks Placed
Unweighted GCSA (variance not accounted for)	.184	.644	.695[c]
Unweighted GCSA (variance accounted for)	.494	.814[d]	.780[c]
Weighted GCSA (variance not accounted for)	.268	.695[c]	.746[c]
Weighted GCSA (variance accounted for)	.594	.915[e]	.932[e]

[a]All Spearman Rank Correlation Coefficients were corrected for tied ranks.

[b]N = 9.

[c]$p < .05$ (two-tailed test).

[d]$p < .01$ (two-tailed test).

[e]$p < .001$ (two-tailed test).

Discussion of Prediction

Table 21 indicates that the "weighted" analysis, where variance was accounted for, was the best predictor of two success categories. None of the predictor values were significantly related to "weeks on

placement" but all of the correlations with this success measure were positive.

The trend of the correlations seems to indicate that the most important aspect of the predictor values is that of "variance accounted for." This is indicated by the fact that "unweighted-variance accounted for" predictor values are more related to the success values than are the "weighted-variance not accounted for" predictor values. In addition, it should be noted both "weighting" and "accounting for variance" combined led to the best prediction of success.

The fact that the predictor values were not significantly related to "weeks on placement" seems to indicate that when the judges were evaluating the potential success of the stimulus-persons, they were generally inclined to use "earning power" (income earned and income earned per weeks placed) as the underlying judgment criterion. Also, these high correlations indicate that the judges, in fact, are able to adequately evaluate the potential success (in specific areas) of mentally retarded adults on vocational placement.

This study does not show that other methods of prediction would not do as well. This problem will be dealt with when the cross-validation and generalization studies are discussed.

The OTC-1 Study

The 15 stimulus-persons used as subjects in the OTC-1 study were evaluated by four judges using the similarity estimations required by the GCSA methodology. The reliability (correlated S_{ij} and S_{ji} similarity estimations) was high (r = .87, df = 103, p <.001).

The dimensional analysis began by averaging the similarity estimations (across judges) and intercorrelating the column vectors of the resultant matrix to form the similarity correlation coefficient matrix. Both matrices are found in Table 22. When the correlation similarity matrix was factor analyzed (principal components) and rotated (varimax), three Judgment-dimensions were produced and are found in Table 23.

These judgment-dimensions were interpreted by the judges and Judgment-dimensions I, II, and III were termed Employment Stability, In appropriate Intellectual Independence and Unpredicted Success Outcomes, respectively. It was determined that high negative factor loadings on all three judgment-dimensions were indicative of vocational success. Consequently, the algebraic signs of all three judgment-dimensions found in Table 24.

The same judgments (similarity estimations) were "weighted" depending upon how familiar each judge was with the stimulus-person. A demonstration of this procedure described in Chapter III (see Table 2, p. 36). The "weighted" mean similarity estimations and the similarity correlation matrix are found in Table 25. When this "weighted" similarity correlation matrix was factor analyzed (principal components) and rotated (varimax), three "weighted" judgment-dimensions were obtained and are found in Table 26. Again, these "weighted" judgment-dimensions were very similar to the unweighted judgment-dimensions and consequently there was no need to have them reinterpreted by the judges. The intercorrelations of "weighted" judgment-dimensions and unweighted judgment-dimensions are found in Table 27.

The only discrepancy between the interpretation of judgment-dimensions was that the "weighting" procedure resulted in the reversing

TABLE 22

ARITHMETIC MEAN (ACROSS JUDGES) SIMILARITY ESTIMATIONS, SIMILARITY CORRELATION COEFFICIENTS, MEAN SIMILARITY, AND STANDARD DEVIATIONS OF THE SIMILARITY ESTIMATIONS FOR THE OTC-1 GROUP

	Stimulus-Persons														
Stimulus-Persons	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	1.00	.37	.33	.57	.52	.51	.46	.48	.55	.49	.44	.46	.55	.48	.52
2	-.54	1.00	.51	.43	.65	.61	.66	.65	.34	.55	.53	.64	.51	.67	.53
3	-.40	-.04	1.00	.36	.48	.33	.36	.39	.55	.41	.31	.29	.34	.36	.41
4	.28	-.21	-.49	1.00	.57	.62	.52	.56	.39	.61	.50	.56	.63	.56	.49
5	-.10	.44	-.18	.20	1.00	.48	.68	.62	.33	.56	.64	.55	.71	.64	.34
6	-.05	.31	-.61	.43	.10	1.00	.57	.62	.32	.71	.64	.61	.64	.69	.55
7	-.26	.57	-.51	.11	.56	.41	1.00	.73	.30	.69	.72	.71	.63	.70	.57
8	-.21	.52	-.45	.21	.48	.49	.76	1.00	.35	.66	.64	.62	.63	.68	.50
9	.24	-.58	.40	-.34	-.58	-.65	-.75	-.67	1.00	.34	.33	.39	.34	.33	.39
10	-.17	.27	-.54	.37	.29	.71	.66	.59	-.68	1.00	.72	.72	.65	.73	.51
11	-.21	.30	-.62	.18	.51	.56	.75	.61	-.66	.77	1.00	.75	.73	.73	.42
12	-.19	.42	-.70	.21	.29	.54	.72	.54	-.59	.73	.81	1.00	.70	.71	.56
13	.07	.12	-.65	.47	.60	.54	.54	.51	-.64	.60	.76	.66	1.00	.66	.48
14	-.23	.55	-.57	.23	.46	.67	.74	.66	-.75	.76	.78	.75	.60	1.00	.55
15	.05	.09	-.24	-.03	-.45	.20	.13	.00	-.22	.05	-.14	.17	-.11	.11	1.00
Mean Similarity	.51	.58	.43	.56	.58	.59	.62	.61	.42	.62	.61	.62	.61	.63	.52
Standard Deviation	.15	.16	.17	.15	.16	.16	.17	.15	.18	.16	.19	.17	.16	.16	.15

Note:

Similarity estimates are above the major diagonal (unity values) and correlational similarities are below the major diagonal.

TABLE 23

ROTATED GCSA FACTOR MATRIX FOR THE OTC-1 GROUP

	Unweighted Judgment-Dimensions			
Stimulus-Persons	I	II	III	h^2
1	.17	-.80	.06	.67
2	-.51	.71	-.05	.77
3	.72	.54	-.15	.84
4	-.36	-.66	-.10	.57
5	-.46	.11	-.78	.84
6	-.77	-.18	.23	.67
7	-.84	.25	-.13	.79
8	-.76	.18	-.18	.64
9	.85	-.16	.08	.76
10	-.85	-.07	-.01	.72
11	-.85	-.02	-.27	.80
12	-.86	.00	.06	.74
13	-.75	-.35	-.36	.81
14	-.90	.12	-.04	.82
15	-.19	.04	.88	.81
Amount of variance accounted for	49%	15%	11%	75%

TABLE 24

THE UNWEIGHTED PREDICTOR-DIMENSIONS FOR THE OTC-1 GROUP

	Unweighted Predictor-Dimensions		
Stimulus-Persons	I	II	III
1	-.17	.80	-.06
2	.51	-.71	.05
3	-.72	-.54	.15
4	.36	.66	.10
5	.46	-.11	.78
6	.77	.18	-.23
7	.84	-.25	.13
8	.76	-.18	.18
9	-.85	.16	-.08
10	.85	.07	.01
11	.85	.02	.27
12	.86	.00	-.06
13	.75	.35	.36
14	.90	-.12	.04
15	.19	-.04	-.88
Amount of variance accounted for	55%	12%	12%

TABLE 25

ARITHMETIC MEAN (ACROSS JUDGES) SIMILARITY ESTIMATIONS, SIMILARITY CORRELATION COEFFICIENTS, MEAN SIMILARITY, AND STANDARD DEVIATIONS OF THE SIMILARITY ESTIMATIONS FOR THE WEIGHTED OTC-1 GROUP

	Stimulus-Persons														
Stimulus-Persons	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	1.00	.43	.30	.57	.59	.53	.49	.52	.56	.49	.46	.49	.55	.50	.56
2	-.35	1.00	.51	.45	.63	.64	.67	.57	.37	.54	.50	.72	.53	.70	.59
3	-.48	-.10	1.00	.38	.43	.35	.36	.32	.58	.45	.29	.29	.35	.38	.49
4	.21	-.16	-.52	1.00	.62	.66	.59	.64	.38	.64	.55	.59	.65	.58	.49
5	.09	.29	-.48	.39	1.00	.57	.70	.62	.36	.60	.69	.63	.77	.68	.40
6	-.03	.36	-.62	.52	.39	1.00	.63	.68	.32	.76	.69	.67	.67	.74	.50
7	-.16	.48	-.63	.32	.61	.59	1.00	.76	.34	.75	.78	.79	.70	.77	.59
8	-.02	.25	-.72	.50	.51	.68	.81	1.00	.35	.70	.70	.71	.70	.71	.51
9	.17	-.50	.47	-.52	-.62	-.75	-.76	-.76	1.00	.34	.35	.36	.30	.36	.45
10	-.20	.21	-.51	.47	.43	.79	.75	.72	-.75	1.00	.75	.76	.71	.76	.49
11	-.13	.18	-.70	.37	.65	.68	.82	.76	-.68	.82	1.00	.79	.78	.74	.38
12	-.13	.50	-.72	.33	.54	.67	.87	.75	-.72	.78	.86	1.00	.76	.75	.51
13	.02	.17	-.67	.54	.78	.64	.73	.73	-.76	.73	.86	.78	1.00	.70	.44
14	-.16	.54	-.60	.31	.57	.77	.84	.75	-.76	.79	.79	.82	.71	1.00	.56
15	.08	.22	.08	-.19	-.46	-.16	-.03	-.14	-.01	-.24	-.42	-.16	-.39	-.08	1.00
Weighted Mean Similarity	.53	.59	.43	.58	.62	.63	.66	.63	.43	.65	.63	.65	.64	.66	.53
Standard Deviation	.15	.15	.18	.15	.16	.16	.17	.17	.18	.17	.20	.18	.18	.16	.14

Note:

Similarity estimates are above the major diagonal (unity values) and correlational similarities are below the major diagonal.

TABLE 26

THE WEIGHTED GCSA FACTOR MATRIX (ROTATED) FOR THE OTC-1 GROUP

	Weighted Judgment-Dimensions			
Stimulus-Persons	I	II	III	h^2
1	.14	.89	.12	.82
2	-.54	-.51	.43	.74
3	.70	-.62	.00	.87
4	-.44	.55	-.25	.57
5	-.63	.10	-.44	.60
6	-.81	.13	-.04	.68
7	-.92	-.05	.01	.85
8	-.85	.20	-.08	.77
9	.89	.03	-.02	.78
10	-.84	-.01	-.23	.76
11	-.85	.05	-.40	.89
12	-.92	-.02	-.05	.85
13	-.82	.20	-.41	.88
14	-.93	-.08	.02	.87
15	.08	.04	.94	.88
Amount of variance accounted for	55%	12%	12%	79%

TABLE 27

CORRELATION COEFFICIENTS (r) BETWEEN UNWEIGHTED AND WEIGHTED GCSA FACTOR MATRICES FOR THE OTC-1 GROUP

	Weighted Judgment-Dimensions		
Unweighted Judgment-Dimensions	1	2	3
1	.989	-.045	.292
2	-.064	-.922	.022
3	.240	.271	.829

judgment-dimension II (as indicated by the high negative correlation found in Table 27). Thus, the algebraic signs were reflected on only judgment-dimensions I and III in order to create the "weighted" predictor-dimensions found in Table 28.

The unweighted and "weighted" predictor dimensions were used to compute the predictor values for the OTC-1 group. These "derived" predictor values and the corresponding ranks are found in Table 29. Three stimulus-persons were eliminated from the study for the reasons listed in Table 29. Again, it was predicted that predictor values determined by the "weighted" GCSA with variance accounted for would be the best predictor of vocational success.

Follow-up of Actual Success

The follow-up information was gathered by the Opportunity Training Center. A summary of this information is found in Table 30. The "weeks placed" and "income earned per week" were used to estimate

TABLE 28

THE WEIGHTED PREDICTOR-DIMENSIONS FOR THE OTC-1 GROUP

	Weighted Predictor-Dimensions		
Stimulus-Persons	I	II	III
1	-.14	.89	-.12
2	.54	-.51	-.43
3	-.70	-.62	.00
4	.44	.55	.25
5	.63	.10	.44
6	.81	.13	.04
7	.92	-.05	-.01
8	.85	.20	.08
9	-.89	.03	.02
10	.84	-.01	.23
11	.85	.05	.40
12	.92	-.02	.05
13	.82	.20	.41
14	.93	-.08	-.02
15	-.08	.04	-.94
Amount of variance accounted for	55%	12%	12%

TABLE 29

PREDICTOR VALUES (DERIVED AND RANK) FROM GCSA MODIFICATIONS FOR THE OTC-1 GROUP

Stimulus-Persons	Predictor Values							
	Unweighted GCSA (variance not accounted for)		Unweighted GCSA (variance accounted for)		Weighted GCSA (variance not accounted for)		Weighted GCSA (variance accounted for)	
	Derived	Rank	Derived	Rank	Derived	Rank	Derived	Rank
1[a]								
2[b]								
3	-.37	12	-.56	11	-.44	12	-.58	11
4	.37	4	.37	10	.41	3	.43	10
5	.38	2.5	.42	9	.39	4	.52	9
6	.24	9.5	.53	7.5	.33	7	.59	8
7	.24	9.5	.57	6	.29	9	.63	5
8	.25	8	.53	7.5	.38	5	.63	5
9	-.26	11	-.58	12	-.28	11	-.61	12
10	.31	5	.60	4	.35	6	.62	7
11	.38	2.5	.64	1	.43	2	.66	1.5
12	.27	6.5	.59	5	.32	8	.65	3
13	.49	1	.63	2	.48	1	.66	1.5
14	.27	6.5	.61	3	.28	10	.63	5
15[c]								

[a]Married and moved out of town, status unknown.

[b]Returned home to parents, status unknown.

[c]Enlisted in the United States Army, status unknown.

TABLE 30

EMPLOYMENT SUCCESS FIGURES FOR THE OTC-1 GROUP

	Follow-up Information			
Stimulus-Persons	Job Placement (general)	Weeks Placed (61 total possible)[a]	Income Earned[b]	Income Earned Per Week[c]
1[d]				
2[e]				
3	Clean up car lot	50	$3200.00	$64.00
4	Lumber yard	61	3050.00	50.00
5	Food service	61	2147.00	35.20
6	Food service	61	3294.00	54.00
7	General labor	47	3008.00	64.00
8	Janitor	61	5002.00	82.00
9[f]	-	0	0.00	0.00
10	Bakery	61	3050.00	50.00
11	Food service	61	3904.00	64.00
12	Food service	61	4026.00	66.00
13	Food service	61	2440.00	40.00
14	Bakery	58	3712.00	64.00
15[g]				

[a]From June 1, 1971 to August 1, 1972.

[b]Estimated figure.

[c]Based on per hour wage multiplied by hours per week worked.

[d]Married and moved out of town, status unknown.

[e]Returned home to parents, status unknown.

[f]Institutionalized for "psychotic" condition at State Hospital.

[g]Enlisted in the United States Army, status unknown.

the total "income earned" for the period of their placement time. Three stimulus-persons were eliminated from the follow-up study since their placement status was unknown.

The follow-up information was ranked and these rank values are found in Table 31. For the 14 month period allowed for vocational adjustments, these ranks indicate the success that each stimulus-person experienced, given the success criterion used. The actual success ranks were correlated (Spearman Rank-Order Correlation Coefficients) with the predictor value ranks to determine the viability of using this MESA methodology for predicting success. The correlations are found in Table 32.

Discussion of Prediction

It appeared that "variance not accounted for" in both unweighted and "weighted" GCSA modifications were the best predictors of "weeks placed," where none of the predictor values were significantly related to the two "income" categories. It should be noted, however, that the highest nonsignificant correlations with all three success value ranks were found with the "weighted-variance accounted for" predictor values (the values hypothesized as being the best predictors).

In this study, it appeared as though "accounting for variance" actually interferred with prediction. This was demonstrated by the fact that "variance accounted for" in both the "weighted" and unweighted cases led to lower relationships with the success values. One possible explanation for this non-hypothesized event is that all three judgment dimensions contained information relevant to successful prediction. Thus, accounting for the factor loadings on all three predictor-

TABLE 31

RANK EMPLOYMENT SUCCESS FOR THE OTC-1 GROUP

Stimulus-Persons	Rank Weeks Placed[a]	Rank Income Earned[a]	Rank Income Earned Per Week
1[b]			
2[c]			
3	10	6	4.5
4	4.5	7.5	8.5
5	4.5	11	11
6	4.5	5	7
7	11	9	4.5
8	4.5	1	1
9	12	12	12
10	4.5	7.5	8.5
11	4.5	3	4.5
12	4.5	2	2
13	4.5	10	10
14	9	4	4.5
15			

[a]From June 1, 1971 to August 1, 1972.

[b]Married and moved out of town, status unknown.

[c]Returned home to parents, status unknown.

[d]Enlisted in the United States Army, status unknown.

TABLE 32

SPEARMAN RANK CORRELATION COEFFICIENTS (r_S)[a] BETWEEN RANK SUCCESS AND PREDICTOR VALUES FOR THE OTC-1 GROUP

	Rank Weeks Placed	Rank Income Earned	Rank Income Earned Per Week
Unweighted GCSA (variance not accounted for)	.699[c]	-.074	-.260
Unweighted GCSA (variance accounted for)	.392	.358	.282
Weighted GCSA (variance not accounted for)	.799[d]	.011	-.185
Weighted GCSA (variance accounted for)	.416	.459	.459

[a]All Spearman Rank Correlation Coefficients were corrected for tied ranks.

[b]N = 12.

[c]$p < .05$ (two-tailed test).

[d]$p < .01$ (two-tailed test).

dimensions equally, produced a reasonably high correlation with at least one success variable (that of "weeks placed").

When the "variance accounted for" computations were correlated with success, Prediction-dimension I appeared to be given much more predictive power than it really should have. For example, since Predictor-dimensions I, II and III accounted for 55 percent, 12 percent and 12 percent of the variance, respectively. Predictor-dimension I

was weighted 4.58 (or .55/.12) where predictor-dimensions II and III were weighted 1.00 (or .12/.12 and .12/.12). Thus, accounting for variance in this study meant counting Predictor-dimension I as being 4.58 times more "important" when, perhaps, that predictor-dimension was not such a powerful predictor of success.

Again, this study does not show that any other prediction method would be better or worse than the one employed; although this problem will be dealt with shortly.

The mean similarity estimation for each judge, the judge similarity intercorrelation matrix, and the Observer Factor Judgment Analysis (OFJA) were all reported for the SS-1 and OTC-1 groups in Hanson (1971). Generally, there was only one judge-dimension for the SS-1 group but there were two judge-dimensions for the OTC-1 group. This point is important for determining the success of prediction as will be discussed in Chapter V.

