DOCUMENT RESUME

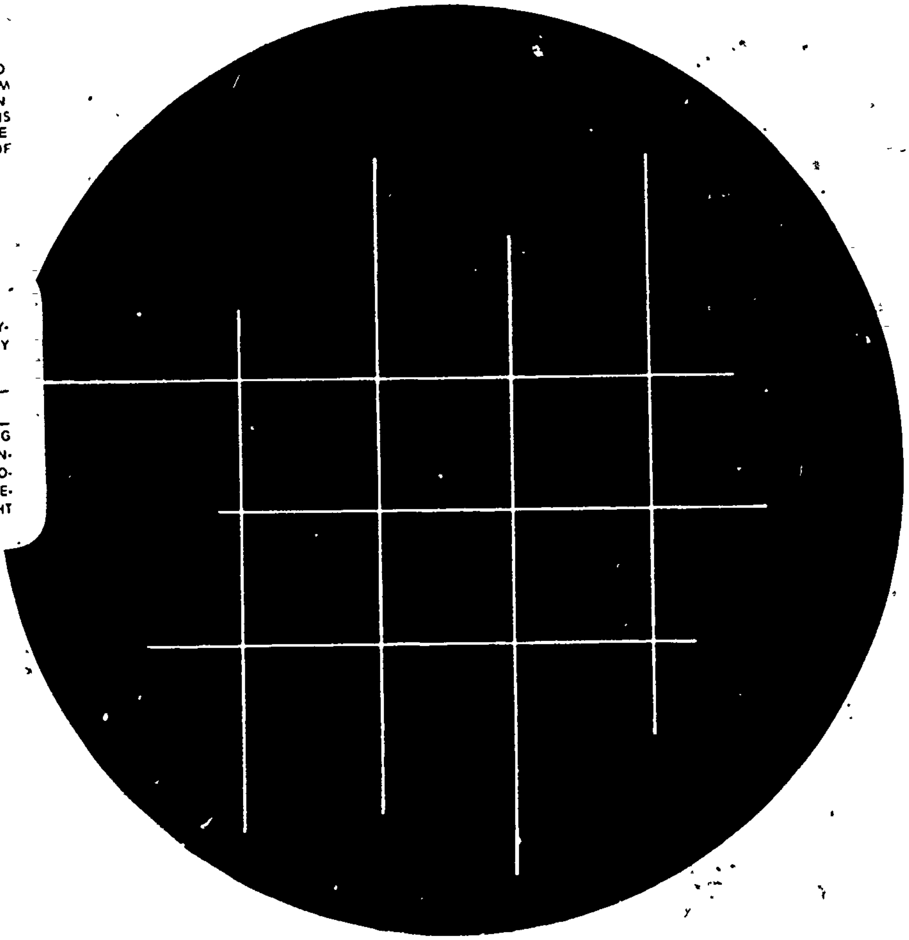
ED 109 155		TM 004 598
AUTHOR	Campbell, Joel T.; And Others	
TITLE	An Investigation of Sources of Bias in the Prediction of Job Performance: A Six-Year Study. Final Project Report.	
INSTITUTION	Civil Service Commission, Washington, D.C.; Educational Testing Service, Princeton, N.J.	
SPONS AGENCY	Ford Foundation, New York, N.Y.	
REPORT NO	PR-73-37	
PUB DATE	Sep 73	
NOTE	336p.; For a related document, see ED 073 121	
AVAILABLE FROM	Educational Testing Service, Rosedale Road, Princeton, New Jersey 08540 ($10.00)	
EDRS PRICE	MF-$0.76 HC-$17.13 PLUS POSTAGE	
DESCRIPTORS	Achievement Rating; Aptitude Tests; Criterion Referenced Tests; Data Analysis; Discriminatory Attitudes (Social); Ethnic Groups; Evaluation Criteria; *Job Analysis; Medical Laboratory Assistants; Minority Groups; Performance Tests; Predictive Validity; Statistical Analysis; *Task Analysis; *Task Performance; *Test Bias; Test Validity	
IDENTIFIERS	Cartographic Technician; Inventory Management Specialist	

ABSTRACT

The fairness of testing practices for selection or promotion of members of different racial and ethnic groups has been questioned. In this study, a search was made of occupations in the Federal Government to find those in various agencies with sufficient ethnic group representation for acceptable sample size, and for which dependable and objective criterion measures might be developed. Research conducted over a six-year period focused on the occupations of Medical Technician, Cartographic Technician, and Inventory Management Specialist. After careful job analyses, aptitude and ability test batteries were selected, several types of criterion measures were developed, and background data and job activities questionnaires were prepared. Data analysis indicated that tests found to be valid for one ethnic group were in general valid for the other ethnic groups included in the study, across occupations. Regression equations developed on majority group data appeared to predict almost equally well for minority groups, and some instances predicted higher criterion scores for minorities. The use of supervisors' ratings as a criterion of job performance in validity studies with ethnic sub-groups was called into question by findings of this study, which disclosed interaction effects ascribed to ethnic group membership of rater and ethnic group membership of ratee. (Author/BJG)

An Investigation of Sources of Bias in the Prediction of Job Performance

A Six-Year Study



FINAL PROJECT REPORT (PR-73-37)

Joel T. Campbell
Lois A. Crooks
Margaret H. Mahoney
Donald A. Rock

EDUCATIONAL TESTING SERVICE, PRINCETON, NEW JERSEY

AN INVESTIGATION OF SOURCES OF BIAS IN THE

PREDICTION OF JOB PERFORMANCE

. . . A SIX-YEAR STUDY . . .

FINAL PROJECT REPORT (PR-73-37)

Joel T. Campbell,
Principal Investigator
Lois A. Crooks
Margaret H. Mahoney
Donald A. Rock

September, 1973

Additional copies of this report may be obtained

from: Educational Testing Service
Rosedale Road
Princeton, New Jersey 08540

Price per copy: $10.00

Foreword

The study described in this report represents a major cooperative effort of the U. S. Civil Service Commission and Educational Testing Service, with support from the Ford Foundation. The initial joint proposal to the Ford Foundation was motivated by the many questions being raised in all sectors as to the fairness of testing practices for selection or promotion of members of different racial and ethnic groups. Dependable and defensible research findings on these questions were few, and contradictory results were being reported from study to study.

For the present study, a search was made of occupations in the Federal Government to find those in various agencies with sufficient ethnic group representation for acceptable sample size, and for which dependable and objective criterion measures might be developed. Research was ultimately carried out in three stages over a six-year period, with separate studies in depth of the occupations of Medical Technician, Cartographic Technician, and Inventory Management Specialist.

Each stage began with careful job analysis by researchers to determine factors necessary for successful job performance. On the basis of the job analyses, aptitude and ability test batteries were selected, several types of criterion measures were developed, and background data and job activities questionnaires were prepared.

The data gathering phases for each occupation were conducted after eliciting the cooperation of supervisory staff and job incumbents in installations of the government agencies agreeing to participate in the project. All testing was administered by ETS research staff.

Data were analyzed for each occupation by ethnic group so that comparisons could be made at every level of statistical treatment. A major and important finding of the study is that tests found to be valid for one ethnic group are in general valid for the other ethnic groups included in the study, across occupations. Regression equations developed on majority group data appeared to predict almost equally well for minority groups, and in some instances predicted higher criterion scores for minorities. The use of supervisors' ratings as a criterion of job performance in validity studies with ethnic subgroups is called into question by findings of this study, which disclosed interaction effects ascribed to ethnic group membership of rater and ethnic group membership of ratee.

Findings of this study have implications in particular for employers, behavioral scientists, and others concerned with social and public policy issues. This research has contributed significantly toward understanding a major concern--the fairness of testing. At the same time, it has raised questions. It is hoped that some of these questions will receive attention in the future.

William A. Gorham

Samuel J. Messick

Preface

Many people have been involved in this study over the past six years, not only in design and direction of the research but in development of the instrumentation, in data collection and analyses, and in the various stages of reporting the results. The project could not have been carried out without the assistance of those in the Federal agencies who facilitated the data collection, and the cooperation of the 1,400 job incumbents who were the subjects.

Members of the Advisory Committee, who were convened periodically for consultation on research design, progress of the study, and implications of the findings, filled an invaluable role. They were:

John K. Hemphill, Far West Laboratory for Educational Research and Development, Chairman

Marvin D. Dunnette, University of Minnesota

Robert M. Guion, Bowling Green State University

S. O. Roberts, Fisk University

Members of the Management Committee, who joined with the Advisory Committee in following the progress of the study, made themselves available for counsel and support on a day-to-day basis. Their names, with period served, follow:

William W. Turnbull, President, Educational Testing Service (until July 1969)

Albert P. Maslow, Chief, Personnel Research and Development Center, U. S. Civil Service Commission (until September 1971, when he joined Educational Testing Service)

Samuel J. Messick, Vice President, Educational Testing Service (from July 1969)

William A. Gorham, Associate Director, Personnel Research and Development Center, U. S. Civil Service Commission (from September 1971)

The Project Staff, from the U. S. Civil Service Commission, included

William A. Gorham (then Chief of Research and Development) until he joined the Management Committee, and Mary L. Tenopyr, Acting Chief of Research (now with American Telephone and Telegraph Company), and the following from Educational Testing Service: (Those marked with asterisks worked on the study from its inception to completion. The other staff members were involved at various stages of the study, as it progressed.)

*Joel T. Campbell, Senior Research Psychologist, Principal Investigator

Ronald L. Flaugher, Senior Research Psychologist

*Donald A. Rock, Senior Research Psychologist

Franklin R. Evans, Research Psychologist

Lewis W. Pike, Research Psychologist

David M. Nolan, Director, Washington, D. C., Office

Lois A. Crooks, Associate Research Psychologist

William S. Hall, Associate Research Psychologist

Lila Norris, Associate Research Psychologist

Charles F. Barr, Assistant Program Director, Occupational and Professional Assessment Programs

Barbara Dynarski, Senior Research Assistant

*Margaret H. Mahoney, Senior Research Assistant

Mary Ellen Parry, Senior Research Assistant

*Harriet Blizzard, Research Assistant

*Virginia Rau, Administrative Assistant

C. Brooke Gruenberg, Secretary

Other ETS staff members, not listed, assisted in data collection.

Chapters of this report were written by the following:

Chapters I, V, VI, VIII, IX, and XII by Joel T. Campbell

Chapters III, IV, and VII by Lois A. Crooks

Chapters II and X by Margaret H. Mahoney

Chapter XI by Donald A. Rock

The overall editing was done by Lois A. Crooks.

Acknowledgements

The quality of this technical report has been enhanced by the suggestions and comments of the speakers and participants at our June 1972 Invitational Conference, where principal findings of the research were first presented. Drafts of many of the chapters were reviewed prior to the Conference by the following speakers not directly involved in the research:

Lewis E. Albright, Director, Manpower Planning and Development, Kaiser Aluminum and Chemical Corporation

Anne Anastasi, Professor of Psychology, Fordham University

Roscoe C. Brown, Jr., Institute of Afro-American Affairs, New York University

Edward J. Casavantes, Executive Director, Association of Psychologists for La Raza

Robert M. Guion, Professor of Psychology, Bowling Green State University

Raymond Jacobson, Director, Bureau of Policies and Standards, U. S. Civil Service Commission

S. Rains Wallace, Professor of Psychology, The Ohio State University

A special debt of gratitude is due the following for their discerning and critical examination of all aspects of this report:

William A. Gorham, Associate Director, Personnel Research and Development Center, U. S. Civil Service Commission

Robert L. Linn, Director, Developmental Research Division, Educational Testing Service

Albert P. Maslow, Director, Occupational and Professional Assessment Programs, Educational Testing Service

Samuel J. Messick, Vice President, Educational Testing Service

Joel T. Campbell
Lois A. Crooks
Margaret H. Mahoney
Donald A. Rock

Table of Contents

Chapter		Page
	Foreword	i
	Preface	iii
	Acknowledgements	v
	Table of Contents	vii
I	Introduction	1
II	Description and Design of the Study: Three Occupations	9
III	Comparison of Ethnic Groups on Aptitude Measures	47
IV	Comparison of Ethnic Groups on Criterion Measures	81
V	Validity of Aptitude Tests for Different Ethnic Groups	125
VI	Regression Analysis of Best Predictors of Job Performance Measures for Different Ethnic Groups	157
VII	Multiple Correlations and Cross-Ethnic Cross-Validation Coefficients for Prediction of Criterion Measures	239
VIII	Biasing Effects of Ethnic Rater-Ratee Interaction	277
IX	Factor Analysis of Predictor and Criterion Variables by Ethnic Group	319
X	Comparisons of Background and Task Variables by Occupation by Ethnic Group	357
XI	An Alternative Approach to Culture Fairness When Tests Are Used in Selection	397
XII	Conclusions and Inferences	423
	References	427
	Appendices	433
	Appendix to Chapter II	435
	Appendix to Chapter III	465
	Appendix to Chapter V	475

Chapter I

Introduction

This is the final report of a study initiated in late 1966 to explore the relationships of selected background and ability measures to various criteria of job performance for several ethnic groups. The project, spanning a six-year period, was a joint effort of the U. S. Civil Service Commission and Educational Testing Service, and was supported by grants from the Ford Foundation.

At the time the proposal to undertake the research was made, concern as to the fairness of testing practices for minority groups was widespread. Tests were perceived as a barrier to selection and promotion of minorities, both in industry and government. The dearth of carefully-done research and the sometimes conflicting findings being reported motivated the joint undertaking.

The study was begun after an initial survey of occupations within the Federal Government to determine those with sufficient ethnic subgroup representation. The occupation of Medical Technician (GS-645) was selected for the pilot project, in which data were obtained on Black and Caucasian job incumbents at specified levels. Experience with the pilot study demonstrated the feasibility of conducting such research, and the findings were of sufficient interest to justify funding for study of two additional occupations: Cartographic Technician (GS-1371) and Inventory Management Specialist (GS-2010). Both of these occupations included sufficient numbers of Mexican-Americans, as well as Blacks and Caucasians, for meaningful comparisons between ethnic groups.

Previous reports published

A number of reports of the Medical Technician pilot study were published: Pike (1969); Flaugher, Campbell, and Pike (1969); Campbell, Pike, and Flaugher (1969); Campbell, Pike, Flaugher, and Mahoney (1970); Rock, Campbell, and Evans (1970); and Parry and Mahoney (1970). A description of the instrumentation, methodology, and sample obtained in the Cartographic Technician phase can be found in Parry (1971) and for the Inventory Management Specialist phase in Crooks and Mahoney (1971). Proceedings of an invitational conference, convened June 22, 1972, to report and critique the principal findings, were also published (Crooks, Ed., 1972).

Related research

Much attention has been directed over the years to differences in performance on tests among different ethnic groups. Until comparatively recently, however, almost all studies were limited to comparisons of mean scores, score ranges, or amount of test overlap. Shuey (1966) catalogued some 380 such studies. With great uniformity, these studies showed that Black samples scored below corresponding Caucasian samples, with a usual difference in means of about one standard deviation.

Beginning in the 1960's, interest and concern shifted from test score comparisons to test validity and the fairness of prediction for different ethnic groups. During the last few years there has been a number of studies of the effectiveness of differential prediction in industrial or occupational settings. These studies have differed in a number of respects: in the type of criterion used; whether the tests were used for selection, administered experimentally, or administered for some ancillary purpose; the size of the samples involved; the type of statistical analysis used; the definition of "test bias," etc. For these and other reasons, there has been some continuing controversy on the interpretation of results from the various studies.

As Boehm (1972) has pointed out, the first study of this type (Lopez, 1966) produced the strongest indications of differential validity yet obtained. However, the statistical procedures used in that study were defective in some respects, which make the findings difficult to interpret. Similarly, another early report (Kirkpatrick, et al., 1968), which included research on a number of jobs and training situations, produced several apparent instances of differential validity. However, the sub-study which produced the largest number of these instances was one where ethnic group and criterion were confounded, again making interpretation difficult.

Boehm (1972), in reviewing differential validation studies, pointed out the difference, apparently frequently confused, between what she calls "single group validity" and "differential validity." Differential validity she defines as a situation where "(a) there is a significant difference between the correlation coefficient of a selection device and a criterion obtained for one ethnic group and the correlation of the same device with the same criterion obtained for the other group, and (b) the validity coefficients are significantly different from zero for one or both groups." Single group validity she defines as a situation where "a given predictor exhibits validity significantly different from zero for one group only, and there is no significant difference between the two validity coefficients."

Many of the studies which have reported instances of supposed "differential validity" were in fact examples of "single group validity." In many of these instances, the minority group sample had a small N and did not reach the level of statistical significance for this reason, even where the validity was in fact larger for the minority group than for the majority group. This was true, for example, in the studies reported by O'Leary, et al. (1970) and Farr, et al. (1971).

In a later paper, Boehm (1972, unpublished) has shown that findings of both "single group validity" and "differential validity" are closely related to sample size. Most instances of such phenomena occur when one or both samples are below 50.

Ruch (1972) reviewed those studies carried out in a business or industrial setting where it was possible to test homogeneity of regression. (A number of studies, including some data from our study of Medical Technicians, were reviewed in both the Boehm and Ruch analyses.) Ruch's conclusions were that (1) there were fewer differences than would be expected by chance regarding dispersion and slope of the regression lines, and (2) there were more differences than would be expected by chance in regression line intercepts, with the intercept for the minority groups falling below that of the majority group.

Stanley (1971), reviewing the literature in the educational field, came to similar conclusions. Guinn, Tupes, and Alley (1970), studying prediction of training criteria in ten Air Force technical schools, also found overprediction of minority performance rather than underprediction.

These findings would indicate (1) that the degree of prediction is as good for minority groups as for majority groups (as was shown by the Boehm review), and (2) that the use of a prediction equation based on a majority sample will tend to overpredict rather than underpredict the job performance of the minority group, findings that are consistent with results of the study being reported here.

In recent years, a number of definitions of "test bias" have been proposed. The one most often applied is the Cleary (1968) definition. Cleary defines a test as being "culture-fair" for populations A and B when the regression equation based on population A neither systematically over- nor underpredicts the level of performance for members of population B. This "regression" model has been applied in the present study.

Thorndike (1971) has proposed an alternative definition which may lead to entirely opposite conclusions as to whether a test is "culture-fair." He suggests that a test may be judged "culture-fair" if the overlap on the criterion scores between groups A and B is essentially equivalent to their overlap on the predictors. (Additional analyses using the Thorndike approach were carried out on data in this study. Theory and results are discussed in Chapter XI.)

Darlington (1971) and Cole (1972) take slightly different approaches to the issue, based on the conditional probability that being selected for a job is the same for minority and majority groups, given satisfactory criterion performance. Linn (1973) discusses the major differences in these formulations. He points out that there is more than one reasonable definition of test fairness and that these definitions are in conflict, so that the choice of definition is a matter of weighing competing values.[1]

Organization of this report

This report is in general organized to reflect the order in which each phase of the project was carried out, from intensive analysis of each occupation and factors in job performance, to selection and development of instruments, to data gathering, to basic analysis of data, and then to increasingly sophisticated treatment of the data to determine other underlying relationships, patterns, and implications. Results from all three occupations are included in each chapter.

A brief description of the content of the chapters follow:

Chapter II describes the design of the study, identification of occupations suitable for study, task analysis of the three occupations chosen, issues in the selection of background and ability measures, considerations in the selection and development of types of criterion measures, how the samples were drawn, and how testing was carried out. For more detailed descriptions of instrumentation development for each phase and samples of some of the instruments, see Pike (1969); Parry (1971); and Crooks and Mahoney (1971).

In Chapter III, aptitude and ability measures selected as the predictor batteries for the three occupations are described in relation to factors observed in job performance in each occupation. Means and

[1] Each of these alternative definitions requires an explicit recognition of ethnic or other group membership. For that reason, use of any one of these definitions in employment decisions may run counter to the principle stated in the Griggs vs. Duke Power decision (1971): "Congress has made [job] qualifications the controlling factor so that race, religion, nationality, and sex become irrelevant."

standard deviations of scores obtained on the measures are compared by ethnic group for each of the three occupations.

In Chapter IV, criterion measures developed for the three occupations are described briefly. Means and standard deviations of criterion scores and intercorrelations of the measures by ethnic group are presented for each occupation studied.

Chapter V deals with the validity of the aptitude tests for each ethnic group as measured by the degree of relationship of aptitude test scores with the various criteria of job performance for each occupation studied.

Chapter VI explores differences in the linear relationships (regression lines) between test scores and criterion measures by ethnic group by occupation.

In Chapter VII, results are described of stepwise multiple regression analyses, in which a best set of predictors was selected by ethnic group for each criterion measure for each occupation. Results of cross-ethnic cross-validation, in which prediction equations derived for each ethnic group were then applied to data for the other ethnic samples, are also presented.

In Chapter VIII, the possible biasing effects of ethnic group membership of raters (supervisors) in interaction with ethnic group membership of ratees (job incumbents) are described for the three occupations studied.

Chapter IX presents factor analyses of predictor and criterion variables to determine whether differential patterns of abilities may be observed from one ethnic group to another. Selected background variables were added by extension.

Chapter X presents comparisons of background and task variables by ethnic group b ccupation. Relationships of selected background variables to measures of aptitude and job performance are compared by ethnic group. Results of task analyses carried out to determine whether there were variations in the type of work done by members of different ethnic groups are presented.

Chapter XI contains a theoretical discussion of the Cleary (1968) and Thorndike (1971) definitions of "culture-fair" tests, and a reanalysis of ethnic sample data from this study using a modified Thorndike approach.

Chapter XII presents the major findings in the study and their implications. Recommendations for future approaches to some of the unresolved problems are given.

Chapter II

Description and Design of the Study: Three Occupations

This chapter deals with identification of occupations suitable for study, analysis of the three occupations chosen, issues in the selection of background and ability measures judged most likely to relate to job performance in these occupations, considerations in the selection and development of the three types of criterion measures used, and how the samples were drawn and testing implemented. For detailed descriptions of the instrumentation development for each phase, see Pike (1969); Parry (1971); and Crooks and Mahoney (1971).

A. Preliminary Study of Occupations

The first phase of the program was a feasibility study to identify the range of technical and administrative issues. A critical first step was to determine whether there were sufficient numbers of minority group members in selected occupations to support a full-scale study. Some of the stipulations in identifying occupations suitable for study were: the occupational groups should be recognizable by occupational classification series or specialty, the incumbents must include sizeable numbers of more than one identifiable ethnic group, written tests must be an acceptable method for measuring job qualifications, it should be possible to obtain objective measures of job performance, and the incumbents should have the same basic skills in common across installations or specialties. An added qualification was that those in an occupation suitable for study should be employed in a number of geographic locations, and include persons with a diversity of backgrounds.

After securing occupational data on minority employment in selected professional, technical and clerical jobs from nine agencies within the Federal government[1]; it appeared that selecting appropriate jobs for study would be no problem at the lower grade levels (GS 1 through 5). At the higher levels, only a few jobs were a possibility.

The process of gaining entry and managing negotiations with the various agencies under consideration was handled by United States Civil Service Commission (USCSC) representatives. The purpose of the study was discussed with representatives of minority groups, stressing the intentions of the research and security of individual data.

Representatives of USCSC also met with representatives of five Federal employees' unions and informed them of the project: American Federation of Government Employees, Government Employees Council, National Alliance of Postal and Federal Employees, National Association of Internal Revenue Employees, and National Federation of Federal Employees. They were asked to make known to their locals that they had been briefed, and approved members' participation in the data collection.

Medical Technicians

The occupation selected for the feasibility study was Medical Technician (GS-645). It was chosen because Blacks and Caucasians were represented in this occupation in sufficient numbers for study, the tasks which the incumbents performed appeared to have highly visible and objective outcomes that would lend themselves to evaluation, and a large proportion of the entire population was employed by a single

[1] U. S. Departments of the Treasury, Commerce, Army, Air Force, Social Security, Veterans Administration, Agriculture, Labor, and Defense Supply Agency

government agency, the Veterans Administration, at various geographic locations. One particular advantage of studying this occupation was that written test scores were not a part of the basis for hiring the majority of technicians. Thus, there should be less restriction of range due to preselection by test scores. Such preselection would tend to lower expected relationships between test scores and job performance criteria.

A count was made by the Veterans Administration of the number of all personnel in the Medical Technician occupation (GS-645) at each installation by grade level, specialty (bacteriology, hematology, histopathology, etc.), and ethnic group. The Veterans Administration asked supervisors to identify employees individually by ethnic group, since Federal regulations prohibit ethnic identification on personnel records.

During Autumn of 1967, five ETS staff members visited nine Veterans Administration Hospital laboratories in several parts of the country[2]. Before each hospital was visited, a letter was sent from the Veterans Administration Washington headquarters to the hospital's director, explaining the purpose of the research program and the necessity for the visit, naming ETS researchers who would be coming, and briefly describing what they proposed to accomplish while they were there and what the later phases of the study would entail. The support and assistance of the director and the staff were encouraged.

In these preliminary visits, every effort was made to gain as

[2] San Francisco and Livermore, Calif., Jackson, Miss., Lyons, N. J., Manhattan, N. Y., Washington, D. C., Dayton, O., Wilkes-Barre, Pa., and Beckley, W. Va.

thorough an understanding of the occupation of Medical Technician as possible, including preliminary education, training, skills, and abilities needed; kinds of tasks involved; and aspects of the job which were thought most important for successful performance. Hospitals of various sizes with personnel and facilities for treating diverse types of disability were selected in different geographic locations. Administrative and personnel officers were interviewed. A brief history of each hospital was obtained. Hiring patterns and area economic conditions were examined. Medical Technicians at representative GS-levels and their supervisors were interviewed and observed at work. Supervisors often were medical technologists or M.D.s specializing in pathology, and from them insights relating to peripheral problems in the laboratories were gained. Particular emphasis was placed on fact-finding for a thorough job analysis, based on interviews, observations, and Civil Service Position Descriptions. Incumbents were asked how they spent their time each day--what was important, what unimportant? How did they learn to operate the equipment they used? How routine was its operation? How highly automated? Were they trained in depth in one specialty or in several? Did they ever attend seminars or workshops where new equipment or techniques were demonstrated? What constituted a serious error, and what were its consequences? What qualifications and personal characteristics did they feel were necessary? Personnel files were examined to determine the kinds of personal history variables that should be documented in the full study. A collection of standard record forms, written instructions, and specifications routinely used in the laboratory was made. The published literature was reviewed for any past research on this occupation. After the selection and design of the actual assessment materials had begun, several return visits were made to the hospital laboratories in order to verify or correct some impressions and to obtain further job information.

The feasibility study demonstrated the enormous technical and logistic difficulties of conducting such research. The value of thorough job analysis in planning the instrumentation for such a study was clearly confirmed. Since performance on the job is multidimensional and complex, it is necessary to acquire knowledge of the job in depth before selecting aptitude tests, constructing rating scales, or developing a work sample. In addition, information about the background of job incumbents is necessary to permit construction of a comprehensive personal history questionnaire.

Cartographic Technicians

The occupation selected for study in the second phase of the project was Cartographic Technician (GS-1371). One of the major reasons this occupation was chosen was that job incumbents included large numbers of both Mexican-Americans and Blacks. Thus, relationships of predictors to job performance could be compared for an additional ethnic group. It was also an important objective to see whether the findings for the Medical Technicians would be replicated for the Cartographic Technicians.

Most Cartographic Technicians are concentrated in three government agencies: U. S. Department of the Army - Corps of Engineers and Topographic Command, U. S. Department of Commerce - Coast &

Geodetic Survey and Bureau of the Census, and U. S. Department of the Interior - Geological Survey. Early in 1968, six ETS staff members began visiting seven government mapping installations[3]. The same general pre-visitation procedures were followed as for the Medical Technicians. A count was obtained through the USCSC of incumbents by grade level, specialty, and ethnic group. Particular care was exercised in informing heads of installations of impending visits so that necessary advance preparations were assured, including letters to supervisory personnel and clearance with union officials.

One obstacle to the study of this occupation concerned government security regulations. Since much military mapping was currently in progress, ETS researchers had to be granted security clearance by the USCSC before they were permitted to visit most of the mapping installations. Often it was impossible to interview and observe the Cartographic Technicians on the job because they were mapping restricted areas or using classified equipment. Researchers were assured, however, that the same basic skills and abilities were needed to perform the classified tasks, and that job analyses based on observations of unclassified projects would be adequate.

After the preliminary visits had been completed, the following conclusions were reached:

(1) the secure nature of the work at some installations would not preclude the inclusion in the study of the technicians working on classified assignments.

(2) the varied nature of the work done by Cartographic

[3] Arlington, Va., Rockville-Silver Spring and Bethesda, Md., Norfolk, Va., Providence, R. I., San Antonio, Tex., and Detroit, Mich.

Technicians would require more than one work sample task.

(3) practical considerations would require excluding Cartographic Technicians from the study at installations employing less that 20 incumbents.

(4) the existence of a substantial number of Mexican-Americans in GS-1371 positions would enable an investigation into the test performance of a second minority group.

(5) Cartographic Technicians employed by the U. S. Department of the Interior - Geological Survey would not be included because of insufficient minority group representation.

Inventory Management Specialists

The third and final phase of the study concerned the occupation of Inventory Management Specialist (GS-2010). These specialists are primarily employed in defense agencies, in a limited number of locations, and include large numbers of Blacks and Mexican-Americans. This classification spanned a higher GS-level than the other two studied, and appeared to require a different set of skills and abilities.

In the Spring of 1968, preliminary visits were made to two Department of Defense agencies in Philadelphia. Approval of funding for this and the other phases of the study was not received until late 1969. In July, 1970, members of the ETS research staff began visiting installations of four agencies of the Department of Defense: Defense Personnel Support Center and Defense Industrial Supply Center, installations of the Defense Supply Agency in Philadelphia, Navy Aviation Supply Office in Philadelphia, Army Tank-Automotive Command in Detroit, and San Antonio Air Materiel Area (Air Force). Again, contacts were made with the agencies by the Civil Service Commission through Washington headquarters. An effort was made to meet with military administrative heads of the installations, as well as personnel officers and civilian representatives of higher management levels, to introduce the purpose of the research and to emphasize the importance of getting their people to cooperate. On-the-job training instructors were interviewed and their lesson materials examined.

A primary area of concern was the route by which individuals entered these jobs. Extensive questioning revealed that incumbents were chosen in a number of different ways. Many had been screened on a written test (Federal Service Entrance Examination) for entry into the career development training program at GS level 5, some had qualified on the basis of experience at lower levels, and others had been transferred or were newly hired from another occupation.

Because of the diversity of items managed, special emphasis was placed during the interviews on determining the areas of possible variation among and between jobs, general activities perceived as common across jobs and agencies and common problem areas, and areas where differences were perceived across jobs or agencies. Again, samples of forms, written procedures, job descriptions, and communications of various kinds were gathered. A number of Inventory Management Specialists interviewed were asked to keep a Daily Activity Log for five days, and to record, at half-hour intervals, activities and interactions with others in the process of accomplishing their work.

B. Selection of Aptitude Tests (Predictors)

In selecting aptitude tests for each of the three occupations studied, researchers utilized their knowledge and familiarity with the job, gained from interviews and work observations. Considerations were finding or developing tests that measured specific abilities observed as needed for the job, tests of general abilities which had been found in previous research to relate to job success, and tests which were purported to measure other possibly job-related ability factors. One important basis for selection was time required to take the tests, since time spent away from the subjects' regular jobs was a necessary consideration, especially for the Medical Technicians. Ease of group administration and face validity were also felt to be important. Other considerations were the availability of tests with two separately timed halves, to facilitate reliability estimation, and the need for speeded versus unspeeded tests.

It was decided to use available tests rather than to develop new ones. After considering a number of alternatives, the French, et al., Kit of Reference Tests for Cognitive Factors became the source of the majority of tests in the predictor batteries for all three occupations. These tests have short time limits, are divided into separately timed halves, have known factorial content, and can be group-administered easily. Since these tests were designed for research purposes, the subjects used in this study probably would not have taken them previously. Four U. S. Civil Service Commission tests were used. For the Cartographic Technicians, the Coordination Test of the Flanagan Industrial Tests series was also selected.