Cross-Validation Studies

The SS-2 Study

The three judges of the SS-2 group, two of whom were involved in the previous SS-1 study, were asked to evaluate 14 stimulus-persons who were being considered as possible vocational trainees within the next year. The reliability estimate of the judges' similarity estimations (determined by correlating the S_{ij} and S_{ji} stimulus-person pairs) was moderately high (r = .72, df = 89, p <.001). This reliability estimate, however, was considerably lower than that found in the other five investigations.

The average (arithmetic mean) similarity estimation produced by each judge is found in Table 33. These values indicate that all three judges appeared to cluster their similarity estimations between 50 and 55. The standard deviations in that table indicate that a similar narrow range of values between judges were used to describe perceived similarity between all possible pair combinations of stimulus-persons.

TABLE 33

MEANS AND STANDARD DEVIATIONS OF THE SIMILARITY ESTIMATIONS MADE BY EACH JUDGE IN THE SS-2 GROUP

Judge	Mean Similarity	Standard Deviation
A	.50	.15
B	.54	.13
C	.55	.17

The similarity estimations made for each stimulus-person pair were intercorrelated for each judge to begin the Observer Factor Judgment Analysis (OFJA). The result of this analysis is found in Table 34 and indicates rather low agreement between the three judges. The OFJA dimensional analysis continued by factor analyzing (principal components) the intercorrelation matrix (Table 34) and two judge-dimensions were produced (Table 35). A total of 75 percent of the judgmental variance was explained by the two judge-dimensions. Judges A and C had high positive loadings on Judge-dimension I where Judge B held the highest positive loading on Judge-dimension II. It should be noted that Judges A and C were directly involved with vocational placements where Judge B was not directly involved with vocational placements (see Table 4, p. 44).

TABLE 34

INTER-JUDGE SIMILARITY CORRELATION COEFFICIENTS FOR THE SS-2 GROUP

	Judges		
Judges	A	B	C
A	1.00	.12	.20
B		1.00	-.05
C			1.00

TABLE 35

UNROTATED OFJA FACTOR MATRIX FOR THE SS-2 GROUP

	Judge-Dimensions		
Judges	I	II	h^2
A	.81	.12	.67
B	.28	.89	.86
C	.69	-.50	.73
Amount of variance accounted for	40%	35%	75%

Prediction of Success

The Group Composition Structure Analysis (GCSA) began for the SS-2 group by averaging (across judges) the similarity estimations made for each stimulus-person pair. The column vectors of this arithmetic mean similarity estimation matrix were intercorrelated to produce the similarity correlation coefficient matrix. Both matrices are found in Table 36.

TABLE 36

ARITHMETIC MEAN (ACROSS JUDGES) SIMILARITY ESTIMATIONS, SIMILARITY CORRELATION COEFFICIENTS, MEAN SIMILARITY, AND STANDARD DEVIATIONS OF THE SIMILARITY ESTIMATIONS FOR THE SS-2 GROUP

	Stimulus-Persons													
Stimulus-Persons	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	1.00	.36	.48	.46	.53	.63	.36	.61	.50	.54	.58	.56	.44	.43
2	-.65	1.00	.49	.57	.50	.36	.67	.53	.46	.52	.51	.48	.53	.57
3	-.16	-.13	1.00	.57	.70	.60	.53	.62	.46	.71	.70	.34	.76	.62
4	-.39	.32	.09	1.00	.48	.43	.69	.53	.46	.45	.58	.41	.58	.55
5	-.04	-.13	.65	-.18	1.00	.56	.55	.54	.44	.72	.64	.43	.68	.60
6	.49	-.66	.13	-.47	.10	1.00	.48	.54	.50	.61	.64	.60	.52	.38
7	-.68	.58	.08	.62	.04	-.42	1.00	.46	.44	.53	.49	.38	.61	.57
8	.31	-.14	.27	-.06	.05	.08	-.32	1.00	.53	.58	.68	.44	.53	.53
9	.08	-.24	-.38	-.29	-.40	.03	-.38	-.03	1.00	.49	.54	.58	.45	.43
10	-.01	-.14	.68	-.27	.70	.21	-.03	.17	-.26	1.00	.64	.40	.71	.60
11	.18	-.25	.58	-.01	.40	.36	-.20	.49	-.06	.43	1.00	.43	.66	.48
12	.27	-.22	-.71	-.45	-.48	.25	-.50	-.28	.36	-.53	-.44	1.00	.35	.47
13	-.35	.05	.79	.22	.62	-.09	.32	.02	-.41	.64	.43	-.73	1.00	.59
14	-.45	.29	.30	.18	.28	-.55	.28	-.07	-.43	.24	-.23	-.34	.33	1.00
Mean Similarity	.53	.54	.61	.55	.60	.56	.55	.58	.52	.61	.61	.49	.60	.56
Standard Deviation	.16	.15	.16	.15	.15	.16	.16	.14	.15	.15	.14	.17	.16	.15

Note:

Similarity estimates are above the major diagonal (unity values) and correlational similarities are below the major diagonal.

The similarity correlation coefficient matrix was factor analyzed (principal components) and rotated (varimax) to produce the four judgment-dimensions found in the rotated GCSA factor matrix (Table 37). The four judgment-dimensions accounted for 76 percent of the total judgmental variance.

Judgment-dimension I accounted for 30 percent of the total judgmental variance and was interpreted by the judges as a Skill Development dimension. Stimulus-persons who received high positive factor loadings were, in general, viewed as people having more work experience, actually doing better work while institutionalized and were also considered as having a more mature attitude toward their skill abilities and limitations. As a result of this interpretation, high positive loadings on Judgment-dimension I were considered indicative of vocational success.

Judgment-dimension II accounted for 19 percent of the total judgmental variance and was interpreted as an Age-Maturity dimension with older stimulus persons having high negative loadings. In addition, the older stimulus-persons were seen as more mature workers where the younger stimulus-persons (loading positively on this dimension) were seen as immature. Indeed, judgment-dimension II correlated moderately high with chronological age ($r = .69$, $df = 12$, $p < .01$) and with months as "client with the agency" ($r = -.60$, $df = 12$, $p = < .05$). As a result of this interpretation, high negative loadings were seen as indicative of vocational success.

Judgment-dimension III accounted for 11 percent of the total judgmental variance (which made this judgment-dimension the least "important" one of the four) and was interpreted as a Personality dimension. Stimulus-person-8 was seen as quiet, withdrawn, and

TABLE 37

ROTATED GCSA FACTOR MATRIX FOR THE SS-2 GROUP

	Unweighted Judgment-Dimensions				
Stimulus-Persons	I	II	III	IV	h^2
1	-.17	.56	.32	.47	.67
2	-.09	-.50	-.10	-.62	.65
3	.87	-.10	.26	.06	.84
4	-.04	-.87	.12	-.12	.79
5	.86	.18	-.05	-.02	.77
6	.14	.37	-.07	.82	.83
7	.17	-.81	-.32	-.25	.86
8	.08	.14	.92	.00	.88
9	-.51	.23	.07	.21	.36
10	.85	.23	.06	.02	.79
11	.49	-.07	.56	.47	.78
12	-.68	.51	-.35	.13	.86
13	.86	-.30	.03	-.03	.82
14	.39	.00	-.07	-.81	.81
Amount of variance accounted for	30%	19%	11%	16%	76%

capable whereas Stimulus-person-12 was described as loud and boastful. It was determined by the judges that a reserved manner of behavior would meet with more success than a loud and boastful manner. As a consequence, high positive loadings on Judgment-dimension III were seen as predictive of vocational success.

Judgment-dimension IV accounted for 16 percent of the total judgmental variance and was revealed as a Vocational Attitude dimension by judge interpretation. Stimulus-person 14 was seen as "determined to succeed" by the judges whereas Stimulus-person-6 was described as not being highly motivated for vocational success. Thus, high negative loadings on Judgment-dimension IV were considered as being more indicative of success. In addition, Judgment-dimension IV correlated moderately high with IQ ($r = -.66$, $df = 12$, $p < .01$) and with Judge Familiarity Ratings ($r = -.55$, $df = 12$, $p < .05$). These relationships seem to indicate that the judges became more familiar with stimulus-persons who showed favorable vocational attitude. Also, stimulus-persons with a higher IQ were generally seen as having higher motivation and were more successful.

The predictor-dimensions (Table 38) were created from the judgment-dimensions so that positive loadings on predictor-dimensions were indicative of success. Since positive loadings on Judgment-dimensions I and III were interpreted as predictors of successful vocational placement, these dimensions remained unchanged in Table 38 and predictor-dimensions. However, high negative loadings on Judgment-dimensions II and IV were viewed as being indicative of success so the algebraic signs were reflected as they were modified to be predictor-dimensions.

TABLE 38

THE UNWEIGHTED PREDICTOR-DIMENSIONS FOR THE SS-2 GROUP

	Unweighted Predictor-Dimensions			
Stimulus-Persons	I	II	III	IV
1	-.17	-.56	.32	-.47
2	-.09	.50	-.10	.62
3	.87	.10	.26	-.06
4	-.04	.87	.12	.12
5	.86	-.18	-.05	.02
6	.14	-.37	-.07	-.82
7	.17	.81	-.32	.25
8	.08	-.14	.92	.00
9	-.51	-.23	.07	-.21
10	.85	-.23	.06	-.02
11	.49	.07	.56	-.47
12	-.68	-.51	-.35	-.13
13	.86	.30	.03	.03
14	.39	.00	-.07	.81
Amount of variance accounted for	30%	19%	11%	16%

The purpose of creating a predictor-dimension factor structure from judgment-dimensions was to increase simplicity as an aid for further analyses. The actual factor loadings and the amount of variance that each judgment-dimension accounted for were not changed in any way (see Table 38). The unweighted predictor-dimensions served as the basis for determining two different sets of predictor values, to be described later.

Earlier in this paper, it was hypothesized that the judges may differ in their ability to predict vocational success and that more accurate predictions would be possible if this factor were accounted for. As before, the familiarity ratings assigned to the stimulus-persons were employed for this purpose. The actual familiarity values assigned to each stimulus-person by each judge are found in Table 39. When these ratings were averaged for each judge, it became clear how familiar each judge felt he was with the stimulus-persons. The method of determining the "weights" for each judge entailed using the lowest judge familiarity average as a base of 1.00 and adjusting the other judges' average ratings accordingly. The results of these calculations are found in Table 40.

Since the original GCSA was completed using a deck of computer cards for each judge (where the similarity estimations have been punched), the "weighted" GCSA was completed by duplicating a judge deck depending on his "weighted" value. In the case of the SS-2 judge group, Judges A, B, and C received "weights" of 1.00, 3.00, and 2.00, respectively. Thus, Judge A was accounted for one time, Judge B three times and Judge C two times in the "weighted" GCSA.

TABLE 39

JUDGE FAMILIARITY RATINGS FOR THE SS-2 GROUP BASED ON A FIVE-POINT SCALE

	Judges			
Stimulus-Persons	A	B	C	Average
1	2	4	4	3.33
2	3	4	5	4.00
3	3	5	5	4.33
4	2	4	2	2.67
5	1	4	1	2.00
6	1	3	3	2.33
7	2	4	4	3.33
8	1	3	2	2.00
9	1	4	3	2.67
10	1	4	2	2.33
11	1	3	1	1.67
12	1	4	5	3.33
13	2	3	2	2.33
14	3	5	5	4.33
Average	1.71	3.86	3.14	(2.90)

TABLE 40

MEAN JUDGE FAMILIARITY RATING AND JUDGE "WEIGHT" VALUE FOR THE SS-2 GROUP

	Judges		
Familiarity Values	A	B	C
Raw mean[a]	1.71	3.86	3.14
Weighted value[b]	1.00	3.00	2.00

[a]Where five indicated "know very well" and one indicated "do not know very well."

[b]Where the lowest "raw mean" value in each judge group was set at 1.00 and the corresponding values for other judges were lowered proportionately and rounded to the nearest whole number.

The "weighted" GCSA analyses for the SS-2 group (now based on six judgmental computer decks) began by averaging (across judges) the similarity estimations. The column vectors of the mean similarity estimation matrix were intercorrelated to produce the similarity correlation coefficient matrix. Both matrices are found in Table 41. The similarity correlation coefficient matrix, in turn, was factor analyzed (principal components) and rotated (varimax) to produce the four "weighted" judgment-dimensions found in Table 42. These "weighted" judgment-dimensions were shown to be very similar to the unweighted judgment-dimensions that were originally produced from the SS-2 group (Table 37). The intercorrelations between unweighted and "weighted" judgment-dimensions are found in Table 43. This table indicates the unweighted Judgment-dimension II became the "weighted" Judgment-dimension I and the unweighted Judgment-dimension I became

TABLE 41

ARITHMETIC MEAN (ACROSS JUDGES) SIMILARITY ESTIMATIONS, SIMILARITY CORRELATION COEFFICIENTS, MEAN SIMILARITY, AND STANDARD DEVIATIONS OF THE SIMILARITY ESTIMATIONS FOR THE WEIGHTED SS-2 GROUP

	Stimulus-Persons													
Stimulus-Persons	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	1.00	.35	.51	.45	.58	.66	.35	.60	.56	.59	.59	.58	.51	.41
2	-.70	1.00	.51	.53	.50	.34	.63	.52	.48	.51	.50	.51	.53	.58
3	-.04	-.16	1.00	.53	.71	.60	.52	.65	.52	.70	.71	.38	.75	.59
4	-.43	.18	-.10	1.00	.43	.42	.69	.52	.48	.41	.53	.43	.55	.51
5	.12	-.18	.63	-.40	1.00	.57	.55	.53	.52	.73	.65	.46	.71	.58
6	.61	-.72	.16	-.45	.17	1.00	.46	.53	.56	.62	.65	.60	.54	.37
7	-.73	.48	-.09	.60	-.07	-.49	1.00	.43	.45	.50	.48	.43	.58	.58
8	.27	-.16	.31	-.12	-.04	.08	-.46	1.00	.58	.55	.69	.48	.52	.50
9	.25	-.28	-.10	-.25	-.11	.21	-.43	.23	1.00	.54	.62	.54	.53	.45
10	.21	-.22	.61	-.48	.73	.30	-.22	.04	-.02	1.00	.61	.42	.70	.57
11	.29	-.32	.58	-.17	.37	.42	-.36	.54	.30	.33	1.00	.44	.65	.45
12	.24	-.12	-.66	-.35	-.41	.20	-.38	-.18	.10	-.43	-.39	1.00	.40	.53
13	-.15	-.04	.74	.00	.64	-.02	.13	-.07	-.16	.60	.35	-.63	1.00	.61
14	-.53	.36	.11	.04	.13	-.62	.30	-.24	-.40	.05	-.42	-.11	.25	1.00
Weighted Mean Similarity	.55	.53	.62	.53	.61	.56	.55	.58	.56	.60	.61	.51	.61	.55
Standard Deviation	.16	.15	.15	.15	.15	.16	.16	.14	.14	.15	.14	.16	.15	.15

Note:

Similarity estimates are above the major diagonal (unity values) and correlational similarities are below the major diagonal.

TABLE 42

THE WEIGHTED GCSA FACTOR MATRIX (ROTATED) FOR THE SS-2 GROUP

	Weighted Judgment-Dimensions				
Stimulus-Persons	I	II	III	IV	h^2
1	.41	-.04	.21	.74	.76
2	-.15	-.12	-.06	-.83	.72
3	-.05	.88	.25	.01	.85
4	-.90	-.16	-.07	-.16	.88
5	.29	.83	-.14	.08	.80
6	.24	.12	.03	.88	.84
7	-.68	.05	-.49	-.40	.86
8	.06	.08	.89	.01	.80
9	.18	-.18	.53	.27	.41
10	.37	.80	-.05	.14	.80
11	-.07	.50	.64	.37	.80
12	.52	-.73	-.19	.15	.86
13	-.11	.88	-.09	-.08	.79
14	.17	.19	-.33	-.76	.75
Amount of variance accounted for	15%	27%	14%	22%	78%

TABLE 43

CORRELATION COEFFICIENTS (r) BETWEEN UNWEIGHTED AND WEIGHTED GCSA FACTOR MATRICES FOR THE SS-2 GROUP

	Weighted Judgment-Dimensions			
Unweighted Judgment-Dimensions	1	2	3	4
1	-.055	.952	.021	.315
2	.970	-.063	.232	.023
3	-.107	.321	.881	.447
4	-.124	.551	.273	.960

the "weighted" Judgment-dimension II. The order of factor extraction remained the same for unweighted and "weighted" judgment-dimensions III and IV. The order of factor extraction for the factor analyses was not important and the high intercorrelations between unweighted and "weighted" judgment-dimensions meant that there was no need for a reinterpretation of the dimensions by the SS-2 judges.

The "weighted" predictor-dimensions were created (Table 44) to reflect in interpretations made regarding the unweighted predictor-dimensions. Tables 38 and 44 show that the "weighted" predictor-dimensions are not radically different from the originally extracted unweighted predictor-dimensions. The "weighted" predictor-dimensions, however, have the hypothesized advantage of accounting for differential predictive ability of judges as assessed by how familiar the judges were with the abilities of the stimulus-persons.

TABLE 44

THE WEIGHTED PREDICTOR-DIMENSIONS FOR THE SS-2 GROUP

	Weighted Predictor-Dimensions			
Stimulus-Persons	I	II	III	IV
1	-.41	-.04	.21	-.74
2	.15	-.12	-.06	.83
3	.05	.88	.25	-.01
4	.90	-.16	-.07	.16
5	-.29	.83	-.14	-.08
6	-.24	.12	.03	-.88
7	.68	.05	-.49	.40
8	-.06	.08	.89	-.01
9	-.18	-.18	.53	-.27
10	-.37	.80	-.05	-.14
11	.07	.50	.64	-.37
12	-.52	-.73	-.19	-.15
13	.11	.88	-.09	.08
14	-.17	.19	-.33	.76
Amount of variance accounted for	15%	27%	14%	22%

Another variable was hypothesized as being important, that of the "importance" of judgment-dimensions. As described in earlier studies, it was hypothesized that the "amount of variance accounted for" by each judgment-dimension directly corresponded to the dimension's "importance" as a predictor of success. Again, the predictor value was derived by using the lowest variance accounted for value (counted as 1.00) and dividing the higher judgment-dimension percentages by the lowest percentage value. Of course, when unweighted and "weighted" predictor-dimensions were used to derive predictor values that did not account for variance, all predictor-dimensions received a weight of 1.00 and thus were equally weighted. The calculation of each predictor value for each stimulus person then followed the generalized formula described earlier (p. 39) and all computations were computer programmed.

The results of these calculations are found in Table 45 and all predictor values have been ranked from highest to lowest. These ranks were later correlated with actual success values that were also converted to ranks.