In the next chapter, the aptitude tests will be described as they relate to specific aspects of each occupation studied.

C. Selection and Development of Measures of Job Performance (Criteria)

For the first two occupations studied (Medical Technician and Cartographic Technician), three types of job performance measures were developed, Criterion Rating Scales, a Work Sample, and a Job Knowledge Test. Since preliminary study of the third occupation (Inventory Management Specialist) yielded wide variations in specific knowledge needed and procedures followed within and among installations, a decision was made not to develop a Job Knowledge Test, but to attempt to measure the variance in performance attributable to job knowledge in the Work Sample and Criterion Rating Scales.

Rating Scales

During the preliminary interviews and observations of each occupation, opinions were solicited concerning job requirements and the qualities that were needed for successful job performance. Information obtained by the various interviewers was then pooled, and a list of job factors that appeared to be important was identified.

Additional visits were made, with emphasis on obtaining a broader sampling of activities, confirming or correcting original impressions, and evaluating the degree of similarity within job classification across laboratories, installations, or agencies. Interviewees were asked to evaluate and comment on the job factors previously identified as being important. A tentative list of factors was decided on as a basis for developing a set of rating scales, and further judgments concerning their suitability were made. Some were combined because the definitions were overlapping, or were eliminated because the factors were judged to be of lesser importance. Descriptions of the job factors included in the sets for all three occupations are given in Chapter IV.

For all three occupations studied, the anchored rating scale format (Smith and Kendall, 1963) was used, in which scales are anchored by examples of specific behavior. Although some of the same characteristics were felt to be important for all occupations being studied, all scales and behavioral examples were defined in terms of the specific job under investigation.

Detailed instructions were given at the beginning of the booklet of rating scales, and a list of incumbents to be rated, along with identification numbers assigned to them, was included at the end of the booklet. The Medical Technicians were rated by co-workers and two supervisors whenever possible. For the Cartographic Technicians and Inventory Management Specialists, only supervisors' ratings were obtained. Peer ratings were dropped because of time pressure, impracticality, and possible adverse reaction. For the latter two occupations, the supervisor was asked to indicate how long each person he rated had been under his jurisdiction.

For all three occupations, raters were also asked to evaluate the performance of a fictitious incumbent whose characteristic behavior was described on the page opposite each scale. This standard stimulus was used in an attempt to disclose stylistic variations or bias in rating behavior. For discussion see Appendix to Chapter II.

Whenever possible, the raters met in groups, with the researcher leading them through the rating procedure. Ratings were made for all ratees by their supervisors on one scale at a time. The purpose was to focus on the definitions of the scales, and to emphasize the experimental and confidential nature of the study and the importance of valid ratings.

For the second and third occupations studied, each supervisor, after completing the ratings, was asked to indicate on a separate form how important the characteristic described on each of the scales was to overall job performance from "Not Important - Irrelevant to Proper Job Performance" at the low end, to "Important - Contributes to Proper Job Performance" at the midpoint, and "Very Important - Essential to Proper Job Performance" at the high end of the scale.

Job Knowledge Test

Job knowledge is one aspect of job performance that can be measured directly. To construct a valid test of job knowledge, collaboration is necessary between persons who know the job and measurement experts, in writing test items that will effectively sample the technical knowledge required for successful job performance. Preliminary testing of the items on a sample of job incumbents, analysis of these data, and selection of the items for the final form of the test which best meet measurement objectives are additional required steps.

In search of likely sources, no appropriate test was found that could be administered to Medical Technicians in the first occupation studied. However, through informal contacts it was discovered that one medical laboratory training school had developed a large number of test items over a period of years. A total of

584 items from the Allegheny General Hospital Medical Laboratory Assistant Training Program in Pittsburgh, Pennsylvania, was made available. After reviews by two supervisors of Medical Technicians, one from a VA hospital and one from a civilian hospital, 147 multiple choice items were selected as most likely to differentiate among technicians on the dimension of job knowledge. These items were pretested and a selection was made for the final form, as described later in this chapter.

For the Cartographic Technicians, a test of basic job knowledge was developed especially for this project. ETS researchers worked with three cartographers from U. S. Army TOPOCOM, one from each of the major divisions where Cartographic Technicians are employed: cartography, photogrammetry, and triangulation. All of these cartographers had knowledge of the work done in the other divisions, and could therefore work together effectively as a team in writing test items. Every effort was made to have the test content pertain to the types of knowledge necessary for successful job performance. Information collected from the task list of the Preliminary Background Questionnaire was helpful in accomplishing this.

About 200 multiple choice items were written. The majority of questions covered the work in the cartographic area, which is the entry-level for this job. The largest number of technicians in the sample to be tested worked in this specialty. About half as many photogrammetric items and about 10 triangulation items were written. ETS staff members edited all the items, randomized the alternatives, and arranged the items into subject matter categories within the three major areas.

The items were then reviewed by three additional cartographers at U. S. Army TOPOCOM. The accuracy of the stem, keyed answer, and inaccuracy of the distractors were checked for each item. Any items felt to be ambiguous were either clarified or omitted, and those where the correct answer could not be agreed upon were eliminated. About 15 new items were written in several areas not adequately covered. The resulting 165 items were put into two test forms and the items divided by subcategory within the major areas of cartography, photogrammetry, and triangulation. A description of the pretesting and development of the final form appears in Section F of this chapter.

Work Samples

Perhaps the most difficult aspect of the project was the construction of an adequate work sample for each of the three occupations, to ensure that the measures were unbiased insofar as possible, that the method of administration was fair to all, and that the tasks selected closely approximated what was being done on the job. Although the development of work samples is costly and administration is time-consuming for both the administrator and the persons taking the test, the Advisory Committee felt that inclusion of such criterion measures was of vital importance in view of the shortcomings of criteria usually used in validity studies.

During the preliminary job analyses, it became evident that many deterrents to validity and reliability stood in the way of developing acceptable measuring instruments. A given job classification is often divided into specialties. Although emphasis is placed on personnel being able to work in more than one specialty, this is not always the case. Especially in the larger installations, many incumbents have tended to develop more skills in one area than in another. Also, a wide variety of tasks within a given specialty make a representative sampling extremely difficult.

Some of the problems that had to be considered were: How representative of the total job is the particular task chosen for the work sample? Is the difficulty level appropriate? Is the task performed the same way at all installations? How much variation is there in equipment needed to perform the task? Does the final product adequately reflect the skills and knowledge employed? What is the optimum testing time that can be allocated to this test?

In developing the work samples for the Medical Technicians, the following guidelines were employed:

(1) limit maximum testing time to 2 to 3 hours.

(2) use standardized specimens having known values, against which to evaluate laboratory test results.

(3) select laboratory tests that tap a variety of task components, and are relevant for technicians in most specialties.

(4) minimize variance due to between-laboratory differences such as available equipment and favored procedures.

The work sample selected for Medical Technicians was a laboratory simulation in which technicians were required to conduct tests (an alkaline phosphatase and a differential cell count) on standardized specimens. These were decided on after extensive consultation with medical laboratory personnel in Veterans Administration and civilian hospitals, and with representatives of pharmaceutical companies engaged in providing quality control materials and services to their laboratories.

The difficulties encountered in the attempt to collect adequate work sample data for the Medical Technicians proved to be greater then anticipated. Both tasks were pretested before the main administration. The pretest appeared to proceed smoothly, and a decision was made to go ahead. However, during the main administration, many unforeseen problems in preparation and implementation arose. The sample of incumbents was very small (N = 34-40) and was chosen from two specialties, chemistry and hematology. Test administrators met with some initial resistance due to lack of information regarding the purpose of the research, but when the intent was clarified subjects were cooperative. Working conditions were very crowded. It was necessary to use some reagents and equipment from the hospitals, causing problems in standardization of specimens. There was too much variance in quality of reagents, equipment, and general physical setting of the laboratories.

The tasks proved to be new, unrepresentative, and too difficult for the majority of subjects. The less competent took a long time to complete the tasks, and asked many questions. It was felt that the results were unreliable and should not be included in analysis of other data. The experience was not in vain. Much was learned that enabled the researchers to design work samples for the other occupations studied which would fulfill their criterion role reliably.

In devising the work sample tasks for the Cartographic Technicians, the aid of the directors of the three U. S. Army TOPOCOM divisions for each mapping specialty was enlisted. They and members of their staff met with ETS representatives and discussed possible criteria and the kinds of work assignments that might be used. A number of tasks was considered, from which three were selected: a logical contouring problem, a "pull-up" from an aerial photograph, and a geometric restitution task. These were chosen because it was felt that they closely approximated what Cartographic Technicians actually do on the job, were familiar to all technicians, could be administered in a standard fashion, and had characteristics that would allow them to be reliably scored. Also, to assure subjects that their own supervisors would not see their work, the selected tasks could be administered away from their work locations and scored without using U. S. Army TOPOCOM staff. Each of the three tasks selected required about an hour to finish.

Because of the complexity of the tasks, a comprehensive scoring method was necessary. A brief description of each task and criteria for scoring follow:

Logical contouring - given the drainage pattern and spot elevations on a plastic master sheet, technicians were asked to compile contour lines at 20-foot intervals on a plastic overlay. Each of the three hilltops depicted on the map was scored for size, location, and number of contour lines. Correct placement of contour lines at the upper right and lower left-hand corners of the area shown was credited. A contour line indicating a 500-foot elevation had to be identified and run continuously through a number of specific points on the map. Contour lines had to be evenly spaced throughout.

Pull-up - subjects were asked to extract the drainage system and cultural detail from one of three vertical aerial photographs, using a hand stereoscope. This was also compiled on a plastic overlay. Scores were based on proper placement and identification of power lines, roads, railroads, ponds, towns and buildings, intersections, and drainage patterns.

Geometric restitution - this task required the technicians to reconstitute information given on an oblique photograph to a vertical format, using a blank plastic worksheet and a small triangle to compile the new details. This was scored by crediting corner ticks, parallelness and accuracy of placement of roads and railroads, and placement, shape, and size of buildings.

Scorers were given extensive preliminary training in scoring methods for the Work Sample tasks. Fifty protocols from each of the tasks were independently evaluated by two scorers as a reliability check.

During the preliminary study of the job of Inventory Management Specialist, it was found that the general activities in inventory management at the different job levels and the knowledges and abilities needed to do the work apply across agencies, despite the fact that wide variations occur in specific procedures. Although there was a number of constraints associated with the selection of a work sample which would be appropriate across installations, a common communication system was found to exist which had been developed to expedite procurement, supply, and control of inventory items within and across agencies of the Department of Defense. This is called the Military Standard Requisitioning and Issue Procedure (MILSTRIP).

It was decided to develop a work sample utilizing the "in-basket" technique (Frederiksen, 1957; Hemphill, Griffiths, & Frederiksen, 1962; Crooks, 1968) to simulate the general and procedural framework of inventory management and the role of an inventory manager. In such a test, the person taking the test is provided with instructions and selected background information. He is asked to work on a set of problems in a specific time period as if he were indeed in the job described in the instructions; that is, actually to make and communicate decisions, to give directions, to respond to requests from others, to ask for information needed to solve problems, to consider priorities in planning and scheduling work, etc. All action taken or planned must be written down in the form of communications to others or notes to self.

In developing the Inventory Management Specialist Work Sample, it was decided that the setting must be unique; that is, different from any existing agency. An hypothetical agency, U. S. Aerospace Research and Development Administration, was created, as of July 1, 1972. The new Inventory Management Specialist working for this agency was to manage items concerned with life support of crews of two space stations (food, clothing, medical items, spare parts for the environmental and waste disposal system, and some items with a repair cycle). The set of 82 items to be managed was somewhat representative of the range of items managed across present agencies, although specifically described for space station use. Forms and stationery were adapted from samples obtained on visits to installations. By establishing a new and different setting, it was possible to extract from the extensive MILSTRIP coding and other procedures only those elements necessary to work on the set of 19 problems provided. Some of the problems were related in content, and information in both background material and other problems could be utilized in taking action. An Action Form was developed which subjects were asked to complete at the end of working time on the problems. On this form, they were to describe briefly the action taken and reasons for their action.

The Work Samples obtained in the pretesting were used in developing the scoring procedure, following an approach described by Crooks (1968, 1972). Scoring dimensions are based on administrative skills and abilities disclosed as important to performance in early study of the inventory management job, and to relevant scoring dimensions used for other "in-basket" exercises. The ten dimensions fall into five general categories: (1) to describe <u>what kind</u> of action was taken, (2) to describe <u>how</u> action was taken, (3) the amount of work accomplished (productivity), (4) the quality of the action taken (appropriateness), and (5) overall effectiveness of performance in the assumed role.

D. Measurement or Description of Background Variables

A necessary step in designing the instrumentation was the inclusion of a personal history questionnaire to determine the composition of the work force. Since it was expected that there would be wide variation in background characteristics of subgroups, some differences in test and job performance for a given Civil Service grade level in a particular occupation might appear to be related to ethnic group membership when they actually reflected dissimilarities in education and background. Unless the background characteristics of majority and minority subgroups were analyzed, a study of the job in question would not yield interpretable results.

Subject matter for the questionnaires was gathered by many personal interviews with incumbents, supervisors and administrative staff members, and, for the Medical Technician study, a search of hospital personnel records. Form and content of the questionnaires were also influenced by a survey of Civil Service Personnel Research Questionnaires, the Student Information Blank from Project TALENT (Flanagan, 1964), Glennon and Albright's Catalog of Life History Items (1966), and similar sources.

Questionnaires were designed to furnish what was judged to be the most pertinent background information: age, sex, length of government service, socioeconomic level, amount of education and where received, nature and length of work experience, job held immediately prior to current position classification, U. S. Civil Service grade level at entry into occupation, and whether test performance had been a basis for obtaining the job. Also covered were the comparative roles of experience and formal training as determiners of Civil Service grade level achieved, and the time taken to achieve it. Many of the questionnaire items were identical for all three occupations studied. Questions dealing with type of work done, however, were different for each job.

The questionnaire developed for the feasibility study was deliberately overinclusive, to reduce the likelihood of overlooking any background variables which might prove relevant in moderating the prediction of job success. With the addition of a third ethnic group in the second and third occupations studied, a number of questions relating to English language facility were also introduced.

A brief preliminary questionnaire was mailed to about 1500 Cartographic Technicians in August, 1968. This number included technicians in U. S. Army TOPOCOM, Coast & Geodetic Survey, Bureau of the Census, and Army Corps of Engineers. Results from these preliminary questionnaires played a large part in the decision to select only those technicians working at U. S. Army TOPOCOM for the main study. Subjects from all three occupations were given the Personal History Questionnaire at the time of the main test administration.

E. Measurement or Description of Occupation by Job Elements or Task List

For all three occupations, incumbents were asked to complete a task information checklist as an addendum to the Personal History Questionnaire. Broadly, its function was to give a systematic check on the researchers' impressions of what personnel in various subgroupings actually do. More specifically, it represented an effort to determine which tasks in a given job classification are performed most frequently, whether there are differences across agencies or grade levels, which tasks are performed in some installations or specialties and not in others, and whether there are differences in the kinds of tasks performed by the three ethnic subgroups studied.

Items for the Medical Technicians' checklist were selected from those developed in a study conducted by Morsh and Christal at the Air Force Personnel Research Laboratory (1966). The task list for Cartographic Technicians was compiled from job specifications, personal observations, and interviews with technicians and supervisors, and was administered by mail as part of the preliminary Personal History Questionnaire, previously described. The Inventory Management Specialist task list was developed from interviews with incumbents and their supervisors, from job descriptions obtained during visits to agencies, from Civil Service Position Classification Standards, and from the Daily Activity Log, previously described.

The task list for Medical Technicians contained 58 items, for Cartographic Technicians, 130 items, and for Inventory Management Specialists, 184 items, reflecting the relative complexity of the jobs. For each task listed, the Medical Technicians and Cartographic Technicians were instructed to check "Often," "Sometimes," "Seldom," or "Never," to indicate how often they performed the given task. The Inventory Management Specialists recorded the frequency with which they performed a task as follows: "5 - Significant part of my job every day," "4 - Substantial part of my job, at least several times a week," "3 - Part of my job, probably once a week or twice a month," "2 - Part of my job, but seldom have to do this, perhaps every month or at some regular time of year," "1 - Not a regularly assigned responsibility, but I sometimes do this," "0 - Definitely not part of my job; does not apply."

F. Pretesting of Materials

For all three phases of the study, newly developed instruments were pretested to determine optimum order of presentation, timing, placement of rest breaks, clarity of instructions, and potential problem areas.

The aptitude test battery and peer rating scales for Medical Technicians were administered to all available subjects at the Veterans Hospitals in Wilkes-Barre, Pennsylvania, and East Orange, New Jersey, prior to the full-scale testing program. Work Sample tasks were pretested at the Lyons, New Jersey, VA hospital.

One hundred forty-seven Job Knowledge Test items were pretested on a total of 26 technicians at two VA hospitals. From timing information gained in these two test administrations, a seventy-five item test was assembled, with only the first sixty-three items to be scored. The remaining twelve items were included as a "filler" to occupy those technicians who completed the test most quickly.

The 165 items written for the Cartographic Technicians Job Knowledge Test were divided into two test forms to shorten time needed for administration. These were pretested at the Providence, Rhode Island, field office of U. S. Army TOPOCOM. Each form was given to 40 Cartographic Technicians, about half of whom were currently working in the cartographic division and half in the photogrammetric division. Based on how well the items discriminated between the high- and low-scoring groups and the total number of technicians getting an item correct, approximately half of all the items contained in both forms were then chosen for the 75-item final form. Items dealing with the triangulation specialty were dropped, since the number of incumbents currently working in that specialty was small.

After the three Cartographic Technician Work Sample tasks had been selected, each was tried out on a small sample of technicians at U. S. Army TOPOCOM in Bethesda, Maryland, and Providence, Rhode Island, to verify its appropriateness and to determine the range of time required for completion. Each of the three tasks selected required about one hour to finish.

The Inventory Management Specialist Work Sample was pretested in preliminary form at the Defense Personnel Support Center in Philadelphia. It was administered to 49 Inventory Management Specialists from the Medical, Subsistence, and Clothing Directorates, and included Black and Caucasian male and female subjects at the GS-9 and -11 salary grade levels. The original time allotted for pretesting was three hours. As a result of the administration, instructions were changed and augmented, and information in several problems was made more specific. The actual working time on the problems was increased from two hours to two and one-half hours, making the total testing time for the work sample exercise three and one-half hours.

The Criterion Rating Scales were administered to supervisors of the Inventory Management Specialists in the pretesting sample. After each supervisor completed his ratings, another task was presented. For each Inventory Management Specialist he rated, he was asked to select the <u>one</u> trait he felt was most important to that person's job. This was not necessarily to be each manager's strongest trait, but rather the one most necessary for his success on <u>his particular</u> job.

As a result of the pretesting, this task was dropped as being difficult for the supervisors to do and not very meaningful in terms of the actual ratings. The Importance Rating Scale, previously described, was given instead in the main administration.

G.. Description of Samples Obtained and Data Gathered

From the pretesting of materials, useful observations were made for testing practices which were incorporated in the main administration. The following general procedures were followed for data collection:

Testing was done on site. All tests, rating scales, and questionnaires were administered by specially trained ETS staff members. This was felt to be necessary in order to maintain the confidential nature of the data, to see that standard directions were followed, and to ensure that maximum rapport was established.

Before the main test administration for each of the three occupations, a letter was sent from the central agency office in Washington to the director of each installation selected, explaining briefly the nature and purpose of the study and the respective roles of ETS and the U. S. Civil Service Commission, and expressing general support for the project. A suggested letter for notifying participating incumbents was sent to each installation, to go out over the director's signature in advance of the actual testing session. In addition to outlining the purpose of the study and the nature of their participation, it emphasized that the research was not identified in any way with formal agency personnel records or practices, and that testing materials and individual results would be handled directly by ETS representatives on a confidential basis, with a guarantee that no information about any individual would be made known to any government official or employee and that only <u>overall</u> results would be reported. Full sanction was to be given to the research effort, including time taken off the job to participate. Although the addressees' personal cooperation and support were solicited, freedom not to comply was to be given.

Every effort was made by ETS researchers to preserve the confidentiality of the data and to communicate this effort to the incumbents being tested. Tight security precautions were taken with both used and unused tests and rating scales. At the testing sites, boxes were stored in locked rooms, and considerable precaution was taken not to leave the materials unattended during an administration. When tests or ratings were completed, subjects personally placed all materials in individual envelopes and sealed them. ETS researchers were solely responsible for removal of testing materials from the premises.

For the Medical Technician study, a group of 33 of the 170 Veterans Administration hospitals was chosen. Most were selected because they employed large numbers of Blacks in their laboratories, a few because they were entirely or predominantly Caucasian, and about five because they were small. An effort was made to include hospitals in various geographic locations. It was hoped to test all Medical Technicians (GS-645) at these hospitals. Tallies indicated that there were 774 technicians, including 261 Blacks. Of these, data were obtained on 465, of whom 168 were Black. The greatest attrition was due to the fact that technicians listed as working in particular hospitals actually worked in affiliated outpatient clinics, research groups, and the like. Second largest cause of attrition was that many of the technicians worked only part time or were on evening hours, and could not reasonably be included in the test sessions. The remaining attrition was due to such expected reasons as vacations, illnesses, being needed in the laboratory, and, in only a few cases, individual refusal to participate in the study.

Total time allotted for the administration of the Aptitude and Job Knowledge tests, Personal History Questionnaire, and Peer Rating Scales, was approximately six hours per technician, with another 2 - 3 hours needed for those participating in the Work Sample tasks. Most researchers took a full week to gather data from one hospital, since the technicians were generally available for testing during the afternoons only, because of morning time pressures on their jobs. Often it was possible to test only about half the personnel from a laboratory at one time, since no more could be spared from their duties.

For the Cartographic Technician study, a sample of approximately 440 was tested from 1,000 in the GS-1371 classification at four U. S. Army TOPOCOM locations. Technicians were selected from those who had filled out the Preliminary Personal History Questionnaire, plus a few newer employees. The total sample consisted of 101 Black, 99 Mexican-American, and 241 Caucasian technicians, and reflected the proportions of technicians working in each of the three major specialties, cartography, photogrammetry, and triangulation.

All available technicians at San Antonio (where nearly all of the Mexican-American technicians were employed) were tested. At the other three locations, all Black technicians available during the testing period were included in the sample. Caucasian technicians were selected who were similar in GS-level to that of the Black technicians, and at a ratio of about two Caucasians for every Black. The resulting sample at the four locations included 167 technicians of 564 employed at Washington, 77 of 146 at Kansas City, 75 of 143 at Louisville, and 122 of 130 at San Antonio.

Time required to administer the aptitude battery, Technical Knowledge Questionnaire, and the Personal History Questionnaire, was about five hours. Each Work Sample task required one hour, making the total testing time about eight hours. The Work Sample was administered on a half-day following the other tests.

A sample of Cartographic Technicians at the Coast & Geodetic Survey was used in a supplemental study. Subjects were given all of the aptitude tests and the Personal History Questionnaire, and supervisors' ratings were obtained, but the Technical Knowledge Questionnaire and the Work Sample were not administered. Most of the work in this agency involves the production of aeronautical and navigational charts, and supervisors, after careful review, advised that these measures were not appropriate for their installation.

For the Coast & Geodetic Survey sample in Silver Spring, Maryland, all available Cartographic Technicians were tested, a total of 98. This number included nine Caucasian deaf-mutes. This group possibly had difficulty in understanding the directions given for the tests, although the directions were interpreted in sign language for them. A preliminary analysis showed that their mean test scores were substantially lower than the means for the others tested, so this group was excluded from further analyses.

For the Inventory Management Specialists, plans were made to test approximately 200 Blacks, 100 Mexican-Americans, and 300 Caucasians at five locations. The sample obtained was smaller than anticipated, as shown in Table II-5 on page 43. When the testing began, much greater resistance was encountered than in the previous phases, although every effort was made in advance contacts to provide information about the project and to encourage participation.

A decision was made to test primarily at grade levels 9 and 11, the journeyman levels in inventory management/after progress through the GS-5 and -7 training periods. (Entry into the 2010 classification is at grade 5, with progress to grade 7 and then to grade 9 within a prescribed period, subject to satisfactory performance.) A few inventory managers in GS-7 were included in order to increase the ethnic samples.

Time required to administer the ~~titude battery and the Personal History Questionnaire was about six and one-quarter hours. The Work Sample exercise, which was administered on a half-day following the other tests, required three and one-half hours, making the total testing time about ten hours.

Table II-1 shows the number, and corresponding percentages of the total samples, of Medical Technicians tested at each grade level by ethnic group, and total number of subjects tested.

Table II-1

Distribution of Medical Technicians Tested, by Grade Level and Ethnic Group

Grade Level	Black	%	Caucasian	%	Total
GS-4 and below	20	11.9	29	9.8	49
5	46	27.4	71	23.9	117
6	61	36.3	123	41.4	184
7	35	20.8	58	19.5	93
8 and above	6	3.6	16	5.4	22
Total	168		297		465

Table II-2 describes the sample of Cartographic Technicians tested, by location, ethnic group, and total.

Table II-2

Distribution of the Sample of Cartographic Technicians Tested, by Ethnic Group and Total

U. S. Army TOPOCOM	Black	Mexican-American	Caucasian	Total
Washington	53	0	114	167
Kansas City	21	0	56	77
Louisville	25	0	50	75
San Antonio	2	99	21	122
Total	101	99	241	441
Coast & Geodetic Survey				
Silver Spring	38	0	60*	98*

* Includes 9 deaf-mutes, excluded from the comparative analyses

Table II-3 shows the distribution of Cartographic Technicians tested by grade level, percentages by ethnic group, and total number of subjects tested, at all TOPOCOM installations included in the study.

Table II-3

Distribution of Cartographic Technicians Tested at U. S. Army TOPOCOM Installations, by Grade Level and Ethnic Group

Grade Level	Black	%	Mexican-American	%	Caucasian	%	Total
GS-5	0		0		12	5.0	12
6	0		0		1	.4	1
7	32	32.3	17	17.3	56	23.4	105
8	10	10.1	0		19	8.0	29
9	51	51.5	81	82.6	131	54.8	263
10	0		0		1	.4	1
11	5	5.0	0		19	8.0	24
12	1	1.0	0		0		1
Total	99		98		239		436

The numbers in Tables II-2 and -3 differ slightly because of missing data.

Table II-4 shows the distribution of Cartographic Technicians tested by grade level, percentages by ethnic group, and total number of subjects tested, at the Coast & Geodetic Survey.

Table II-4

Distribution of Cartographic Technicians Tested

at the Coast & Geodetic Survey,

by Grade Level and Ethnic Group

Grade Level	Black	%	Caucasian	%	Total
GS-2	0		2	3.3	2
3	1	2.6	6	10.0	7
4	0		5	8.3	5
5	1	2.6	7	11.7	8
6	10	26.3	9	15.0	19
7	9	23.7	5	8.3	14
9	13	34.2	18	30.0	31
11	4	10.5	8	13.3	12
Total	38		60		98

Table II-5 describes the sample of Inventory Management Specialists tested, by location, ethnic group, and total.

Table II-5

Description of the Sample of Inventory Management Specialists Tested, by Ethnic Group and Total

Installation	Black	Mexican-American	Caucasian	Total
Army Tank-Automotive Command, Detroit	42	0	49	91
Defense Electronic Supply Center, Dayton	36	0	44	80
Defense Industrial Supply Center, Philadelphia	14	0	31	45
Navy Aviation Supply Office, Philadelphia	6	0	15	21
San Antonio Air Materiel Area, San Antonio (Air Force)	18	75	75	168
Total	116	75	214	405

Table II-6 shows the distribution of Inventory Management Specialists tested by grade level, percentages by ethnic group, and total number of subjects tested.

Table II-6

Distribution of Inventory Management Specialists Tested, by Grade Level and Ethnic Group

Grade Level	Black	%	Mexican-American	%	Caucasian	%	Total
GS-7	7	6.1	14	18.9	8	3.9	29
9	76	66.7	50	67.6	136	66.3	262
11	24	21.0	6	8.1	41	20.0	71
Other (or missing data)	7	6.1	4	5.4	20	9.8	31
Total	114		74		205		393

The grade level data in Table II-6 are taken from the Personal History Questionnaire, not completed by a small number of subjects, which explains the difference in the Ns in the two tables.

H. Reporting Individual Test Results to Subjects

Individual test scores were not made available to participants in the Medical Technicians study. Inventory Management Specialists were given the option of receiving their scores on the aptitude tests and Cartographic Technicians on the aptitude tests and Job Knowledge Test. Supervisors' Ratings and Work Sample Test scores were not reported to any individuals.

A letter outlining factors that might influence test scores and giving a brief description of each test and ~~what it~~ was intended to measure accompanied the individual test results. Score reports were sent to home addresses given by those who elected to receive a report as a guarantee of confidentiality.

Scores were reported as percentiles, based on the total sample tested at all installations. An interpretation of percentile rankings was given, using the following guidelines:

Percentile	Interpretation
15 and below	Low
16 - 30	Below Average
31 - 69	Average
70 - 84	Above Average
85 and above	High

Chapter III

Comparison of Ethnic Groups on Aptitude Measures

The aptitude and ability measures selected as the predictor batteries for the three occupations studied are described in detail in the instrumentation reports previously cited (Pike, 1969; Parry, 1971; Crooks and Mahoney, 1971). In this chapter, the measures will be described briefly in relation to factors observed as directly related to job performance in the three occupations and to other factors of general research interest, as background for discussion of the results. The means and standard deviations of the scores obtained on the measures will be compared by ethnic group for each of the three studies.

Medical Technicians

Nine tests were selected to comprise the predictor battery for study of the occupation of Medical Technician:

Factor	Name of Test
Number Facility	Subtraction & Multiplication
Facility in performing basic arithmetic computations	
Perceptual Speed	Number Comparison
Accurate performance of clerical tasks under time pressure	
Flexibility of Closure	Hidden Figures
Ability to identify known configurations when they are presented with perceptual distractions	
Speed of Closure	Gestalt Completion
Ability to organize and recognize an apparently disparate field as a single percept	

Both Flexibility of Closure and Speed of Closure are suggested as related to examination of laboratory specimens.

Factor	Name of Test
Visualization	Paper Folding

Related to mechanical ability and comprehension in setting up, calibrating, adjusting, and maintaining laboratory equipment

Fine Finger Dexterity	Pin-Dexterity (USCSC Test No. 26)

Skill at hand manipulation of instruments and materials under time pressure

Verbal Comprehension	Extended Range Vocabulary

Ability related to successful performance in a wide range of jobs (tests of this factor have been criticized as biased against minorities)

Associative (Rote) Memory	Picture-Number

A general ability to be explored as related to job performance of Medical Technicians

General Reasoning	Necessary Arithmetic Operations

A general ability factor to be explored as related to job performance

Ethnic Group Comparisons (Black - Caucasian)

Table III-1 shows the means and standard deviations of these measures for the Medical Technicians by ethnic group (Black and Caucasian). On all measures, the Black group scored lower on the average. On all but one of the nine measures, mean score differences were significant at the .01 level. On the other test, Pin-Dexterity (measuring fine finger dexterity), the mean difference was significant at the .05 level. It will be seen that all but two differences were approximately one-half standard deviation in size. The mean difference on the Picture-Number Test (a measure of rote, short-term memory) was less than one half standard deviation. The mean difference on the Necessary Arithmetic

Operations Test (a test of general reasoning) was approximately one standard deviation.

The intercorrelations of the tests in the predictor battery are shown in Appendix Table III-A, by ethnic group, if of interest. Some difference in size of correlations can be observed between the two ethnic groups, which should be kept in mind in considering the results of further analyses described in later chapters.

Cartographic Technicians

Thirteen tests were selected to comprise the predictor battery for the study of the occupation of Cartographic Technician (map maker):

Factor	Name of Test
Hand-Arm Movement Coordination	Flanagan Coordination
Facility in coordinating hand-arm movements in drawing and drafting	
Flexibility of Closure	Hidden Figures
Ability to extract detail from aerial photographs of terrain with indistinct or distracting background	
Verbal Comprehension	Vocabulary (speeded)
	Extended Range Vocabulary (unspeeded)
Ability to understand written instructions and specifications	
Associative (Rote) Memory	Object-Number
Ability to form and remember new associations; memory for specifications	
Spatial Orientation	Card Rotations
Ability to perceive spatial patterns and maintain orientation of objects in space	

Factor	Name of Test
Number Facility	Arithmetic (USCSC Test No. 24)

Ability to make numerical computations (i.e., measuring distance to scale)

Spatial Scanning	Map Planning (speeded)
	Maze Tracing Speed (unspeeded)

Ability to explore visually a wide or complicated spatial field

Visualization	Surface Development

Ability to manipulate the image of spatial patterns into other visual arrangements

Following Oral Directions	Following Oral Directions (USCSC Test No. 135)

Ability to carry out simple and complex instructions accurately when given orally and under time pressure

Perceptual Speed	Identical Pictures

Ability to carry out close, accurate visual tasks under time pressure

General Reasoning	Necessary Arithmetic Operations

A general reasoning ability found to be important to performance in many jobs

The results for the two groups of Cartographic Technicians (TOPOCOM and Coast & Geodetic Survey) are shown and discussed separately. The work performed by the technicians in these agencies differs considerably, and it was felt that the results should be analyzed independently. Table III-2 shows the means and standard deviations of the scores on the predictor battery for the TOPOCOM Cartographic Technicians by ethnic group (Black, Mexican-American, and Caucasian). Table III-3 shows the results for the Coast & Geodetic technicians (Black and Caucasian only).

Ethnic Group Comparisons (Black - Caucasian)

In comparing the mean scores of Blacks and Caucasians for the TOPOCOM sample, it will be seen that mean score differences were significant at the .01 level on seven of the thirteen tests, and at the .05 level on four of the remaining tests. The mean scores of the Blacks were in all cases lower. On most of these tests, the differences ranged from one-fourth to one-half standard deviation in size. On two tests, Vocabulary and Extended Range Vocabulary, the mean differences were small and nonsignificant.

In comparing the mean scores of Blacks and Caucasians for the Coast & Geodetic Survey sample, it will be seen that mean score differences were significant at the .01 level on eleven of the thirteen tests and at the .05 level on the other two. The Black sample scored lower on all tests. The mean differences ranged from one-half standard deviation to one standard deviation in size. The largest differences were on Vocabulary, CS Arithmetic, Following Oral Directions, and Necessary Arithmetic Operations, all approximating one standard deviation in size.

Some differences in the two samples of Black Cartographic Technicians are apparent. The Coast & Geodetic Survey sample scored lower on the average on every test than the TOPOCOM sample, while the Caucasian sample on the average scored approximately at the same level. Differences in background variables such as level of education and experience for the two Black samples may account for some of the variation.

In selecting aptitude tests for the study of Cartographic Technicians, it was suggested that score differences between ethnic groups usually found in testing situations might be attributable at least in part to "testwiseness" or practice in taking tests in favor of the majority group, particularly advantageous to this group under speeded conditions. To test the validity of this suggestion, it was decided to select two pairs of tests measuring common factors, one of each pair to be administered under regular, speeded conditions and the other to be made essentially unspeeded by extending the usual time limit. One such pair was chosen under the Spatial Scanning factor, with the Map Planning Test given as a speeded test and the Maze Tracing Speed Test given as essentially unspeeded. The relative mean score differences on these two tests for the TOPOCOM Black and Caucasian samples do not show any notable change under speeded and unspeeded conditions. The mean score differences and the standard deviations are almost identical for the two groups on the two tests.

The other pair of such tests measures a Verbal Comprehension factor. The Vocabulary Test was given under speeded conditions and the Extended Range Vocabulary Test was given as essentially unspeeded by extending the time limit. As seen from Table III-3, the mean scores and standard deviations were not significantly different under speeded and unspeeded conditions.

For the Coast & Geodetic Survey sample, the mean scores of the Black sample on the tests given under speeded and unspeeded conditions. for the Spatial Scanning factor do not appear to differ in relationship to the mean scores of the Caucasian sample (Table III-4). Each set of scores differs by more than one-half standard deviation. On the Vocabulary (speeded) and Extended Range Vocabulary (unspeeded)

Tests, the mean score under the unspeeded condition appears to favor the Black sample. The mean score of the Black sample under the speeded condition was slightly more than one standard deviation below the Caucasian mean score, while on the Extended Range Vocabulary Test, essentially unspeeded, the mean difference was closer to one-half standard deviation.

Ethnic Group Comparisons (Mexican-American - Caucasian)

In comparing the mean scores and standard deviations of Mexican-American and Caucasian technicians employed by TOPOCOM, it will be seen that the mean score for the Mexican-American sample was higher than for the Caucasian sample on one test of the thirteen in the battery, the Object-Number Test (measuring the ability to form and remember new associations), although this difference was small and not significant. On one other test, Identical Pictures (measuring the ability to carry out close, accurate visual tasks), the mean difference, favoring the Caucasian sample, was also small and not significant. The mean score differences on the remaining eleven tests were significant at the .01 level. In all cases the Caucasian mean was higher. On eight of the tests, the differences in mean scores ranged in magnitude from about one-third to one-half standard deviation. On three of the tests, Following Oral Directions, Extended Range Vocabulary, and Necessary Arithmetic Operations, the size of the mean score differences was approximately one standard deviation.

On these three tests, the Mexican-Americans may have scored lower because they are less fluent in the English language. One hundred percent of the Mexican-Americans reported on the Personal History Questionnaire that a foreign language was spoken at home (50% said "almost all of the time"). The Mexican-Americans also scored lower on the average than the Black Cartographic Technicians on both Vocabulary tests, as well as on the Following Oral Directions Test (where quick listening comprehension is needed), although they scored as well or better on most of the other tests not involving language to as great a degree.

On the paired speeded-unspeeded tests, the speeded condition appeared to make no appreciable difference in the scores on the Spatial Scanning factor. On the Extended Range Vocabulary Test, given under unspeeded conditions, the mean score of the Mexican-American sample was somewhat lower than on the speeded Vocabulary Test. This probably reflects the wider range of the difficulty of the Extended Range Vocabulary Test, a handicap to those less fluent in English which would not be overcome by extending the time limit.

Intercorrelations of the predictor battery for the TOPOCOM and Coast & Geodetic Survey samples are included in the appendix as Appendix Tables III-B and III-C. The correlations of the speeded and unspeeded tests for the Spatial Scanning factor range from .43 to .61, and on the Verbal Comprehension factor from .77 to .86 for the separate samples and ethnic groups.

Inventory Management Specialists

Twelve aptitude and ability measures were selected for the predictor battery for the occupation of Inventory Management Specialist:

Factor	Name of Test
Perceptual Speed	Number Comparison

Speed and accuracy in comparing sets of numbers and identifying differences

Factor	Name of Test
Flexibility of Closure	Hidden Figures

Ability to identify known configurations when they are presented with perceptual distractions, suggested as a measure of analytical functioning

Verbal Comprehension	Vocabulary	(speeded)
	Extended Range Vocabulary	(unspeeded)

Ability to understand the English language (Inventory Managers issue and receive both oral and written instructions of varying complexity)

Associative (Rote) Memory	Object-Number

Ability to remember bits of unrelated material (memory for large amounts of information is important to Inventory Managers)

Induction	Letter-Sets

Ability to form and try out hypotheses to find general concepts to fit sets of data

Syllogistic Reasoning	Inference	(speeded)
	Nonsense Syllogisms	(unspeeded)

Ability to reason from stated premises to their necessary conclusions

Number Facility	Subtraction & Multiplication

Ability to do arithmetical operations with speed and accuracy

General Reasoning	Necessary Arithmetic Operations

General reasoning ability, related to performance on a wide range of jobs

Following Oral Directions	Following Oral Directions (USCSC Test No. 135)

Ability to carry out simple and complex instructions accurately when given orally and under time pressure

Factor	Name of Test
General Reasoning	Portions of the Federal Service Entrance Examination (USCSC Test No. 170)

Verbal ability, abstract reasoning ability, and numerical reasoning

Table III-4 shows the means and standard deviations of the aptitude test scores for the Inventory Management Specialists by ethnic group (Black, Mexican-American, and Caucasian).

Ethnic Group Comparisons (Black - Caucasian)

The mean score differences on nine of the 12 tests in the battery were significant at the .01 level, approximately one-half standard deviation in size. On one test, Necessary Arithmetic Operations (a test of general reasoning), the difference approached one standard deviation. For the other three tests, Number Comparison, Object-Number, and Nonsense Syllogisms, the mean differences were small and nonsignificant. However, the mean scores of the Black sample were lower on all tests.

For the speeded pair of tests under the Syllogistic Reasoning factor (Inference and Nonsense Syllogisms), the speeded condition appeared to make a difference, with the Black sample scoring significantly lower on the average than the Caucasian sample on the Inference Test, while on the Nonsense Syllogisms Test, given essentially unspeeded, the mean difference was small and nonsignificant.

Ethnic Group Comparisons (Mexican-American - Caucasian)

The mean scores of Mexican-American Inventory Management Specialists were higher on two of the 12 aptitude measures: the Number Comparison Test (measuring perceptual speed), where the mean difference was not significant, and the Object-Number Test (measuring associative (rote) memory), where the mean difference was significant at the .01 level. On the other ten tests, the Caucasians had higher mean scores, although the mean difference on the Hidden Figures Test was not significant. For three tests, Letter Sets, Subtraction & Multiplication, and Necessary Arithmetic Operations, the mean differences were significant at the .05 level, and for six tests, the mean differences were significant at the .01 level. The size of the significant differences ranged from one-fourth to one-half standard deviation.

On the speeded-unspeeded pairs of tests under the Syllogistic Reasoning factor and the Verbal Comprehension factor, the mean differences were about the same under both conditions. It may be noted that the Mexican-Americans were again shown to be at a disadvantage on tests involving language, as seen in the study of Cartographic Technicians.

Intercorrelations of the predictor battery for Inventory Management Specialists are included in the appendix as Appendix Table III-D. The intercorrelations of the speeded-unspeeded pairs of tests range from .30 to .56 for the Syllogistic Reasoning factor and from .73 to .88 on the Verbal Comprehension factor.

Cross-Study Comparisons

Three tests were included in the batteries for all three occupations: Necessary Arithmetic Operations, Hidden Figures, and Vocabulary. Table III-5 shows the means and standard deviations for these tests by occupation, by ethnic group. Tables III-6a and III-6b show the means and standard deviations for the five tests included in the predictor batteries for two occupations by ethnic group.

The mean scores on these tests vary somewhat across the three occupations, perhaps reflecting individual differences in aptitudes or abilities needed from occupation to occupation, background factors such as amount of education or special training required, GS level, and restriction of range as a result of preselection or self-selection. Later analyses will show how these differences relate to measures of performance.

Reliabilities

Tables III-7, III-8, and III-9 show the alternate form reliabilities of the aptitude tests and as corrected for attenuation by the Spearman-Brown formula for the occupational samples by ethnic group. In general, the reliabilities of the tests do not differ to any marked degree by ethnic group and are of acceptable size.

Summary

Black - Caucasian Comparisons

For all three occupations, Caucasians consistently scored higher on the average than Blacks on the aptitude measures. For the Medical Technician samples, all mean score differences were statistically significant. For the Cartographic Technician samples, the mean score differences of 11 of the 13 tests were statistically significant. For the Inventory Management Specialist samples, the mean score differences on nine of the 12 tests were statistically significant. Differences ranged from one-fourth to one-half standard deviation on most tests, with the exception of the Necessary Arithmetic Operations Test (a test of general reasoning). On this test, for all three occupations, the differences approached one standard deviation.

The differences in the test scores of the three samples of Blacks and Caucasians are smaller t found in other populations, and it is suggested that these samples represent selected groups of Blacks and/ or Caucasians. The differences in the test results vary among the three samples, also suggesting that the special aptitudes or skills needed for the three occupations may result in greater selectivity among the Blacks for some jobs than for others. There was little, if any, explicit pre-employment selection on tests for the Medical Technicians and Cartographic Technicians. Thus, the fact that these groups appear to be somewhat selected is incidental to the other selection procedures or to retention factors on the job, or both.

Mexican-American - Caucasian Comparisons

For the two occupations in which Mexican-Americans are represented, Caucasians scored higher on the average than Mexican-Americans on most of the aptitude tests. However, the Mexican-Americans scored higher on the average than the Caucasians on tests of rote memory and perceptual speed. The other differences between Mexican-Americans and Caucasians, where significant, were of about the same magnitude as for Blacks and Caucasians, with the exception of tests involving vocabulary, where differences approached one standard deviation. Blacks also scored higher than Mexican-Americans on these tests.

Note: Differences between mean scores of ethnic groups on aptitude tests and on selected criterion measures are shown graphically in Chapter IV in Figures IV-1, IV-2, IV-3, and IV-4, in terms of standard deviation units from the Caucasian mean.

Table III-1

Means and Standard Deviations of

Aptitude Test Scores by Ethnic Group

Medical Technicians

	Black N=168		Caucasian N=297	
Test	Mean	S.D.	Mean	S.D.
Subtraction and Multiplication	46.6**	17.1	57.0	19.6
Vocabulary	19.2**	7.7	25.7	9.1
Hidden Figures	5.7**	4.0	7.8	4.9
Necessary Arithmetic Operations	10.4**	3.7	14.0	4.8
Pin-Dexterity	22.5*	14.7	29.1	15.8
Number Comparison	36.5**	9.7	42.4	9.8
Gestalt Completion	10.3**	5.2	12.1	5.6
Picture-Number	17.5**	8.8	20.8	9.4
Paper Folding	6.8**	3.3	9.1	3.8

*Significantly different from Caucasian mean at .05 level

**Significantly different from Caucasian mean at .01 level

Table III-2

Means and Standard Deviations of

Aptitude Test Scores by Ethnic Group

Cartographic Technicians (TOPOCOM)

	Black N=101		Mexican-American N=101		Caucasian N=240	
Test	Mean	S.D.	Mean	S.D.	Mean	S.D.
Coordination	5.5*	1.8	5.0**	2.0	6.0	2.1
Hidden Figures	10.5*	5.6	9.4**	5.6	11.9	6.3
Vocabulary	19.0	7.4	15.4**	6.2	19.9	7.3
Object-Number	10.0**	4.9	12.4	6.3	12.1	6.7
Card Rotations	112.7*	39.8	108.9**	41.5	124.0	41.5
CS Arithmetic	35.2**	9.1	37.2**	8.8	40.9	9.0
Map Planning	16.9*	6.9	16.5**	7.0	18.9	6.8
Surface Development	24.0**	14.7	22.8**	14.0	31.3	15.0
Maze Tracing Speed	30.1**	8.1	30.8**	8.4	34.7	8.9
Following Oral Directions	15.6**	8.2	11.0**	7.4	20.2	8.2
Identical Pictures	61.2**	13.3	62.5	12.7	65.3	13.0
Extended Range Vocabulary	18.6	9.6	12.8**	7.6	19.9	9.2
Necessary Arithmetic Operations	10.3**	5.0	10.6**	5.0	14.6	5.4

*Significantly different from Caucasian mean at .05 level

**Significantly different from Caucasian mean at .01 level

Table III-3

Means and Standard Deviations of

Aptitude Test Scores by Ethnic Group

Cartographic Technicians (Coast & Geodetic Survey)

	Black N=38		Caucasian N=51	
Test	Mean	S.D.	Mean	S.D.
Coordination	5.2*	1.9	6.1	2.2
Hidden Figures	7.6*	5.5	10.0	6.4
Vocabulary	13.9**	8.3	20.8	6.4
Object-Number	9.1**	5.9	13.9	7.3
Card Rotations	82.0**	47.4	116.4	41.3
CS Arithmetic	26.3**	10.1	38.2	11.3
Map Planning	13.3**	7.6	17.4	6.9
Surface Development	17.7**	11.7	29.9	16.3
Maze Tracing Speed	23.8**	10.6	30.9	9.8
Following Oral Directions	9.6**	7.5	18.8	9.4
Identical Pictures	51.0**	12.5	59.8	12.7
Extended Range Vocabulary	14.5**	10.2	20.0	9.4
Necessary Arithmetic Operations	7.6**	4.8	13.1	6.3

*Significantly different from Caucasian mean at .05 level

**Significantly different from Caucasian mean at .01 level

Table III-4

Means and Standard Deviations of

Aptitude Test Scores by Ethnic Group

Inventory Management Specialists

Test	Black N=112 Mean	Black N=112 S.D.	Mexican-American N=72 Mean	Mexican-American N=72 S.D.	Caucasian N=194 Mean	Caucasian N=194 S.D.
Number Comparison	39.8	11.1	43.6	10.3	40.9	11.7
Hidden Figures	5.2**	5.3	7.0	4.8	7.5	5.7
Vocabulary	22.3**	7.1	21.2**	5.9	25.6	6.4
Object-Number	10.8	6.4	14.5**	7.1	11.2	6.0
Letter Sets	12.8**	6.0	14.2*	5.5	15.8	6.3
Nonsense Syllogisms	7.7	7.3	6.3**	6.1	9.0	8.2
Subtraction and Multiplication	63.9**	19.2	70.7*	20.1	76.6	21.6
Extended Range Vocabulary	24.2**	10.1	21.0**	8.8	27.7	10.4
Necessary Arithmetic Operations	11.1**	5.6	13.7*	5.4	15.5	5.4
Following Oral Directions	14.6**	7.4	14.9**	7.4	18.1	8.5
Inference	9.4**	4.5	9.2**	4.1	11.7	4.9
Federal Service Entrance Examination	53.1**	18.8	57.6**	14.5	65.1	18.7

*Significantly different from Caucasian mean at .05 level

**Significantly different from Caucasian mean at .01 level

Table III-5

Means and Standard Deviations of Aptitude Test Scores in Predictor Batteries

for Three Occupations by Ethnic Group

Occupation	Necessary Arithmetic Operations						Hidden Figures						Vocabulary					
	Black		Mexican-American		Caucasian		Black		Mexican-American		Caucasian		Black		Mexican-American		Caucasian	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Medical Technician	10.4	3.7	--	--	14.0	4.8	5.7	4.0	--	--	7.8	4.9	19.2	7.7	--	--	25.7	9.1
Cartographic Technician																		
TOPOCOM	10.3	5.0	10.6	5.0	14.6	5.4	10.5	5.6	9.4	5.6	11.9	6.3	19.0	7.4	15.4	6.2	19.9	7.3
Coast & Geodetic	7.6	4.8	--	--	13.1	6.3	7.6	5.5	--	--	10.0	6.4	13.9	8.3	--	--	20.8	6.4
Inventory Management Specialists	11.1	5.6	13.7	5.4	15.5	5.4	5.2	5.3	7.0	4.8	7.5	5.7	22.2	7.1	21.2	5.9	25.6	6.4

Table III-6a

Means and Standard Deviations of Aptitude Test Scores in Predictor Batteries

for Three Occupations by Ethnic Group

	Object-Number						Subtraction and Multiplication						Number Comparison					
	Black		Mexican-American		Caucasian		Black		Mexican-American		Caucasian		Black		Mexican-American		Caucasian	
Occupation	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Medical Technician	--	--	--	--	--	--	46.6	17.1	--	--	57.0	19.6	36.5	9.7	--	--	42.4	9.8
Cartographic Technician																		
TOPOCOM	10.0	4.9	12.4	6.3	12.1	6.7	--	--	--	--	--	--	--	--	--	--	--	--
Coast & Geodetic	9.1	5.9	--	--	13.9	7.3	--	--	--	--	--	--	--	--	--	--	--	--
Inventory Management Specialist	10.8	6.4	14.5	7.1	11.2	6.0	63.9	19.2	70.7	20.1	76.6	21.6	39.8	11.1	43.6	10.3	40.9	11.7

Table III-6b

Means and Standard Deviations of Aptitude Test Scores in Predictor Batteries

for Three Occupations by Ethnic Group

	Extended Range Vocabulary						Following Oral Directions					
	Black		Mexican-American		Caucasian		Black		Mexican-American		Caucasian	
Occupation	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Medical Technician	--	--	--	--	--	--	--	--	--	--	--	--
Cartographic Technician												
TOPOCOM	18.6	9.6	12.8	7.6	19.9	9.2	15.6	8.2	11.0	7.4	20.2	8.2
Coast & Geodetic	14.5	10.2	--	--	20.0	9.4	9.6	7.5	--	--	18.8	9.4
Inventory Management Specialist	24.2	10.1	21.0	8.8	27.7	10.4	14.6	7.4	14.9	7.4	18.1	8.5

Table III-7

Alternate Form and Spearman-Brown Corrected Reliabilities

of Aptitude Tests by Ethnic Group

Medical Technicians

	Black		Caucasian	
	r_{AB}	r_{SB}	r_{AB}	r_{SB}
Subtraction & Multiplication	.89	.94	.90	.95
Vocabulary	.77	.87	.83	.91
Hidden Figures	.71	.83	.62	.77
Necessary Arithmetic Operations	.56	.72	.71	.83
Number Comparison	.83	.91	.82	.90
Gestalt Completion	.74	.85	.78	.88
Picture-Number	.65	.79	.70	.82
Paper Folding	.51	.68	.69	.82

Table III-8

Alternate Form and Spearman-Brown Corrected Reliabilities

of Aptitude Tests by Ethnic Group

Cartographic Technicians

	TOPOCOM						COAST & GEODETIC SURVEY			
	Black		Mexican-American		Caucasian		Black		Caucasian	
	r_{AB}	r_{SB}	r_{AB}	r_{SB}	r_{AB}	r_{SB}	r_{AB}	r_{SB}	r_{AB}	r_{SB}
Hidden Figures	.48	.65	.42	.59	.62	.77	.69	.82	.72	.84
Vocabulary	.77	.87	.74	.85	.77	.87	.78	.88	.83	.91
Object-Number	.40	.57	.63	.77	.72	.84	.56	.72	.67	.80
Card Rotations	.77	.87	.82	.90	.79	.88	.77	.87	.88	.94
Map Planning	.70	.82	.64	.78	.63	.77	.74	.85	.71	.83
Surface Development	.83	.91	.78	.88	.82	.90	.70	.82	.84	.91
Maze Tracing Speed	.87	.93	.85	.92	.89	.94	.96	.98	.94	.97
Extended Range Vocabulary	.79	.88	.76	.86	.85	.92	.79	.88	.90	.95
Necessary Arithmetic Operations	.62	.77	.68	.81	.66	.80	.63	.77	.67	.80

Table III-9

Alternate Form and Spearman-Brown Corrected Reliabilities

of Aptitude Tests by Ethnic Group

Inventory Management Specialists

	Black		Mexican-American		Caucasian	
	r_{AB}	r_{SB}	r_{AB}	r_{SB}	r_{AB}	r_{SB}
Number Comparison	.85	.92	.79	.88	.83	.91
Hidden Figures	.54	.70	.39	.56	.51	.68
Vocabulary	.71	.83	.60	.75	.70	.82
Object-Number	.75	.86	.69	.82	.61	.76
Letter Sets	.73	.84	.67	.80	.76	.86
Nonsense Syllogisms	.56	.72	.32	.48	.50	.67
Subtraction & Multiplication	.85	.92	.91	.95	.92	.96
Extended Range Vocabulary	.83	.91	.75	.86	.85	.92
Necessary Arithmetic Operations	.68	.81	.75	.86	.68	.81
Inference	.59	.74	.51	.68	.68	.81

Chapter IV

Comparison of Ethnic Groups on Criterion Measures

A description of the selection and development of the criterion measures for the three occupations is given in detail in the instrumentation reports previously cited (Pike, 1969; Parry, 1971; Crooks and Mahoney, 1971). In this chapter, the measures will be described only briefly as background for discussion of the results. The means and standard deviations of the scores and the intercorrelations of the measures will be presented by ethnic group for each of the three studies.

Medical Technicians

Two types of criterion measures were used in this study: Supervisors' Ratings (eight scales describing aspects of performance on the job and a rating of overall effectiveness) and a Job Knowledge Test. Peer ratings were obtained but not used in the analyses.

Rating Scales

The rating scales are described below in behavioral terms:

Flexibility

Ability to adapt readily to a wide variety of tasks, to changes in procedures and equipment, and to emergency situations

Organization

Skill in planning and organizing work, assigning priorities, and anticipating needs

Interest

Interest in w[illegible] and initiative in improving performance

Learning Ability

Willingness and ability to learn on the job; quick to understand

Job Knowledge (Technical Knowledge)

Knowledge and comprehension of the body of technical information and procedures surrounding the job

Technique

Quality of technique in the laboratory in handling equipment, materials, and tasks

Low Need for Supervision

Degree of responsibility assumed on the job without direct supervision

Communication

Ability to communicate with others

Overall Rating

Overall effectiveness as a Medical Technician

Job Knowledge Test

The Job Knowledge Test was developed to measure the knowledge of the Medical Technicians about their field (information and techniques).

Ethnic Group Comparisons

Table IV-1 shows the means and standard deviations of the criterion measures for the Black and Caucasian samples in the Medical Technician study.

Supervisors' Ratings: For all but one of the rating scales, the mean ratings of Caucasians were slightly higher. The mean rating on the Flexibility scale was very slightly higher for the Black sample. The only significant difference in mean ratings (at the .05 level) was on the Low Need for Supervision scale ("low need" at the upper end of the scale).

Job Knowledge Test: On the Job Knowledge Test, the mean score difference between the Black and Caucasian samples was significant at the .01 level. The mean of the Caucasian sample was approximately one-half standard deviation higher. (It may be noted that the means of the supervisors' ratings on the Job Knowledge scale for the Black and Caucasian samples differed only slightly.)

Figure IV-1 shows graphically the differences between Caucasian and Black mean scores for aptitude tests (see Chapter III) and selected criterion measures for Medical Technicians (supervisors' ratings of Learning Ability, Job Knowledge, and Overall effectiveness; and the Job Knowledge Test). The differences between the means are plotted in terms of standard deviation units, showing the Caucasian mean as point of departure.

Table IV-2 shows the intercorrelations of the criterion measures for the Medical Technicians by ethnic group. The intercorrelations of the ratings for both Black and Caucasian samples are quite high (ranging from .65 to .84 for the Black sample and from .55 to .77 for the Caucasian sample). The separate rating scales also correlate highly with the overall rating (R_9), ranging from .70 to .81 for the Black sample and from .65 to .77 for the Caucasian sample. It was hoped that the "halo" effect usually found in such ratings would be less through the use of behavioral descriptions to anchor the scales. The Job Knowledge Test correlations with the rating scales range from .38 to .58 for the Black sample and from .22 to .35 for the Caucasian sample. The highest correlations of ratings with the Job Knowledge Test for both samples are with the scales for Job Knowledge and Learning Ability, which might be considered a kind of validation of the Job Knowledge Test. The reasons for the difference in the magnitude of the intercorrelations of the ratings with the Job Knowledge Test for the Black and Caucasian samples are obscure, and later analyses may be enlightening. The differences suggest, however, that the two groups are being evaluated on somewhat different bases.

Cartographic Technicians

Three types of criterion measures were used in the study of Cartographic Technicians: Supervisors' Ratings (seven rating scales describing aspects of performance and a rating of overall effectiveness), a Job Knowledge Test, and a Work Sample with three separate tasks. All three were included for the TOPOCOM sample, but only Supervisors' Ratings were obtained for the Coast & Geodetic Survey sample. The Job Knowledge Test and the Work Sample tasks were not considered appropriate for this group, for reasons discussed in Chapter II.

Rating Scales

The rating scales used to obtain supervisors' ratings for the Cartographic Technicians are described below in terms of behavior on the job:

Accuracy

Ability to produce accurate work, whether simple or complex

Interest

Takes interest and pride in the work; improves through study and practice

Learning Ability

Ability to understand new procedures and ideas with a minimum of instruction or explanation

Job Knowledge (Technical Knowledge)

Knowledge and comprehension of the body of technical information and procedures in the cartography field

Dexterity

Ability to work with tools and instruments with precision and skill; eye-hand coordination

Low Need for Supervision

Ability to carry out and show initiative in work with minimum direction or supervision

Perseverance

Patience and stick-to-itiveness in working on painstaking tasks to completion

Overall Rating

Overall effectiveness as a Cartographic Technician

Job Knowledge Test

The Job Knowledge Test was developed to measure the knowledge of the Cartographic Technicians about their field (information and techniques relating specifically to the cartographic and photogrammetric aspects considered to be the primary areas of work).

Work Samples

Three work sample tasks were developed for th Cartographic Technician study: a geometric restitution task, a logical contouring problem, and a "pull-up" from an aerial photograph, as described more fully in Chapter II. The scores for the separate tasks were standardized so they could be combined to provide a Work Sample Composite score.

Ethnic Group Comparisons (TOPOCOM)

Table IV-3 shows the means and standard deviations of the criterion measures for TOPOCOM installations, which included Black, Mexican-American, and Caucasian samples.

Black - Caucasian Comparisons

Supervisors' Ratings: Caucasians received higher mean ratings on all scales. On one scale, Interest (in the work), the mean difference was significant at the .01 level. On three other scales, Accuracy, Learning Ability, and Low Need for Supervision ("low need" at the upper end of the scale), the mean difference was significant at the .05 level. The differences were all less than one-half standard deviation in size. On three scales and on the Overall Rating, the mean difference was not significant.

Job Knowledge Test: The mean score of the Black sample on the Job Knowledge Test was lower than for the Caucasian sample. The mean score difference was significant at the .01 level, and was about one-half standard deviation in size.

Work Sample: The mean scores of the Black sample on all three tasks in the Work Sample were lower than for the Caucasian sample. The mean differences were significant at the .05 level, and were about one-third standard deviation in size. On the Work Sample Composite, the effect of standardizing and combining the scores was to make the Composite score mean difference significant at the .01 level.

Mexican-American - Caucasian Comparisons

Supervisors' Ratings: On one scale, Perseverance, the mean for the Mexican-American sample was higher than that of the Caucasian sample, the difference significant at the .05 level. The mean rating of the Mexican-Americans was also higher on the Dexterity scale, but not significantly so. The Caucasian mean rating for Learning Ability was slightly higher, but this and other differences were trivial.

Job Knowledge Test: On the Job Knowledge Test, the Caucasians scored higher on the average than the Mexican-Americans. The mean score difference was significant at the .01 level, and was almost one standard deviation in size. The Mexican-Americans also stored almost one-half standard deviation below the Black sample on this test. At least part of this difference can probably be attributed to the language handicap noted on the Vocabulary Test scores in the prediction batteries.

Work Sample: On one of the Work Sample tasks, Logical Contouring, the mean score of the Mexican-Americans was significantly lower (at the .01 level) than for the Caucasians. The mean difference was almost one standard deviation in size. The Cartographic Technicians at San Antonio, the majority of whom were Mexican-Americans, indicated that the Logical Contouring task was less familiar to them. On the other two tasks, while the Mexican-American sample means were slightly lower than the Caucasian sample means, the mean difference was not significant. On the Work Sample Composite, the effect of standardizing and combining the scores was to make the mean difference significant at the .01 level.

Figure IV-2 shows graphically the differences between Caucasian mean scores and Black and Mexican-American mean scores for aptitude tests (see Chapter III) and selected criterion measures for Cartographic Technicians - TOPOCOM (supervisors' ratings of Learning Ability, Job Knowledge, and Overall effectiveness; Job Knowledge Test; and Work Sample Composite). The differences between the means are plotted in terms of standard deviation units, showing the Caucasian mean as point of departure.

Table IV-4 shows the intercorrelations of the criterion measures for the Cartographic Technicians (TOPOCOM). The intercorrelations for the ethnic group samples are shown on the same table for comparison. It will be seen that the rating scales, including the overall rating, are highly intercorrelated for all three samples, ranging from .59 to .93. The correlations of the ratings with the Job Knowledge Test range from .13 to .51 across the three ethnic groups. The correlations of the supervisors' ratings on the Job Knowledge scale with the Job Knowledge Test scores are .39, .37, and .38 for the respective ethnic groups, remarkably consistent, and somewhat higher than validity coefficients usually found where supervisors' ratings are the criterion. The intercorrelations of the ratings and the Work Sample tasks are low, in general, ranging from -.02 to .49 across the ethnic groups. The correlations of the Job Knowledge Test scores with the Work Sample tasks range from .22 to .44, showing some overlap of knowledge and performance criteria.