It was clear that successful prediction by predictor values derived in the manner described above would not exclude the possibility that another prediction method could be as good or better. The most straight forward method used for this comparison (as a prediction method) was that of probability ratings. As described in Chapter III, the judges were asked to estimate "probable success" (using values from zero to 100) on vocational judgment. The results of this rating are

TABLE 45

PREDICTOR VALUES (DERIVED AND RANK) FROM GCSA MODIFICATIONS FOR THE SS-2 GROUP

Stimulus-Persons	Unweighted GCSA (variance not accounted for)		Unweighted GCSA (variance accounted for)		Weighted GCSA (variance not accounted for)		Weighted GCSA (variance accounted for)	
	Derived	Rank	Derived	Rank	Derived	Rank	Derived	Rank
1	-.22	9.5	-.26	10	-.24	10.5	-.26	11
2	.23	5.5	.21	7	.20	5	.21	3
3	.29	2	.39	2	.29	1	.36	1
4	.27	4	.24	6	.21	3.5	.15	8
5[a]								
6	-.28	11	-.22	9	-.24	10.5	-.25	10
7	.23	5.5	.28	4.5	.16	6	.17	6
8[a]								
9	-.22	9.5	-.29	11	-.02	9	-.08	9
10	.16	7.5	.28	4.5	.06	8	.16	7
11	.16	7.5	.19	8	.21	3.5	.20	4
12	-.42	12	-.47	12	-.40	12	-.43	12
13	.30	1	.43	1	.24	2	.33	2
14	.28	3	.31	3	.11	7	.19	5

[a]Placed but status unknown.

found in Table 46 and the individual probability judgments have been averaged.

Since it was probable that the individual judges used a different method of assessing probability, and, since the standard deviations of each judge vary, the probability estimations for each judge were standardized (mean = .50, standard deviation = .10) and averaged (see Table 47). In addition, the standardized probability averages were "weighted" in the identical manner that the unweighted and "weighted" predictor-dimensions were computed. These values are also found in Table 47. There were seven predictor values that were then compared to the actual success values which were gathered in the follow-up study.

Follow-up of Actual Success

The follow-up information was dated 56 weeks following the initial judgments (both similarity estimations and probability ratings). The specific information required for follow-up was not designated since no information was collected specifically for use in this study. The follow-up information is found in Table 48 and two stimulus-persons were eliminated since their vocational status was unknown. The values of the three success categories were then ranked and these are found in Table 49.

It is not claimed that this follow-up information completely represents all the factors that may be considered in representing "vocational success." However, it is assumed that much of what is considered successful adjustment is included in these success values. The three success values were correlated (Spearman Rank-Order

TABLE 46

STIMULUS-PERSON SUCCESS PROBABILITY RATINGS FOR THE SS-2 JUDGE GROUP

	Judges			
Stimulus-Persons	A	B	C	Average
1	.45	.80	.45	.57
2	.80	.60	.85	.75
3	.50	.40	.65	.52
4	.60	.50	.60	.57
5	.50	.60	.70	.60
6	.50	.60	.30	.47
7	.50	.50	.50	.50
8	.50	.40	.40	.43
9	.20	.50	.50	.40
10	.70	.50	.50	.57
11	.70	.50	.53	.58
12	.30	.60	.40	.43
13	.70	.60	.75	.68
14	.50	.50	.98	.66
Average	.53	.54	.58	(.55)
Standard Deviation	.16	.10	.18	

TABLE 47

STIMULUS-PERSON SUCCESS PROBABILITY RATINGS (STANDARDIZED) FOR THE SS-2 JUDGE GROUP

	Judges				
Stimulus-Persons	A	B	C	Average	Weighted Average
1	.45	.76	.43	.547	.598
2	.67	.56	.65	.627	.608
3	.48	.35	.54	.457	.435
4	.54	.46	.51	.503	.490
5	.48	.56	.57	.537	.550
6	.48	.56	.35	.463	.477
7	.48	.46	.46	.467	.463
8	.48	.35	.40	.410	.388
9	.29	.46	.46	.403	.427
10	.61	.46	.46	.510	.427
11	.61	.46	.47	.513	.488
12	.35	.56	.40	.437	.472
13	.61	.56	.59	.587	.578
14	.48	.46	.72	.553	.550
Average	.50	.50	.50	(.500)	(.500)
Standard Deviation	.10	.10	.10		

TABLE 48

EMPLOYMENT SUCCESS FIGURES FOR THE SS-2 GROUP

	Follow-up Information			
Stimulus-Persons	Job Placement Location	Total Weeks Placed (56 possible)[a]	Present Monthly Wage[b]	Total Income Earned[c]
1	Training Center	26	$113.33	$215.44
2	Training Center	44	121.00	606.54
3	Training Center	44	84.00	74.79
4	Training Center	44	200.00	800.00
5[d]				
6	Not referred[e]	0	0.00	0.00
7	Not referred[e]	0	0.00	0.00
8[f]				
9	Training Center[g]	0+	0.00+	0.00+
10	Not referred[e]	0	0.00	0.00
11	Training Center[g]	0+	0.00+	0.00+
12	Not referred[e]	0	0.00	0.00
13	Direct Placement	5	50.00	110.00
14	Training Center	56	15.04	370.67

[a]From June 1, 1972 to July 1, 1973.

[b]Based on most current monthly report (June, 1973).

[c]Actual income earned, regardless of most current monthly wage.

[d]Training status unknown.

[e]"Not referred" indicates that this person was not considered for any kind of placement to date.

[f]Training status unknown.

[g]Stimulus-persons 9 and 11 were recently referred for placement but have not had the opportunity to work. These persons were given a success score of "0+" since such a status can be considered better than "not referred."

TABLE 49

RANK EMPLOYMENT SUCCESS FIGURES FOR THE SS-2 GROUP

	Follow-up Information		
Stimulus-Persons	Rank Total Weeks Placed[a]	Rank Monthly Wage[b]	Rank Total Income Earned[c]
1	5	3	4
2	3	2	2
3	3	4	6
4	3	1	1
5[d]			
6	10.5	10.5	10.5
7	10.5	10.5	10.5
8[e]			
9	7.5	7.5	7.5
10	10.5	10.5	10.5
11	7.5	7.5	7.5
12	10.5	10.5	10.5
13	6	5	5
14	1	6	3

[a]From June 1, 1972 to July 1, 1973.

[b]Based on the most current montly report (June, 1973).

[c]Actual income earned, regardless of most current monthly wage.

[d]Training status unknown.

[e]Training status unknown.

Correlation Coefficients) with the seven predictor values and the results of these correlations are found in Table 50.

Discussion of Prediction

The only GCSA modification predictor values that significantly correlated with a success value (rank weeks placed) was the unweighted GCSA predictor values where variance was not accounted for. This relationship was not expected and is difficult to explain, short

TABLE 50

SPEARMAN RANK CORRELATION COEFFICIENTS (r_S)[a] BETWEEN RANK SUCCESS AND PREDICTOR VALUES FOR THE $\overline{SS}$-2 GROUP[b]

	Rank Weeks Placed	Rank Monthly Wage	Rank Total Income Earned
Unweighted GCSA (variance not accounted for)	.646[c]	.523	.556
Unweighted GCSA (variance accounted for)	.392	.225	.261
Weighted GCSA (variance not accounted for)	.505	.508	.447
Weighted GCSA (variance accounted for)	.474	.349	.335
Probability Rating	.564	.532	.625[c]
Standardized Probability Rating	.478	.474	.581[c]
Standardized Probability Rating (weighted)	.554	.636[c]	.707[d]

[a]All Spearman Correlation Coefficients were corrected for tied ranks.

[b]$N = 9$.

[c]$p < .05$ (two-tailed test).

[d]$p < .01$ (two-tailed test).

of considering it a spurious statistically significant relationship. The efficacy of "weighting" by judge familiarity is shown by the fact that the "weighted" standardized probability value correlated with rank total income earned and produced the highest computed correlation coefficient. In short, it appeared as though the most direct form of prediction (that of rating probable success) was a superior method of predicting vocational success (rank total income earned) for the SS-2 group.

It is interesting to note that unweighted Judgment-dimension IV correlated (Pearson product-moment) moderately high ($r = -.611$, $df = 7$, $p < .05$) with "weeks placed." This correlation was derived from values that excluded the two stimulus-persons from the follow-up study. It may help explain the fact that the unweighted Judgment-dimension IV correlated significantly with weeks placed. Since variance was not accounted for in that relationship, the low variance weight value of Judgment-dimension IV was not obscured by accounting for variance and consequently obscuring the predictor value of Judgment-dimension IV. This explanation is substantiated by the fact that the "weighted" GCSA predictor value, where variance was "not accounted for," correlated higher (though not significantly) with rank weeks-placed than did the GCSA predictor values where variance was accounted for.

Another explanation of the disarray shown by the modified GCSA predictor values may be found with the predictive ability of the judges. It was shown (Table 35, p. 93) that Judge B loaded on a dimension by himself, thus demonstrating an underlying judgment criterion that differed from the other two judges. It was

also documented that Judge-B was given the largest "weight" (3.00) and yet was not as directly involved with placement procedures as were Judges A and C. In addition, the rather low reliability estimate between stimulus-pair presentations reported earlier may indicate that this group of judges probably did not tediously evaluate each stimulus-person pair combination. With lower reliability in judging, one would expect lowered predictive reliability.

The OTC-2 Study

The four judges, of the OTC-2 study, all of which were used in the earlier OTC-1 study, were asked to evaluate (using similarity estimations) 11 stimulus-persons. The reliability of the judges' similarity estimations (estimated by the correlation between S_{ij} and S_{ji} stimulus-person pairs) was very high ($r = .94$, $df = 53$, $p < .001$).

The average (arithmetic mean) similarity estimation and the standard deviation of each made by the judges is found in Table 51. These values indicate that the four judges made similarity estimations having somewhat variable means and vastly diverse standard deviations. For example, Judge-D had a smaller degree of judgment variation (SD = .14) than Judge-B, who had the greatest variability of similarity estimates (SD = .35).

The Observer Factor Judgment Analysis (OFJA) began by intercorrelating the similarity estimations made by each judge. A table of these intercorrelations is found in Table 52. These correlations are similarly high and, as would be predicted in such a case, a single judge-dimension (Table 53) was extracted. This judge-

TABLE 51

MEANS AND STANDARD DEVIATIONS OF THE SIMILARITY ESTIMATIONS MADE BY EACH JUDGE IN THE OTC-2 GROUP

Judge	Mean Similarity	Standard Deviation
A	.54	.28
B	.46	.35
C	.45	.27
D	.53	.14

TABLE 52

INTER-JUDGE SIMILARITY CORRELATION COEFFICIENTS FOR THE OTC-2 GROUP

	Judges			
Judges	A	B	C	D
A	1.00	.83	.85	.78
B		1.00	.88	.79
C			1.00	.84
D				1.00

TABLE 53

UNROTATED OFJA FACTOR MATRIX FOR THE OTC-2 GROUP

	Judge-Dimension	
Judges	I	h^2
A	.93	.86
B	.94	.88
C	.96	.92
D	.91	.83
Amount of variance accounted for	87%	

dimension was a product of a principal components factor analysis of the intercorrelation matrix and is monopolar. Such an analysis indicates that the four judges were using a similar underlying judgment criterion in assessing the vocational potential of the OTC-1 stimulus-persons. Such a finding is quite different from the diversity of judgment criterion used by the judges in the SS-2 group described earlier.

Prediction of Success

The Group Composition Structure Analysis (GCSA) began by averaging (arithmetic mean across judges) the similarity estimations made for each stimulus-person pair. The column vectors of this mean similarity estimation matrix (Table 54) were intercorrelated to produce the similarity correlation coefficient matrix also found in Table 54. The similarity correlation coefficient matrix was then

TABLE 54

ARITHMETIC MEAN (ACROSS JUDGES) SIMILARITY ESTIMATIONS, SIMILARITY CORRELATION COEFFICIENTS, MEAN SIMILARITY, AND STANDARD DEVIATIONS OF THE SIMILARITY ESTIMATIONS FOR THE OTC-2 GROUP

	Stimulus-Persons										
Stimulus-Persons	1	2	3	4	5	6	7	8	9	10	11
1	1.00	.41	.63	.61	.46	.71	.44	.55	.51	.51	.52
2	-.55	1.00	.55	.51	.78	.49	.69	.48	.23	.71	.68
3	.21	.09	1.00	.53	.57	.66	.63	.59	.44	.68	.61
4	.27	-.32	-.19	1.00	.50	.68	.48	.76	.52	.61	.54
5	-.45	.88	.14	-.35	1.00	.49	.75	.46	.20	.68	.69
6	.60	-.42	.27	.45	-.40	1.00	.57	.71	.53	.58	.59
7	-.44	.77	.30	-.39	.84	-.22	1.00	.48	.21	.69	.71
8	.18	-.50	-.06	.71	-.55	.52	-.49	1.00	.61	.56	.55
9	.30	-.88	-.27	.28	-.93	.24	-.91	.52	1.00	.36	.28
10	-.35	.71	.35	-.13	.66	-.22	.68	-.33	-.71	1.00	.71
11	-.29	.69	.25	-.24	.72	-.15	.76	-.34	-.83	.68	1.00
Mean Similarity	.58	.59	.62	.61	.60	.64	.60	.61	.44	.64	.62
Standard Deviation	.17	.21	.14	.15	.21	.14	.21	.16	.24	.16	.18

Note:

Similarity estimates are above the major diagonal (unity values) and correlational similarities are below the major diagonal.

factor analyzed (principal components) and rotated (varimax) to produce the three judgment-dimensions found in Table 55. The three judgment-dimensions accounted for 81 percent of the total judgmental variance.

TABLE 55

ROTATED GCSA FACTOR MATRIX FOR THE OTC-2 GROUP

	Unweighted Judgment-Dimensions			
Stimulus-Persons	I	II	III	h^2
1	.40	.73	.13	.71
2	-.86	-.28	-.23	.88
3	-.31	.77	-.17	.71
4	.13	.01	.92	.86
5	-.87	-.19	-.29	.88
6	.15	.69	.54	.79
7	-.88	.00	-.27	.85
8	.32	.08	.85	.83
9	.94	-.01	.22	.93
10	-.84	.02	-.05	.71
11	-.86	.05	-.10	.76
Amount of variance accounted for	45%	16%	20%	81%

Judgment-dimension I accounted for 45 percent of the total judgmental variance and was interpreted by the judges as an Employment Success dimension. Those stimulus-persons with high negative loadings were described as the "best prospects" and the most proven workers of

all the stimulus-persons. Stimulus-person-9 (with the highest positive loading) was described as the "poorest prospect" for vocational placement success. Thus, high negative factor loadings on Judgment-dimension I were considered as being predictive of vocational success. It appeared that those who were with the agency longer were considered less successful since Judgment-dimension I correlated moderately high with "months as client of agency" ($r = .67$, $df = 9$, $p < .05$).

Judgment-dimension II accounted for 16 percent of the judgmental variance and was interpreted as an Emotional Secondary Handicap dimension. It was clear that those stimulus-persons with high positive loadings had handicaps secondary to mental retardation, which were primarily emotional (psychiatric) difficulties. Aside from lack of vocational skills, these stimulus-persons were considered to be poor vocational prospects due to the emotional difficulty which would interfere with their working ability. Those with high negative loadings were considered indicative of vocational success.

Judgment-dimension III accounted for 20 percent of the total judgmental variance and was interpreted as a Family Influence dimension. Those stimulus-persons with high positive loadings were characterized as having families who greatly influenced their training program and future success. Those with high negative loadings had no family problems and thus stimulus-persons with high negative loadings on Judgment-dimension III were considered better prospects for vocational placement success.

The unweighted Predictor-dimensions (Table 56) were created to clarify the relationship between positive and negative factor loadings and the interpretations. Only the algebraic signs of the judgment-

TABLE 56

THE UNWEIGHTED PREDICTOR-DIMENSIONS FOR THE OTC-2 GROUP

	Unweighted Predictor-Dimensions		
Stimulus-Persons	I	II	III
1	-.40	-.73	-.13
2	.86	.28	.23
3	.31	-.77	.17
4	-.13	-.01	-.92
5	.87	.19	.29
6	-.15	-.69	-.54
7	.88	.00	.27
8	-.32	-.08	-.85
9	-.94	.01	-.22
10	.84	-.02	.05
11	.86	-.05	.10
Amount of variance accounted for	45%	16%	20%

dimensions in Table 55 were changed in order to create predictor-dimensions (Table 56). In reviewing the interpretations given the three unweighted judgment-dimensions, it was determined that negative factor loadings on all three judgment-dimensions were indicative of attributes that were considered predictive of future vocational success. As a result, the algebraic signs of all three judgment-dimensions were reflected to create the predictor-dimensions which were later used to calculate predictor values.

The judges' familiarity ratings for the OTC-2 group are found in Table 57. As in the studies described earlier, the average familiarity rating (where one indicated "do not know very well" and five indicated "know very well") were used to determine the "weight" of each judge. In short, it was assumed that those judges who were more familiar with the stimulus-persons would be "more informed" and consequently better judges of future vocational success.

TABLE 57

JUDGE FAMILIARITY RATINGS FOR THE OTC-2 GROUP BASED ON A FIVE-POINT SCALE

Stimulus-Persons	A	B	C	D	Average
1	4	4	4	4	4.00
2	4	4	3	4	3.75
3	2	3	1	3	2.25
4	3	3	4	4	3.50
5	2	4	2	3	2.75
6	5	4	5	5	4.75
7	2	4	4	3	3.25
8	5	3	5	5	4.50
9	3	3	4	4	3.50
10	3	3	4	4	3.50
11	4	4	4	4	4.00
Average	3.36	3.55	3.64	3.90	(3.61)

As in the earlier studies, the lowest mean familiarity value was assigned a value of 1.00 and the other three judge familiarity means were adjusted accordingly. These values are found in Table 58. Thus, the similarity estimations of Judge D were counted twice that of the other judges when the "weighted" GCSA was computed. In other words, five decks of computer cards were analyzed by the MESA methodology, two of which represented the similarity estimations made by Judge D.

TABLE 58

MEAN JUDGE FAMILARITY RATING AND JUDGE "WEIGHT" VALUE FOR THE OTC-2 GROUP

	Judges			
Familiarity Values	A	B	C	D
Raw mean[a]	3.36	3.55	3.64	3.90
Weighted mean[b]	1.00	1.00	1.00	2.00

[a]Where five indicated "know very well" and one indicated "do not know very well."

[b]Where the lowest "raw mean" value in each judge group was set at 1.00 and the corresponding values for other judges were lowered proportionately and rounded to the nearest whole number.

The "weighted" GCSA was computed in the manner described several times before and the mean similarity correlation coefficients are found in Table 59. The factor analysis (principal components) and rotation (varimax) of the similarity correlation coefficient matrix resulted in the "weighted" judgment-dimensions found in Table 60.