Examination of the intercorrelations for the three ethnic groups shows that the intercorrelations of the Mexican-American sample tend to be slightly higher than for the Black and Caucasian samples, as seen in the following summary.

Range of Correlations

	Blacks	Mexican-Americans	Caucasians
Ratings with Ratings	.61 - .90	.68 - .93	.59 - .89
Ratings with JKT	.12 - .39	.37 - .51	.25 - .45
Ratings with Work Sample Restitution	-.01 - .13	.06 - .24	.06 - .23
Ratings with Work Sample Logical Contouring	-.02 - .13	.11 - .38	.02 - .14
Ratings with Work Sample Pull-up	.01 - .18	.31 - .49	.15 - .27

One reason for the higher intercorrelations of the criterion measures for the Mexican-American sample may be the relative homogeneity of the sample. All but one or two subjects in this sample were from one installation. It was possible to obtain at least two ratings for each subject, and in some cases three or four. These ratings were averaged, probably contributing to higher reliability. Later analyses may throw more light on these differences.

Ethnic Group Comparisons (Coast & Geodetic Survey)

The sole criterion measure used for this sample was Supervisors' Ratings. Table IV-5 shows the means and standard deviations for the Black and Caucasian samples.

Black - Caucasian Comparisons

The Black sample received lower mean ratings on all scales, including the Overall Rating. On one scale, Job Knowledge, the difference was significant at the .01 level. On two scales, Accuracy and Learning Ability, and on the Overall Rating, the difference was significant at the .05 level. Differences ranged from one-half to two-thirds standard deviation in size. Differences on the other scales were not significantly large.

Figure IV-3 shows graphically the differences between Caucasian and Black mean scores for aptitude tests (see Chapter III) and selected criterion measures for Cartographic Technicians - Coast & Geodetic Survey (supervisors' ratings of Learning Ability, Job Knowledge, and Overall effectiveness) The differences between the means are plotted in terms of standard deviation units, showing the Caucasian mean as point of departure.

Table IV-6 shows the intercorrelations of the ratings for the two ethnic groups in the Coast & Geodetic Survey samples. It can be seen that the intercorrelations for the Black sample tend to be higher on the whole than for the Caucasian sample. It appears that in ratings of Caucasians there was more differentiation on the basis of the behavior defined by the separate scales, whereas supervisors tended to rate the Blacks on a common factor across scales.

Inventory Management Specialists

Two types of criterion measures were used in this study: Supervisors' Ratings (nine scales describing aspects of performance on the job and a rating of overall effectiveness) and a Work Sample.

Rating Scales

The rating scales are described below in behavioral terms relating to job performance:

Organization

Ability to organize work, plan ahead, and assign priorities

Interest

Interest and pride in work and initiative in improving performance

Learning Ability

Ability to absorb and understand new materials and techniques

Communication

Ability to communicate effectively in person and in writing

Technical Knowledge (Job Knowledge)

Knowledge and comprehension of the body of technical information and procedures surrounding the job

Stability/Adaptability

Ability to adapt to new procedures and conditions; stability in emergencies

Dependability

Reliability in carrying out assignments with minimum supervision

Judgment

Ability to use practical judgment in making decisions and solving problems

Cooperation

Ability to initiate and maintain effective interpersonal relationships with others on the job

Overall Rating

Overall effectiveness as an Inventory Management Specialist

Work Sample

As described more fully in Chapter II, the Work Sample is a simulation of the job of an Inventory Management Specialist. In the test, the subject is presented with a set of problems like those usually encountered in the job. The scoring dimensions are described below:

Takes Leading Action

Takes action toward solving the problem (asks for information, decision, advice; initiates necessary documents)

Uses New Procedures

Uses new procedures in taking action on problems (refers to or uses background information, procedures and codes, forms provided in Work Sample)

Shows Inventory Management Knowledge

Shows inventory management knowledge (drafts messages, shows understanding of MILSTRIP, takes actions not specific to new setting based on present knowledge of inventory management)

Analyzes Problems

Analyzes problems (takes into consideration background information and information in other problems to analyze situation, make recommendations, or arrive at decisions beyond what is specifically asked for in problem)

Organizes Systematically

Organizes systematically (recognizes priorities)

Maintains Control

Maintains control (makes provision for follow-up; records information on inventory records)

Follows Directions

Follows directions (does what is asked in problem)

Productivity

Amount of work accomplished (number of usual actions taken, number of unusual actions taken, number of problems worked on)

Quality of Actions

Quality of Actions taken (from scoring key developed from pooled judgments of appropriate and inappropriate actions)

Rating of Overall Performance

Rating of overall performance (scorer's judgment on a 7-point scale of how effective subject would be in the assumed role, based on initial handling of a sample of the work)

Ethnic Group Comparisons

Tables IV-7 and IV-8 show the means and standard deviations for the Supervisors' Ratings and the Work Sample for Inventory Management Specialists by ethnic group.

Black - Caucasian Comparisons

Supervisors' Ratings: The mean supervisors' rating for the black sample was significantly lower (at the .05 level) on only one scale (Technical Knowledge). The size of the differences was less than one-third standard deviation. Caucasians received very slightly higher mean ratings on all other scales with the exception of the Cooperation scale, where the mean rating for the Black sample was slightly higher than for the Caucasian sample.

Work Sample: For the Work Sample, all mean scores of the Black sample were lower than for the Caucasian sample. The mean score differences were significant at the .01 level. The size of the differences was about two-thirds standard deviation on most scores and almost one standard deviation on the Quality of Actions score. The standard deviations of the Black sample were somewhat smaller on all scores than for the Caucasian sample, indicating less dispersion of scores.

Mexican-American - Caucasian Comparisons

Supervisors' Ratings: The mean supervisors' rating for the Mexican-American sample was significantly lower (at the .05 level) on only one of the rating scales (Communication). All other mean differences were trivial, in some cases favoring the Mexican-American sample.

Work Sample: None of the mean score differences between the Mexican-American sample and the Caucasian sample was significant. On some of the scoring dimensions, the small mean differences favored the Mexican-Americans; on others, the Caucasians.

Figure IV-4 shows graphically the differences between Caucasian mean scores and Black and Mexican-American mean scores on aptitude tests (see Chapter III) and selected criterion measures for Inventory Management Specialists (supervisors' ratings of Learning Ability, Job

Knowledge, and Overall effectiveness; and Work Sample Overall Performance). The differences between the means are plotted in terms of standard deviation units, showing the Caucasian mean as point of departure.

Table IV-9 shows the intercorrelations of the criterion measures by ethnic group.

Ratings with Ratings Comparisons

The intercorrelations of the rating scales range from .46 to .84 for the Black sample, from .34 to .83 for the Mexican-American sample, and from .41 to .85 for the Caucasian sample, perhaps again indicating a tendency noted before for the Blacks to be rated on a common factor across rating scales. In looking at a distribution of the correlations by ethnic group, the intercorrelations for the Mexican-American sample tended to spread more widely over the range than for either of the other groups. Some slight indication of the emphasis given the various characteristics described by the scales by the supervisors can be noted from the intercorrelations of the rating scales with Overall Rating. For the Blacks, the highest correlations of scales with Overall Rating were Organization, Interest, Dependability, and Judgment; the lowest correlation was with Cooperation. For the Mexican-Americans, the highest correlation with Overall Rating were Organization, Interest, and Stability/Adaptability; the lowest correlations were with Communication, Technical Knowledge, and Cooperation. For the Caucasians, the highest correlations with Overall Rating were Organization, Stability/Adaptability, and Judgment; the lowest was with Cooperation. All correlations with Overall Rating were fairly high, ranging from .54 to .85 across ethnic groups, so the differences are not very meaningful.

Ratings with Work Sample Comparisons

In general, the intercorrelations of ratings with Work Sample were low, ranging from .07 to .35 for the Black sample, from -.07 to .42 for the Mexican-American sample, and from -.08 to .34 for the Caucasian sample. Correlations of rating scales with Work Sample Overall Rating for the Black sample, significant at the .01 level, were Organization, Learning Ability, Technical Knowledge, Stability/Adaptability, Dependability, Judgment, and Cooperation; for the Mexican-Americans, Learning Ability and Communication; and for the Caucasians, all scales except Dependability and Cooperation. The intercorrelations across rating scales with the Work Sample Overall Rating for the Black and Caucasian samples again may be evidence of a common factor underlying supervisors' ratings, but the reason the Mexican-American sample has different relationships is not understood. Later analyses may throw more light on these differences.

Summary

For the Medical Technicians and Inventory Management Specialists, the mean supervisors' ratings are only slightly different for the ethnic groups. The mean supervisors' ratings for the Cartographic Technicians again are quite similar, but differences do appear among the groups on various scales, suggesting that some aspects of job performance for the ethnic groups were seen differentially. For the Cartographic Technician sample, multiple ratings were obtained at some installations, perhaps increasing their reliability when averaged.

On the Job Knowledge Tests, the Black Medical Technicians and the Black Cartographic Technicians scored lower by about one-half standard deviation than the Caucasian groups. The Mexican-American Cartographic Technicians scored lower than the Caucasians by almost one standard deviation (and one-half standard deviation lower than the Black Cartographic Technicians). In the latter instance, the Mexican-American results on the Job Knowledge Test may reflect bias due to a language handicap. (Blacks scored higher on Vocabulary tests and some other tests involving language in both the Cartographic Technician and the Inventory Management Specialist studies, although Mexican-Americans scored higher than Blacks on rote memory, perceptual speed, number fluency, and tests of general reasoning.)

On the Work Sample for the Cartographic Technicians, the mean scores of the Black Cartographic Technicians were lower on all three tasks than the Caucasian mean scores (at about the same magnitude). However, the mean scores of the Mexican-American Cartographic Technicians on two of the tasks were not different from the mean scores of the Caucasian sample. On the Logical Contouring task, the Mexican-Americans were one-half standard deviation lower than the Caucasians and one-third standard deviation lower than the Blacks, suggesting possible differential experience on this task among the groups.

On the Work Sample for the Inventory Management Specialists, the Black sample scored consistently lower than both the Mexican-Americans and the Caucasians. This Work Sample, while highly verbal in content, was developed around the actual work performed by Inventory Managers, and the language handicap of Mexican-Americans, reflected in their lower scores on aptitude tests involving the use of language (in which Mexican-Americans scored lower than Blacks as well as Caucasians), appeared not to be a factor.

Of the three kinds of criterion measures, the supervisors' ratings appear to reflect least well the differences among the ethnic groups shown by aptitude tests and other criterion measures. Further examination of rating variations and sources of rating bias will be made in Chapter VIII.

Black subjects scored less well than Caucasians on the Job Knowledge Tests in both the Medical Technician and Cartographic Technician studies. The Mexican-American Cartographic Technicians appeared to do less well on the Job Knowledge Test than their performance on other measures would indicate. Whether this is attributable in part to language handicap is not clear.

The Work Sample, with the exception of the Logical Contouring task for the Mexican-Americans, appeared to elicit somewhat consistent results from the ethnic groups in the Cartographic Technician study. On the Work Sample for the Inventory Management Specialist study, the Blacks performed consistently less well than the Mexican-Americans and Caucasians, but they also had lower aptitude test scores in general (with the exception of the Vocabulary tests mentioned above).

Table IV-1

SUPERVISORS' RATINGS AND JOB KNOWLEDGE TEST SCORES

MEANS AND STANDARD DEVIATIONS

BY ETHNIC GROUP

Medical Technicians

Criterion Measures	Black N = 166		Caucasian N = 290	
	Mean	S.D.	Mean	S.D.
Rating Scales				
Flexibility	5.5	1.8	5.4	1.8
Organization	5.7	1.8	6.0	1.8
Interest	5.6	1.8	5.9	1.7
Learning Ability	5.8	1.9	6.0	1.8
Job Knowledge	5.2	1.9	5.3	1.7
Technique	5.9	1.7	6.0	1.7
Low Need for Supervision	5.7*	2.0	6.1	1.9
Communication	5.5	1.8	5.8	1.8
Overall Rating	5.7	1.9	5.9	2.0
Job Knowledge Test	31.4**	10.4	35.7	9.4

*Significantly different from Caucasian mean at .05 level

**Significantly different from Caucasian mean at .01 level

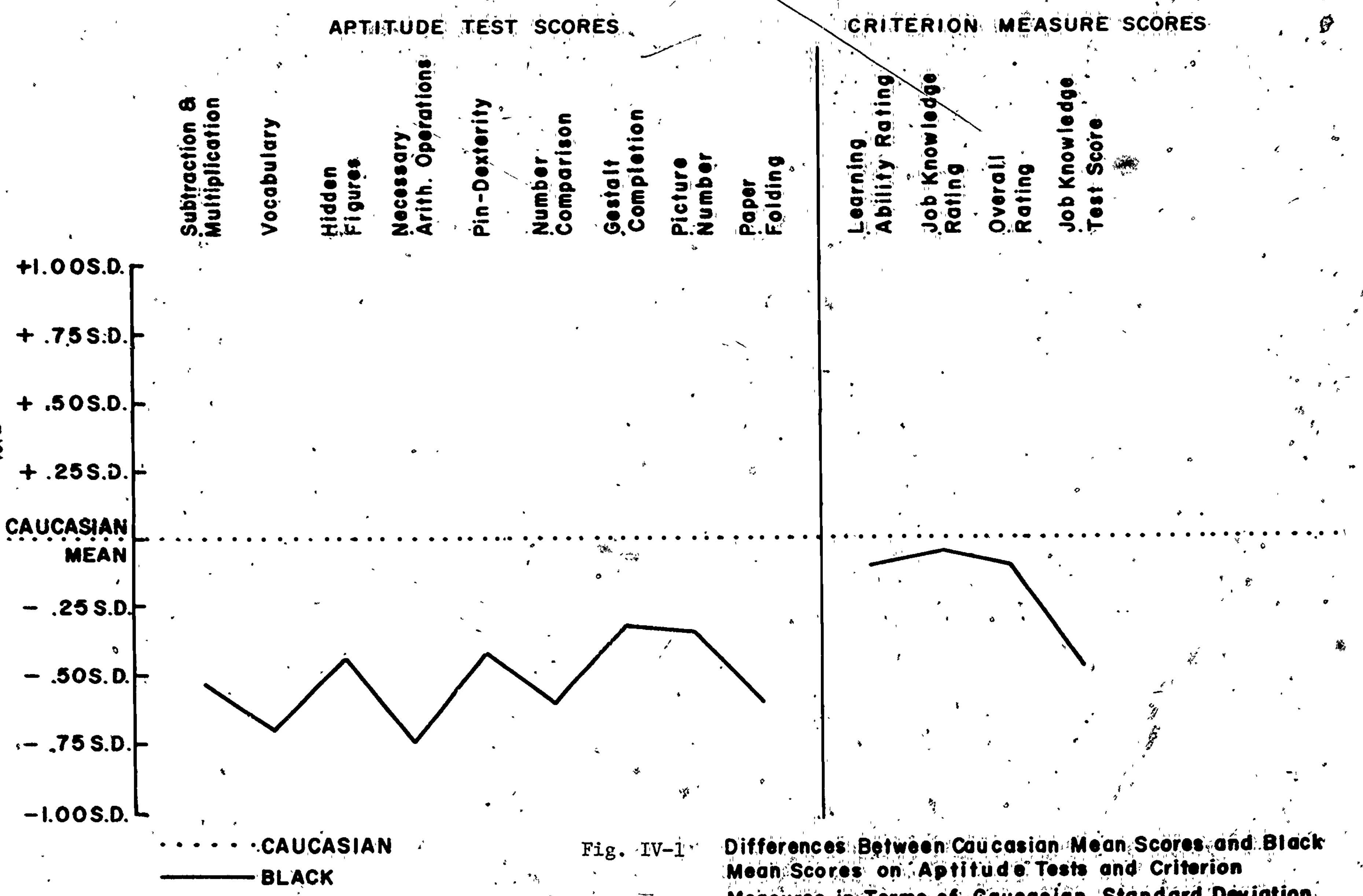


Fig. IV-1 Differences Between Caucasian Mean Scores and Black Mean Scores on Aptitude Tests and Criterion Measures in Terms of Caucasian Standard Deviation Units MEDICAL TECHNICIANS

Table IV-2

INTERCORRELATIONS OF CRITERION MEASURES BY ETHNIC GROUP

(first line Black sample, second line Caucasian sample)

Medical Technicians

Rating Scales	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	R_9	Job Knowledge Test
R_1 Flexibility	1.00	.79	.75	.80	.77	.76	.70	.65	.77	.50
	1.00	.71	.61	.73	.55	.66	.65	.57	.69	.31
R_2 Organization		--	.74	.77	.72	.76	.78	.68	.78	.45
			.70	.71	.64	.74	.76	.61	.74	.27
R_3 Interest			--	.73	.70	.73	.69	.67	.74	.38
				.70	.63	.64	.72	.61	.73	.22
R_4 Learning Ability				--	.84	.82	.79	.75	.80	.58
					.72	.75	.71	.66	.74	.35
R_5 Job Knowledge					--	.79	.75	.78	.77	.58
						.69	.73	.70	.72	.35
R_6 Technique						--	.79	.72	.81	.49
							.73	.65	.75	.23
R_7 Low Need for Supervision							--	.73	.79	.43
								.71	.77	.27
R_8 Communication								--	.70	.50
									.65	.27
R_9 Overall Rating									--	.45
										.25
Job Knowledge Test										--

Significance	r
Black sample	.16 = significant at .01 level
	.21 = significant at .05 level
Caucasian sample	.11 = significant at .01 level
	.15 = significant at .05 level

Table IV-3

SUPERVISORS' RATINGS, JOB KNOWLEDGE TEST AND WORK SAMPLE SCORES

MEANS AND STANDARD DEVIATIONS BY ETHNIC GROUP

Cartographic Technicians (TOPOCOM Sample)

	Black N=101		Mexican-American N=101		Caucasian N=241	
Rating Scales	Mean	S.D.	Mean	S.D.	Mean	S.D.
Accuracy	5.4*	1.8	5.8	1.7	5.9	1.7
Interest	5.3**	1.7	5.9	1.7	5.9	1.6
Learning Ability	5.4*	1.7	5.6	1.7	5.8	1.7
Job Knowledge	5.3	1.7	5.6	1.8	5.7	1.8
Dexterity	5.9	1.6	6.3	1.7	6.1	1.6
Need for Supervision	5.1*	1.8	5.5	1.8	5.6	1.8
Perseverance	6.1	1.6	6.6*	1.5	6.2	1.7
Overall Rating	5.5	1.9	5.8	1.6	5.9	1.7
Job Knowledge Test	35.4**	11.4	30.6**	10.0	40.3	11.6
Work Sample						
Restitution	17.3*	6.9	18.8	7.5	19.1	7.6
Logical Contouring	8.8*	3.4	7.2**	4.1	9.6	2.9
Pull-up	11.3*	5.4	12.2	5.3	12.6	5.1
Composite#	18.3**	9.2	17.9**	10.7	21.6	8.9

*Significantly different from Caucasian mean at .05 level

**Significantly different from Caucasian mean at .01 level

#Obtained by standardizing and combining the raw scores for the separate tasks

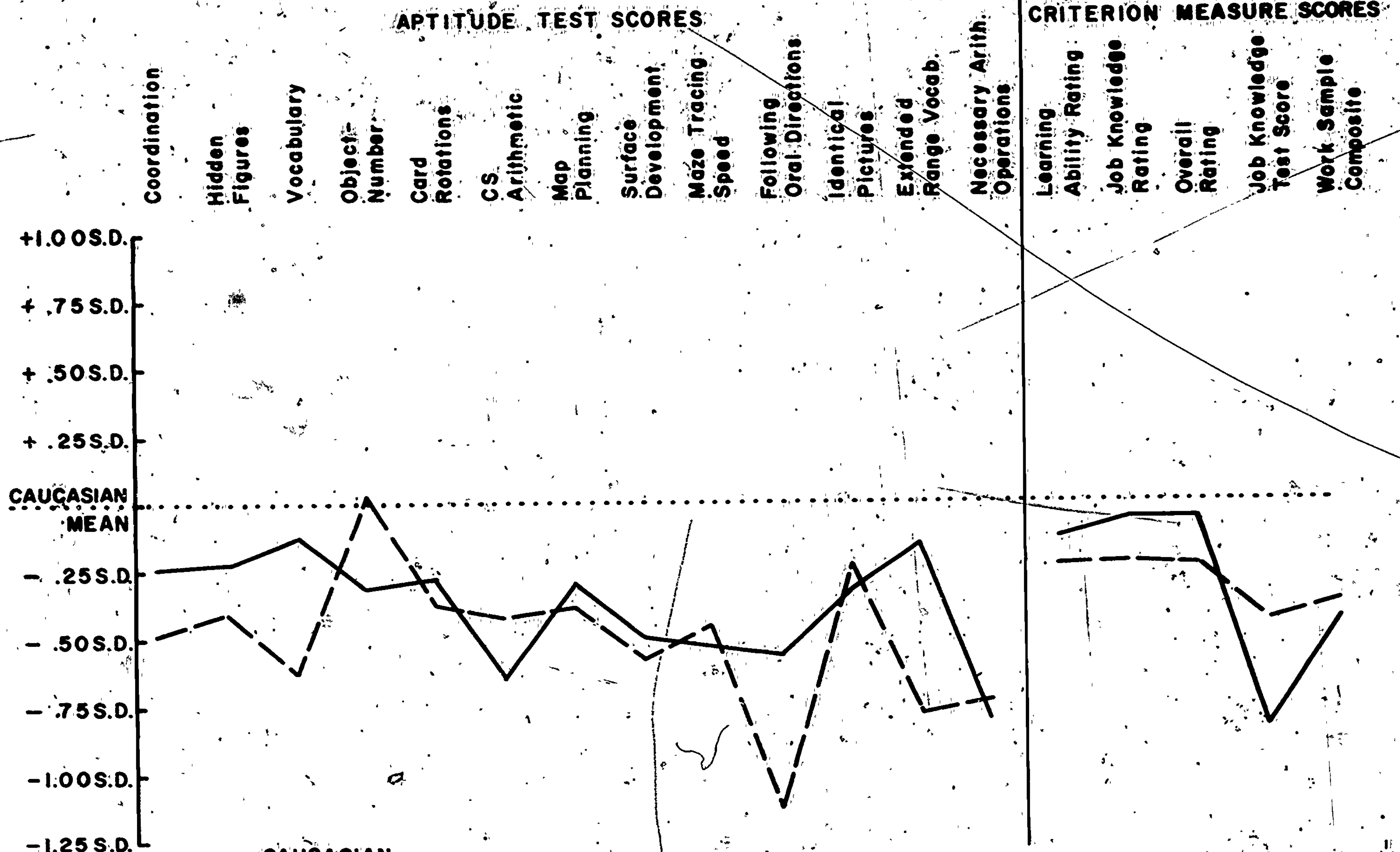


Fig. IV-2 Differences Between Caucasian Mean Scores and Black and Mexican-American Mean Scores on Aptitude Tests and Criterion Measures in Terms of Caucasian Standard Deviation Units CARTOGRAPHIC TECHNICIANS (TOPOCOM)

Table IV-4

Intercorrelations of Criterion Measures by Ethnic Group

Cartographic Technicians (TOPOCOM)

(Black - first line; Mexican-American - second line; Caucasian - third line)

Rating Scales	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	JKT	WS_1	WS_2	WS_3	Comp
R_1 Accuracy	1.00	.74	.84	.84	.77	.86	.69	.89	.22	.09	.05	.14	.13
	1.00	.89	.89	.90	.78	.85	.75	.90	.42	.22	.29	.38	.41
	1.00	.78	.79	.86	.77	.84	.75	.88	.33	.20	.08	.16	.22
R_2 Interest		--	.74	.72	.61	.75	.78	.79	.15	-.01	.02	.01	.01
			.87	.83	.75	.83	.86	.87	.47	.11	.21	.32	.29
			.77	.79	.65	.78	.77	.77	.34	.10	.11	.26	.23
R_3 Learning Ability			--	.87	.75	.85	.62	.83	.35	.11	.13	.18	.21
				.92	.79	.91	.75	.88	.51	.24	.27	.42	.42
				.85	.73	.85	.64	.83	.45	.23	.14	.27	.31
R_4 Job Knowledge				--	.73	.90	.68	.85	.39	.13	.13	.18	.21
					.78	.93	.72	.88	.37	.17	.21	.38	.34
					.74	.89	.70	.85	.38	.17	.13	.23	.25
R_5 Dexterity					--	.75	.68	.76	.12	.10	-.02	.07	.07
						.81	.68	.85	.38	.09	.38	.49	.44
						.76	.59	.80	.27	.17	.04	.17	.19
R_6 Need for Supervision						--	.73	.86	.31	.10	.11	.14	.17
							.73	.89	.44	.16	.24	.44	.38
							.68	.86	.39	.22	.14	.25	.30
R_7 Perseverance							--	.76	.13	.01	.02	.05	.04
								.77	.38	.06	.11	.31	.21
								.75	.25	.06	.02	.15	.11
R_8 Overall Rating								--	.28	.02	.11	.15	.14
									.42	.15	.29	.36	.37
									.36	.17	.07	.16	.19
Job Knowledge Test									--	.22	.31	.44	.47
										.45	.27	.39	.50
										.41	.30	.42	.55
WS_1 Restitution										--	.16	.08	
											.29	.17	
											.10	.27	
WS_2 Logical Contouring											--	.40	
												.43	
												.27	
WS_3 Pull-up												--	
WS Composite													

Significance	r
Black & Mexican-American samples	.25 = significant at .01 level .20 = significant at .05 level
Caucasian sample	.18 = significant at .01 level .14 = significant at .05 level

Table IV-5

SUPERVISORS' RATINGS

MEANS AND STANDARD DEVIATIONS BY ETHNIC GROUP

Cartographic Technicians (Coast & Geodetic Survey Sample)

	Black N=38		Caucasian N=50	
Rating Scales	Mean	S.D.	Mean	S.D.
Accuracy	5.4*	1.9	6.1	1.6
Interest	5.5	2.2	6.1	1.9
Learning Ability	5.1*	2.1	6.1	1.8
Job Knowledge	4.9**	2.1	6.0	1.7
Dexterity	6.1	2.0	6.3	1.8
Need for Supervision	5.3	2.1	5.9	1.8
Perseverance	6.0	2.3	6.6	1.8
Overall Rating	5.4*	2.1	6.2	1.6

*Significantly different from Caucasian mean at .05 level

**Significantly different from Caucasian mean at .01 level

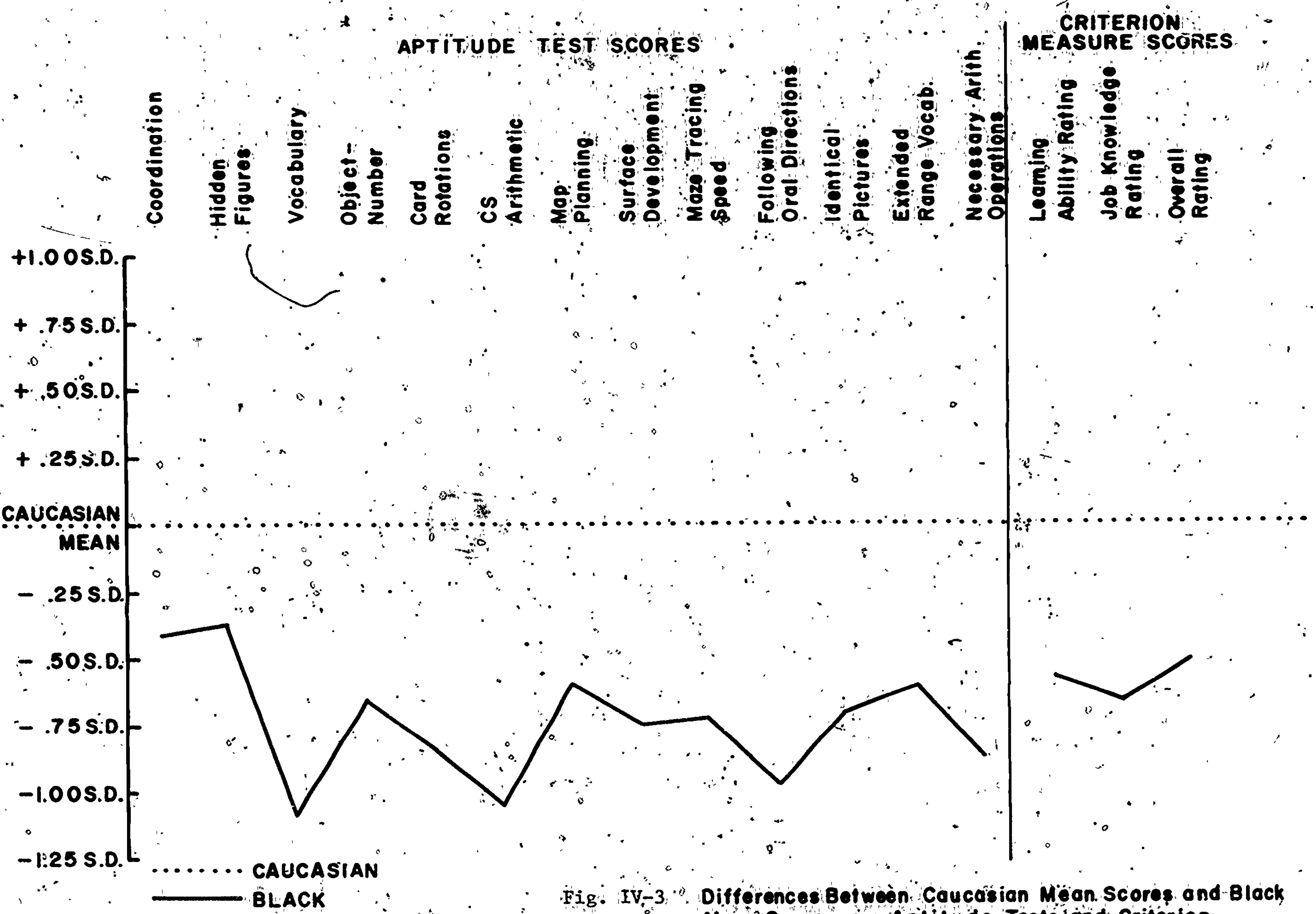


Fig. IV-3 Differences Between Caucasian Mean Scores and Black Mean Scores on Aptitude Tests and Criterion Measures in Terms of Caucasian Standard Deviation Units CARTOGRAPHIC TECHNICIANS (Coast & Geodetic Survey)

Table IV-6

INTERCORRELATIONS OF SUPERVISORS' RATINGS BY ETHNIC GROUP

Cartographic Technicians (Coast & Geodetic Survey Sample)

(Black - first line, Caucasian - second line)

	1	2	3	4	5	6	7	8
1. Accuracy	1.00	.89	.85	.87	.87	.82	.84	.91
	1.00	.78	.82	.79	.76	.78	.66	.88
2. Interest		--	.81	.87	.80	.79	.93	.93
			.69	.70	.56	.70	.75	.82
3. Learning Ability			--	.92	.81	.84	.72	.85
				.86	.72	.86	.55	.86
4. Job Knowledge				--	.86	.90	.80	.89
					.64	.76	.62	.78
5. Dexterity					--	.83	.73	.83
						.72	.47	.78
6. Need for Supervision						--	.74	.86
							.57	.84
7. Perseverance							--	.89
								.73
8. Overall Rating								--

Significance	r
Black sample	.42 = significant at .01 level
	.32 = significant at .05 level
Caucasian sample	.29 = significant at .01 level
	.37 = significant at .05 level

Table IV-7

MEANS AND STANDARD DEVIATIONS FOR SUPERVISORS' RATINGS

BY ETHNIC GROUP

Inventory Management Specialists

	Black N=114		Mexican-American N=74		Caucasian N=205	
Rating Scales	Mean	S.D.	Mean	S.D.	Mean	S.D.
Organization	6.0	1.9	6.4	1.7	6.2	1.9
Interest	6.0	1.9	6.3	1.8	6.1	1.8
Learning Ability	6.0	1.9	6.1	1.8	6.1	1.9
Communication	5.8	1.7	5.5*	1.9	6.0	1.8
Technical Knowledge	5.8*	1.9	6.3	1.6	6.3	1.8
Stability/ Adaptability	6.0	1.9	6.1	1.6	6.0	1.9
Dependability	6.3	2.0	6.6	1.6	6.5	1.8
Judgment	6.2	1.8	6.3	1.6	6.3	1.8
Cooperation	7.2	1.6	6.9	1.6	7.0	1.5
Overall Rating	6.2	1.7	6.4	1.7	6.3	1.7

*Significantly different from Caucasian mean at .05 level

Table IV-8

MEANS AND STANDARD DEVIATIONS OF WORK SAMPLE DATA

BY ETHNIC GROUP

Inventory Management Specialists

Work Sample Dimensions	Black N=100		Mexican-American N=59		Caucasian N=171	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Takes Leading Action	2.9**	2.3	4.5	2.9	4.4	2.8
Uses New Procedures	7.4**	3.8	10.0	5.1	10.2	5.2
Shows I. M. Knowledge	3.6**	2.5	5.1	3.1	5.1	3.0
Analyzes Problems	1.5**	1.4	2.2	1.9	2.6	2.3
Organizes Systematically	4.2**	3.1	6.0	3.6	6.2	4.0
Maintains Controls	1.8**	1.6	3.2	2.3	2.6	2.2
Follows Directions	3.9**	2.1	5.5	3.1	5.6	3.2
Productivity	44.6**	22.3	58.8	33.4	57.3	34.9
Quality of Actions	10.2**	5.2	14.0	6.4	14.5	7.0
Rating of Overall Performance	6.6**	3.4	9.0	4.1	8.9	4.5

**Significantly different from Caucasian mean at .01 level

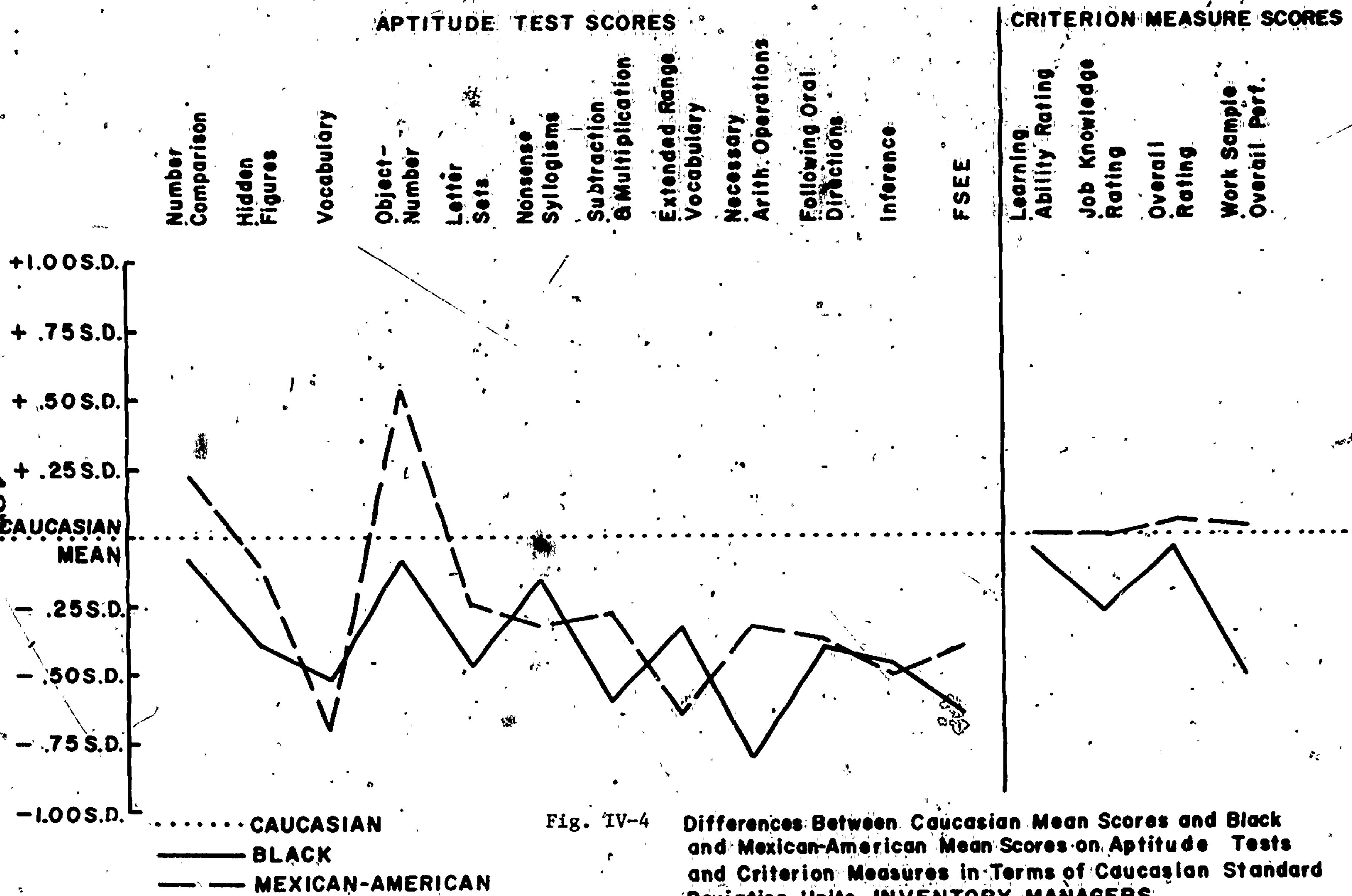


Fig. IV-4 Differences Between Caucasian Mean Scores and Black and Mexican-American Mean Scores on Aptitude Tests and Criterion Measures in Terms of Caucasian Standard Deviation Units INVENTORY MANAGERS

Table IV-9

INTERCORRELATIONS OF CRITERION MEASURES BY ETHNIC GROUP

Inventory Management Specialists

(Black - first line; Mexican-American - second line; Caucasian - third line)

Rating Scales	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	R_9	R_{10}	WS_1	WS_2	WS_3	WS_4	WS_5	WS_6	WS_7	WS_8	WS_9	WS_{10}
R_1 Organization	1.00	.80	.77	.72	.75	.71	.79	.76	.55	.83	.20	.23	.21	.18	.26	.16	.26	.15	.28	.25
	1.00	.76	.72	.67	.52	.71	.75	.65	.47	.83	.18	.02	-.03	.34	.20	.04	.09	.10	.06	.16
	1.00	.75	.82	.71	.73	.70	.73	.76	.45	.80	.14	.21	.17	.14	.25	.19	.21	.11	.17	.19
R_2 Interest		--	.73	.70	.72	.74	.77	.76	.57	.80	.16	.16	.15	.13	.23	.11	.21	.13	.20	.22
			.69	.59	.57	.72	.76	.74	.53	.83	.11	.02	-.05	.31	.26	.04	.11	00	-.01	.01
			.70	.69	.74	.73	.72	.73	.56	.75	.08	.16	.17	.11	.23	.19	.20	.10	.16	.18
R_3 Learning Ability			--	.68	.78	.73	.71	.76	.46	.75	.24	.31	.27	.22	.30	.20	.28	.23	.35	.33
				.79	.72	.72	.60	.76	.34	.73	.25	.24	.05	.40	.28	.06	.26	.13	.27	.33
				.70	.79	.76	.72	.79	.41	.78	.21	.32	.20	.17	.34	.22	.30	.13	.26	.26
R_4 Communication				--	.75	.66	.59	.78	.47	.71	.14	.12	.22	.15	.14	.16	.15	.14	.16	.20
					.60	.66	.47	.66	.39	.65	.32	.26	.10	.42	.29	.08	.29	.21	.27	.34
					.72	.69	.65	.74	.55	.74	.06	.21	.17	.13	.25	.23	.23	.17	.18	.20
R_5 Technical Knowledge					--	.82	.67	.79	.50	.77	.22	.20	.20	.14	.23	.14	.23	.13	.21	.26
						.70	.54	.73	.37	.64	.24	.15	-.07	.33	.22	.01	.14	.08	.14	.20
						.76	.69	.81	.47	.79	.10	.20	.18	.14	.26	.21	.24	.17	.19	.22
R_6 Stability/ Adaptability						--	.74	.74	.58	.76	.24	.24	.27	.08	.17	.07	.23	.16	.24	.26
							.72	.77	.58	.75	.13	.12	-.01	.30	.24	-.03	.19	.08	.09	.13
							.74	.80	.64	.81	.09	.21	.16	.13	.26	.11	.23	.05	.20	.21
R_7 Dependability							--	.79	.58	.84	.22	.22	.24	.18	.20	.28	.23	.15	.31	.25
								.70	.54	.73	.09	.06	.07	.30	.10	.07	.12	.03	.06	.09
								.79	.51	.78	.06	.15	.18	.12	.19	.16	.18	.07	.14	.17
R_8 Judgment								--	.54	.83	.23	.24	.28	.23	.24	.21	.25	.19	.29	.29
									.48	.72	.24	.22	.11	.37	.24	.14	.22	.05	.18	.24
									.52	.85	.09	.22	.19	.12	.24	.18	.23	.07	.22	.22
R_9 Cooperation									--	.60	.20	.24	.29	.17	.08	.25	.23	.17	.28	.24
										.62	.18	.22	.06	.29	.32	.07	.33	.19	.21	.16
										.54	-.08	.01	.04	-.01	.10	-.01	-.02	-.01	00	.03

(continued on next page)

Table IV-9 (cont.)

Rating Scales	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	R_9	R_{10}	WS_1	WS_2	WS_3	WS_4	WS_5	WS_6	WS_7	WS_8	WS_9	WS_{10}
R_{10} Overall Rating										--	.18 .12 .11	.20 .04 .24	.27 -.05 .22	.20 .28 .14	.17 .19 .30	.22 .00 .24	.21 .12 .21	.26 .11 .09	.25 .03 .23	.24 .09 .24
Work Sample																				
WS_1 Takes Leading Action											--	.78 .64 .75	.60 .56 .56	.47 .43 .53	.68 .58 .72	.32 .43 .45	.54 .59 .63	.24 .14 .35	.63 .64 .65	.70 .61 .63
WS_2 Uses New Procedures												--	.66 .70 .70	.59 .45 .60	.68 .61 .75	.34 .53 .56	.78 .82 .86	.41 .24 .50	.84 .87 .88	.85 .85 .80
WS_3 Shows I. M. Knowledge													--	.30 .26 .49	.34 .34 .51	.53 .67 .71	.44 .61 .61	.27 .09 .47	.60 .73 .69	.59 .71 .63
WS_4 Analyzes Problems														--	.45 .38 .49	.18 .28 .46	.53 .42 .60	.43 .44 .59	.62 .56 .76	.68 .55 .76
WS_5 Organizes Systematically															--	.16 .26 .51	.69 .71 .64	.21 .20 .45	.66 .63 .66	.69 .54 .65
WS_6 Maintains Controls																--	.20 .40 .48	.22 .40 .50	.31 .55 .52	.33 .59 .57
WS_7 Follows Directions																	--	.39 .27 .53	.78 .84 .84	.76 .75 .78
WS_8 Productivity																		--	.37 .26 .55	.43 .32 .61
WS_9 Quality of Actions																			--	.88 .91 .90
WS_{10} Rating of Overall Performance																				--

Significance	r
Black sample	.25 = significant at .01 level .20 = significant at .05 level
Mexican-American sample	.30 = significant at .01 level .23 = significant at .05 level
Caucasian sample	.18 = significant at .01 level .14 = significant at .05 level

Chapter V

Validity of Aptitude Tests for Different Ethnic Groups

This chapter deals with the validity of the aptitude tests described in Chapter III, as measured by the degree of their relationship with the criteria of job performance described in Chapter IV. The questions to be explored are: 1) are the selected aptitude tests valid predictors of job performance for the occupations studied; 2) do the validity coefficients obtained differ for the separate ethnic groups; and 3) do the validity coefficients differ for the various criterion measures? In the following sections, the relationships between aptitude test scores and 1) supervisors' ratings, 2) job knowledge test scores, and 3) work sample measures will be discussed by occupation and by ethnic group.

Validity of Aptitude Tests Against Supervisors' Ratings

In this section, correlations of aptitude tests with supervisors' ratings on two rating scales used for each occupation studied are presented. These rating scales are: 1) Learning Ability (rating of ability to learn on the job) and 2) Overall (rating of overall effectiveness on the job). The first of these was chosen since ability to learn on the job would appear to be closely related to the purpose of most of the aptitude tests. Correlations of the aptitude tests with the ratings of overall performance will give an indication of the extent to which the abilities measured by the tests relate to effective job performance.[1]

Table V-1 shows the correlations of aptitude tests with supervisors' ratings for Medical Technicians by ethnic group. It can be seen that (1)

[1] Validity coefficients for aptitude tests against supervisors' ratings on all scales for the three occupations studied are shown in Appendix Tables V-A, V-B, V-C, and V-D.

test validities are generally higher when the criterion is learning ability than when it is overall performance and (2) test validities are higher for Blacks than for Caucasians on tests measuring traditional academic areas (verbal and mathematical skills), but higher for Caucasians than for Blacks on tests of spatial perception, finger dexterity, and rote memory. None of the correlations for Blacks is significantly different from the corresponding coefficient for Caucasians.

Table V-2 shows the correlations between aptitude tests and supervisors' ratings for Cartographic Technicians at TOPOCOM installations by ethnic group. Again, the highest coefficients are those against the ratings on learning ability. The vocabulary tests appear to have little or no validity for Caucasians or Mexican-Americans. While the coefficients for Blacks are higher and consistently positive, they do not reach the level of significance. In contrast, the tests of spatial and perceptual abilities, arithmetic, and reasoning have significantly high positive coefficients for all three groups against the learning ability ratings, while against the overall ratings these coefficients are consistently positive, although a number do not reach the significance level. In no instance is a minority group coefficient significantly different from the corresponding coefficient for Caucasians.

Table V-3 shows the correlations between the aptitude tests and the two rating scales for Cartographic Technicians (Coast and Geodetic Survey). The patterns of relationships are quite similar to those found for the TOPOCOM technicians, although the sizes of the correlation coefficients are, in most instances, a little higher. In every instance except one, there is no significant difference between the coefficients for the

Caucasian group and the coefficients for the two minority groups. The exception is the correlation between the Extended Range Vocabulary Test and Learning Ability rating. For Caucasians, the coefficient is .35, for Blacks, -.11. The difference in coefficients is significant at the .01 level.

Table V-4 shows the correlations between aptitude tests and supervisors' ratings for Inventory Management Specialists by ethnic group. Again, correlations of aptitude tests with the Learning Ability rating are generally higher than correlations with the Overall rating, but the differences are not very large for this occupation. Although all of the coefficients are positive, the coefficients for the two vocabulary tests and for the Object-Number (short-term memory) Test are quite low. While much of the work of Inventory Management Specialists is verbal in nature, it appears that beyond some level, additional increments of verbal ability or short-term memory do not affect supervisors' evaluations of job performance. The other aptitude measures have substantial relationships with supervisors' evaluations. In general, the relationships are higher for Blacks and Mexican-Americans than for Caucasians.

Validity of Aptitude Tests Against Job Knowledge Test Scores

For two of the occupations studied, Medical Technician and Cartographic Technician, it was possible to obtain or construct a paper-and-pencil test of job knowledge. In both of these fields, this kind of knowledge is considered essential to job performance, and can be acquired on the job as well as through education or other training. A justification for using a written test format is the fact that use of written materials and communication in writing is an essential part of the job.

Table V-5 shows the correlations between aptitude tests and Job Knowledge Test scores for Medical Technicians. All of the coefficients are positive, and all coefficients for the Black sample are higher than the corresponding coefficients for the Caucasian sample. However, none of the differences between pairs of coefficients reaches the level of significance.

Table V-6 shows the correlations between aptitude tests and Job Knowledge Test scores for Cartographic Technicians at TOPOCOM installations. All but one of the coefficients are positive. Two of the coefficients for the Caucasian sample are significantly different, at the .05 level, from the corresponding coefficient for one of the minority samples. One of these instances is for the Coordination Test, where the coefficient for Caucasian technicians is .21 and the coefficient for Black technicians is -.06. The other is for the Arithmetic Test, where the coefficient is .43 for Caucasian technicians and .62 for Black technicians.

Validity of Aptitude Tests Against Work Sample Measures

Work sample problems, as described in Chapter II, were developed for the Cartographic Technician and Inventory Management Specialist occupations.

Table V-7 shows the correlations between aptitude tests and the Work Sample Composite score (composite of standard scores on three individual problems) for Cartographic Technicians at TOPOCOM installations. All but one of the coefficients are positive. The Object-Number Test, however, has very low validity for all three groups, and the two vocabulary tests have essentially zero validity for the Mexican-American group. For the first vocabulary test, the coefficients for Caucasian and Mexican-American samples are significantly different from each other at the .05 level.

Table V-8 shows the correlations between aptitude tests and the Work

Sample overall score for Inventory Management Specialists. All of these coefficients are positive. However, the Object-Number Test appears to have low validities for all three groups. For two of the tests, Extended Range Vocabulary and Necessary Arithmetic Operations, the coefficients for the Caucasian and Mexican-American groups are significantly different from each other at the .05 level.

Comparison of Validities Across Different Criteria

For comparison of validities across different criterion measures, correlation coefficients between aptitude tests and selected criterion measures have been plotted by occupation and ethnic group. The points have been joined to form what are essentially validity profiles.

Figure V-1 shows profiles of validity coefficients for Medical Technicians for all aptitude tests against the Learning Ability rating and Job Knowledge Test scores for the Black and Caucasian samples. It is readily apparent that the patterns are similar, with only small differences between the two groups at most points. The Black line lies above the Caucasian line at a number of points on the Learning Ability graph and at all points on the Job Knowledge Test graph, indicating higher test validities for the Black group wherever this occurs. In general, however, where a test is valid for one group it is also valid for the other group, and validities for both types of criterion measures are comparable in degree.

Figure V-2 shows profiles of validity coefficients for Cartographic Technicians (TOPOCOM) for all aptitude tests against the Learning Ability rating, Job Knowledge Test scores, and Work Sample Composite scores for the Black, Mexican-American, and Caucasian samples. Here, again, the patterns are similar for all three groups, with few large differences. As pointed out in previous sections, validities for Blacks and Mexican-Americans are higher than for Caucasians in many instances, and tests valid for one group are also valid for the other groups, with few exceptions. The differences in test validity across criterion measures may be interpreted in view of the aspects of job performance being measured.[10] For example, the vocabulary tests have higher validities when the criterion is the Job Knowledge Test than when the Learning Ability rating or Work Sample Composite score are the criteria. A measure of rote memory (Object-Number) has more relationship with the Learning Ability rating and the Job Knowledge Test than with the Work Sample Composite score.

Figure V-3 shows profiles of validity coefficients for Cartographic Technicians (Coast and Geodetic Survey) for all aptitude tests against the Learning Ability rating and Overall Performance rating for the Black and Caucasian samples (supervisors' ratings were the only criteria for this group). The larger differences in validities between the two ethnic samples pointed out in a previous section are easily seen in this figure. In general, validities for the Black sample are higher than for the Caucasian sample.

Figure V-4 shows profiles of validity coefficients for Inventory Management Specialists for all aptitude tests against the Learning Ability rating and the Work Sample Overall score for Black, Mexican-American, and Caucasian samples. The validity profiles for the three samples are very similar for the Learning Ability rating. For the Work Sample Overall score, the Black and Caucasian lines follow each other closely, while the Mexican-American profile reflects the higher validities obtained in general for this sample. The same generalizations may be made, however, that tests valid for one group tend to be valid for another and that validities are generalizable across different types of criteria.

Summary

In the studies of Medical Technicians, Cartographic Technicians, and Inventory Management Specialists, aptitude tests, selected after careful study of the job, were correlated with three different kinds of criteria: Supervisors' Ratings, scores on Job Knowledge tests, and scores on Work Sample tasks. These correlations were computed separately for Black, Mexican-American, and Caucasian job incumbents.

In a few instances, particular tests showed little or no validity for one or more ethnic groups. However, most validities were substantially positive for all ethnic groups, for all criterion measures. The degree of relationship was essentially the same for all three ethnic groups. Out of 229 comparisons between correlation coefficients for the Caucasian sample and the corresponding coefficients for the minority groups on the tables in this Chapter, one such difference was significant at the .01 level and six were significant at the .05 level. This is fewer than would be expected by chance. In five of these seven instances, the higher coefficient was for the minority group.

Table V-1

Correlations Between Aptitude Tests and Supervisors' Ratings for Medical Technicians by Ethnic Group

	Learning Ability Rating		Overall Performance Rating	
Test	Black N=168	Caucasian N=285	Black N=168	Caucasian N=285
Subtraction & Multiplication	.36**	.24**	.29**	.13*
Vocabulary	.20**	.07	.09	.06
Hidden Figures	.02	.12*	.02	.05
Necessary Arithmetic Operations	.34**	.30**	.17*	.16**
Pin Dexterity	.20**	.26**	.10	.12*
Number Comparison	.26**	.16**	.17*	.04
Gestalt Completion	.18*	.19**	.05	.11
Picture Number	.06	.20**	-.02	.15**
Paper Folding	.21**	.28**	.08	.11

* significant at .05 level
** significant at .01 level

Table V-2

Correlations Between Aptitude Tests and Supervisors' Ratings for Cartographic Technicians by Ethnic Group (TOPOCOM Sample)

Test	Learning Ability Rating Black N=101	Learning Ability Rating Mexican-American N=99	Learning Ability Rating Caucasian N=240	Overall Performance Rating Black N=101	Overall Performance Rating Mexican-American N=99	Overall Performance Rating Caucasian N=240
Coordination	.15	.17	.21**	.04	.05	.18**
Hidden Figures	.29**	.41**	.25**	.21*	.29**	.21**
Vocabulary	.17	.01	.03	.19	-.02	.01
Object-Number	.21*	.12	.04	.19	.01	.02
Card Rotations	.28**	.19	.31**	.16	.04	.26**
CS Arithmetic	.42**	.34*	.25**	.31**	.21*	.24**
Map Planning	.33**	.39**	.40**	.24**	.23*	.30**
Surface Development	.41**	.35**	.34**	.28**	.21*	.28**
Maze Tracing Speed	.20*	.33**	.32**	.14	.15	.27**
Following Oral Directions	.32**	.32**	.33**	.18	.15	.25**
Identical Pictures	.33**	.26**	.20**	.21*	.18	.14*
Extended Range Vocabulary	.16	.07	-.05	.17	.03	-.07
Necessary Arithmetic Operations	.32**	.36**	.29**	.25**	.22*	.19**

* significant at .05 level
** significant at .01 level

Table V-3

Correlations Between Aptitude Tests and Supervisors' Ratings for Cartographic Technicians by Ethnic Group (Coast and Geodetic Survey)

Test	Learning Ability Rating: Black N=38	Learning Ability Rating: Caucasian N=50	Overall Performance Rating: Black N=38	Overall Performance Rating: Caucasian N=50
Coordination	.26	.12	.13	.00
Hidden Figures	.40*	.51**	.30	.42**
Vocabulary	.28	.02	.09	.07
Object-Number	.44**	.09	.35*	.07
Card Rotation	.51**	.39**	.37*	.30*
CS Arithmetic	.61**	.48**	.51**	.44**
Map Planning	.44**	.45**	.31	.32*
Surface Development	.53**	.47**	.38*	.40**
Maze Tracing	.21	.45**	.11	.33*
Following Oral Directions	.43**	.52**	.32*	.46**
Identical Pictures	.51**	.52**	.44**	.36**
Extended Range Vocabulary	.35*	-.11	.21	-.01
Necessary Arithmetic Operations	.38*	.51**	.33*	.48**

* significant at .05 level
** significant at .01 level

Table V-4

Correlations Between Aptitude Tests and Supervisors' Ratings for Inventory Managers by Ethnic Group

	Learning Ability Rating			Overall Performance Rating		
	Black	Mexican-American	Caucasian	Black	Mexican-American	Caucasian
Test	N=112	N=72	N=191	N=112	N=72	N=191
Number Comparison	.34**	.33**	.34**	.29**	.10	.28**
Hidden Figures	.37**	.32**	.08	.30**	.25*	.10
Vocabulary	.10	.15	.13	.14	.04	.12
Object-Number	.03	.06	.06	.04	.00	.07
Letter Sets	.31**	.24*	.25**	.32**	-.04	.21**
Nonsense Syllogisms	.26**	.40**	.13	.17	.33**	.10
Subtraction & Multiplication	.35**	.36**	.31**	.39**	.25*	.32**
Extended Range Vocabulary	.04	.15	.18*	.03	.10	.14
Necessary Arithmetic Operations	.36**	.39**	.28**	.34**	.25*	.25**
Following Oral Directions	.35**	.34**	.18*	.22*	.25*	.13
Inference	.30**	.22	.21**	.23*	.13	.20**
FSEE (VA & QR)	.30**	.32**	.28**	.24*	.24*	.22**

* significant at .05 level
** significant at .01 level

Table V-5

Correlations Between Aptitude Tests and Job Knowledge Test for Medical Technicians by Ethnic Group

Test	Black N=168	Caucasian N=297
Subtraction & Multiplication	.34**	.23**
Vocabulary	.32**	.27**
Hidden Figures	.15	.14*
Necessary Arithmetic Operations	.46**	.34**
Pin-Dexterity	.28**	.17**
Number Comparison	.23**	.14*
Gestalt Completion	.25**	.17**
Picture Number	.21**	.16**
Paper Folding	.22**	.21**

* significant at .05 level
** significant at .01 level

Table V-6

Correlations Between Aptitude Tests and Job Knowledge Test for Cartographic Technicians by Ethnic Group (TOPOCOM Sample)

Test	Black N=101	Mexican-American N=99	Caucasian N=241
Coordination	-.06	.18	.21**
Hidden Figures	.41**	.45**	.40**
Vocabulary	.46**	.22*	.39**
Object-Number	.28**	.10	.17**
Card Rotations	.34**	.34**	.33**
CS Arithmetic	.62**	.44**	.43**
Map Planning	.35**	.42**	.36**
Surface Development	.58**	.55**	.54**
Maze Tracing	.36**	.46**	.37**
Following Oral Directions	.48**	.50**	.54**
Identical Pictures	.31**	.40**	.29**
Extended Range Vocabulary	.50**	.23*	.39**
Necessary Arithmetic Operations	.66**	.44**	.54**

* significant at .05 level
** significant at .01 level

Table V-7

Correlations Between Aptitude Tests and Work Sample Composite for Cartographic Technicians by Ethnic Group (TOPOCOM Sample)

Test	Black N=99	Mexican-American N=97	Caucasian N=236
Coordination	.13	.25*	.05
Hidden Figures	.35**	.43**	.29**
Vocabulary	.22*	-.02	.22**
Object-Number	.05	.03	.11
Card Rotations	.26**	.43**	.34**
CS Arithmetic	.30**	.34**	.11
Map Planning	.29**	.39**	.24**
Surface Development	.31**	.38**	.41**
Maze Tracing	.30**	.45**	.34**
Following Oral Directions	.17	.20	.41**
Identical Pictures	.31**	.37**	.26**
Extended Range Vocabulary	.18	.05	.23**
Necessary Arithmetic Operations	.28**	.39**	.35**

* significant at .05 level
** significant at .01 level

Table V-8

Correlations Between Aptitude Tests and Work Sample Overall Score for Inventory Management Specialists by Ethnic Group

Test	Black N=99	Mexican-American N=58	Caucasian N=167
Number Comparison	.17	.36**	.34**
Hidden Figures	.21*	.29*	.30**
Vocabulary	.32**	.41**	.37**
Object-Number	.04	.20	.06
Letter Sets	.28**	.49**	.29**
Nonsense Syllogisms	.29**	.38**	.13
Subtraction & Multiplication	.08	.37**	.13
Extended Range Vocabulary	.28**	.58**	.32**
Necessary Arithmetic Operations	.33**	.60**	.35**
Following Oral Directions	.36**	.41**	.42**
Inference	.39**	.56**	.34**
FSEE (VA + QR)	.37**	.60**	.40**