The "weighted" judgment-dimensions were found to be very similar to the unweighted judgment-dimensions and thus a reinterpretation

TABLE 59

ARITHMETIC MEAN (ACROSS JUDGES) SIMILARITY ESTIMATIONS, SIMILARITY CORRELATION COEFFICIENTS, MEAN SIMILARITY, AND STANDARD DEVIATIONS OF THE SIMILARITY ESTIMATIONS FOR THE WEIGHTED OTC-2 GROUP

	Stimulus-Persons										
Stimulus-Persons	1	2	3	4	5	6	7	8	9	10	11
1	1.00	.39	.59	.60	.45	.68	.45	.50	.49	.49	.51
2	-.50	1.00	.52	.47	.76	.44	.64	.41	.20	.70	.62
3	.12	.15	1.00	.48	.57	.61	.66	.52	.40	.65	.61
4	.27	-.44	-.37	1.00	.44	.66	.44	.77	.55	.56	.49
5	-.40	.86	.26	-.53	1.00	.47	.71	.41	.19	.67	.68
6	.51	-.47	.09	.48	-.42	1.00	.59	.73	.50	.54	.57
7	-.32	.66	.49	-.51	.77	-.13	1.00	.44	.21	.69	.70
8	.10	-.61	-.32	.77	-.67	.59	-.57	1.00	.66	.51	.49
9	.23	-.86	-.40	.47	-.92	.25	-.88	.68	1.00	.34	.25
10	-.35	.71	.40	-.32	.67	-.30	.68	-.48	-.72	1.00	.67
11	-.22	.60	.38	-.40	.71	-.13	.75	-.48	-.82	.62	1.00
Weighted Mean Similarity	.56	.56	.60	.59	.58	.62	.59	.58	.44	.62	.60
Standard Deviation	.17	.22	.15	.17	.22	.16	.21	.18	.25	.17	.18

Note:

Similarity estimates are above the major diagonal (unity values) and correlational similarities are below the major diagonal.

TABLE 60

THE WEIGHTED GCSA FACTOR MATRIX (ROTATED) FOR THE OTC-2 GROUP

	Weighted Judgment-Dimensions			
Stimulus-Persons	I	II	III	h^2
1	-.32	.77	.12	.71
2	.78	-.38	-.32	.86
3	.43	.65	-.26	.68
4	-.29	-.06	.86	.82
5	.83	-.21	-.38	.87
6	-.06	.58	.72	.86
7	.88	.11	-.24	.84
8	-.43	-.05	.84	.89
9	-.93	-.01	.28	.94
10	.81	-.11	-.16	.69
11	.85	.08	-.14	.75
Amount of variance accounted for	44%	14%	23%	81%

of the "weighted" judgment-dimensions was not necessary. This similarity was demonstrated by intercorrelating the two sets of dimensions and the results are found in Table 61. "Weighted" Judgment-dimension I was extracted with reversed algebraic signs (compared to unweighted Judgment-dimension I) but otherwise correlated with its counterpart, as did Judgment-dimensions II and III.

TABLE 61

CORRELATION COEFFICIENTS (r) BETWEEN UNWEIGHTED AND WEIGHTED GCSA FACTOR MATRICES FOR THE OTC-2 GROUP

	Weighted Judgment-Dimensions		
Unweighted Judgment-Dimensions	1	2	3
1	-.991	-.294	-.734
2	.388	.978	.302
2	.717	.213	.986

These relationships indicated that only "weighted" Judgment-dimensions II and III had to be reversed to create the three predictor-dimensions found in Table 62. Again, the predictor-dimensions were created to clarify the meaning of the factor loadings and high positive loadings were determined to be predictive of vocational success.

The nature of the computations required to produce "derived" and "rank" predictor values have been described several times earlier and will not be repeated here. The unweighted and "weighted" judgment-dimensions served as the basis for producing predictor values where variance was and was not accounted for. These predictor values (both derived and rank) are found in Table 63. When the follow-up information was gathered, it was discovered that the status of Stimulus-persons 1 and 5 could not be determined so these persons were excluded from the study. The four predictor values (Table 63) were later correlated with rank success values collected in the follow-up procedures.

TABLE 62

THE WEIGHTED PREDICTOR-DIMENSIONS FOR THE OTC-2 GROUP

	Weighted Predictor-Dimensions		
Stimulus-Persons	I	II	III
1	-.32	-.77	-.12
2	.78	.38	.32
3	.43	-.65	.26
4	-.29	.06	-.86
5	.83	.21	.38
6	-.06	-.58	-.72
7	.88	-.11	.24
8	-.43	.05	-.84
9	-.93	.01	-.28
10	.81	.11	.16
11	.85	-.08	.14
Amount of variance accounted for	44%	14%	23%

TABLE 63

PREDICTOR VALUES (DERIVED AND RANK) FROM GCSA MODIFICATIONS FOR THE OTC-2 GROUP

Stimulus-Persons	Predictor Values							
	Unweighted GCSA (variance not accounted for)		Unweighted GCSA (variance accounted for)		Weighted GCSA (variance not accounted for)		Weighted GCSA (variance accounted for)	
	Derived	Rank	Derived	Rank	Derived	Rank	Derived	Rank
1[a]								
2	.46	1	.59	1.5	.49	1	.58	2
3	-.10	6	.06	6	.01	3	.19	6
4	-.35	7	-.30	7	-.36	7	-.39	9
5[b]								
6	-.46	11	-.35	8	-.45	11	-.34	7.5
7	.38	3	.56	3	.34	5	.53	3
8	-.42	9.5	-.40	9.5	-.41	10	-.46	10
9	-.38	8	-.57	11	-.40	8.5	-.58	11
10	.29	5	.48	5	.36	4	.50	4
11	.30	4	.49	4	.30	6	.49	5

[a] Returned home, status unknown.

[b] Part time worker, but status unknown.

The OTC-2 judges were also required to estimate the probable success of all the stimulus-persons using estimates from zero to 100. Three sets of predictor values were derived from these estimates and this procedure was described earlier. The average probability estimate that was assigned each stimulus-person is found in Table 64.

TABLE 64

STIMULUS-PERSON SUCCESS PROBABILITY RATINGS FOR THE OTC-2 JUDGE GROUP

	Judges				
Stimulus-Persons	A	B	C	D	Average
1	.50	.50	.10	.40	.38
2	.85	.95	.80	.90	.88
3	.80	.90	.40	.70	.70
4	.85	.50	.70	.30	.59
5	.90	.85	.85	.60	.80
6	.75	.60	.50	.30	.54
7	.90	.90	.90	.60	.83
8	.80	.50	.50	.10	.48
9	.40	.10	.30	.10	.23
10	.80	.80	.70	.50	.70
11	.70	.75	.75	.70	.73
Average	.75	.67	.59	.47	(.62)
Standard Deviation	.15	.24	.24	.25	

These probability estimates were standardized (with a mean of .50 and standard deviation .10) and "weighted" by values derived from Table 58. The results (Table 65) were also used as predictor values to be correlated with success value.

TABLE 65

STIMULUS-PERSON SUCCESS PROBABILITY RATINGS (STANDARDIZED) FOR THE OTC-2 JUDGE GROUP

	Judges					
Stimulus-Persons	A	B	C	D	Average	Weighted Average
1	.34	.43	.30	.47	.385	.402
2	.57	.62	.59	.67	.613	.624
3	.53	.59	.42	.59	.533	.544
4	.57	.43	.55	.43	.495	.482
5	.60	.57	.61	.55	.583	.576
6	.50	.47	.46	.43	.465	.458
7	.60	.59	.63	.55	.593	.584
8	.53	.43	.46	.35	.443	.424
9	.27	.27	.38	.35	.317	.324
10	.53	.55	.55	.51	.535	.530
11	.47	.53	.57	.59	.540	.550
Average	.50	.50	.50	.50	(.500)	(.500)
Standard Deviation	.10	.10	.10	.10		

Follow-up of Actual Success

After a total of 59 weeks had passed since the original similarity estimations were made, data regarding the success of the stimulus-persons was collected. Again, there was no special effort made by the agency to collect follow-up information specifically for this study. As a result, the follow-up information was a product of what was available and that which could be represented in quantitative terms.

The information that reflected vocational success is found in Table 66 and is not mean to reflect other, perhaps more important, measures of vocational success that were not available. The information in Table 66 was ranked and the results are found in Table 67. These ranked success values were correlated (Spearman Rank Correlation Coefficients) with the seven predictor values and the results are found in Table 68. As shown, the best predictor of vocational success was the "weighted" GCSA where the variance (or "importance") of the judgment-dimensions had been accounted for. Such a result was hypothesized earlier.

Discussion of Prediction

It is of interest to note that the success of this OTC-2 study may have been due to the absence of confounding variables that were found in the SS-2 study, which was not successful in terms of the informal hypotheses. In the SS-2 study, the judge that received the greatest "weight" value was not involved with direct vocational placement contact and had used a different underlying judgment criterion (see Tables 4 and 35). In the OTC-2 study, by contrast, the judge

TABLE 66

EMPLOYMENT SUCCESS FIGURES FOR THE OTC-2 GROUP

	Follow-up Information			
Stimulus-Persons	Job Placement Location	Total Weeks Placed (59 possible)[a]	Present Hourly Wage[b]	Estimated Income Earned[c]
1[d]				
2	Lumber yard	46	$1.60	$2944.00
3	Car washer	23	1.25	1150.00
4	Not placed	0	0.00	0.00
5[e]				
6	Messenger	41	1.60	2624.00
7	Laborer	52	1.60	3328.00
8	Shop clerk	20	1.00	800.00
9	Not placed	0	0.00	0.00
10	Laborer	59	1.60	3776.00
11	Laundry	59	1.60	3776.00

[a]From May 15, 1972 to July 1, 1973.

[b]Based on 40 hour work week.

[c]Derived from estimated earnings per week times total weeks placed.

[d]Returned home, status unknown.

[e]Part time worker but status unknown.

TABLE 67

RANK EMPLOYMENT SUCCESS FIGURES FOR THE OTC-2 GROUP

	Follow-up Information		
Stimulus-Persons	Rank Total Weeks Placed[a]	Rank Present Hourly Wage[b]	Rank Estimated Income Earned[c]
1[d]			
2	4	3	4
3	6	6	6
4	8.5	8.5	8.5
5[e]			
6	5	3	5
7	3	3	3
8	7	7	7
9	8.5	8.5	8.5
10	1.5	3	1.5
11	1.5	3	1.5

[a]From May 15, 1972 to July 1, 1973.

[b]Based on 40 hour week.

[c]Derived from estimated earning per week times total weeks employed.

[d]Returned home, status unknown.

[e]Part time worker but status unknown.

TABLE 68

SPEARMAN RANK CORRELATION COEFFICIENTS (r_s)[a] BETWEEN RANK SUCCESS AND PREDICTOR VALUES FOR THE OTC-2 GROUP[b]

	Weeks Placed	Rank Present Hourly Wage	Rank Estimated Income earned
Unweighted GCSA (variance not accounted for)	.622	.523	.622
Unweighted GCSA (variance accounted for)	.740[c]	.725[c]	.740[c]
Weighted GCSA (variance not accounted for)	.647	.523	.647
Weighted GCSA (variance accounted for)	.798[d]	.826[d]	.798[d]
Probability Rating	.705[c]	.700[c]	.705[c]
Standardized Probability Rating	.740[c]	.725[c]	.740[c]
Standardized Probability Rating (weighted)	.664	.670[c]	.664

[a]All Spearman Correlation Coefficients were corrected for tied ranks.

[b]$N = 9$.

[c]$p < .05$ (two-tailed test).

[d]$p < .01$ (two-tailed test).

with the most "weight" was in greater vocational training contact (Judge D was the training shop supervisor) than the other judges. Also, all four judges appeared to use similar underlying judgment criterion (see Tables 6 and 53), where OFJA produced only one judge-dimension.

It appears as though "variance accounted for" allowed for improved prediction as shown by the high predictive value of unweighted GCSA predictor values. This would indicate that accounting for the "importance" of the judgment-dimensions will increase predictive potency as opposed to using the "weighting" procedure alone.

Evidence illustrating the limitation of the "weighting" procedure is revealed by considering the standardized probability rating (weighted) correlations with success values. These correlations are lower than those found with absence of "weighting" and standardization.

Generalization Studies

The SWS Study

The four judges used in the SWS group were asked to evaluate (using similarity estimations) eight stimulus-persons who were being considered for vocational placements within the next year. The reliability estimate of the judges' similarity estimations (determined by correlating S_{ij} and S_{ji} stimulus-person pairs as described earlier) was very high (r = .96, df = 28, p <.001).

The average (arithmetic mean) similarity estimation assigned to each stimulus-person pair by each judge is found in Table 69,

along with the standard deviations of the similarity estimates. It is evident that the four judges had varying means and used various ranges (variability) in making their similarity estimations. Judge C used a greater range of similarity estimation within the bounds of zero to 100.

TABLE 69

MEANS AND STANDARD DEVIATIONS OF THE SIMILARITY ESTIMATIONS MADE BY EACH JUDGE IN THE SWS GROUP

Judge	Mean Similarity	Standard Deviation
A	.61	.19
B	.57	.15
C	.60	.24
D	.43	.19

The Observer Factor Judgment Analysis (OFJA) began by intercorrelating the similarity estimations made by each judge. The results are found in Table 70 and when this matrix was factor analyzed (principal components), a single judge-dimension was produced (Table 71). This single judge-dimension indicated that all four judges used a similar underlying judgment criterion.

Prediction of Success

The Group Composition Structure Analysis (GCSA) began by averaging (arithmetic mean) similarity estimations across judges. The column vectors of the resultant matrix were intercorrelated to produce the similarity correlation coefficient matrix. Both matrices are found in Table 72. Factor analysis (principal components) of

TABLE 70

INTER-JUDGE SIMILARITY CORRELATION COEFFICIENTS FOR THE SWS GROUP

	Judges			
Judges	A	B	C	D
A	1.00	.48	.11	.26
B		1.00	.27	.66
C			1.00	.21
D				1.00

TABLE 71

UNROTATED OFJA FACTOR MATRIX FOR THE SWS GROUP

	Judge-Dimension	
Judges	I	h^2
A	.65	.42
B	.90	.81
C	.46	.21
D	.80	.64
Amount of variance accounted for	52%	

TABLE 72

ARITHMETIC MEAN (ACROSS JUDGES) SIMILARITY ESTIMATIONS, SIMILARITY CORRELATION COEFFICIENTS, MEAN SIMILARITY, AND STANDARD DEVIATIONS OF THE SIMILARITY ESTIMATIONS FOR THE SWS GROUP

	Stimulus-Persons							
Stimulus-Persons	1	2	3	4	5	6	7	8
1	1.00	.22	.41	.78	.75	.76	.73	.30
2	-.93	1.00	.72	.24	.23	.19	.21	.69
3	-.66	.78	1.00	.32	.44	.32	.28	.49
4	.90	-.94	-.79	1.00	.81	.84	.78	.32
5	.88	-.92	-.62	.92	1.00	.75	.75	.29
6	.89	-.95	-.77	.95	.89	1.00	.76	.25
7	.86	-.93	-.80	.92	.88	.91	1.00	.25
8	-.83	.79	.43	-.82	-.86	-.85	-.85	1.00
Mean Similarity	.62	.44	.50	.63	.63	.61	.59	.45
Standard Deviation	.27	.32	.25	.29	.27	.31	.30	.27

Note:
Similarity estimates are above the major diagonal (unity values) and correlational similarities are below the major diagonal.

the similarity correlation matrix and rotation (varimax) produced the single judgment-dimension found in Table 73. This dimension is bipolar and accounted for a large portion (86 percent) of the total judgmental variance.

Judgment-dimension I was interpreted by the judges as a Probable Success dimension. Those stimulus-persons with high positive factor

TABLE 73

UNROTATED GCSA FACTOR MATRIX FOR THE SWS GROUP

	Unweighted Judgment-Dimension	
Stimulus-Persons	I	h^2
1	.94	.88
2	-.98	.96
3	-.78	.61
4	.97	.95
5	.94	.89
6	.97	.94
7	.96	.93
8	-.87	.75
Amount of variance accounted for	86%	

loadings were expected to be successful when placed in a vocational position. These persons shared higher motivation, higher intellect, and more adequate job skills to provide a cohesive group of trainees with a high probability of success. Those with high negative loadings did not share these qualities and, in fact, had qualities that would make future vocational success improbable.

The predictor-dimension (Table 74) was identical to the judgment-dimension. This occurred because the original extraction of the judgment-dimension represented persons with high success probability in terms of high positive factor loadings.

TABLE 74

THE UNWEIGHTED PREDICTOR-DIMENSION FOR THE SWS GROUP

	Unweighted Predictor-Dimension
Stimulus-Persons	I
1	.94
2	-.98
3	-.78
4	.97
5	.94
6	.97
7	.96
8	-.87
Amount of variance accounted for	86%

The "weights" for each judge were determined by averaging the familiarity ratings for each judge that they originally assigned to each stimulus-person. The results of this procedure are found in Table 75. By setting the lowest mean to 1.00 and proportionately lowering the other means, the "weighted" values were created by rounding off the proportionate value to the nearest whole number. Thus, the similarity estimations for Judge A were counted three times, Judge B and C's twice, and Judge D's only once (see Table 76). In terms of duplicating judgmental computer card decks, there were a total of eight decks for the "weighted" analysis where the original GCSA had used only four, one for each judge.

TABLE 75

JUDGE FAMILIARITY RATINGS FOR THE SWS GROUP BASED ON A FIVE-POINT SCALE

	Judges				
Stimulus-Persons	A	B	C	D	Average
1	5	4	4	4	4.25
2	4	3	4	3	3.50
3	5	2	4	2	3.25
4	4	4	4	2	3.50
5	5	5	4	3	4.25
6	5	5	4	3	4.25
7	4	5	4	3	4.00
8	5	4	4	4	4.25
Average	4.63	4.00	4.00	3.00	(3.91)

TABLE 76

MEAN JUDGE FAMILIARITY RATINGS AND JUDGE "WEIGHT" VALUE FOR THE SWS GROUP

	Judges			
Familiarity Values	A	B	C	D
Raw mean[a]	4.63	4.00	4.00	3.00
Weighted value[b]	3.00	2.00	2.00	1.00

[a]Where five indicated "know very well" and one indicated "do not know very well."

[b]Where the lowest "raw mean" value in each judge group was set at 1.00 and the corresponding values for other judges were lowered proportionately and rounded to the nearest whole number.

The "weighted" GCSA matrices are found in Table 77 and the factor analysis (principal component) of the similarity correlation coefficient matrix produced the single "weighted" judgment-dimension found in Table 78. This dimension was very similar to the unweighted judgment-dimension and the correlation between the unweighted and "weighted" judgment-dimensions was very high (see Table 79).

TABLE 77

ARITHMETIC MEAN (ACROSS JUDGES) SIMILARITY ESTIMATIONS, SIMILARITY CORRELATION COEFFICIENTS, MEAN SIMILARITY, AND STANDARD DEVIATIONS OF THE SIMILARITY ESTIMATIONS FOR THE WEIGHTED SWS GROUP

	Stimulus-Persons							
Stimulus-Persons	1	2	3	4	5	6	7	8
1	1.00	.19	.41	.78	.77	.78	.75	.28
2	-.95	1.00	.74	.20	.20	.16	.20	.71
3	-.66	.79	1.00	.33	.46	.33	.30	.48
4	.91	-.96	-.77	1.00	.82	.85	.80	.29
5	.90	-.93	-.60	.93	1.00	.78	.78	.27
6	.92	-.97	-.75	.97	.91	1.00	.78	.24
7	.90	-.94	-.77	.94	.91	.93	1.00	.22
8	-.86	.81	.42	-.84	-.88	-.87	-.87	1.00
Weighted Mean Similarity	.62	.42	.51	.64	.64	.61	.60	.44
Standard Deviation	.29	.34	.24	.31	.29	.32	.31	.28

Note:
Similarity estimates are above the major diagonal (unity values) and correlational similarities are below the major diagonal.