* significant at .05 level
** significant at .01 level

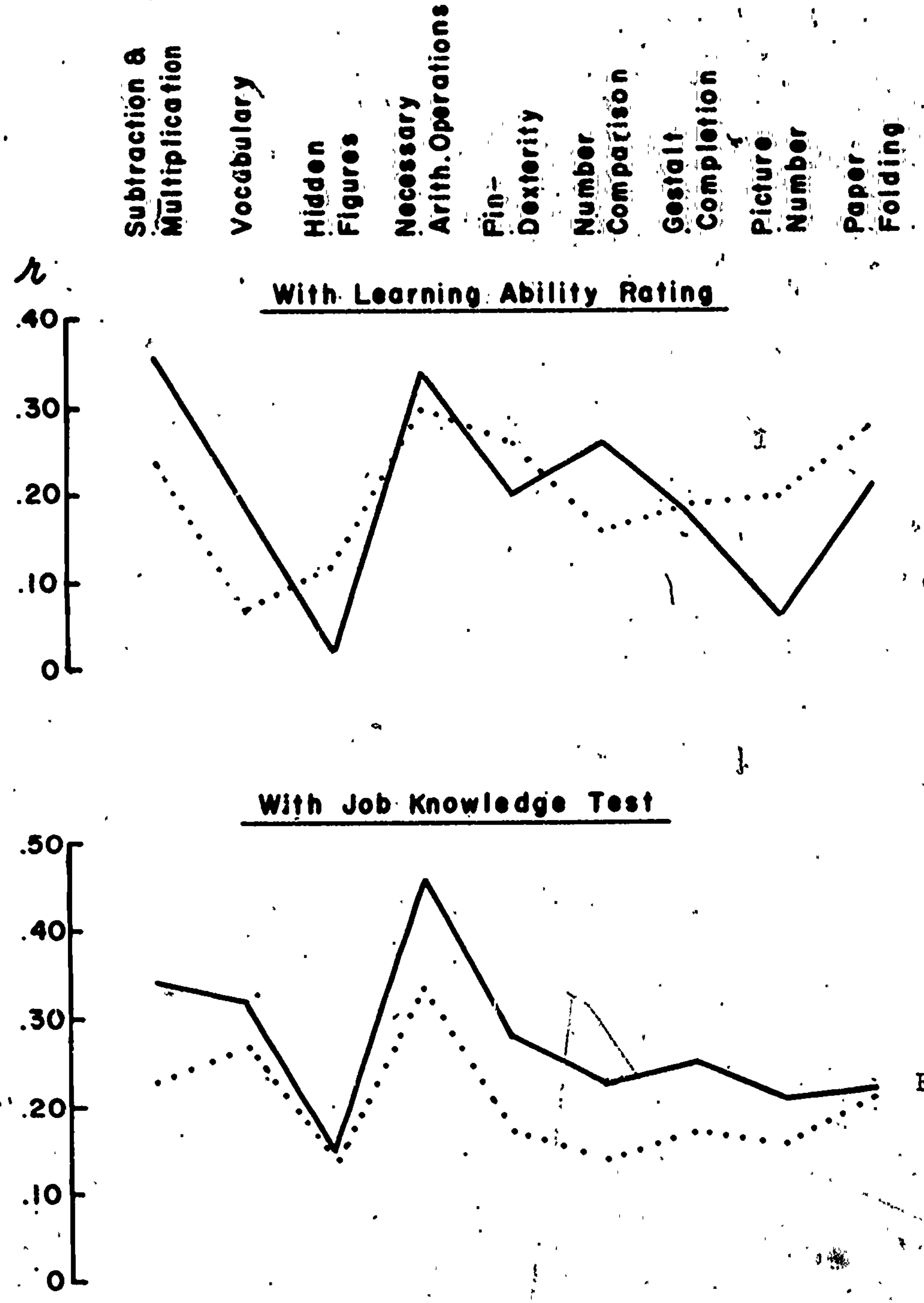


Fig. V-I Validity Coefficients
MEDICAL TECHNICIANS

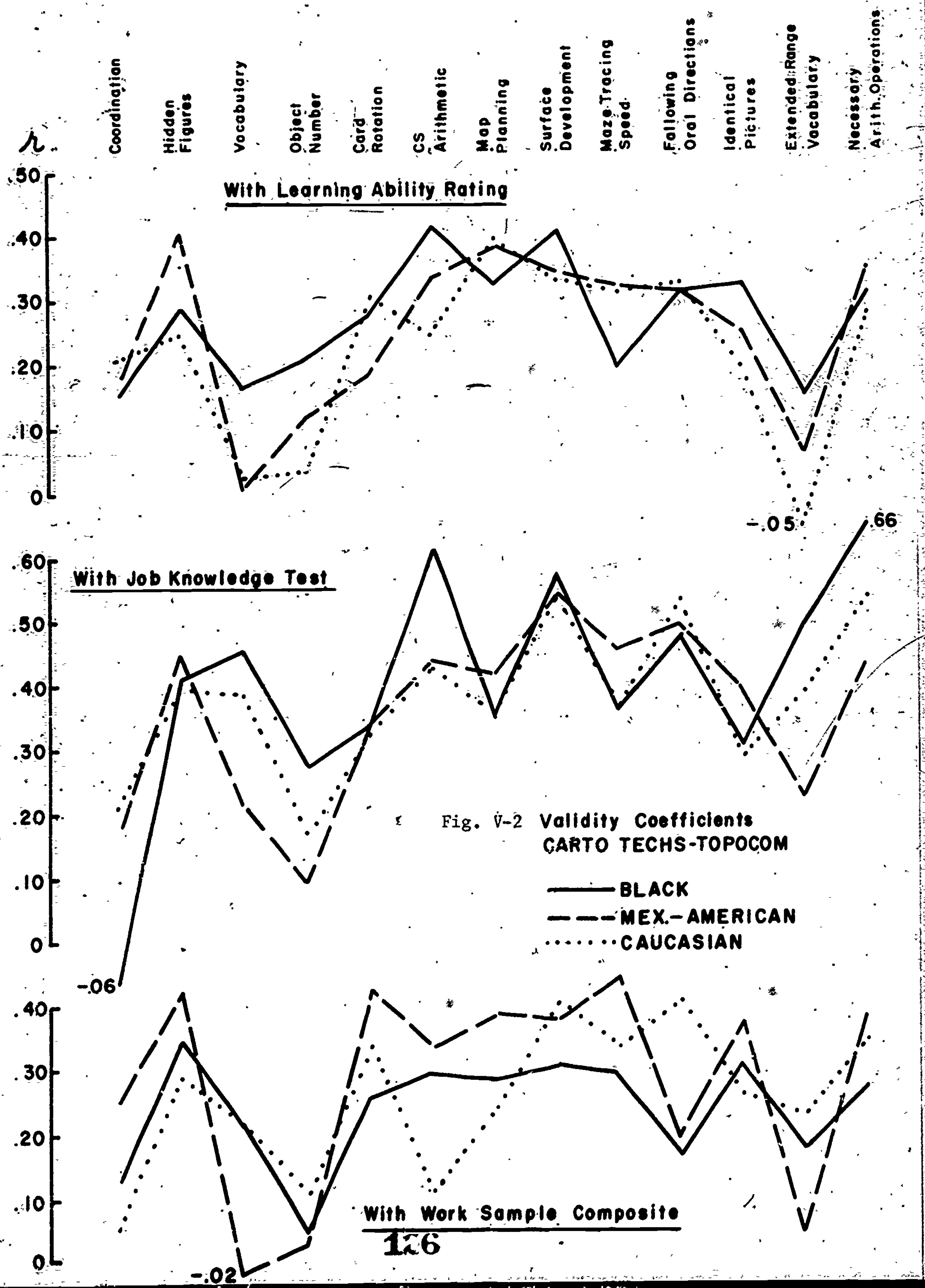


Fig. V-2 Validity Coefficients CARTO TECHS-TOPOCOM

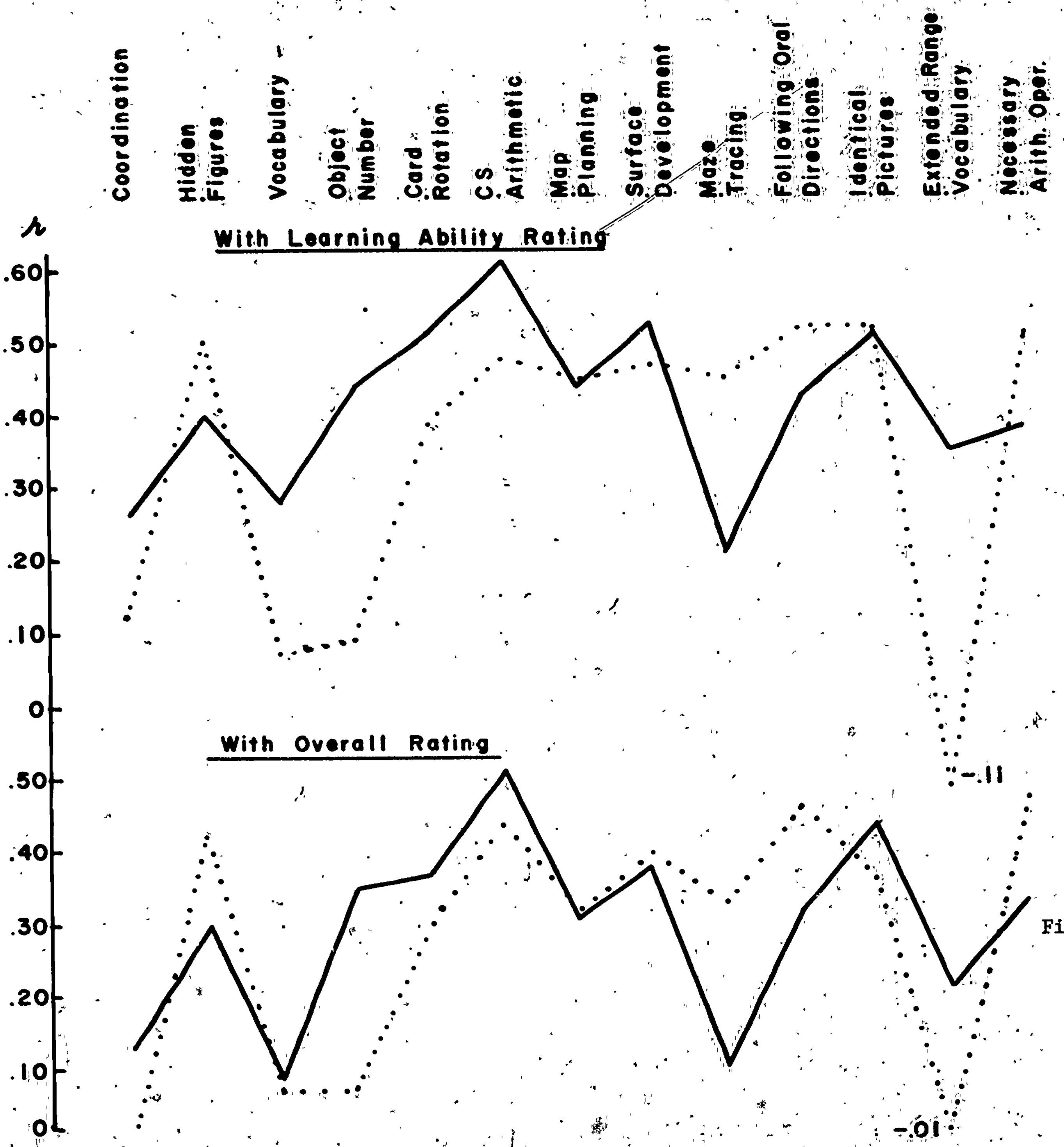


Fig. V-3 Validity Coefficients
CARTO TECHS
(Coast & Geodetic)

——— BLACK
· · · · · · CAUCASIAN

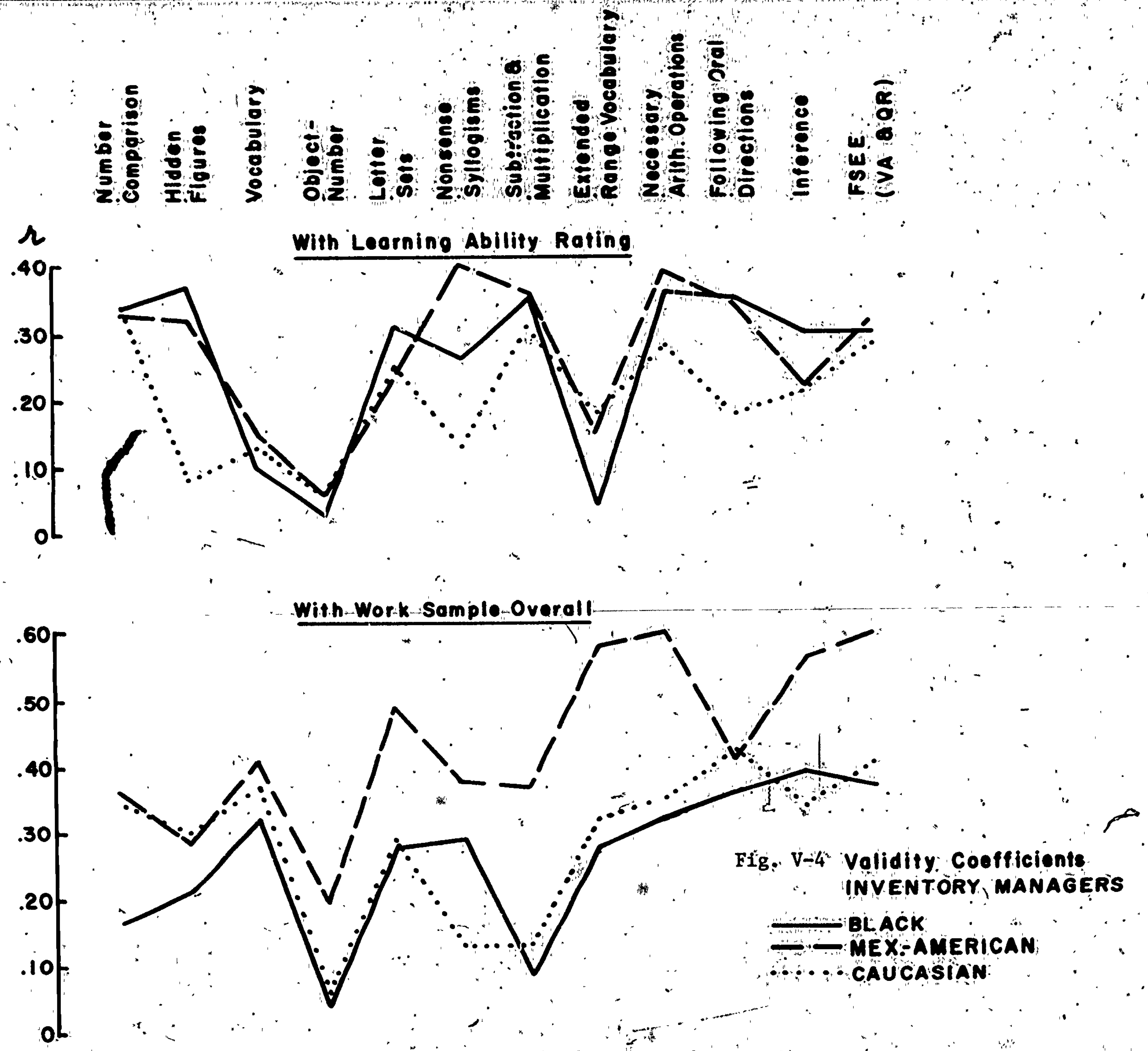


Fig. V-4 Validity Coefficients INVENTORY MANAGERS

Chapter VI

Regression Analysis of Best Predictors of Job Performance Measures for Different Ethnic Groups

Results reported in the previous chapter show that aptitude tests have substantially equal validity for the three ethnic groups, regardless of the type of criterion used. However, the question of test fairness demands that other aspects of the relationships between test scores and measures of job performance be considered as well. Specifically, the regression lines for different ethnic groups must be compared. These lines show the linear relationships between test scores and criterion measures, and if the lines for two ethnic groups differ significantly, this indicates that the same test score is predicting different criterion scores for the two groups.

Significance Tests for Regression Lines

In comparing two regression lines, three different aspects of the relationship are important:

(1) How much scatter is there around the regression lines? (Do the standard errors of estimate differ?)

(2) Do the slopes of the regression lines differ? (Is there a closer relationship for one group than for another? This is almost, but not quite, the same question considered in the previous chapter, since the standard deviations of both variables, as well as the correlation coefficients affect the slope.)

(3) Do the intercepts differ? (Is one regression line generally higher than the other?)

Table VI-1 summarizes the results of this kind of comparison for the three studies in this project. The Gulliksen-Wilks tests for significance of differences between pairs of regression lines involve three successive tests: (1) for differences in scatter or <u>dispersion</u> around the regression lines; (2) if test (1) is not significant, then for differences in <u>slope</u> of the regression lines; and (3) if test (2) is also not significant, for differences in <u>intercept</u>, or general level of the regression line.

Table VI-1 shows that when the criterion is Supervisors' Overall Rating, there are very few differences in regression lines. Out of 72 pairs of regression lines, comparing the line for Caucasians with the line for each minority group separately, four are significantly different at the .05 level or better. Three of these differences are in slope and one in dispersion.

When the criterion is the Job Knowledge Test, on the other hand, out of 35 comparisons, 31 pairs are significantly different. All but two of these are differences in intercept.

With a work sample as the criterion, 38 pairs out of 50 comparisons show significant differences. Of these, 22 differences are in dispersion, three in slope, and 13 in intercept.

In the case of the Inventory Manager samples, a further step was taken. For the first two occupations, investigation of response patterns to the task list items showed that the different ethnic groups were doing essentially the same job. In the case of the Inventory Managers, while comparison of Blacks and Caucasians showed few significant differences in task list responses, comparison of Mexican-Americans and Caucasians showed a substantial number of significant differences.

However, when the Mexican-American Inventory Managers, who were all located at San Antonio, were compared with Caucasians at San Antonio, there were very few task list items showing significant differences in response patterns. Excluding the San Antonio Inventory Managers, comparison of Blacks and Caucasians at the other four installations showed fewer task list items with significant differences than would have been expected by chance. (These findings will be described more fully in Chapter X .) Thus, it seemed appropriate to compare regression lines for Caucasian and Mexican-American Inventory Managers at San Antonio separately from the Caucasian and Black Inventory Managers at the other four installations.

Table VI-2 shows these comparisons. With Supervisors' Overall Rating as the criterion, only two out of 12 pairs of regression lines for Blacks and Caucasians show significant differences at the .05 level or better, both of these in the intercept. There are no significant differences in the regression lines for Mexican-Americans and Caucasians.

With the Work Sample Overall score as the criterion, all 12 pairs of regression lines for Blacks and Caucasians show significant differences, all in dispersion around the regression lines. This reflects the fact that the standard deviation for the Work Sample Overall score is 3.4 for Black Inventory Managers and 4.4 for Caucasian Inventory Managers.

Comparison of Regression Lines by Criterion Measure

Too many regression lines are involved to consider them all individually in this chapter. Figures are presented which show regression lines for those aptitude tests with the highest validity for each type of criterion measure for each ethnic group for the three occupations studied. Some charts will be included where Gulliksen-Wilks tests referred to in the previous section found no significant differences between the pairs of regression lines, as well as some where there were significant differences. (It will be apparent that visual inspection of the regression lines is not sufficient to determine which pairs of regression lines are significantly different.)

Supervisors' Overall Rating as Criterion

Figure VI-1 shows the regression lines for the Necessary Arithmetic Operations Test against Supervisors' Overall Rating for Black and Caucasian Medical Technicians. This test had the highest correlation with Supervisors' Overall Rating for Caucasian Medical Technicians. Since there were no significant differences between the Black and Caucasian regression lines, it can be concluded that a given test score will predict essentially the same criterion score for both groups.

Figure VI-2 shows similar regression lines for the Subtraction and Multiplication Test, which had the highest validity for Black Medical Technicians. Here, the slopes of the two regression lines show a difference significant at the .05 level. If the Caucasian regression line were to be used in selecting Blacks, criterion scores for those with high test scores would be underestimated, while criterion scores would be overestimated for those with low test scores.

Figure VI-3 shows the regression lines for the Map Planning Test against Supervisors' Overall Rating for Black, Mexican-American, and Caucasian Cartographic Technicians. This test had the highest validity for the Caucasian sample. Here, neither the Black nor the Mexican-American regression line is significantly different from the Caucasian regression line, so that it would appear that various test scores have the same meaning in relation to the rating criterion for all three groups.

The same statement can be made about Figure VI-4, which shows the regression lines for the Civil Service Arithmetic Test, the most valid test for Blacks, and about Figure VI-5, which shows the regression lines for the Hidden Figures Test, the most valid test for Mexican-Americans.

Figure VI-6 shows the regression lines for the Subtraction and Multiplication Test against Supervisors' Overall Rating for Black, Mexican-American, and Caucasian Inventory Managers. This test had the highest validity for Blacks and Caucasians. Again, there is no significant difference between the pairs of regression lines, so that the test scores appear to have the same meaning for the various groups.

Figure VI-7 shows the regression lines for the Nonsense Syllogisms Test, which had the highest validity for Mexican-American Inventory Management Specialists. Again, there are no significant differences between the Caucasian regression line and either of the minority group regression lines.

Job Knowledge Test as Criterion

The Necessary Arithmetic Operations Test had the highest correlations with Job Knowledge Test scores for both Black and Caucasian

Medical Technicians. These regression lines are shown in Figure VI-8. In this instance, the slopes of the two lines are significantly different at the .01 level. Thus, if the Caucasian regression line were used to predict Job Knowledge Test scores for Blacks, criterion scores for those Blacks with high test scores would be slightly underestimated, while criterion scores for those with low test scores would be overestimated to a somewhat greater extent.

For Caucasian Cartographic Technicians, three tests had equally high validities against Job Knowledge Test scores. These three tests were the Surface Development Test, which also had the highest validity for Mexican-Americans, Following Oral Directions, and Necessary Arithmetic Operations, which also had the highest validity for Blacks.

Figure VI-9 shows the regression lines for the Surface Development Test for Black, Mexican-American, and Caucasian Cartographic Technicians (TOPOCOM). In this instance, there are no significant differences between the lines for Blacks and Caucasians. There is a significant difference in the intercept between Mexican-Americans and Caucasians, that is, the general level of the regression line above the base line. Thus, if the Caucasian regression line were used to predict criterion scores for Mexican-Americans, their performance would be somewhat overestimated.

Figure VI-10 shows the regression lines for the Necessary Arithmetic Operations Test. Here again, there are no significant differences between the Black and Caucasian regression lines; but between the Mexican-American and Caucasian regression lines there is a difference in intercepts significant at the .01 level. Again, use of the

Caucasian regression line for Mexican-Americans would overestimate their criterion performance.

A similar situation obtains for the Following Oral Directions Test, shown in Figure VI-11. Here again, the Caucasian regression line would overestimate the Mexican-American criterion score.

Work Samples as Criteria

The Hidden Figures Test had the highest correlation with the Work Sample Composite score for Black Cartographic Technicians; the Maze Tracing Speed Test had the highest correlation for Mexican-Americans; and two tests, Surface Development and Following Oral Directions, had equally high correlations for Caucasians.

Figure VI-12 shows the regression lines for the Hidden Figures Test for Black, Mexican-American, and Caucasian Cartographic Technicians. There is a difference in slopes between the Mexican-American and Caucasian regression lines, significant at the .05 level; and a difference in intercepts between the Black and Caucasian regression lines, significant at the .01 level. As can be seen, the Caucasian regression line lies above the regression lines for both minority groups.

Figure VI-13 shows the regression lines for the Maze Tracing Speed Test. Again, the line for Caucasians lies above the lines for the two minority groups. There are no significant differences between the lines for Blacks and Caucasians, but there is a difference in intercepts, significant at the .05 level, between the Mexican-American and Caucasian lines.

Figure VI-14 shows the regression lines for the Surface Development Test. Again, the line for Caucasians lies above the lines for the two minority groups. There are no significant differences between the Black and Caucasian regression lines, but there is a difference in dispersion, significant at the .01 level, between the lines for the Mexican-American and Caucasian samples.

Figure VI-15 shows the regression lines for the Following Oral Directions Test. There is a difference in slope between the Caucasian and Black regression lines, significant at the .05 level. Between the Mexican-American and Caucasian regression lines, there is a difference in dispersion, significant at the .01 level. Although use of the Caucasian regression equation would overestimate the criterion scores for those minority incumbents with high test scores, the criterion scores of low-scoring minorities would be somewhat underestimated.

The most valid test for Black Inventory Managers against Work Sample Overall Performance was the Inference Test; for Mexican-Americans the Necessary Arithmetic Operations Test and the Verbal and Quantitative Reasoning sections of the Federal Service Entrance Examination had equally high validities; and for Caucasians the Following Oral Directions Test was most valid.

Figure VI-16 shows the regression lines for the Inference Test. There are differences in dispersion between the Caucasian regression line and both the Mexican-American (significant at the .05 level) and the Black (significant at the .01 level) regression lines. Of perhaps more interest is the fact that the Mexican-American regression line lies almost entirely above the Caucasian regression line. In this instance, then, the use of the Caucasian regression line for Mexican-Americans would seriously underpredict their criterion scores.

A somewhat similar situation is seen in Figure VI-17, which shows the regression lines for the Federal Service Entrance Examination (V + Q). Again, the differences in dispersion between the Caucasian regression line and the two minority regression lines are significant, at the .01 level between the Blacks and Caucasians and at the .05 level between the Mexican-Americans and Caucasians. And again, the Mexican-American regression line lies almost entirely above the Caucasian regression line.

Figure VI-18 shows the regression lines for the Necessary Arithmetic Operations Test. Again, the regression line for the Mexican-American group lies above that for the Caucasian group, which, in turn, is above that for the Black group. There are significant differences in dispersion, at the .05 level between the Caucasian and Mexican-Americans and at the .01 level between Caucasians and Blacks.

Figure VI-19 shows the regression lines for the Following Oral Directions Test. The relative placement of the regression lines is similar to the configuration in the previous figure. In this instance, there are no significant differences between the regression lines for Caucasians and Mexican-Americans, but between the Blacks and Caucasians there is, again, a difference significant at the .01 level.

It appears, then, from this analysis, that in predicting Work Sample Overall Performance, use of a Caucasian prediction equation would discriminate against one of the minority groups.

Additional Analyses for Inventory Managers by Installation

As indicated earlier, separate regression lines and Gulliksen-Wilks comparisons were made for four subgroups: Mexican-American and

Caucasian Inventory Management Specialists at San Antonio, and Black and Caucasian Inventory Management Specialists at the Philadelphia, Dayton, and Detroit installations.

With Supervisors' Overall Rating as the criterion, the Nonsense Syllogisms Test was most predictive for Mexican-Americans, the Necessary Arithmetic Operations Test was most predictive for San Antonio Caucasians, and the Subtraction and Multiplication Test was most valid for both Blacks and Caucasians at the other installations.

Figure VI-20 shows the regression lines for the four groups on the Necessary Arithmetic Operations Test. There are no significant differences between the Mexican-American and the San Antonio Caucasian regression lines. The regression lines for Blacks and Caucasians at other installations have intercepts significantly different at the .05 level. Here, the Caucasian regression line would underestimate the criterion scores for Blacks.

Figure VI-21 shows the regression lines for the Nonsense Syllogisms Test. Here, there are no significant differences between the regression lines for Mexican-Americans and San Antonio Caucasians or the lines for Blacks and Caucasians at other installations.

Figure VI-22 shows the regression lines for the Subtraction and Multiplication Test. The Mexican-American and San Antonio Caucasian lines are not significantly different. The lines for Blacks and Caucasians at other installations have intercepts different at the .05 level. Again, the Caucasian regression line would underestimate the Supervisors' Overall Ratings received by the Black Inventory Management Specialists.

With Work Sample Overall Performance as the criterion, the Necessary Arithmetic Operations Test and the Federal Service Entrance Examination had equally high validity for the Mexican-American Inventory Managers, the Necessary Arithmetic Operations Test and the Vocabulary Test had equally high validity for the San Antonio Caucasians, the Inference Test had highest validity for Blacks from other installations, and the Following Oral Directions Test had the highest validity for Caucasians from other installations.

Figure VI-23 shows the regression lines for the Inference Test. There are no significant differences between the Mexican-American and San Antonio Caucasian regression lines. The regression lines for Blacks and Caucasians from other installations had dispersions significantly different at the .01 level. In this instance, the Caucasian regression line would overpredict the Black criterion performance.

Figure VI-24 shows the regression lines for the Necessary Arithmetic Operations Test. Again, there are no significant differences between the Mexican-American and San Antonio Caucasian regression lines, and a difference in dispersion, significant at the .01 level, between the regression lines for Blacks and Caucasians at other installations. And again, the use of the Caucasian regression line would overestimate the criterion performance of Black Inventory Managers.

Figure VI-25 shows the regression lines for the Following Oral Directions Test. Again, there are no significant differences in the lines for the Mexican-American and San Antonio Caucasians. There is a difference in dispersion between the regression lines for the Blacks and Caucasians at other installations, significant at the .05 level. The Caucasian regression line would overestimate the criterion performance of Black Inventory Managers.

Figure VI-26 shows the regression lines for the Federal Service Entrance Examination, and Figure VI-27 the regression lines for the Vocabulary Test. For both of these, there are no significant differences between the Mexican-American and the San Antonio Caucasian regression lines. The regression lines for Blacks and Caucasians from other installations have dispersions significantly different at the .01 level. Again, the Caucasian regression line would overpredict the Black criterion performance.

In most of the regression line comparisons that have been made, use of Caucasian regression lines to predict criterion measures for minority groups would result in bias <u>in favor of</u> the minorities; that is, higher criterion scores would be predicted by use of Caucasian regression lines than those predicted from their own regression lines. However, when regression line comparisons are made between Mexican-American Inventory Managers and Caucasian Inventory Managers, with Work Sample Overall Performance as the criterion, the bias is in the other direction. This bias disappears when the comparison is limited to the regression lines for Mexican-Americans and Caucasians from San Antonio.

However, the comparisons between Black and Caucasian Inventory Managers at the Philadelphia, Dayton, and Detroit installations, with Supervisors' Overall Rating as the criterion, now shows a similar bias against the Blacks. This difference reached the significance level for only two of the 12 regression lines.

Contingency Tables for Selected Tests

There is another aspect of the problem of differential prediction which should be considered. A great many, if not most, validity studies are done with supervisors' ratings as the sole criterion, and with a largely, if not exclusively, Caucasian group as the validation sample. What happens, then, when a test, established as valid in that situation, is used to make predictions for other ethnic groups or is compared with other types of criteria?

Tables VI-3, VI-4, VI-5, and VI-6 are responsive to these questions. In each of these tables, the score range for the test in question has been divided into four intervals. Succeeding columns show the mean criterion score for those persons in each of the four intervals. Table VI-3 shows the data for the Medical Technician sample. The Necessary Arithmetic Operations Test, which best predicts the Supervisors' Overall Rating for the Caucasian sample, also produced valid discrimination on that criterion for the Black sample. With Job Knowledge Test scores as the criterion, there is again valid discrimination for both samples.

Table VI-4 shows similar data for Cartographic Technicians from TOPOCOM installations. Here, the Map Planning Test produces valid discrimination on Supervisors' Overall Rating for Black and Mexican-American technicians, as well as for Caucasian technicians, although those in the fourth test score interval have slightly higher supervisors' ratings than those in the third.

Table VI-5 shows the relationship between the Map Planning Test scores and the Supervisors' Overall Rating for Cartographic Technicians in the Coast and Geodetic Survey sample. On the whole, valid discrimination was produced for both the Caucasian and Black technicians. There is, however, one very substantial reversal (for Black technicians in the highest score interval), but there were only two individuals in this category.

Table VI-6 shows comparable information for Inventory Management Specialists. Here, the most valid test for Caucasian Inventory Managers against the Supervisors' Overall Rating was the Subtraction and Multiplication Test. Generally valid discriminations are made for all three ethnic groups and for both criteria. There are, however, five reversals. One involves the Supervisors' Overall Rating and is relatively small. The other four are for the Work Sample, two for Blacks, and one each for Mexican-Americans and Caucasians. Three of these reversals are fairly large.

Despite these discrepancies, however, the general picture is of valid discrimination, regardless of ethnic group or criterion used.

Summary

Regression lines between aptitude tests and three different kinds of criteria have been compared for Caucasians and the two minority groups separately. With Supervisors' Overall Rating as the criterion, very few of the regression lines were significantly different; that is, essentially the same predictions were made for all groups.

There were more instances of significant differences in regression lines when the Job Knowledge Tests or the Work Samples were used as criteria. The "bias" in almost all instances was in favor of the minority groups rather than against them.

The general picture that emerges is one of tests either making unbiased predictions or of showing bias in favor of the minority groups.

Table VI-1

Comparison of Regression Lines for Different Pairs of Ethnic Groups in Three Occupations by Criterion

Criterion	Occupation	Samples Compared	Number of Regression Lines: No Difference	Number of Regression Lines: Difference*	Location of Difference
Supervisors' Overall Rating	Med. Tech.	Black/Caucasian	8	1	Slope
Supervisors' Overall Rating	Carto. Tech.	Black/Caucasian (TOPOCOM)	12	1	Slope
Supervisors' Overall Rating	Carto. Tech.	Mexican-American/ Caucasian (TOPOCOM)	13	0	
Supervisors' Overall Rating	Carto. Tech.	Black/Caucasian (C & G)	11	1 1	Slope Dispersion
Supervisors' Overall Rating	Inv. Mgr.	Black/Caucasian	12	0	
Supervisors' Overall Rating	Inv. Mgr.	Mexican-American/ Caucasian	12	0	
Job Knowledge Test	Med. Tech.	Black/Caucasian	0	1 8	Slope Intercept
Job Knowledge Test	Carto. Tech.	Black/Caucasian (TOPOCOM)	4	9	Intercept
Job Knowledge Test	Carto. Tech.	Mexican-American/ Caucasian (TOPOCOM)	0	1 12	Dispersion Intercept
Work Sample Composite	Carto. Tech.	Black/Caucasian (TOPOCOM)	3	9 1	Intercept Slope
Work Sample Composite	Carto. Tech.	Mexican-American/ Caucasian (TOPOCOM)	0	7 2 4	Dispersion Slope Intercept
Work Sample Overall	Inv. Mgr.	Black/Caucasian	0	12	Dispersion
Work Sample Overall	Inv. Mgr.	Mexican-American/ Caucasian	9	3	Dispersion

* Significant at .05 level or better

Table VI-2

Comparison of Regression Lines for Inventory Management Specialists

by Installation by Criterion

Criterion	Occupation	Samples Compared	Number of Regression Lines: No Difference	Number of Regression Lines: Difference*	Location of Difference
Supervisors' Overall Rating	Inventory Manager	Black and Caucasian at Detroit, Phila., and Dayton Installations	10	2	Intercept
Supervisors' Overall Rating	Inventory Manager	Mexican-American and Caucasian at San Antonio	12	0	
Work Sample Overall	Inventory Manager	Black and Caucasian at Detroit, Phila., and Dayton Installations	0	12	Dispersion
Work Sample Overall	Inventory Manager	Mexican-American and Caucasian at San Antonio	12	0	

* Significant at .05 level or better

	Error of Estimate	Slope	Intercept
Black w/ Cauc.	0.02	0.04	0.00

Medical Technicians

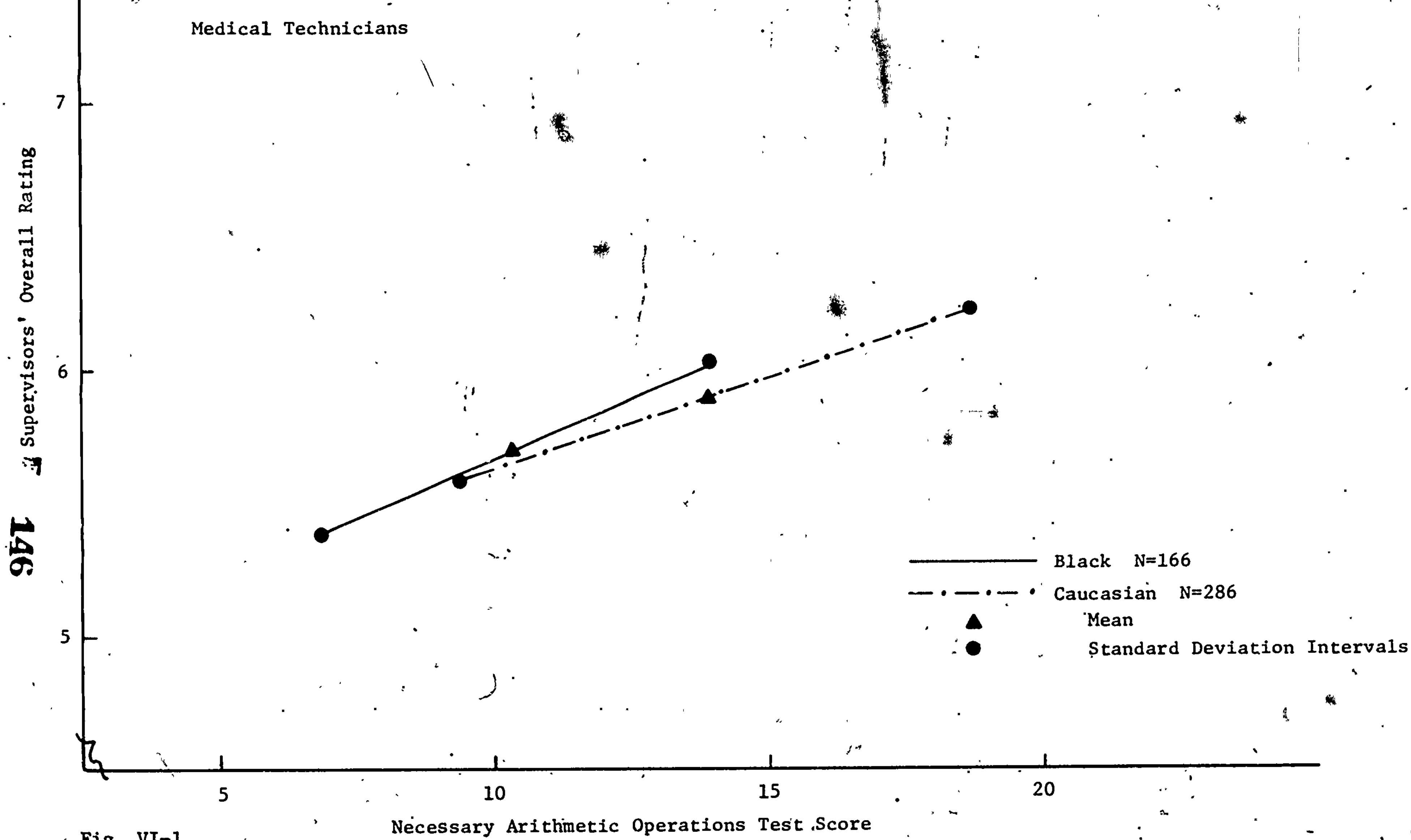


Fig. VI-1.

	Error of Estimate	Slope	Intercept
Black w/ Cauc.	0.09	3.99*	0.11

* $p < .05$

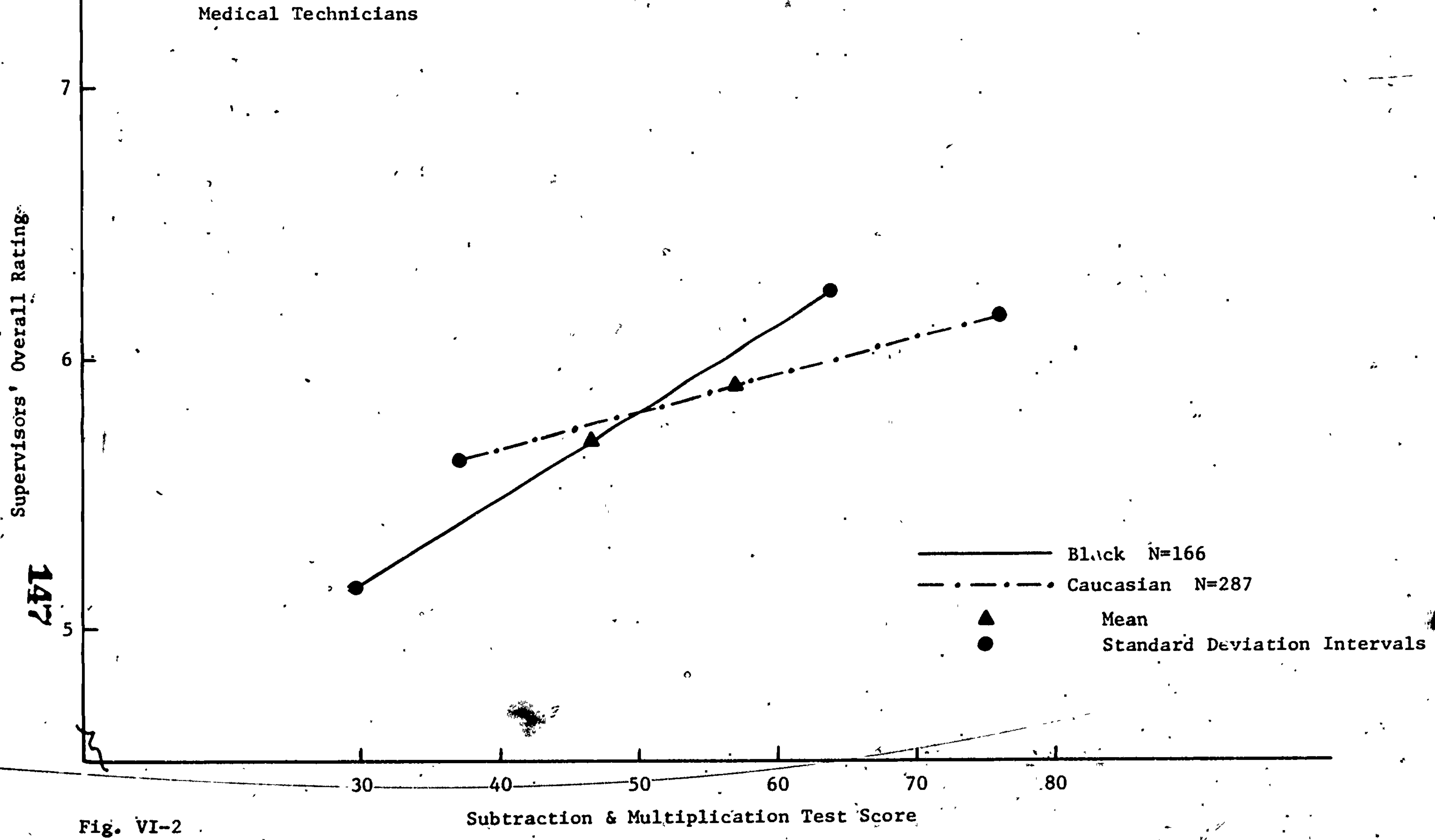


Fig. VI-2

Cartographic Technicians

	Error of Estimate	Slope	Intercept
Black w/ Cauc.	0.99	0.18	1.84
M-A w/ Cauc.	0.21	0.52	0.10

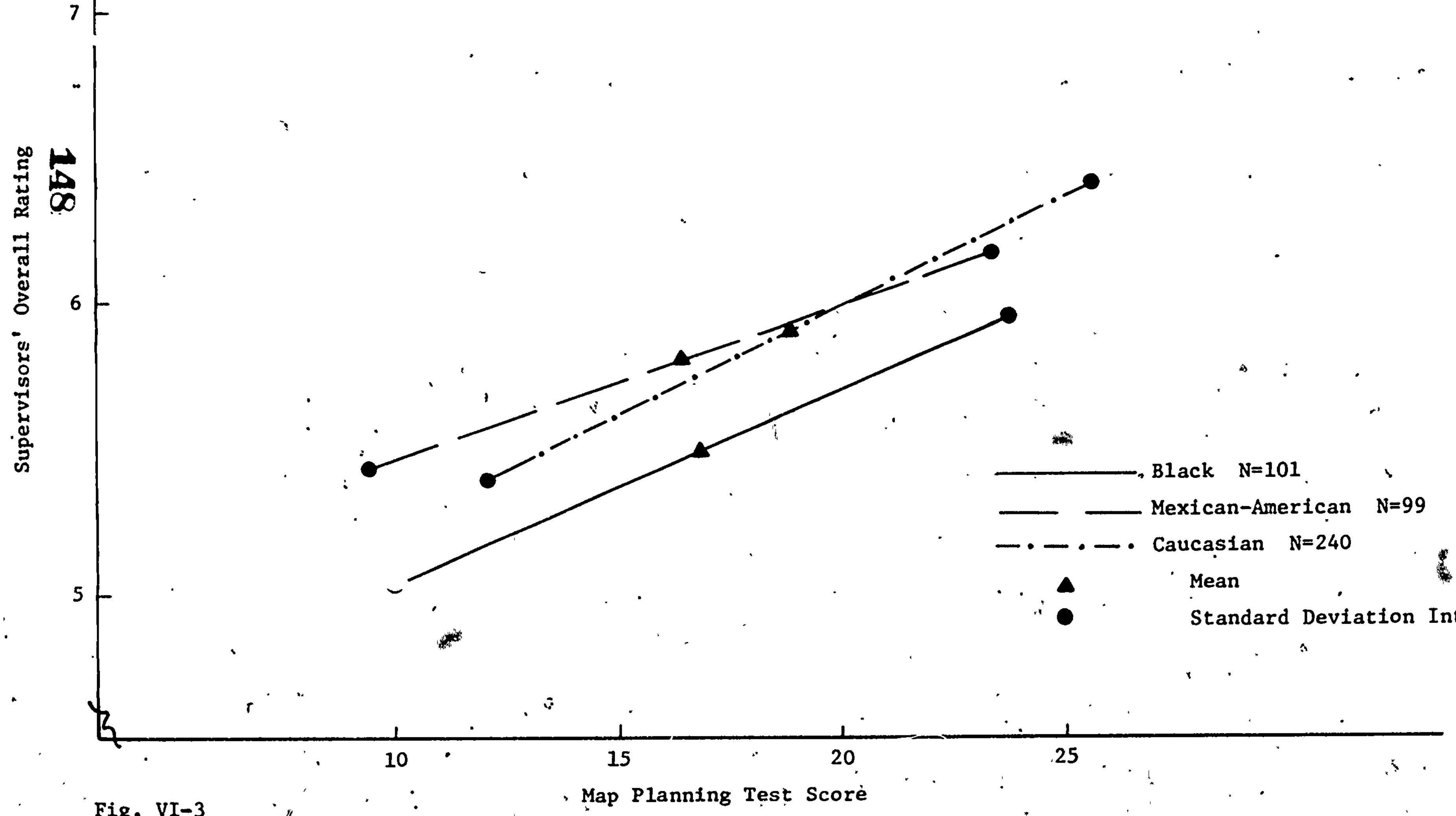


Fig. VI-3

	Error of Estimate	Slope	Intercept
Black w/ Cauc.	0.28	0.61	0.39
M-A w/ Cauc.	0.36	0.07	0.06

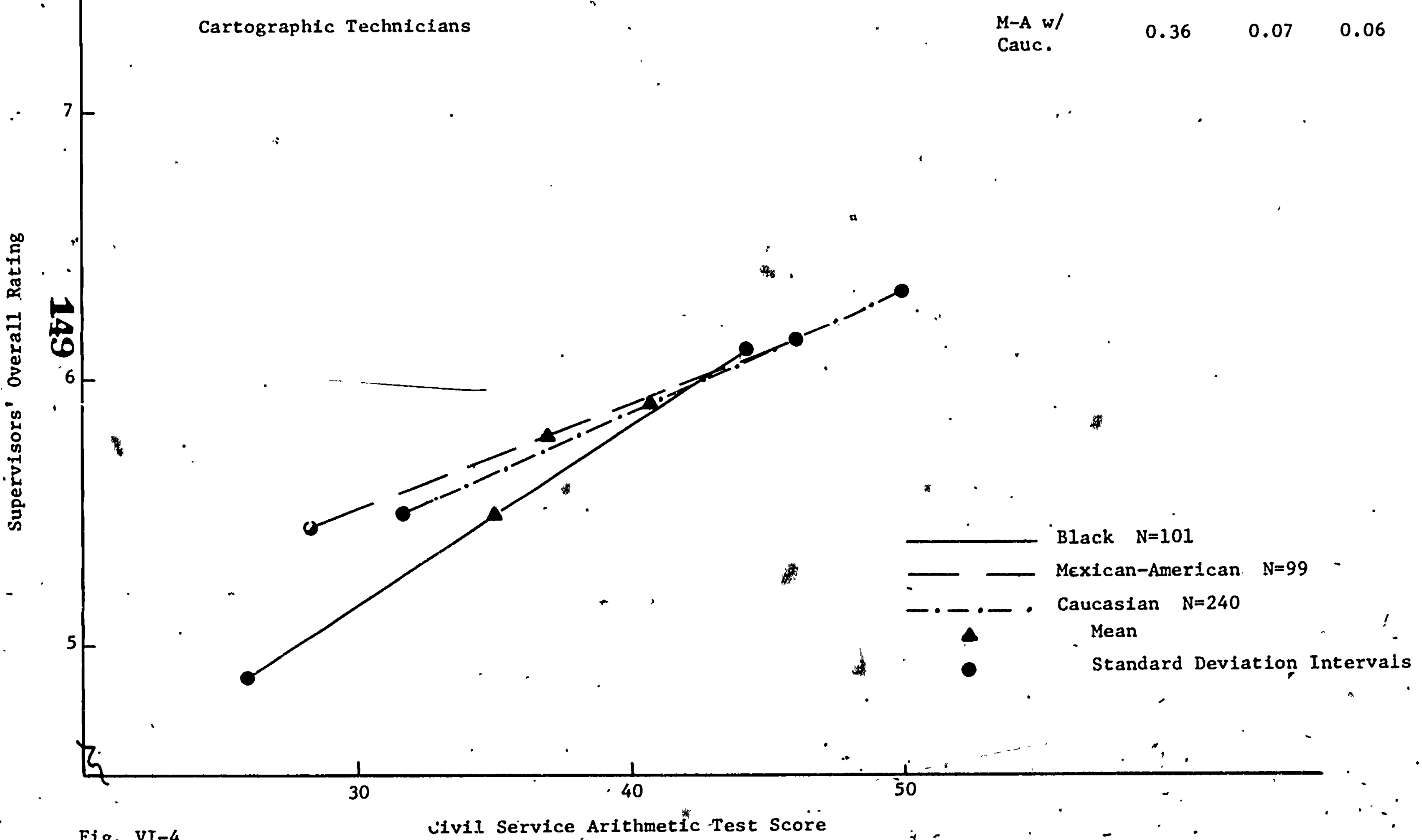


Fig. VI-4

Cartographic Technicians

	Error of Estimate	Slope	Intercept
Black w/ Cauc.	0.68	0.25	2.92
M-A w/ Cauc.	0.97	1.16	0.03

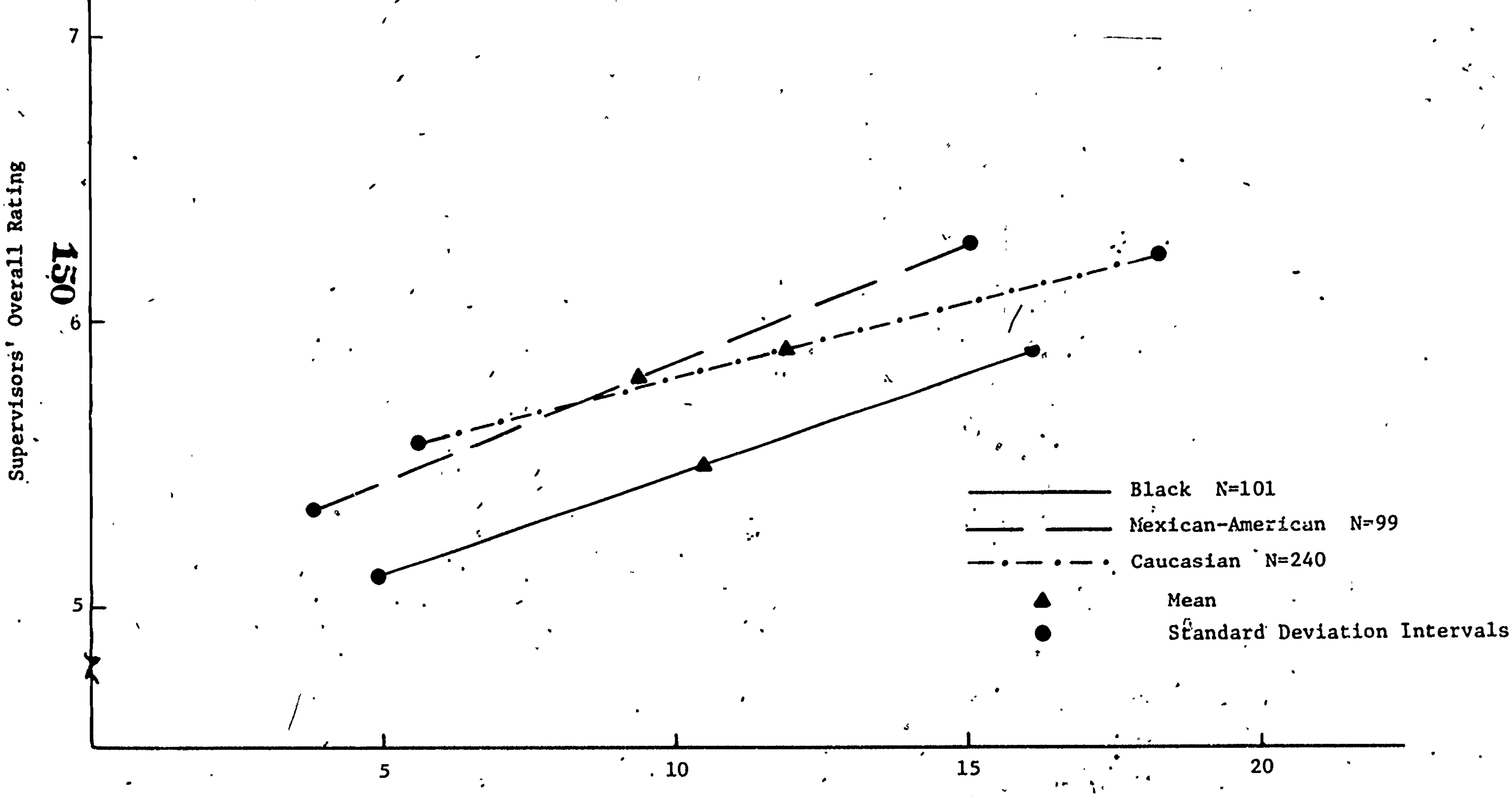


Fig. VI-5

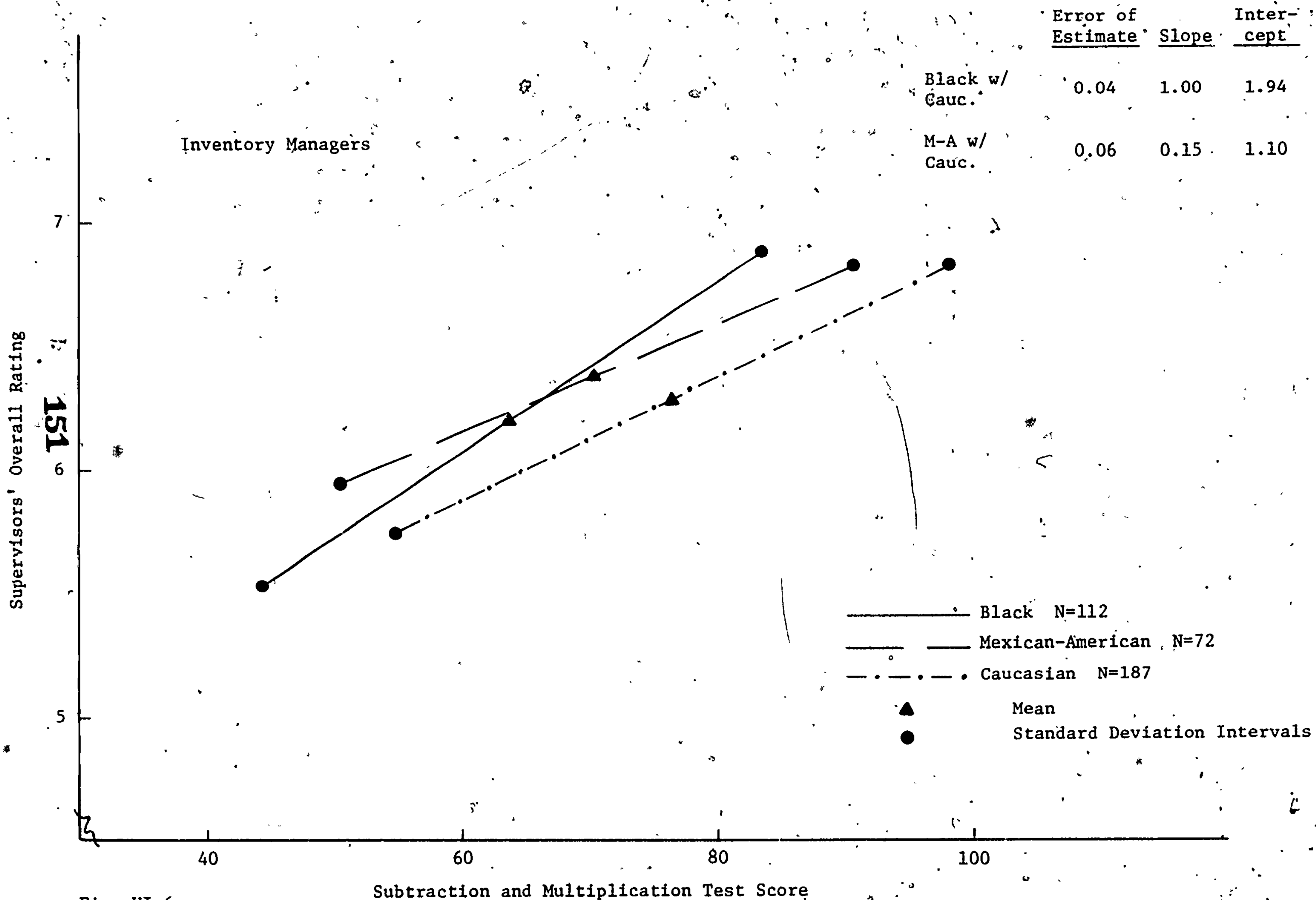

Inventory Managers
Error of Estimate
Slope
Intercept
Black w/ Cauc.
0.04
1.00
1.94
M-A w/ Cauc.
0.06
0.15
1.10
Black N=112
Mexican-American N=72
Caucasian N=187
Mean
Standard Deviation Intervals
Supervisors' Overall Rating
7
6
5
40
60
80
100
Subtraction and Multiplication Test Score

Fig. VI-6

	Error of Estimate	Slope	Intercept
Black w/ Cauc.	0.00	0.50	0.07
M-A w/ Cauc.	0.26	3.79	0.58

Inventory Managers

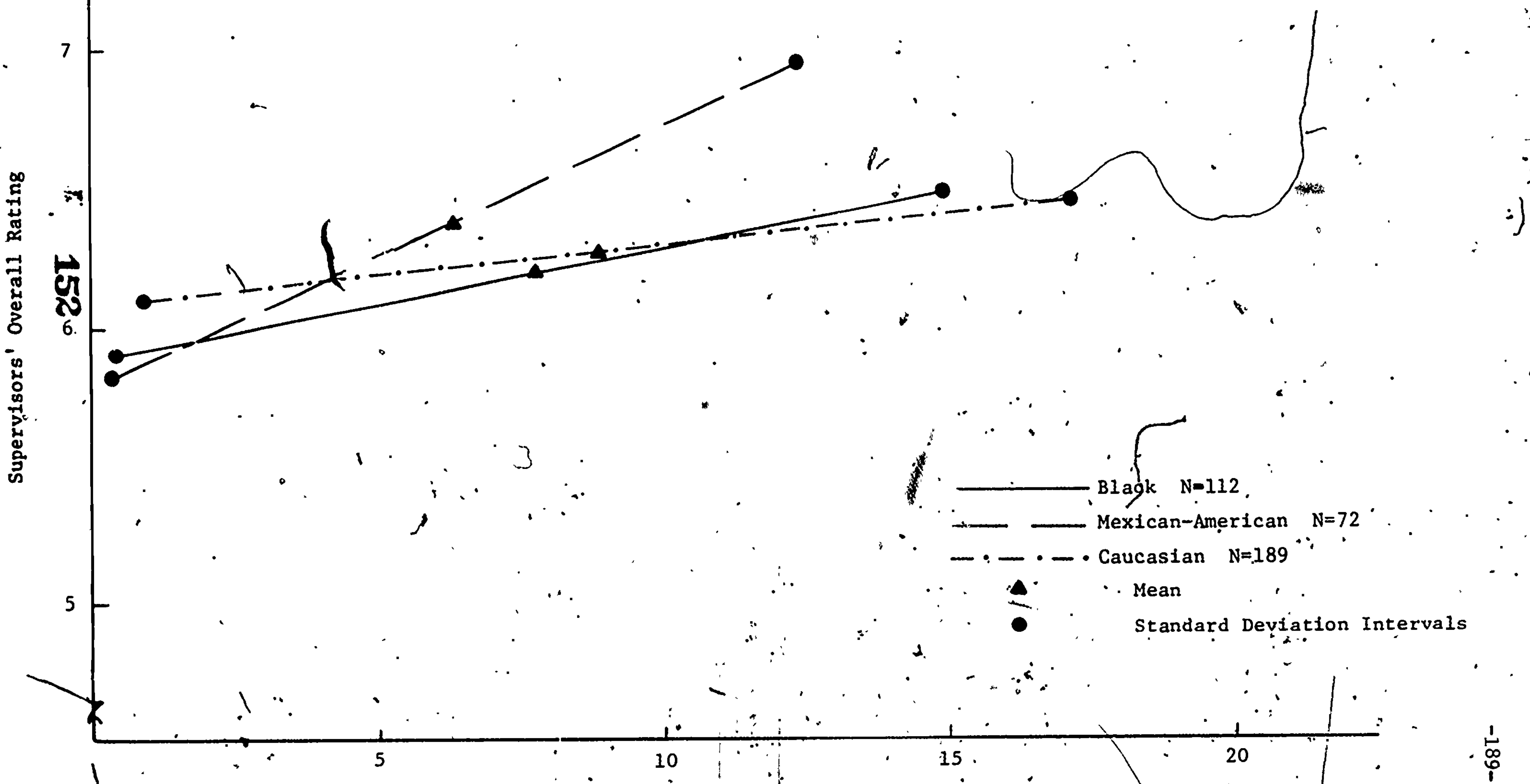


Fig. VI-7.