TABLE 78

THE WEIGHTED GCSA FACTOR MATRIX FOR THE SWS GROUP

	Weighted Judgment-Dimension	
Stimulus-Persons	I	h^2
1	.95	.91
2	-.99	.97
3	-.76	.58
4	.98	.96
5	.95	.91
6	.98	.96
7	.97	.94
8	-.89	.78
Amount of variance accounted for	88%	

TABLE 79

CORRELATION COEFFICIENT (r) BETWEEN UNWEIGHTED AND WEIGHTED GCSA FACTOR MATRICES FOR THE SWS GROUP

	Weighted Judgment-Dimension
Unweighted Judgment-Dimension	1.00

As in the SWS unweighted condition, the "weighted" predictor-dimension (Table 80) was identical to the single "weighted" judgment-dimension since the algebraic sign was originally such that positive loadings were indicative of success.

TABLE 80

THE WEIGHTED PREDICTOR-DIMENSION FOR THE SWS GROUP

	Weighted Predictor-Dimension
Stimulus-Persons	I
1	.95
2	-.99
3	-.76
4	.98
5	.95
6	.98
7	.97
8	-.89
Amount of variance accounted for	88%

The SWS study was unique in that the GCSA produced only a single judgment-dimension and consequently did not require rotation. In addition, only one judgment-dimension was producing allowing no means of "accounting for variance" among several dimensions as in the other studies.

The results of this analysis produced only two predictor values and the derived values and their ranks are found in Table 81. As shown, Stimulus-person-7 was excluded since, upon collecting follow-up information, it was discovered that this person was continuing training at another agency and comparable success data could not be obtained.

TABLE 81

PREDICTOR VALUES (DERIVED AND RANK) FROM GCSA MODIFICATIONS FOR THE SWS GROUP

Stimulus-Persons	Unweighted GCSA		Weighted GCSA	
	Derived	Rank	Derived	Rank
1	.94	3.5	.95	3.5
2	-.98	7	-.99	7
3	-.78	5	-.76	5
4	.97	1.5	.98	1.5
5	.94	3.5	.95	3.5
6	.97	1.5	.98	1.5
7[a]				
8	-.87	6	-.89	6

[a]Placed at Vocational Technical School for continued training and thus was excluded from the study.

The SWS judges were also asked to rate the probable success (from zero to 100) they thought each stimulus-person would have. These estimations and the judges' averages are found in Table 82. The values were standardized (mean = .50; SD = .10), and averaged where the "weighted" average was dependent on the judge "weight" discussed earlier. These values are found in Table 83. Thus three additional predictor values were used to correlate with success values and were to be ultimately compared with the unweighted and "weighted" GCSA predictor values.

TABLE 82

STIMULUS-PERSON SUCCESS PROBABILITY RATINGS FOR THE SWS JUDGE GROUP

	Judges				
Stimulus-Persons	A	B	C	D	Average
1	.80	.80	.60	.78	.75
2	.10	.10	.10	.40	.18
3	.50	.30	.10	.40	.33
4	.85	.80	.70	.80	.79
5	.79	.50	.70	.80	.70
6	.85	.90	.90	.90	.89
7	.85	.70	.80	.90	.81
8	.15	.30	.10	.43	.25
Average	.61	.55	.50	.68	(.58)
Standard Deviation	.30	.27	.32	.21	

TABLE 83

STIMULUS-PERSON SUCCESS PROBABILITY RATINGS (STANDARDIZED) FOR THE SWS JUDGE GROUP

	Judges					
Stimulus-Persons	A	B	C	D	Average	Weighted Average
1	.56	.59	.53	.55	.558	.559
2	.33	.34	.38	.37	.355	.350
3	.46	.41	.38	.37	.405	.416
4	.58	.59	.56	.56	.573	.575
5	.56	.48	.56	.56	.540	.540
6	.58	.63	.62	.61	.610	.606
7	.58	.55	.59	.61	.583	.579
8	.35	.41	.38	.38	.380	.376
Average	.50	.50	.50	.50	(.500)	(.500)
Standard Deviation	.10	.10	.10	.10		

Follow-up of Actual Success

A total of 58 weeks elapsed after the similarity estimations were made and before follow-up information was collected. Again, there was no special attempt on the part of the agency to collect information specifically for this study. The follow-up information collected is found in Table 84 and, when ranked (Table 85), was correlated with the five predictor values of interest.

TABLE 84

EMPLOYMENT SUCCESS FIGURES FOR THE SWS GROUP

	Follow-up Information	
Stimulus-Persons	Weeks Placed (58 possible)[a]	Present Hourly Wage[b]
1	28	$2.30
2[c]	0	.35[d]
3	0	.50[d]
4	39	2.45
5	56	2.30
6	56	2.45
7[e]		
8[c]	0	.45[d]

[a]From May 22, 1972 to July 1, 1973.

[b]As of July 1, 1973.

[c]Not placed as of July 1, 1973.

[d]Commensurate salary (piece rate) paid by the training agency.

[e]Placed at Vocational Technical School for continued training and thus excluded from the follow-up study.

TABLE 85

RANK EMPLOYMENT SUCCESS FOR THE SWS GROUP

	Rank Employment Success	
Stimulus-Person	Rank Weeks Placed[a]	Rank Present Hourly Wage[b]
1	4	3.5
2[c]	6	7[d]
3[c]	6	5[d]
4	3	1.5
5	1.5	3.5
6	1.5	1.5
7		
8[c]	6	6[d]

[a]From May 22, 1972 to July 1, 1973.

[b]As of July 1, 1973.

[c]Not placed as of July 1, 1973.

[d]Commensurate salary (piece rate) paid by training agency.

[e]Placed at Vocational Technical School for continued training and thus excluded from the follow-up study.

The Spearman-Rank-Order Correlations were calculated between the five predictor values and the two success values. The resulting correlations are found in Table 86. As shown, all of the correlations were high and statistically significant. The "weighting" procedure for the GCSA did not have any differential affect on prediction as was originally hypothesized.

TABLE 86

SPEARMAN RANK CORRELATION COEFFICIENTS (r_S)[a] BETWEEN RANK SUCCESS AND PREDICTOR VALUES FOR THE SWS GROUP

	Rank Weeks Placed	Rank Present Hourly Wage
Unweighted GCSA	.897[c]	1.000[c]
Weighted GCSA	.897[c]	1.000[c]
Probability Rating	.805[d]	.982[c]
Standardized Probability Rating	.805[d]	.982[c]
Standardized Probability Rating (weighted)	.805[d]	.982[c]

[a] All Spearman rank correlation coefficients were corrected for tied ranks.

[b] N = 7.

[c] $p < .01$ (two-tailed test).

[d] $p < .05$ (two-tailed test).

Discussion of Prediction

The GCSA modifications for prediction of vocational success correlated highly with actual success as hypothesized, except for the differential "weighting" procedure. Again, it should be noted that all four judges were apparently using a similar underlying judgment criterion which had proven successful in the OTC-2 study.

The fact that the GCSA modifications were shown somewhat superior to probability ratings is a point that should not be exaggerated.

The "N" size for the study was quite small and variability of both predictor and success values was small. It appears as though there were two distinct groups of stimulus-persons. One group were obviously good vocational prospects and one group were obviously poor prospects. This lack of normality in the distribution may have lead to artifically high correlations. These relationships may have been altered had a "middle" probable success and actual success group been present in the study.

The PU Study

The nine judges of the PU group were asked to evaluate (using similarity estimations) 14 stimulus-persons who were admitted to the psychiatric unit and qualified under the selection criteria described earlier. The reliability estimate of the judges' similarity estimations (determined by correlating S_{ij} and S_{ji} stimulus-person pairs) was moderately high ($r = .79$, $df = 89$, $p < .001$).

The average (arithmetic mean) similarity estimation assigned to the stimulus-persons by each judge (and the variability of the estimates) are found in Table 87. These values indicate that a wide range of similarity estimations were used by the nine judges.

The Observer Factor Judgment Analysis (OFJA) began by intercorrelating the similarity estimations made by each judge and these correlations are found in Table 88. The intercorrelations are quite low and initially indicated that the judges were using differential underlying judgmental criterion. This impression was confirmed when the intercorrelation matrix was factor analyzed (principal components) and extracted four judge-dimensions (Table 89). These four judge-

TABLE 87

MEANS AND STANDARD DEVIATIONS OF THE SIMILARITY ESTIMATIONS MADE BY EACH JUDGE IN THE PU GROUP

Judge	Mean Similarity	Standard Deviation
A	.49	.20
B	.63	.15
C	.58	.15
D	.44	.24
E	.60	.15
F	.44	.25
G	.49	.17
H	.55	.28
I	.52	.20

TABLE 88

INTER-JUDGE SIMILARITY CORRELATION COEFFICIENTS FOR THE PU GROUP

	Judges								
Judges	A	B	C	D	E	F	G	H	I
A	1.00	.04	.04	-.02	-.08	.06	.37	.44	-.08
B		1.00	.35	.32	.11	-.14	.22	.11	.10
C			1.00	.18	.16	-.19	.23	.07	.04
D				1.00	.11	.10	.21	-.01	.05
E					1.00	.09	.03	-.11	-.05
F						1.00	.08	-.02	-.01
G							1.00	.37	.01
H								1.00	-.02
I									1.00

TABLE 89

UNROTATED OFJA FACTOR MATRIX FOR THE PU GROUP

	Judge-Dimensions				
Judges	I	II	III	IV	h^2
A	.55	-.59	.03	-.08	.65
B	.58	.49	-.14	.06	.60
C	.53	.46	-.23	-.23	.60
D	.42	.45	.37	.22	.56
E	.10	.42	.46	-.43	.59
F	-.04	-.20	.83	.25	.80
G	.75	-.19	.16	.05	.62
H	.59	-.52	-.13	.02	.64
I	.05	.21	-.17	.83	.76
Amount of variance accounted for	22%	17%	13%	12%	64%

dimensions accounted for 64 percent of the total judge variation and the dimensions were rather difficult to interpret by some of the judges. However, some degree of meaning was given each of the four judge-dimensions in terms of the orientation used by the judges when dealing with a psychiatric population.

Judge-dimension I was termed a Diagnostic dimension where high positive loadings indicated that these judges were diagnostically oriented (medical model) in their approach to psychiatric patients. Judge-dimension II was termed a Functional dimension since judges with high negative loadings were viewed as assessing psychiatric patients

in terms of how they may function in the future. Judge-dimension III was termed a Psychological Adjustment dimension since the judge with the highest factor loading was considered psychologically oriented and concerned about environmental influences on patients. Judge-dimension IV was termed a Sociability dimension since the judge that had the high positive loading was viewed as structuring psychiatric problems in terms of social involvement versus social withdrawal. In view of the variety of psychiatric orientation represented, it may have been expected that such a diversity of views, when averaged, would lead to rather poor prediction as a result of erratic and inconsistent group consensus.

The Group Composition Structure Analysis (GCSA) for the PU group began by averaging (across judges) the similarity estimations assigned to each stimulus-person pair. When the column vectors of the mean similarity estimation matrix were intercorrelated, a similarity correlation coefficient matrix was created and both matrices are found in Table 90.

The dimensional analysis began by factor analyzing (principal components) and rotating (varimax) the similarity correlation matrix. Four judgment-dimensions (Table 91) were extracted accounting for 73 percent of the total judgmental variation. These unweighted judgment-dimensions were interpreted, with some difficulty, by several judges involved in the study.

Judgment-dimension I was interpreted as an Adjustment Potential dimension as those stimulus-persons with high positive loadings were characterized as having the "best adjustment" potential of all the stimulus-persons. Those with high negative loadings were characterized

TABLE 90

ARITHMETIC MEAN (ACROSS JUDGES) SIMILARITY ESTIMATIONS, SIMILARITY CORRELATION COEFFICIENTS, MEAN SIMILARITY, AND STANDARD DEVIATIONS OF THE SIMILARITY ESTIMATIONS FOR THE PU GROUP

	Stimulus-Persons													
Stimulus-Persons	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	1.00	.56	.41	.54	.69	.46	.60	.64	.51	.72	.53	.48	.47	.38
2	-.17	1.00	.56	.49	.42	.69	.58	.50	.55	.44	.55	.57	.43	.60
3	-.63	.26	1.00	.51	.43	.64	.46	.44	.54	.35	.45	.54	.54	.63
4	.15	-.31	-.17	1.00	.70	.50	.52	.61	.50	.57	.43	.53	.46	.54
5	.63	-.58	-.58	.55	1.00	.47	.60	.66	.56	.69	.51	.46	.55	.43
6	-.45	.62	.47	-.26	-.55	1.00	.61	.52	.58	.46	.53	.61	.42	.63
7	.30	.04	-.38	-.10	.22	.07	1.00	.60	.59	.59	.59	.53	.49	.44
8	.53	-.34	-.58	.28	.55	-.35	.27	1.00	.54	.69	.58	.54	.49	.43
9	-.19	-.01	.03	-.23	-.07	.06	.09	-.15	1.00	.52	.52	.49	.60	.56
10	.75	-.43	-.74	.23	.69	-.49	.31	.67	-.13	1.00	.56	.45	.45	.35
11	.10	.03	-.34	-.37	-.07	-.06	.26	.21	-.09	.18	1.00	.57	.45	.42
12	-.31	.25	.16	-.10	-.43	.37	-.09	-.15	-.23	-.34	.15	1.00	.42	.54
13	-.16	-.40	.04	-.19	.08	-.41	-.16	-.15	.30	-.14	-.24	-.41	1.00	.44
14	-.65	.38	.57	-.04	-.55	.50	-.41	-.57	.10	-.71	-.39	.18	-.20	1.00
Mean Similarity	.57	.57	.54	.56	.58	.58	.59	.59	.58	.56	.55	.55	.51	.53
Standard Deviation	.16	.14	.16	.14	.16	.15	.13	.14	.13	.17	.14	.14	.15	.16

Note:

Similarity estimates are above the major diagonal (unity values) and correlational similarities are below the major diagonal.

TABLE 91

ROTATED GCSA FACTOR MATRIX FOR THE PU GROUP

	Unweighted Judgment-Dimensions				
Stimulus-Persons	I	II	III	IV	h^2
1	.83	-.03	.09	.01	.70
2	-.29	.76	-.17	.00	.70
3	-.82	.11	.10	.00	.70
4	.25	-.11	.81	-.26	.80
5	.73	-.33	.43	.08	.84
6	-.43	.77	-.07	-.01	.78
7	.57	.41	-.25	.29	.63
8	.75	-.11	.06	-.17	.61
9	-.13	.04	-.11	.78	.64
10	.89	-.17	.08	-.02	.83
11	.32	.03	-.77	-.26	.77
12	-.29	.28	-.26	-.65	.64
13	-.21	-.72	-.11	.52	.85
14	-.76	.34	.29	-.01	.78
Amount of variance accounted for	34%	16%	12%	11%	73%

as having poor adjustment potential. This dimension accounted for the greatest amount of total judgmental variation (34 percent) and thus was considered the most "important" dimension in the formulation of predictor values. Also, high positive loadings were considered predictive of community adjustment success.

Judgment-dimension II accounted for 16 percent of the total judgmental variation and was interpreted as a Therapy Motivation dimension. Those stimulus-persons with high positive loadings were characterized as more motivated in the therapeutic efforts provided by the staff. The stimulus-person with the highest negative loading was characterized as poorly motivated for therapy. Thus, high positive loadings were considered predictive of future community adjustment.

Judgment-dimension III accounted for 12 percent of the judgmental variance and was interpreted as a Psychological Mindedness dimension. The stimulus-person with a high positive loading was characterized as a very insightful person whereas the stimulus-person with the highest negative loading was viewed as having little insight into his difficulties. Thus, higher positive loadings on this dimension were considered predictive of successful community adjustment.

The fourth Judgment-dimension (IV) accounted for 11 percent of the total judgmental variation and was revealed as a Community Tolerance dimension. The two stimulus-persons who obtained high positive loadings were viewed as behaving in hostile and extrapunitive ways that would alienate themselves from persons in the community. The stimulus-person with the high negative loading was characterized as being more tolerated by the community even though his psychiatric problems interfered with community adjustment generally. Thus, high negative factor loadings

were considered indicative of future community adjustment. It should be noted that all four judgment-dimensions were difficult to interpret and the interpretations cited here are all tenuous.

Since high positive loadings on Judgment-dimensions I, II, and III were predictive of community adjustment success, the algebraic signs were not reflected when the Predictor-dimensions were created. Judgment-dimension IV was changed to create Prediction-dimension IV by reflecting the algebraic signs since high negative loadings were indicative of community adjustment success. All four Predictor-dimensions are found in Table 92.

TABLE 92

THE UNWEIGHTED PREDICTOR-DIMENSIONS FOR THE PU GROUP

	Unweighted Predictor-Dimensions			
Stimulus-persons	I	II	III	IV
1	.83	-.03	.09	-.01
2	-.29	.76	-.17	.00
3	-.82	.11	.10	.00
4	.25	-.11	.81	.26
5	.73	-.33	.43	-.08
6	-.43	.77	-.07	.01
7	.57	.41	-.25	-.29
8	.75	-.11	.06	.17
9	-.13	.04	-.11	-.78
10	.89	-.17	.08	.02
11	.32	.03	-.77	.26
12	-.29	.28	-.26	.65
13	-.21	-.72	-.11	-.52
14	-.76	.34	.29	.01
Amount of variance accounted for	34%	16%	12%	11%

To determine the "weight" values for each judge, the judge familiarity ratings (Table 93) were employed and the method used was identical to those described earlier. These "weight" values are found in Table 94.

TABLE 93

JUDGE FAMILIARITY RATINGS FOR THE PU GROUP BASED ON A FIVE-POINT SCALE

	Judges									
Stimulus-Persons	A	B	C	D	E	F	G	H	I	Average
1	2	2	5	1	1	1	1	1	1	1.66
2	4	3	1	5	2	3	3	3	3	3.00
3	4	3	5	2	4	3	2	3	4	3.33
4	3	2	5	1	1	1	1	1	5	2.22
5	4	4	1	4	2	2	2	3	4	2.89
6	5	4	3	3	5	3	3	4	5	3.89
7	4	2	1	3	5	5	2	5	4	3.44
8	2	2	5	5	5	2	2	2	2	3.00
9	3	4	5	3	1	1	5	2	1	2.78
10	5	1	1	1	1	3	5	2	3	2.44
11	2	1	1	1	1	5	1	1	1	1.56
12	2	1	2	1	1	2	1	2	5	1.89
13	4	3	5	5	3	3	3	5	5	4.00
14	5	4	5	5	5	5	4	5	5	4.78
Average	3.5	2.6	3.2	2.9	2.6	2.8	2.5	2.8	3.4	(2.92)

TABLE 94

MEAN JUDGE FAMILIARITY RATINGS AND JUDGE "WEIGHT VALUE FOR THE PU GROUP

	Judges								
Familiarity Values	A	B	C	D	E	F	G	H	I
Raw mean[a]	3.5	2.6	3.2	2.9	2.6	2.8	2.5	2.8	3.4
Weighted mean[b]	2.0	1.0	2.0	1.0	1.0	1.0	1.0	1.0	2.0

[a]Where five indicated "know very well" and one indicated "do not know very well."

[b]Where the lowest "raw mean" value in each judge group was set at 1.00 and the corresponding values for the other judges were lowered proportionately and rounded to the nearest whole number.

When the three additional decks of computer cards for Judges A, C, and I were added for the "weighted" GCSA computations, the mean similarity estimation and similarity correlation coefficient matrices were produced and are found in Table 95. The similarity correlation matrix was factor analyzed (principal components) and rotated (varimax) to produce the four "weighted" judgment-dimensions found in Table 96. A total of 72 percent of the "weighted" judgmental variation was explained by the four judgment-dimensions.

These "weighted" judgment-dimensions were very similar to the unweighted judgment-dimensions with one reversal in the computer order of extraction as seen in Table 97. Unweighted Judgment-dimension IV was extracted third in the "weighted" Judgment-dimension III. Of course, unweighted Judgment-dimension III became "weighted" Judgment-

TABLE 95

ARITHMETIC MEAN (ACROSS JUDGES) SIMILARITY ESTIMATIONS, SIMILARITY CORRELATION COEFFICIENTS, MEAN SIMILARITY, AND STANDARD DEVIATIONS OF THE SIMILARITY ESTIMATIONS FOR THE WEIGHTED PU GROUP

	Stimulus-Persons													
Stimulus-Persons	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	1.00	.58	.42	.56	.69	.47	.60	.62	.54	.72	.51	.49	.47	.40
2	-.09	1.00	.58	.46	.44	.67	.57	.49	.55	.47	.52	.56	.43	.59
3	-.58	.28	1.00	.53	.45	.64	.47	.43	.55	.36	.45	.53	.51	.63
4	.17	-.40	-.11	1.00	.71	.46	.53	.59	.54	.56	.40	.52	.52	.56
5	.59	-.55	-.52	.58	1.00	.46	.60	.63	.59	.69	.51	.47	.60	.46
6	-.41	.56	.46	-.39	-.57	1.00	.60	.52	.57	.46	.52	.58	.42	.63
7	.32	.04	-.37	-.09	.20	.09	1.00	.60	.59	.61	.56	.52	.48	.45
8	.44	-.33	-.57	.20	.42	-.31	.28	1.00	.53	.68	.57	.54	.50	.44
9	-.09	-.04	.03	-.06	.08	.02	.06	-.15	1.00	.54	.51	.50	.61	.56
10	.74	-.37	-.73	.19	.64	-.47	.35	.61	-.06	1.00	.55	.47	.47	.38
11	.03	-.02	-.32	-.46	-.13	-.05	.18	.17	-.15	.16	1.00	.54	.43	.42
12	-.25	.17	.10	-.14	-.39	.25	-.12	-.08	-.26	-.28	.10	1.00	.41	.54
13	-.12	-.41	-.04	.09	.27	-.43	-.19	-.08	.32	-.08	-.30	-.45	1.00	.43
14	-.56	.32	.56	.01	-.47	.47	-.39	-.50	.06	-.64	-.38	.14	-.27	1.00
Weighted Mean Similarity	.58	.56	.54	.58	.59	.57	.58	.58	.58	.57	.53	.55	.52	.53
Standard Deviation	.15	.14	.15	.14	.15	.15	.13	.14	.12	.17	.14	.14	.15	.16

TABLE 96

THE WEIGHTED GCSA FACTOR MATRIX (ROTATED) FOR THE PU GROUP

	Weighted Judgment-Dimensions				
Stimulus-Persons	I	II	III	IV	h^2
1	.83	.01	.04	.16	.72
2	-.21	.79	-.04	-.12	.68
3	-.79	.23	.02	.16	.70
4	.18	-.34	-.11	.81	.82
5	.65	-.43	.21	.42	.83
6	-.38	.77	-.06	-.10	.74
7	.61	.40	.21	-.16	.60
8	.69	-.21	-.20	.00	.56
9	-.09	.05	.76	-.01	.59
10	.89	-.19	.01	.04	.82
11	.25	-.05	-.24	-.82	.79
12	-.23	.14	-.70	-.15	.59
13	-.17	-.65	.63	.01	.84
14	-.69	.36	-.10	.36	.75
Amount of variance accounted for	30%	17%	12%	13%	72%

TABLE 97

CORRELATION COEFFICIENTS (r) BETWEEN UNWEIGHTED AND WEIGHTED GCSA FACTOR MATRICES FOR THE PU GROUP

	Weighted Judgment-Dimensions			
Unweighted Judgment-Dimensions	1	2	3	4
1	.997	-.372	.048	-.016
2	-.467	.976	-.292	-.104
3	.031	-.387	.080	.977
4	-.007	-.241	.985	.043

dimension IV. As a consequence, the algebraic signs of Judgment-dimension III were reversed to create four Predictor-dimensions (Table 98) where positive loadings were hypothesized as being predictive of community adjustment success.

The Predictor-dimensions (both unweighted and "weighted") were then used to determine predictor values (Table 99) in a manner described several times previously where "variance accounted for" was considered. As shown in Table 99, the derived predictor values were ranked and later correlated with rank follow-up variables.

The nine judges were asked to rate probable community adjustment success using values from zero to 100 (where 100 indicated absolute probability of success). These ratings (and the average for each stimulus-person) are found in Table 100. To equalize apparent variation in the use of probability estimates among the judges, the probability estimates were standardized (with a mean of .50 and a standard deviation of .10)

TABLE 98

THE WEIGHTED PREDICTOR-DIMENSIONS FOR THE PU GROUP

	Weighted Predictor-Dimensions			
Stimulus-Persons	I	II	III	IV
1	.83	.01	-.04	.16
2	-.21	.79	.04	-.12
3	-.79	.23	-.02	.16
4	.18	-.34	.11	.81
5	.65	-.43	-.21	.42
6	-.38	.77	.06	-.10
7	.61	.40	-.21	-.16
8	.69	-.21	.20	.00
9	-.09	.05	-.76	-.01
10	.89	-.19	-.01	.04
11	.25	-.05	.24	-.82
12	-.23	.14	.70	-.15
13	-.17	-.65	-.63	.01
14	-.69	.36	.10	.36
Amount of variance accounted for	30%	17%	12%	13%

TABLE 99

PREDICTOR VALUES (DERIVED AND RANK) FROM GCSA MODIFICATIONS FOR THE PU GROUP

	Predictor Values							
	Unweighted GCSA (variance not accounted for)		Unweighted GCSA (variance accounted for)		Weighted GCSA (variance not accounted for)		Weighted GCSA (variance accounted for)	
Stimulus-Person	Derived	Rank	Derived	Rank	Derived	Rank	Derived	Rank
1	.22	2.5	.39	1.5	.26	1.5	.38	1
2	.07	8.5	.00	8	.10	7	.07	8
3	-.15	12	-.34	13	-.09	12	-.24	14
4	.30	1	.26	6	.13	6	.12	6
5	.19	5	.33	4	.21	3	.28	4
6	.07	8.5	-.04	10	.06	9	.00	9
7	.11	6	.27	5	.26	1.5	.35	2
8	.22	2.5	.36	3	.07	8	.20	5
9	-.24	13	-.19	11	.18	5	.10	7
10	.20	4	.39	1.5	.19	4	.33	3
11	-.04	11	.07	7	-.21	13	-.10	10
12	.09	7	-.02	9	-.23	14	-.21	13
13	-.39	14	-.35	14	-.04	11	-.12	11
14	-.03	10	-.23	12	-.02	10	-.15	12

TABLE 100

STIMULUS-PERSON SUCCESS PROBABILITY RATINGS FOR THE PU JUDGE GROUP

	Judges									
Stimulus-Persons	A	B	C	D	E	F	G	H	I	Average
1	.70	.75	.95	.75	.40	.90	.50	.90	.90	.75
2	.40	.40	.20	.60	.50	.40	.30	.30	.60	.41
3	.40	.60	.50	.10	.50	.60	.10	.20	.55	.39
4	.50	.70	.80	.70	.40	.80	.40	.25	.75	.59
5	.90	.85	.90	.95	.60	.80	.70	.99	.80	.83
6	.40	.60	.65	.65	.70	.50	.30	.15	.65	.51
7	.70	.65	.60	.65	.80	.80	.50	.90	.75	.71
8	.75	.75	.55	.60	.80	.80	.50	.95	.70	.71
9	.75	.65	.55	.20	.60	.50	.30	.85	.60	.56
10	.70	.85	.65	.65	.70	.80	.70	.95	.85	.76
11	.90	.65	.40	.70	.60	.50	.50	.60	.50	.59
12	.80	.60	.45	.65	.55	.50	.50	.40	.50	.55
13	.70	.50	.45	.10	.40	.50	.70	.80	.40	.51
14	.70	.45	.50	.25	.40	.50	.30	.15	.70	.44
Average	.66	.64	.58	.54	.57	.64	.45	.60	.66	(.59)
Standard Deviation	.17	.13	.19	.25	.14	.16	.17	.33	.14	

and "weighted" (using the "weights" derived by the familiarity ratings). These values are found in Table 101. Thus, three additional predictor values were created in an attempt to compare a more straightforward predictive approach with the rather complex approach that multidimensional scaling methods offer.

Follow-up of Actual Success

As was true with the five follow-up studies described previously, there was no special attempt to collect follow-up data for the PU group

TABLE 101

STIMULUS-PERSON SUCCESS PROBABILITY RATINGS (STANDARDIZED) FOR THE PU GROUP

Stimulus-Persons	Judges A	B	C	D	E	F	G	H	I	Average	Weighted Average
1	.52	.58	.69	.58	.38	.66	.53	.59	.67	.578	.590
2	.34	.31	.30	.52	.45	.36	.41	.41	.46	.396	.388
3	.34	.47	.46	.33	.45	.48	.30	.38	.42	.403	.404
4	.40	.54	.61	.56	.38	.60	.47	.39	.56	.501	.507
5	.64	.66	.66	.66	.52	.60	.65	.62	.60	.623	.626
6	.34	.47	.54	.54	.60	.42	.41	.36	.49	.463	.462
7	.52	.51	.51	.54	.67	.60	.53	.59	.56	.559	.552
8	.55	.58	.48	.52	.67	.60	.53	.61	.53	.563	.553
9	.55	.51	.48	.37	.52	.42	.41	.58	.46	.478	.483
10	.52	.66	.54	.54	.60	.60	.65	.61	.64	.596	.588
11	.64	.51	.41	.56	.52	.42	.53	.50	.38	.497	.492
12	.58	.47	.43	.54	.49	.42	.53	.44	.38	.476	.473
13	.52	.39	.43	.33	.38	.42	.65	.56	.31	.443	.438
14	.52	.35	.46	.39	.38	.42	.41	.36	.53	.424	.444
Average	.50	.50	.50	.50	.50	.50	.50	.50	.50	(.500)	(.500)
Standard Deviation	.10	.10	.10	.10	.10	.10	.10	.10	.10		

study. Since the agency did not systematically collect follow-up data of its own, relevant information regarding successful community adjustment was not available. As a result, three measures were used which, in several cases, did not appear to reflect community adjustment at all.

Two variables reflecting contacts with the agency included the total number of contacts with agency and total number of months in contact with the agency. A third rating was derived from two mental health workers at the agency who were not originally involved in the PU judge group. These two judges were asked to rate each stimulus-person on a scale from one to five (where one indicated "poor adjustment to community" and five indicated "excellent adjustment to community"). These ratings were standardized and averaged to best represent the opinion of the two judges. The three contact and success values are found in Table 102. These values were ranked (Table 103) where more contacts and more time with agency were given the high ranks and high judge success ratings were given high ranks.

It was not intended to promote a large "number of contacts" and "time with agency" as indicative of success since the prediction of success could not be made on the basis of these values. For example, a large number of contacts may indicate a person with serious problems or a person doing very well in long term therapy. On the other hand, a small number of contacts may indicate a person whose problems were not serious or a person who became discouraged and terminated contacts prematurely with a poor success prognosis.

The only measure that directly reflected community success adjustment was the ratings made by the two judges. All three rank

TABLE 102

TOTAL CONTACTS, TOTAL TIME WITH AGENCY AND SUCCESS VALUES FOR THE PU GROUP

	Total Number of Contacts With Agency	Total Time (in months) With Agency	Average Standardized Success Ratings
Stimulus-Persons			
1	9	2	3.77
2	8	8	.44
3	21	6	1.99
4	5	2	2.53
5	8	1	4.45
6	20	1	3.22
7	17	9	1.99
8	11	6	2.53
9	11	3	3.08
10	21	5	3.77
11	28	17	3.77
12	6	1	2.53
13	42	4	2.53
14	110	10	4.45

TABLE 103

RANK TOTAL CONTACTS, RANK TOTAL TIME WITH AGENCY AND RANK SUCCESS VALUES FOR THE PU GROUP

	Rank Total Number of Contacts With Agency	Rank Total Time (in months) With Agency	Rank Average Standardized Success Ratings
Stimulus-Persons			
1	10	10.5	4
2	11.5	4	14
3	4.5	5.5	12.5
4	14	10.5	9.5
5	11.5	13	1.5
6	6	13	6
7	7	3	12.5
8	8.5	5.5	9.5
9	8.5	9	7
10	4.5	7	4
11	3	1	4
12	13	13	9.5
13	2	8	9.5
14	1	2	1.5

rank values were correlated with the seven predictor values and the results are found in Table 104.

Discussion of Prediction

Only one significant correlation was found and it could be argued that this was a spurious correlation. The unweighted GCSA (variance not accounted for) predictor values correlated significantly with "total number of contacts" indicating that those with a fewer number of contacts were related to this set of predictor values.

The most important finding seems to be that predictor values correlated negatively with number of contacts and total time where the predictor values correlated positively with the two judge success ratings. Since only one correlation was significant, no conclusions can be drawn from these relationships.

The most obvious reason for predictive failure was probably due to the poor quality of success values that were derived from the follow-up study. In addition, the OFJA judge-dimensions (Table 89) indicated that there was little uniformity in the underlying judgmental criterion used by the nine judges. Since all similarity estimations were averaged, it appeared that little predictive success was obtained by such group consensus.

TABLE 104

SPEARMAN RANK CORRELATION COEFFICIENTS (r_S)[a] BETWEEN RANK FOLLOW-UP VARIABLES AND PREDICTOR VALUES FOR THE PU GROUP

	Rank Total Number of Contacts With Agency	Rank Total Time (in months) With Agency	Rank Average Standardized Success Ratings
Unweighted GCSA (variance not accounted for)	-.595[c]	-.286	.107
Unweighted GCSA (variance accounted for)	-.412	-.099	.244
Weighted GCSA (variance not accounted for)	-.325	-.152	.145
Weighted GCSA (variance accounted for)	-.335	-.122	.179
Probability Rating	-.280	-.238	.449
Standardized Probability Rating	-.285	-.281	.456
Standardized Probability Rating (weighted)	-.304	-.270	.492

[a] All Spearman Correlation Coefficients were corrected for tied ranks.

[b] $N = 14$.

[c] $p < .05$ (two-tailed test)

CHAPTER V

DISCUSSION

The Prediction of Vocational Success

Three of the five studies using mentally retarded persons demonstrated that the MESA modifications can, indeed, be rather sensitive predictors of vocational placement success. The sixth study conducted with psychiatric patients was not successful. There appeared to be specific reasons for failure in the three studies where prediction was inadequate as contrasted with the three studies showing good predictive potential.

The SS-1, OTC-2, and SWS studies demonstrated the efficacy of using MESA modifications as a predictor methodology where the OTC-1, SS-2, and PU studies demonstrated cases where the methodology failed or partially failed. The most important difference between these two sets of investigations (successful and unsuccessful prediction) was found in the Observer Factor Judgment Analysis (OFJA). In all three unsuccessful studies, "commonness" of underlying judgmental criterion was demonstrably absent as shown by the OFJA analysis.

The OFJA of the SS-2 group extracted two judge-dimensions, the OFJA of the OTC-1 group extracted two judge-dimensions, and the OFJA of the PU group extracted four judge-dimensions. More than one judge-dimension is indicative of more than one judgment orientation and, as demonstrated, seemed to lead to confusing predictor values. The three

successful studies (SS-1, OTC-2, and SWS) were characterized by having only one judge-dimension extracted by OFJA and indicated that judges were using a common underlying judgmental basis in each study.

The value of computing an OFJA in this predictive methodology is apparent. When the OFJA computations are completed and produce more than one judge-dimension, caution should be used in formulating predictor values. There are several schemes for determining which judge or judges should be used for further analysis to gain adequate prediction. The investigator may decide that the judge-dimension accounting for the most judge variation should be used, thus using only the similarity estimations of judges who loaded high on unrotated Judge-dimension I. Another strategy may be to use the estimations of the judges who were most familiar with the stimulus-persons, provided that the two judges most familiar with the stimulus-persons are found on the same Judge-dimension.

The decision to use either strategy, however, would be based on speculation rather than empirical evidence. A difficult decision, for example, would have to be made in the SS-2 study if both strategies were considered. Two of the judges (A and C) loaded high on the primary Judge-dimension I where the judge most familiar with the stimulus-persons loaded solely on Judge-dimension II (see Tables 35 and 39). The option of using the two judges or of using the single judge would be arbitrary.

A more reasonable method designed to settle this question of strategy is empirical. Since individual judge GCSA computations were provided by the computer (though these were not reported in this paper),

it would not be difficult to determine predictor values for each judge and correlate them with the success values. Then, those judges who were the best predictors (via similarity estimations) would be employed for future predictor studies. A judge who previously was considered a poor predictor could be used only when that judge loaded on the same Judge-dimension as judges previously proven as good predictors. Demonstration of such a strategy would require further analysis of the two SS and two OTC studies.

Of the three studies that demonstrated predictive success, two of them (OTC-2 and SWS) also used a predictive methodology that was radically different from the modified GCSA methodology. The use of the probability ratings allowed a comparative evaluation of the modified MESA predictor values. It was clearly demonstrated that when both probability ratings and modified MESA predictor values were somewhat predictive, the modified MESA technique was demonstrated as producing the more sensitive and accurate predictions. The study does not demonstrate, however, that scaling techniques other than the probability estimates may be more sensitive than the modified MESA technique.

The Predictive Methodology

The use of mean familiarity ratings as "weights" for judges deserves mention. Other information about the judges was available (i.e., years experience, duties, and loadings on judge-dimensions) and could have been incorporated into the weighting procedure. Further studies might incorporate this information about the judges (and perhaps gather additional information) to determine the most effective

way to "weight" the judges' similarity estimations when "weighted" GCSA and predictor values are computed.

When "variance accounted for" was considered, a very straightforward method was used which considered "importance" of the dimension as dependent upon the amount of variance it accounted for. There may be other means to determine this "importance" of judgment-dimensions and further research could illuminate the effect of these.

The difference between unweighted and "weighted" judgment-dimensions in each investigation was found to be rather minimal. If, however, a highly "weighted" judge were to have produced similarity estimations that were radically different from the other judges, there may be less similarity between unweighted and "weighted" judgment-dimensions. If this were to happen, the "weighted" judgment-dimensions would have to be interpreted by the judges. This would not present any problems, however, provided that the new judgment-dimensions could be assessed in terms of one pole indicating positive (or negative) attributes of the stimulus-persons.

In future studies, the possibility of using at least interval level measurement (rather than ordinal) for the correlations may be realized. Since the predictor values consistently ranged from -1.00 to +1.00 (with a variety of variabilities depending on whether or not they were "weighted" or variance was "accounted for"), success values forming a normal distribution would be, in a strict sense, required for Pearson product-moment correlations. The success values used in this study did not distribute in a normal distribution, which further justifies the use of rank measurement. If, however, success values are transformed or "forced" to follow a normal distribution, the

success data could be transformed to values ranging from -1.00 to +1.00. In this case, there would be more justification for the use of interval or ratio level measurement.

It is obvious that any prediction scheme employing MESA must be of the "within group" prediction type. This means that the group of stimulus-persons must be a well defined group about whom a group of judges assesses perceived similarity between the stimulus-persons regarding a specific judgmental question. Since each stimulus-person is compared with every other stimulus-person at one judgmental setting, there is no way to assess potential success of individuals not in that specific group. Such a restriction is not found with the probability rating approach. If this method were demonstrated as a successful way to determine future success, each stimulus-person could be evaluated (by judges who have shown predictive expertise) individually.

Another difference between the MESA prediction method and probability prediction method is the absolute value of a success rating or predictor value. When probability ratings were used, the judge was attempting to assess the individual's absolute probability of succeeding. With MESA, however, the similarity estimation is determined by comparing two stimulus-persons. For example, in a case where 10 stimulus-persons are to be evaluated and all 10 are known to have excellent prognosis for placement, it may be expected that the two methods would produce different results. One might expect that the judges would assign probability ratings in the 80's and 90's for such a potentially successful group. Using MESA with the same subjects, however might produce some predictor values that are high negative (due to bipolar judgment-dimensions). Thus, when the

stimulus-person with the highest probable success is compared with the stimulus-person with the lowest probable success (even though such a stimulus-person is viewed as being potentially successful), the assigned similarity estimation may, when averaged, be quite low (such as "20"). If this potentially successful group of stimulus-persons is viewed comparatively in such a manner, then a wide range of similarity estimations could be expected. If this happens, then bipolar judgment-dimensions can be expected which would result in predictor values ranging from +1.00 to -1.00.

In short, the meaning of negative predictor values is difficult to determine and they only have meaning when compared (such as using ranks) within the group. In this example, where probability ratings were used, even the stimulus-person with the lowest probable success expectation would receive an average probability rating that was high (perhaps in the 80's), since the whole group was hypothesized as being successful in this example.

Clinical Judgment

One interesting aspect of the MESA methodology is the proposed quality of the judgment-dimensions. When these dimensions are interpreted by the judges, they tend to represent rather unusual attributes that seemingly are seldom measured by other instruments. Many times, the judges were surprised that such unusual dimensions were extracted since they did not knowingly consider such attributes when they made their judgments.

It may be that MESA taps "clinical judgment" in such a manner that a multitude of variables impinge on any given similarity estimation

judgment. If this is the case, then "clinical judgment" has been demonstrated as a rather sensitive "instrument" when combined with the MESA methodology in assessing potential vocational success of mentally retarded adults. However, certain demonstrable conditions must be met and were described earlier.

In the past, clinical judgment has often been considered inadequate to the task of prediction, possibly resulting from clinical judgments being handled in rather inadequate ways. Of course, this MESA methodology was not compared to highly efficient multiple regression formulations which certainly could be explored in future studies.

Of the six investigations described, five dealt with mentally retarded persons and one with psychiatric patients. It would appear possible to use this MESA prediction methodology with a seemingly infinite number of different groups. Limitations exist only in regard to the number of stimulus-persons used (where a large number makes the judgmental task too large for each judge) and in finding a group of judges that are familiar with each stimulus-person with reference to some judgmental question.

The MESA prediction methodology seems ideal for situations that require specific outcome information from a group of judges. Arriving at a staff consensus regarding some judgmental question is easily handled by this method. In these situations, outspoken staff (though not necessarily more knowledgeable) may have more influence than "softspoken" staff when a staff consensus is desired. With the MESA prediction methodology, such consensus does not allow for "group dynamics" and, in fact, can account for judges who are more knowledgeable provided some valid criterion is determined to assess the

predictive ability of the judges. It seems clear that the MESA prediction methodology may have a valid use as a descriptor of staff consensus. It has been shown that such staff consensus is a rather sensitive measure of vocational placement outcome with mentally retarded adults.

CHAPTER VI

SUMMARY

Six investigations were conducted to demonstrate the feasibility of using a modification of the MESA methodology to predict vocational success for five groups of mentally retarded adults and successful community adjustment for one group of psychiatric patients. The MESA modifications included accounting for the differential predictive ability of the judges and accounting for the predictive importance of the MESA judgment-dimensions using the "amount of variance accounted for." Predictor values were determined in a rather complex manner based on original MESA judgment-dimensions and "weighted" judgment-dimensions. Also, a probability prediction method was used in four of the six studies to compare the effectiveness of the MESA prediction methodology with an unrelated methodology. Success values were gathered more than one year after the similarity estimations and probability judgments were made. Spearman Rank-Order Correlation Coefficients were computed between predictor and success values to determine the predictive ability of the MESA prediction methodology.

Three of the six investigations conducted demonstrated the excellent predictive effectiveness of the modified MESA methodology. There were specific demonstrable reasons that were discovered partially explaining the failure of prediction in the three unsuccessful studies.

Qualifications and ramifications of the MESA predictive methodology were described for reference in further studies. The use of this methodology for prediction was endorsed given the restrictive conditions discussed. Many questions remained regarding this methodology, most of which appeared suitable for further investigations.

APPENDIX A

INSTRUCTIONS TO SS-1 JUDGE GROUP

INSTRUCTIONS TO JUDGE

This study is concerned with the characteristics needed for retarded adults to be successful on vocational placement. As you probably know, professionals working in the area of mental retardation do not have reliable measures on which to base predictions of vocational "success-failure." Presently, they often use "clinical judgment," a process by which professionals use their experience and "intuition" to assess probable vocational success or failure. Since you are a professional in this area, you have made numerous "clinical judgments."

In the present study, your task as a judge is to estimate the general degree of similarity between each pair of retarded adults presented regarding their potential success on vocational placement. The scale you will use ranges from "0" to "100," where "0" means no similarity and "100" means identity (i.e., identical pair). You may use any number between "0" and "100" to estimate the general degree of similarity. For example, I might estimate the similarity of a Cadillac and a Rolls Royce as being "85," since both cars are big and expensive, but yet quite different. However, I may well estimate the similarity of the same Cadillac and a Volkswagen "bug" as being "20," since they are both cars, but yet quite different kinds of cars. You may have considered somewhat different similarity estimates than the ones I have used since there is no "correct" answer.

Accompanying these instructions is a deck of computer cards. All possible pair-combinations of some retarded adults that you are professionally acquainted with are printed on the cards. Please

write your similarity estimate for each pair near the lower left-hand corner of the computer card. Each pair will be evaluated twice and your judgment may differ somewhat even for an identical pair. This is to be expected and there is no need to look back and view your first judgment of that particular pair.

It is important for you to know that there is no "correct" similarity estimate. I am interested in your opinion as a professional judge. Please do not discuss any of your impressions with your colleagues before you have completed your judgments as this may influence your opinion. Also, it would be preferable if you made all your judgments in "one sitting," rather than doing a small number of judgments at different times.

Remember, you are to judge the generalized similarity between each pair based on your knowledge of their potential vocational placement success. You may use any number between "0" and "100," where "0" indicates no similarity and "100" indicates identity. Please feel free to erase and "improve" your judgment if you find this necessary.

I would like you to begin making those judgments now before you complete the information sheets on the following pages.

Thank you for giving your time and effort. It is greatly appreciated.

DAVID L. HANSON

JUDGE INFORMATION AND RATING SHEET

Please complete your similarity estimations before completing this information and rating sheet.

The information asked for below is needed in order to report a "Description of Judges" section for this study. When this study is written, you will be presented in the form of averages. Please be as accurate as possible. Will you please PRINT your responses.

NAME:_____ SEX:_____
AGENCY OF PRESENT EMPLOYMENT:_____
OFFICIAL OCCUPATIONAL TITLE:_____
SHORT DESCRIPTION OF DUTIES:_____
LENGTH OF TIME IN ABOVE OCCUPATION: YEARS:_____ MONTHS:_____
PREVIOUS EXPERIENCE IN OCCUPATION OTHER THAN ABOVE WHERE YOU PRIMARILY WORKED WITH RETARDED ADULTS:_____
YEARS:_____ MONTHS:_____

RATINGS OF STIMULUS-PERSONS

The following scale will allow this investigator to gain knowledge of how well you know the individuals you have judged. Be sure to rate each on "how well you know him" and not on how successful-unsuccessful you think he might be when given a vocational placement.

	DO NOT KNOW VERY WELL				KNOW VERY WELL
	1	2	3	4	5
[Stimulus-1].....	_____	_____	_____	_____	_____
[Stimulus-2].....	_____	_____	_____	_____	_____
[Stimulus-3].....	_____	_____	_____	_____	_____
[Stimulus-4].....	_____	_____	_____	_____	_____
[Stimulus-5].....	_____	_____	_____	_____	_____
[Stimulus-6].....	_____	_____	_____	_____	_____
[Stimulus-7].....	_____	_____	_____	_____	_____
[Stimulus-8].....	_____	_____	_____	_____	_____
[Stimulus-9].....	_____	_____	_____	_____	_____
[Stimulus-10].....	_____	_____	_____	_____	_____
[Stimulus-11].....	_____	_____	_____	_____	_____
[Stimulus-12].....	_____	_____	_____	_____	_____
[Stimulus-13].....	_____	_____	_____	_____	_____
[Stimulus-14].....	_____	_____	_____	_____	_____

PROBABILITY ESTIMATION SHEET

Please complete your Information and Rating Sheet before completing this Probability Estimation Sheet.

Your task here envolves the estimation of a probability figure indicating the degree of success each stimulus-person will have in the future. For example, if you feel a given individual has 9 chances out of 10 of being successful, you would place a "90" in his blank space. You may feel that another individual has 56 chances out of 100 of being successful on vocational placement so you would put the number "56" in his space. A poor prospect may receive something like a "12," indicating that you feel he has 12 chances in 100 of being successful.

You may use any number from 0 to 100 to indicate your probability estimation for each stimulus-person.

Name	Probability Estimate
[Stimulus-1]	_____
[Stimulus-2]	_____
[Stimulus-3]	_____
[Stimulus-4]	_____
[Stimulus-5]	_____
[Stimulus-6]	_____
[Stimulus-7]	_____
[Stimulus-8]	_____
[Stimulus-9]	_____
[Stimulus-10]	_____
[Stimulus-11]	_____
[Stimulus-12]	_____
[Stimulus-13]	_____
[Stimulus-14]	_____

APPENDIX B

INSTRUCTIONS TO OTC-2 JUDGE GROUP

INSTRUCTIONS TO JUDGE

This study is concerned with the characteristics needed for retarded adults to be successful on vocational placement. As you probably know, professionals working in the area of mental retardation do not have reliable measures on which to base predictions of vocational "success-failure." Presently, they often use "clinical judgment," a process by which professionals use their experience and "intuition" to assess probable vocational success or failure. Since you are a professional in this area, you have made numerous "clinical judgments."

In the present study, your task as a judge is to estimate the general degree of similarity between each pair of retarded adults presented regarding their potential success on vocational placement. The scale you will use ranges from "0" to "100," where "0" means no similarity and "100" means identity (i.e., identical pair). You may use any number between "0" and "100" to estimate the general degree of similarity. For example, I might estimate the similarity of a Cadillac and a Rolls Royce as being "85," since both cars are big and expensive, but yet quite different. However, I may well estimate the similarity of the same Cadillac and a Volkswagen "bug" as being "20," since they are both cars, but yet quite different kinds of cars. You may have considered somewhat different similarity estimates than the ones I have used since there is no "correct" answer.

Accompanying these instructions is a deck of computer cards. All possible pair-combinations of some retarded adults that you are professionally acquainted with are printed on the cards. Please write your similarity estimate for each pair near the lower left-

hand corner of the computer card. Each pair will be evaluated twice and your judgment may differ somewhat even for an identical pair. This is to be expected and there is no need to look back and view your first judgment of that particular pair.

It is important for you to know that there is no "correct" similarity estimate. I am interested in your opinion as a professional judge. Please do not discuss any of your impressions with your colleagues before you have completed your judgments as this may influence your opinion. Also, it would be preferable if you made all your judgments in "one sitting," rather than doing a small number of judgments at different times.

Remember, you are to judge the generalized similarity between each pair based on your knowledge of their potential vocational placement success. You may use any number between "0" and "100," where "0" indicates no similarity and "100" indicates identity. Please feel free to erase and "improve" your judgment if you find this necessary.

I would like you to begin making those judgments now before you complete the information sheets on the following pages.

Thank you for giving your time and effort. It is greatly appreciated.

DAVID L. HANSON

JUDGE INFORMATION AND RATING SHEET

Please complete your similarity estimations before completing this information and rating sheet.

The information asked for below is needed in order to report a "Description of Judges" section for this study. When this study is written, you will be presented in the form of averages. Please be as accurate as possible. Will you please PRINT your responses.

NAME_____SEX:_____
AGENCY OF PRESENT EMPLOYMENT:_____
OFFICIAL OCCUPATIONAL TITLE:_____
SHORT DESCRIPTION OF DUTIES_____
LENGTH OF TIME IN ABOVE OCCUPATION: YEARS:_____MONTHS:_____
PREVIOUS EXPERIENCE IN OCCUPATION OTHER THAN ABOVE WHERE YOU PRIMARILY WORKED WITH RETARDED ADULTS:_____
YEARS:_____ MONTHS:_____

RATINGS OF STIMULUS-PERSONS

The following scale will allow this investigator to gain knowledge of how well you know the individuals you have judged. Be sure to rate each on "how well you know him" and not on how successful-unsuccessful you think he might be when given a vocational placement.

	DO NOT KNOW VERY WELL				KNOW VERY WELL
	1	2	3	4	5
[Stimulus-1].....	_____	_____	_____	_____	_____
[Stimulus-2].....	_____	_____	_____	_____	_____
[Stimulus-3].....	_____	_____	_____	_____	_____
[Stimulus-4].....	_____	_____	_____	_____	_____
[Stimulus-5].....	_____	_____	_____	_____	_____
[Stimulus-6].....	_____	_____	_____	_____	_____
[Stimulus-7].....	_____	_____	_____	_____	_____
[Stimulus-8].....	_____	_____	_____	_____	_____
[Stimulus-9].....	_____	_____	_____	_____	_____
[Stimulus-10].....	_____	_____	_____	_____	_____
[Stimulus-11].....	_____	_____	_____	_____	_____

PROBABILITY ESTIMATION SHEET

Please complete your Information and Rating sheet before completing this Probability Estimation sheet.

Your task here envolves the estimation of a probability figure indicating the probable degree of success each stimulus-person will have in the future. For example, if you feel a given individual has 9 chances out of 10 of being successful, you would place a "90" in his blank space. You may feel that another stimulus-person has 56 chances out of 100 of being successful on vocational placement so you would put the number "56" in his space. A poor prospect may receive something like a "12," indicating a probability of 12 of of 100 chances.

You may use any number from 0 to 100 to indicate your probability estimation for each stimulus-person.

Name	Probability Estimate
[Stimulus-1].....	_____
[Stimulus-2].....	_____
[Stimulus-3].....	_____
[Stimulus-4].....	_____
[Stimulus-5].....	_____
[Stimulus-6].....	_____
[Stimulus-7].....	_____
[Stimulus-8].....	_____
[Stimulus-9].....	_____
[Stimulus-10].....	_____
[Stimulus-11].....	_____

APPENDIX C

INSTRUCTIONS TO SWS JUDGE GROUP

INSTRUCTIONS TO JUDGE

This study is concerned with the characteristics needed for retarded adults to be successful on vocational placement. As you probably know, professionals working in the area of mental retardation do not have reliable measures on which to base predictions of vocational "success-failure." Presently, they often use "clinical judgment," a process by which professionals use their experience and "intuition" to assess probable vocational success or failure. Since you are a professional in this area, you have made numerous "clinical judgments."

In the present study, your task as a judge is to estimate the general degree of similarity between each pair of retarded adults presented regarding their potential success on vocational placement. The scale you will use ranges from "0" to "100," where "0" means no similarity and "100" means identity (i.e., identical pair). You may use any number between "0" and "100" to estimate the general degree of similarity. For example, I might estimate the similarity of a Cadillac and a Rolls Royce as being "85," since both cars are big and expensive, but yet quite different. However, I may well estimate the similarity of the same Cadillac and a Volkswagen "bug" as being "20," since they are both cards, but yet quite different kinds of cars. You may have considered somewhat different similarity estimates than the ones I have used since there is no "correct" answer.

Accompanying these instructions is a deck of computer cards. All possible pair-combinations of some retarded adults that you are professionally acquainted with are printed on the cards. Please write your similarity estimate for each pair near the lower left-

hand corner of the computer card. Each pair will be evaluated twice and your judgment may differ somewhat even for an identical pair. This is to be expected and there is no need to look back and view your first judgment of that particular pair.

It is important for you to know that there is no "correct" similarity estimate. I am interested in your opinion as a professional judge. Please do not discuss any of your impressions with your colleagues before you have completed your judgments as this may influence your opinion. Also, it would be preferable if you made all your judgments in "one sitting," rather than doing a small number of judgments at different times.

Remember, you are to judge the generalized similarity between each pair based on your knowledge of their potential vocational placement success. You may use any number between "0" and "100," where "0" indicates no similarity and "100" indicates identity. Please feel free to erase and "improve" your judgment if you find this necessary.

I would like you to begin making those judgments now before you complete the information sheets on the following pages.

Thank you for giving your time and effort. It is greatly appreciated.

DAVID L. HANSON

JUDGE INFORMATION AND RATING SHEET

Please complete your similarity estimations before completing this information and rating sheet.

The information asked for below is needed in order to report a "Description of Judges" section for this study. When this study is written, you will not be identified by name and most of the information will be presented in the form of averages. Please be as accurate as possible. Will you please PRINT your responses.

NAME_____ SEX:_____
AGENCY OF PRESENT EMPLOYMENT:_____
OFFICIAL OCCUPATIONAL TITLE:_____
SHORT DESCRIPTION OF DUTIES:_____
LENGTH OF TIME IN ABOVE OCCUPATION: YEARS:_____ MONTHS:_____
PREVIOUS EXPERIENCE IN OCCUPATION OTHER THAN ABOVE WHERE YOU PRIMARILY WORKED WITH RETARDED ADULTS:_____
YEARS: _____ MONTHS: _____

RATINGS OF STIMULUS-PERSONS

The following scale will allow this investigator to gain knowledge of how well you know the individuals you have judged. Be sure to rate each on "how well you know him" and not on how successful-unsuccessful you think he might be when given a vocational placement.

	DO NOT KNOW VERY WELL				KNOW VERY WELL
	1	2	3	4	5
[Stimulus-1].....	____	____	____	____	____
[Stimulus-2].....	____	____	____	____	____
[Stimulus-3].....	____	____	____	____	____
[Stimulus-4].....	____	____	____	____	____
[Stimulus-5].....	____	____	____	____	____
[Stimulus-6].....	____	____	____	____	____
[Stimulus-7].....	____	____	____	____	____
[Stimulus-8].....	____	____	____	____	____

PROBABILITY ESTIMATION SHEET

Please complete your Information and Rating sheet before completing this Probability Estimation sheet.

Your task here envolves the estimation of a probability figure indicating the probable degree of success each stimulus-person will have in the future. For example, if you feel a given individual has 9 chances out of 10 of being successful, you would place a "90" in his blank space. You may feel that another stimulus-person has 56 chances out of 100 of being successful on vocational placement so you would put the number "56" in his space. A poor prospect may receive something like a "12," indicating a probability of 12 out of 100 chances.

You may use any number from 0 to 100 to indicate your probability estimation for each stimulus-person.

Name	Probability Estimate
[Stimulus-1].....	_____
[Stimulus-2].....	_____
[Stimulus-3].....	_____
[Stimulus-4].....	_____
[Stimulus-5].....	_____
[Stimulus-6].....	_____
[Stimulus-7].....	_____
[Stimulus-8].....	_____

APPENDIX D

INSTRUCTIONS TO PU JUDGE GROUP

INSTRUCTIONS TO JUDGE

This study is concerned with the characteristics needed for retarded adults to be successful on vocational placement. As you probably know, professionals working in the area of mental retardation do not have reliable measures on which to base predictions of vocational "success-failure." Presently, they often use "clinical judgment," a process by which professionals use their experience and "intuition" to assess probable vocational success or failure. Since you are a professional in this area, you have made numerous "clinical judgments."

In the present study, your task as a judge is to estimate the general degree of similarity between each pair of retarded adults presented regarding their potential success on vocational placement. The scale you will use ranges from "0" to "100," where "0" means no similarity and "100" means identity (i.e., identical pair). You may use any number between "0" and "100" to estimate the general degree of similarity. For example, I might estimate the similarity of a Cadillac and a Rolls Royce as being "85," since both cars are big and expensive, but yet quite different. However, I may will estimate the similarity of the same Cadillac and a Volkswagen "bug" as being "20," since they are both cards, but yet quite different kinds of cars. You may have considered somewhat different similarity estimates than the ones I have used since there is no "correct" answer.

Accompanying these instructions is a deck of computer cards. All possible pair-combinations of some retarded adults that you are professionally acquainted with are printed on the cards. Please write your similarity estimate for each pair near the lower left-

hand corner of the computer card. Each pair will be evaluated twice and your judgment may differ somewhat even for an identical pair. This is to be expected and there is no need to look back and view your first judgment of that particular pair.

It is important for you to know that there is no "correct" similarity estimate. I am interested in your opinion as a professional judge. Please do not discuss any of your impressions with your colleagues before you have completed your judgments as this may influence your opinion. Also, it would be preferable if you made all your judgments in "one sitting," rather than doing a small number of judgments at different times.

Remember, you are to judge the generalized similarity between each pair based on your knowledge of their potential vocational placement success. You may use any number between "0" and "100," where "0" indicates no similarity and "100" indicates identity. Please feel free to erase and "improve" your judgment if you find this necessary.

I would like you to begin making those judgments now before you complete the information sheets on the following pages.

Thank you for giving your time and effort. It is greatly appreciated.

DAVID L. HANSON

JUDGE INFORMATION AND RATING SHEET

Please complete your similarity estimations before completing this information and rating sheet.

The information asked for below is needed in order to report a "Description of Judges" section for this study. When this study is written, you will not be identified by name and most of the information will be presented in the form of averages. Please be as accurate and as informative as possible. Will you please PRINT your responses.

NAME:_____ SEX:_____
AGENCY OF PRESENT EMPLOYMENT:_____
OFFICIAL OCCUPATIONAL TITLE:_____
SHORT DESCRIPTION OF DUTIES:_____
LENGTH OF TIME IN ABOVE OCCUPATION: YEARS:_____ MONTHS:_____
PREVIOUS EXPERIENCE IN OCCUPATION OTHER THAN ABOVE WHERE YOU PRIMARILY WORKED WITH PSYCHIATRIC PATIENTS:_____
YEARS: _____ MONTHS:_____

RATINGS OF STIMULUS-PERSONS

The following scale will allow this investigator to gain knowledge of how well you know the individuals you have judged. Be sure to rate each on "how well you know him" and not on how successful-unsuccessful you think he might be regarding community adjustment.

	DO NOT KNOW VERY WELL				KNOW VERY WELL
	1	2	3	4	5
[Stimulus-1].....	_____	_____	_____	_____	_____
[Stimulus-2].....	_____	_____	_____	_____	_____
[Stimulus-3].....	_____	_____	_____	_____	_____
[Stimulus-4].....	_____	_____	_____	_____	_____
[Stimulus-5].....	_____	_____	_____	_____	_____
[Stimulus-6].....	_____	_____	_____	_____	_____
[Stimulus-7].....	_____	_____	_____	_____	_____
[Stimulus-8].....	_____	_____	_____	_____	_____
[Stimulus-9].....	_____	_____	_____	_____	_____
[Stimulus-10].....	_____	_____	_____	_____	_____
[Stimulus-11].....	_____	_____	_____	_____	_____
[Stimulus-12].....	_____	_____	_____	_____	_____
[Stimulus-13].....	_____	_____	_____	_____	_____
[Stimulus-14].....	_____	_____	_____	_____	_____

PROBABILITY ESTIMATION SHEET

Please complete your Information and Rating sheet before completing this Probability Estimation sheet.

Your task here envolves the estimation of a probability figure indicating the probable degree of success each stimulus-person will have in the future. For example, if you feel a given individual has 9 chances out of 10 of being successful, you would place a "90" in his blank space. You may feel that another stimulus-person has 56 chances out of 100 of being successful in community adjustment so you would put the number "56" in his space. A poor prospect may receive something like a "12," indicating a probability of 12 out of 100 chances.

You may use any number from 0 to 100 to indicate your probability estimation for each stimulus-person.

Name	Probability Estimate
[Stimulus-1]	_____
[Stimulus-2]	_____
[Stimulus-3]	_____
[Stimulus-4]	_____
[Stimulus-5]	_____
[Stimulus-6]	_____
[Stimulus-7]	_____
[Stimulus-8]	_____
[Stimulus-9]	_____
[Stimulus-10]	_____
[Stimulus-11]	_____
[Stimulus-12]	_____
[Stimulus-13]	_____
[Stimulus-14]	_____

References

Bae, A. Y. Factors influencing vocational efficiency of institutionalized retardates in different training programs. American Journal of Mental Deficiency, 1968, 72, 871-875.

Barrett, A. M., Relos, R., & Eisle, J. Vocational success and attitudes of mentally retarded toward work and money. American Journal of Mental Deficiency, 1965, 70, 102-107.

Bernstein, O. A study of the relationship between rehabilitation readiness and employability in mentally retarded adults. Dissertation Abstracts International, 1969, 30(1-B), 366-367.

Bieri, J., Atkins, A., Briar, S., Leaman, R., Miller, H., & Tripodi, T. Clinical and social judgment: The discrimination of behavioral information. New York: John Wiley and Sons, Inc., 1966.

Bolton, B. F., Butler, A. J., & Wright, G. N. Clinical versus statistical prediction of client feasibility. Wisconsin Studies in Vocational Rehabilitation Monograph, 1968, No. 7.

Braginsky, D. D. & Braginsky, B. M. Hansels and Gretels: Studies of children in institutions for the retarded. New York: Holt, Rinehart, and Winston, 1971.

Brolin, D. Value of rehabilitation services and correlates of vocational success with the mentally retarded. American Journal of Mental Deficiency, 1972, 76, 644-651.

Clark, A. T. & Foster, J. A. Objective measures and occupational success. Mental Retardation, 1970, 8, 41-44.

Cobb, H. V. Predicting vocational adjustment in the mentally retarded. Proceedings from a seminar held at Pennsylvania Rehabilitation Center, Johnstown, Penn., January, 1967.

Cobb, H. V. The predictive assessment of the adult retarded for social and vocational adjustment: a review of research. Part II, analysis of the literature. Social and Rehabilitation Service, Department of Health, Education and Welfare, Washington, D. C., 1968.

Cobb, H. V. & Epir, S. Predictive studies of vocational adjustment. In S. G. DiMichael (Ed.), New vocational pathways for the mentally retarded. Washington, D. C.: American Personnel and Guidance Association, 1966. Pp. 5-19.

Cohen, J. An analysis of vocational failures of mental retardates placed in the community after period of institutionalization. American Journal of Mental Deficiency, 1960, 65, 371-375.

Coles, G. J. Methodological note: an application of three-way analysis of variance in the estimation of interjudge reliability. Perceptual and Motor Skills, 1970, 31, 569-570.

Coles, G. J. & Stone, L. A. A new methodological revision of Ekman's "content" model of multidimensional similarity analysis. Multivariate Behavioral Research, 1972, 7, 85-107.

Comstock, J. A. The relationship between clinically derived scores of employability and employability scores predicted by an employability model for mentally retarded adolescents. Dissertation Abstracts International, 1971, 32(i-A), 271-232.

Conley, R. A benefit-cost analysis of the vocational rehabilitation program. Michigan: Division of Vocational Rehabilitation, Michigan Department of Education, 1971.

Daniels, L. K. & Steward, J. A. Development of a vocational adjustment rating scale for the mentally retarded. Training School Bulletin, 1971, 68, 10-14.

DiMichael, S. G. Vocational diagnosis and counsiling of the retarded insheltered workshops. American Journal of Mental Deficiency, 1960, 64, 652-657.

Downie, N. M. The San Francisco Vocational Competency Scale. Clinical Psychologist, 1969, 22, 154.

Eagle, E. Prognosis and outcome of community placement of institutionalized retardates. American Journal of Mental Deficiency, 1967, 72, 232-243.

Elkin, L. Predicting performance of the mentally retarded on sheltered workshop and non-institutional jobs. American Journal of Mental Deficiency, 1968, 72, 533-539.

Engel, A. M. Employment of the mentally retarded. American Journal of Mental Deficiency, 1952, 57, 243-267.

Engelhardt, G. M. Predicting rehabilitation of socially maladjusted boys. Journal of Counseling Psychology, 1970, 17, 546-549.

Erickson, R. C. A cooperative job training program for retarded youth. Part II, the relationship between selected variables and success of the retardate in the cooperative work-study program, an analysis of predictive power. Indiana: Purdue University, NIMH Project #1139, 1966.

Fields, D. L. & Gibson, D. Initial parole hypothesis and predictor cue valence. American Journal of Mental Deficiency, 1972, 77, 85-89.

Fry, L. M. A predictive measure of work success of high-grade mental defectives. American Journal of Mental Deficiency, 1956, 61, 402-408.

Gardner, J. M. & Giampa, F. L. Behavior competence and social and emotional behavior in mental retardates. American Journal of Mental Deficiency, 1970, 75, 168-169.

Gibson, D. & Fields, D. L. Rehabilitation forecast in mental retardation: the configuial search strategy. American Journal of Mental Deficiency, 1970, 74, 558-562.

Goldstein, H. Social and occupational adjustment. In H. A. Stevens and R. Heber (Eds.), Mental Retardation. Chicago: University of Chicago Press, 1964. Pp. 214-258.

Hanson, D. L. Miutidimensional perceptions regarding actual and potential job placement success of adult retardates. Unpublished Master's Thesis, University of North Dakota, 1971.

Hanson, D. L. & Stone, L. A. Multidimensional perceptions regarding actual and potential job-placement success of mentally retarded adults. Perceptual and Motor Skills, 1974, 38, 247-254 (in press).

Hartzler, E. A ten-year survey of girls discharged from the Laurelton State Village. American Journal of Mental Deficiency, 1953, 57, 512-517.

Hendel, D. D. Test format and administration variables as related to the performance of mentally retarded adults on multifactor tests of vocational abilities. Proceedings of the 79th Annual Convention of the American Psychological Association, 1971, 6(pt. 2), 615-616.

Hoffman, J. L. An investigation of factors contributing to successful and non-successful adjustment of discharged retardates. Washington, D. C., Social and Rehabilitation Service, Department of Health, Education and Welfare, Project RD-1606-P, 1969.

Huber, W. G. & Soferenko, A. Z. Factors contributing to the vocational success or non-success of the institutionalized retardate. Training School Bulletin, 1963, 60. 43-51.

Jackson, S. K. & Butler, A. J. Prediction of successful community placement of institutionalized retardates. American Journal of Mental Deficiency, 1963, 68, 211-217.

Jastek, J. F. Letter. Science, 1967, 157, 577-578.

Johnson, R. C. Prediction of independent functioning and of problem behavior from measures os IQ and SQ. American Journal of Mental Deficiency, 1970, 74, 591-596.

Justman, J. The counselor's use of measurement in prediction. National Catholic Guidance Conference Journal, 1968, 12, 145-133.

Kaiser, H. F. The varimax criterion for analytic rotation in factor analysis. Psychometrika, 1958, 23, 187-200.

Kaufman, H. I. diagnostic indices of employment with the mentally retarded. American Journal of Mental Deficiency, 1970, 74, 777-779.

Komisar, D. D. Key political and cultural factors influencing services for the retarded in Western Europe and the United States. In S. G. DiMichael (ed.), New vocational pathways for the mentally retarded. Washington, D. C.: American Personnel and Guidance Association, 1966. Pp. 46-52.

Kraus, J. Supervised living in the community and residential and employment stability of retarded male juveniles. American Journal of Mental Deficiency, 1972, 77, 283-290.

Larson, K. H. The characteristics of vocationally successful mentally retarded youth as described by two types of intelligence. Dissertation Abstracts International, 1964, 25(4), 2815.

Levine, S. & Elzey, F. San Francisco vocational competency scale. New York: The Psychological Corporation, 1968.

Lui, M. C. Changing trends in the care of the subnormal. American Journal of Mental Deficiency, 1963, 68, 345-353.

Madison, H. L. Work placement success for the mentally retarded. American Journal of Mental Deficiency, 1964, 69, 50-53.

Magnifico, L. X. & Doll, E. A. Out of school and self-supporting. Overview, 1962, 3. 32-33.

Mars, L. I. An exploratory cost-benefits analysis of vocational rehabilitation. Washington, D. C., Division of Statistic and Studies, Vocational Rehabilitation Administration, U. S. Department of Health, Education and Welfare, 1967.

Masland, R. L., Sarason, S. B. & Gladwin, T. Mental subnormality: biological, psychological and cultural factors. New York: Basic Books, 1958.

Michal-Smith, H. J. A study of the personal characteristics desirable for the vocational success of the mentally deficient. American Journal of Mental Deficiency, 1950, 55, 139-143.

Nihira, K. Factorial dimensions of adaptive behavior in adult retardates. American Journal of Mental Deficiency, 1969, 73, 868-878.

Nixon, R. A. Impact of automation and technological change on employability of the mentally retarded. American Journal of Mental Deficiency, 1970, 75, 152-155.

Robinson, H. B. & Robinson, N. M. The mentally retarded child: a psychological approach. New York: McGraw-Hill Co., 1965

Rosen, M., Kivitz, M. S. Clark, G. R. & Floor, L. Prediction of postinstitutional adjsutment of mentally retarded adults. American Journal of Mental Deficiency, 1970, 74, 726-734.

Sawyer, J. Measurement and prediction, clinical and statistical. Studies in Personnel Psychology, 1970, 2, 74-105.

Shafter, A. J. Criteria for selecting institutionalized mental defectives for vocational placement. American Journal of Mental Deficiency, 1957, 61, 599-616.

Skaarbrevik, K. J. A follow-up study of educable mentally retarded in Norway. American Journal of Mental Deficiency, 1971, 75, 560-565.

Song, A. Y. & Song, R. H. Prediction of job efficiency of institutionalized retardates in the community. American Journal of Mental Deficiency, 1969, 73, 567-571.

Stephens, W. B. Success of young adult retardates. Ann Arbor, Michigan: University Microfilms, Inc., 1964. Cited by H. V. Cobb, The predictive assessment of the adult retarded for social and vocational adjustment: a review of research. Part II, analysis of the literature. Socail and Rehabilitation Service, Department of Health, Education and Welfare, Washington, D. C., 1968

Stephens, W. B., Peck, J. R. & Veldman, D. J. Personality and success profiles characteristic of young adult male retards. American Journal of Mental Deficiency, 1968, 73, 405-414.

Stevens, S. S. On the operation known as judgment. American Scientist, 1966, 54, 385-401.

Stone, L. A. Multidimensional scaling of persons in groups: progress report. Perceptual and Motor Skills, 1972, 35, 35-42.

Stone, L. A. & Coles, G. J. Correlational similarity: the basis for a new revised method of similarity analysis. Studia Psychologica (Bratislava), 1970, 12, 258-264.

Stone, L. A., Coles, G. J. Psychology graduate students' multidimensional perceptions of their psychology faculty. Acta Psychology, 1971, 35, 364-377.

Stone, L. A., Coles, G. J., & Lindem, A. C. Multidimensional evaluation structure analysis (MESA): A complete multidimensional scaling system usable for a multiplicity of purposes. Grand Forks, North Dakota: Judgmetrics, 1970.

Tarjan, G., Dingman, H. F., Eyman, R. K., & Brown, S. J. Effectiveness of hospital release programs. American Journal of Mental Deficiency, 1960, 64, 609-617.

Tarjan, G., Wright, S. W., Eyman, R. K., & Keeran, C. V. National history of mental retardation: Some aspects of epidemiology. American Journal of Mental Deficiency, 1973, 77, 369-379.

Thorne, F. C. Clinical judgment: A clinician's viewpoint. Journal of Clinical Psychology, 1960, 26, 128-134.

Warren, F. Ratings of employed and unemployed mentally retarded handicapped males on personality and work factors. American Journal of Mental Deficiency, 1961, 65, 629-633.

Wechsler, D. The measurement and appraisal of adult intelligence. Baltimore: Williams and Wilkins Co., 1958.

Windle, C. Prognosis for mental subnormals. American Journal of Mental Deficiency 1962, 66, (Monogr. Suppl. 5).