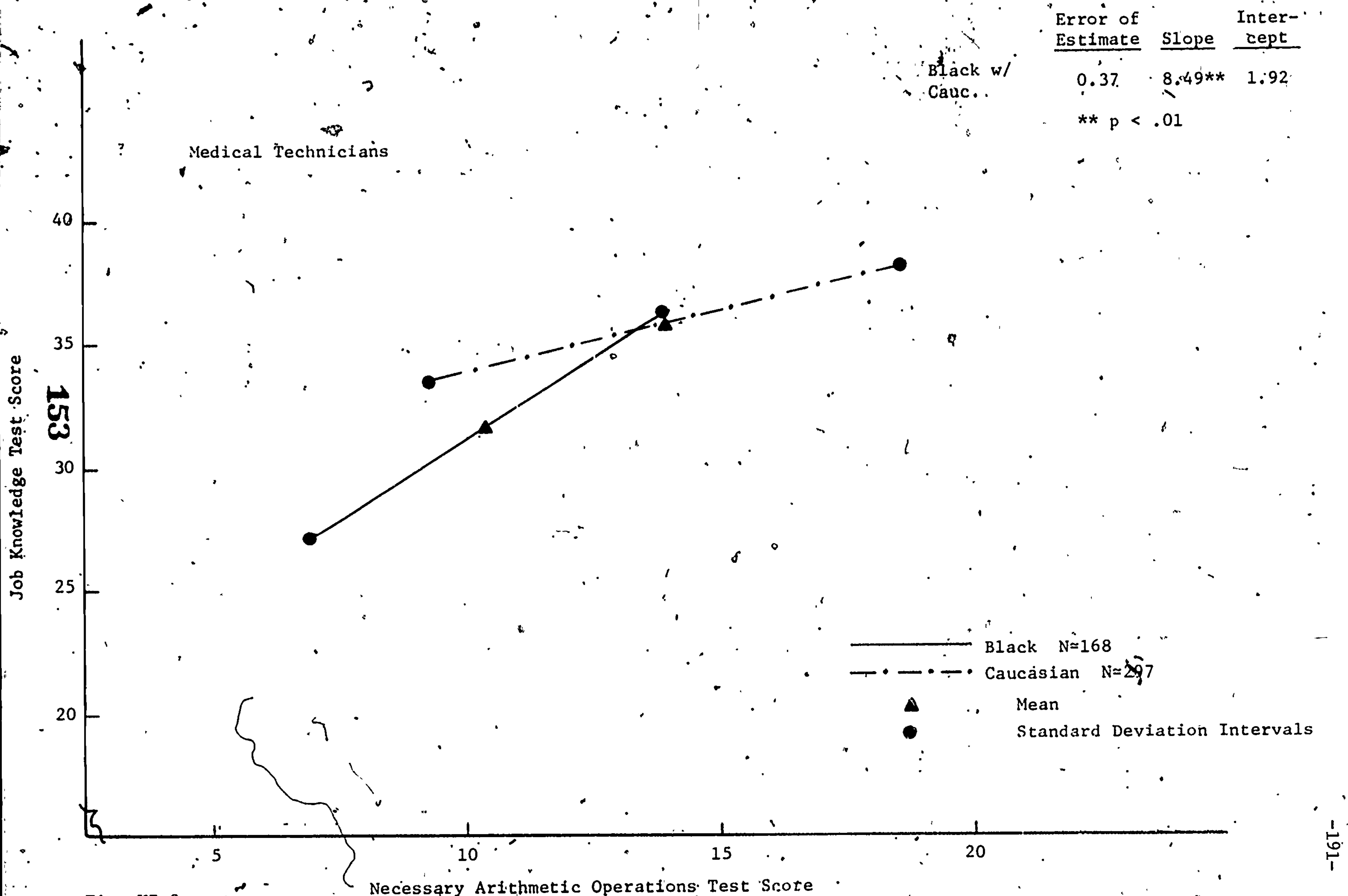

Medical Technicians
Error of Estimate
Slope
Intercept
Black w/ Cauc.
0.37
8.49**
1.92
** p < .01
Black N=168
Caucasian N=297
Mean
Standard Deviation Intervals
Job Knowledge Test Score
40
35
30
25
20
5
10
15
20
Necessary Arithmetic Operations Test Score

Fig. VI-8

Cartographic Technicians

	Error of Estimate	Slope	Intercept
Black w/ Cauc.	0.18	0.21	2.65
M-A w/ Cauc.	2.90	0.00	26.46**

** $p < .01$

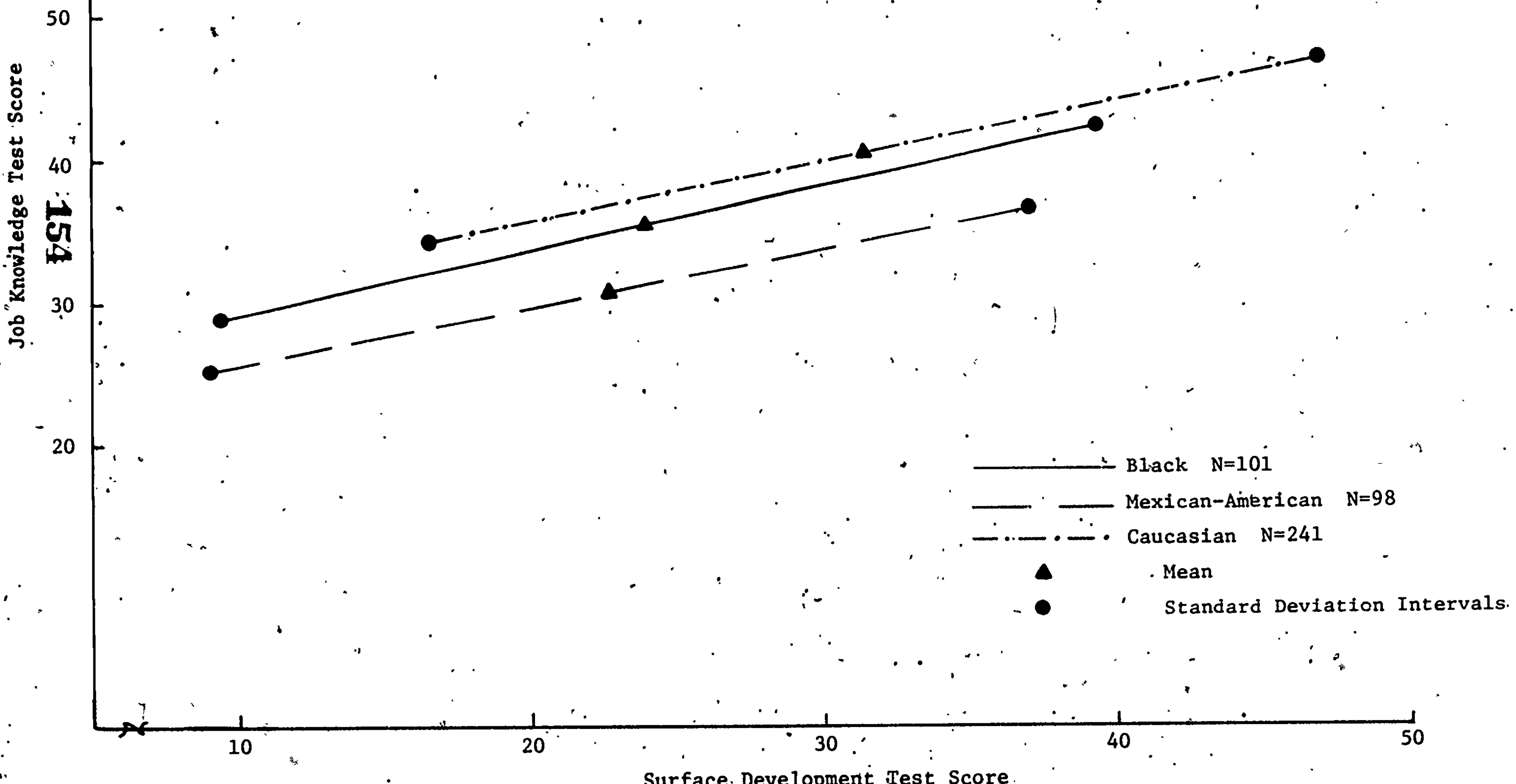


Fig. VI-9

Cartographic Technicians

	Error of Estimate	Slope	Intercept
Black w/ Cauc.	2.24	3.17	0.05
M-A w/ Cauc.	0.80	1.26	20.16**

** $p < .01$

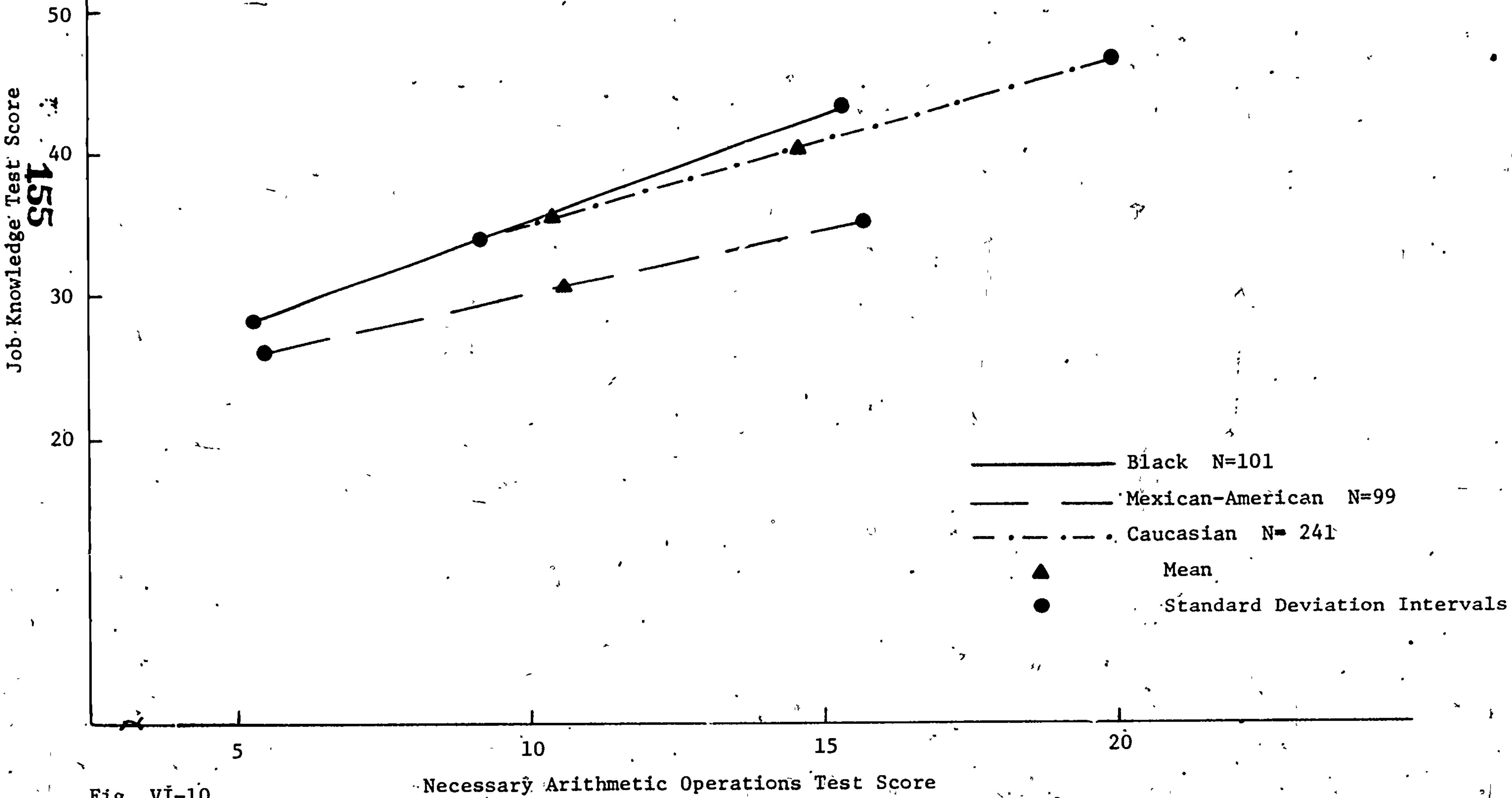


Fig. VI-10

Cartographic Technicians

	Error of Estimate	Slope	Intercept
Black w/ Cauc.	0.17	0.23	1.84
M-A w/ Cauc.	1.23	0.31	4.42*

* $p < .05$

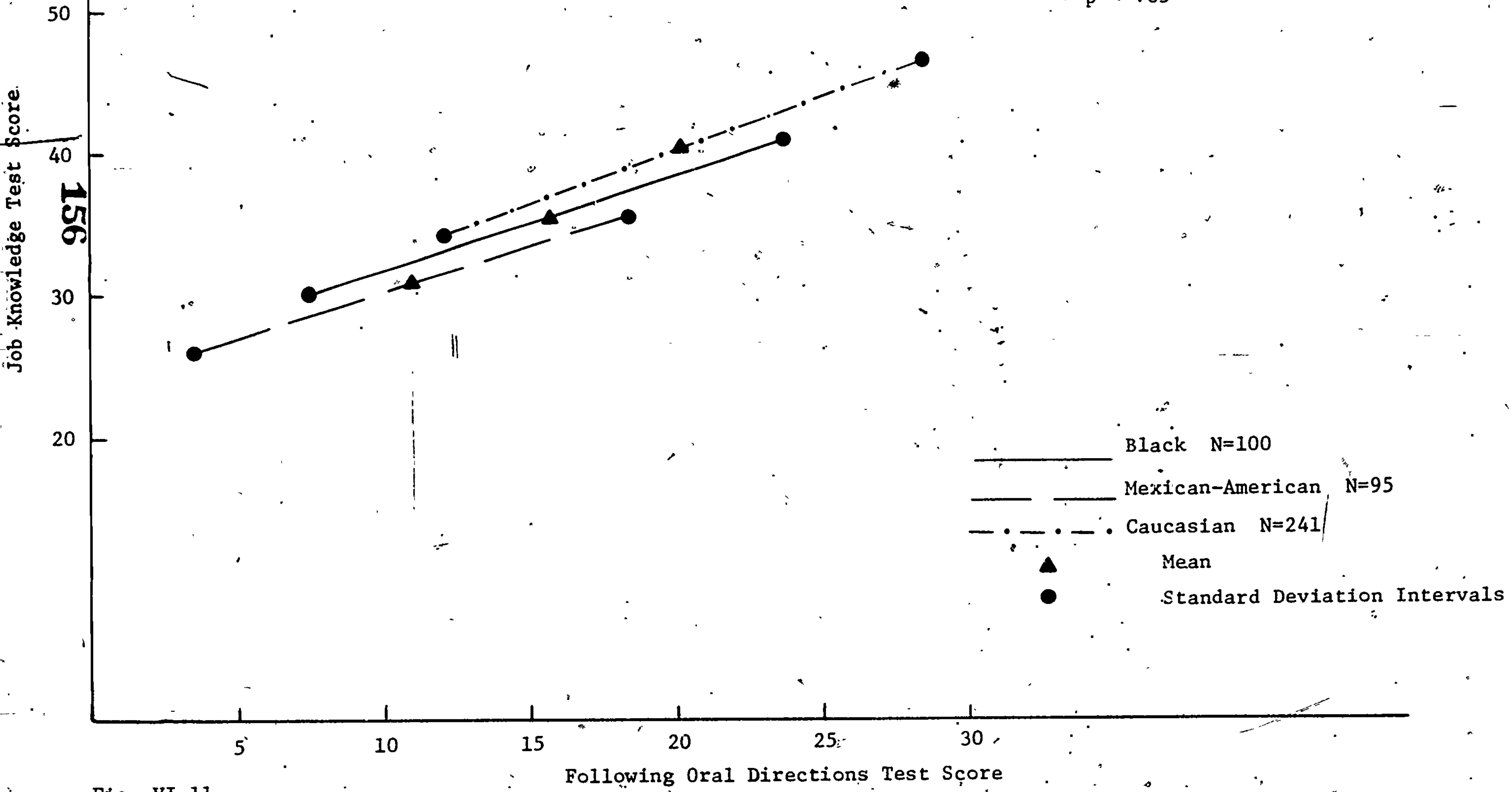


Fig. VI-11

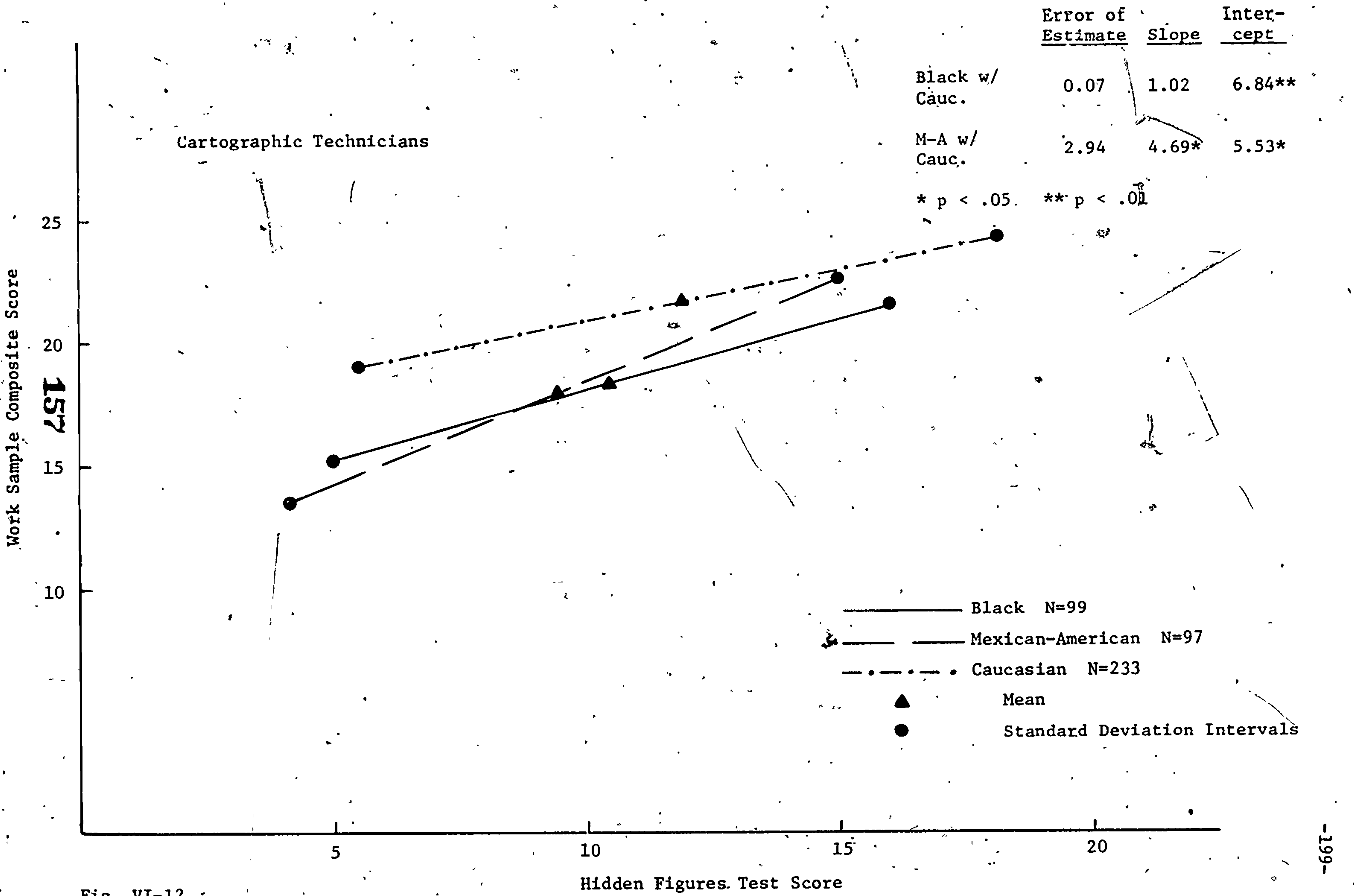
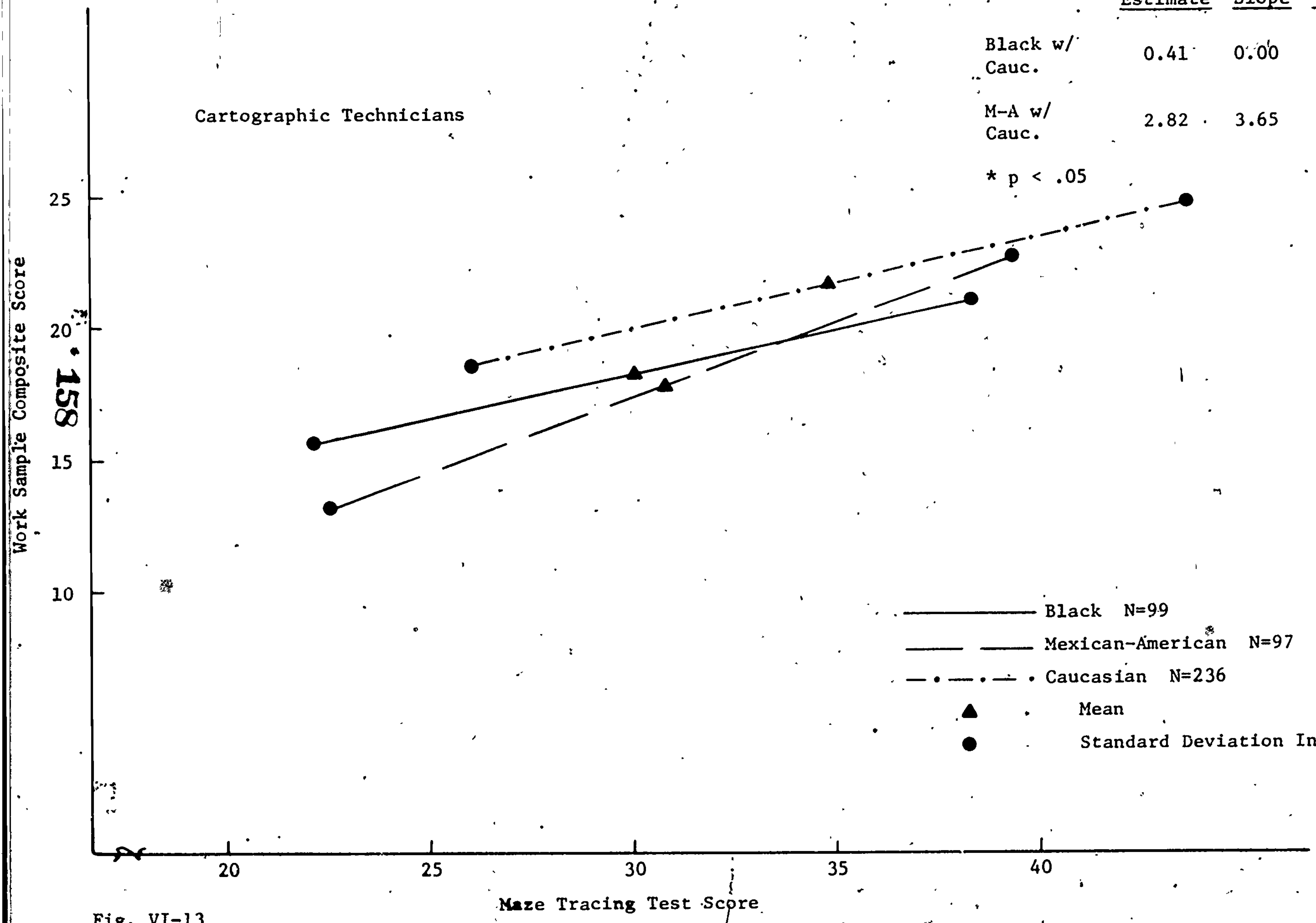

Cartographic Technicians
Error of Estimate
Slope
Inter-cept
Black w/ Cauc. 0.07 1.02 6.84**
M-A w/ Cauc. 2.94 4.69* 5.53*
* p < .05 ** p < .01
Work Sample Composite Score
25
20
15
10
5
10
15
20
Hidden Figures Test Score
Black N=99
Mexican-American N=97
Caucasian N=233
Mean
Standard Deviation Intervals

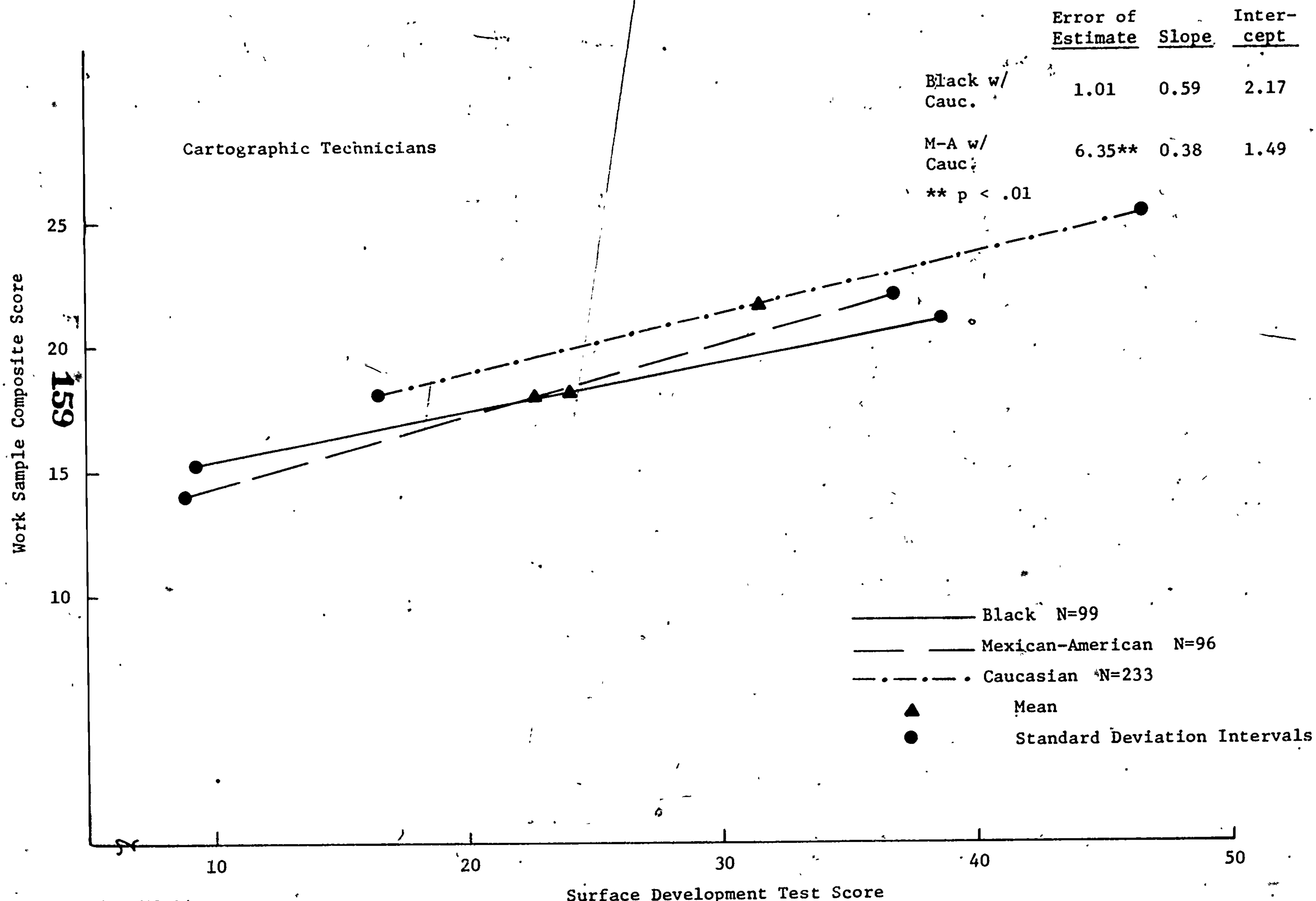
Fig. VI-12



	Error of Estimate	Slope	Inter-cept
Black w/ Cauc.	0.41	0.00	2.62
M-A w/ Cauc.	2.82	3.65	4.46*

* $p < .05$

Fig. VI-13



Cartographic Technicians
Error of Estimate
Slope
Intercept
Black w/ Cauc.
1.01
0.59
2.17
M-A w/ Cauc.
6.35**
0.38
1.49
** p < .01
25
20
15
10
Work Sample Composite Score
10
20
30
40
50
Surface Development Test Score
Black N=99
Mexican-American N=96
Caucasian N=233
Mean
Standard Deviation Intervals

Fig. VI-14

Cartographic Technicians

	Error of Estimate	Slope	Inter-cept
Black w/ Cauc.	1.32	4.28*	1.57
M-A w/ Cauc.	10.47**	1.41	0.00

* $p < .05$ ** $p < .01$

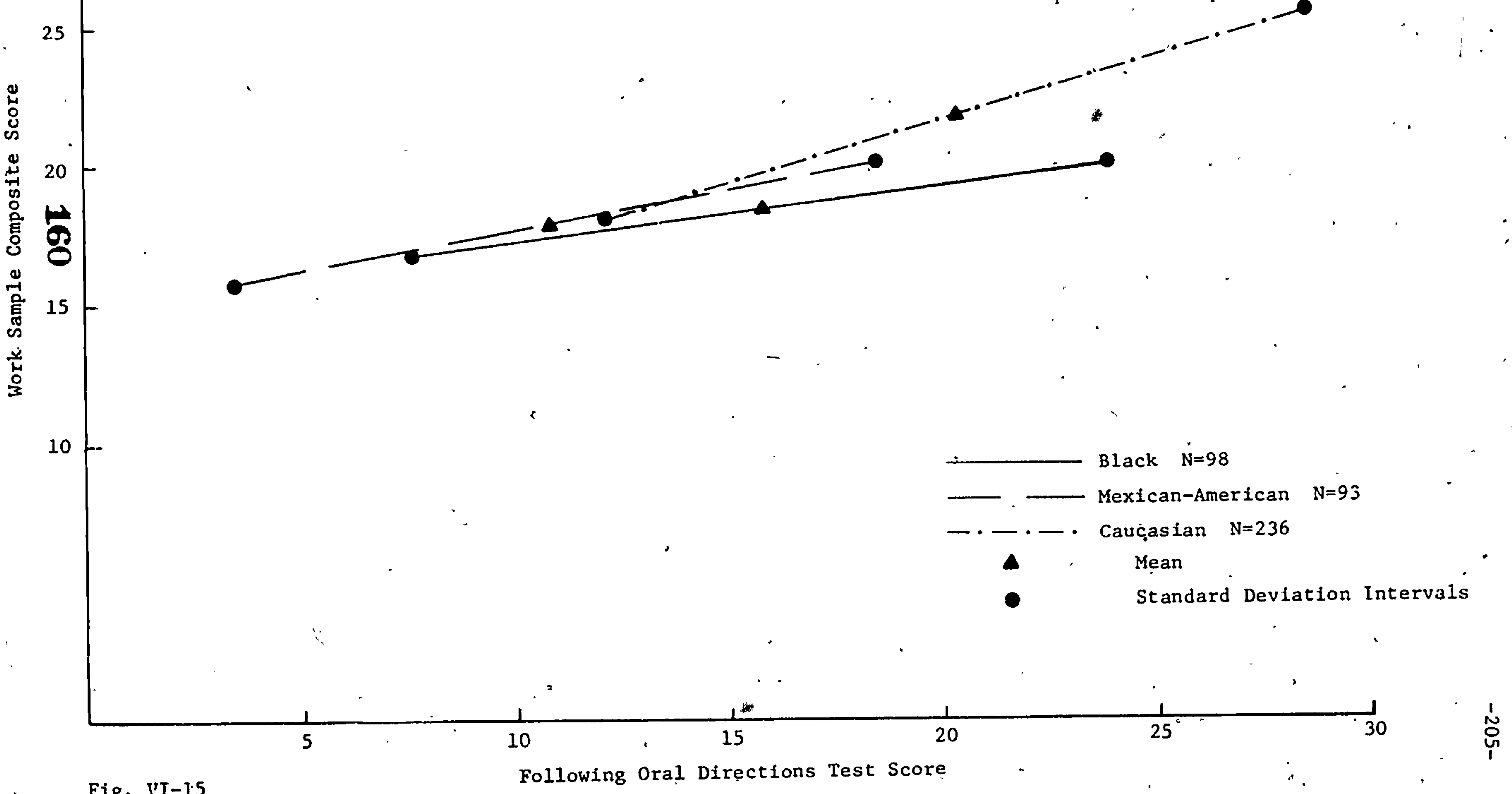


Fig. VI-15

Inventory Managers

	Error of Estimate	Slope	Inter-cept
Black w/ Cauc.	11.40**	0.04	7.27**
M-A w/ Cauc.	3.79*	2.90	2.63

* p < .05 ** p < .01

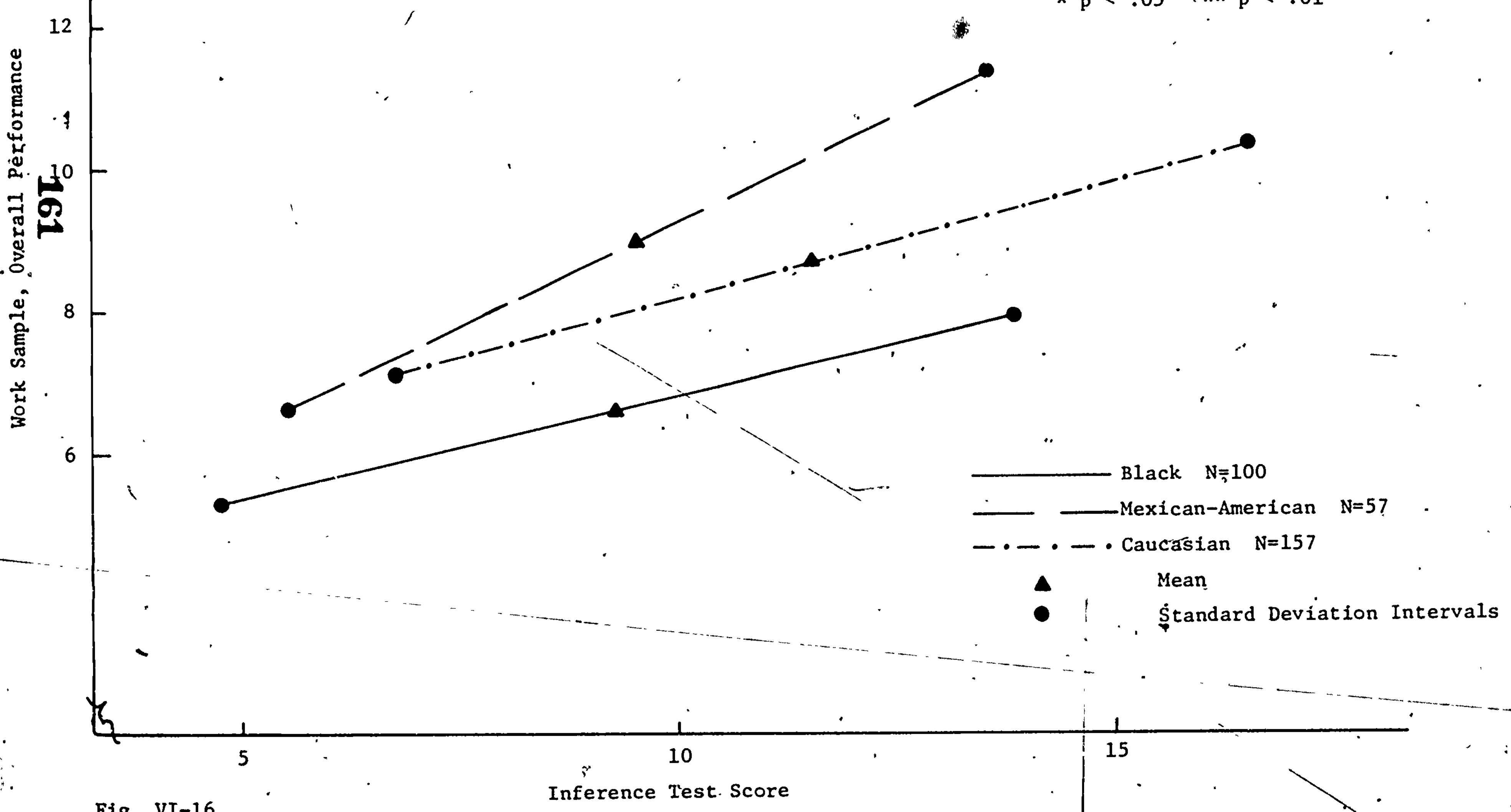


Fig. VI-16

	Error of Estimate	Slope	Intercept
Black w/ Cauc.	9.48**	0.80	5.15*
M-A w/ Cauc.	4.08*	3.28	3.74

* p < .05 ** p < .01

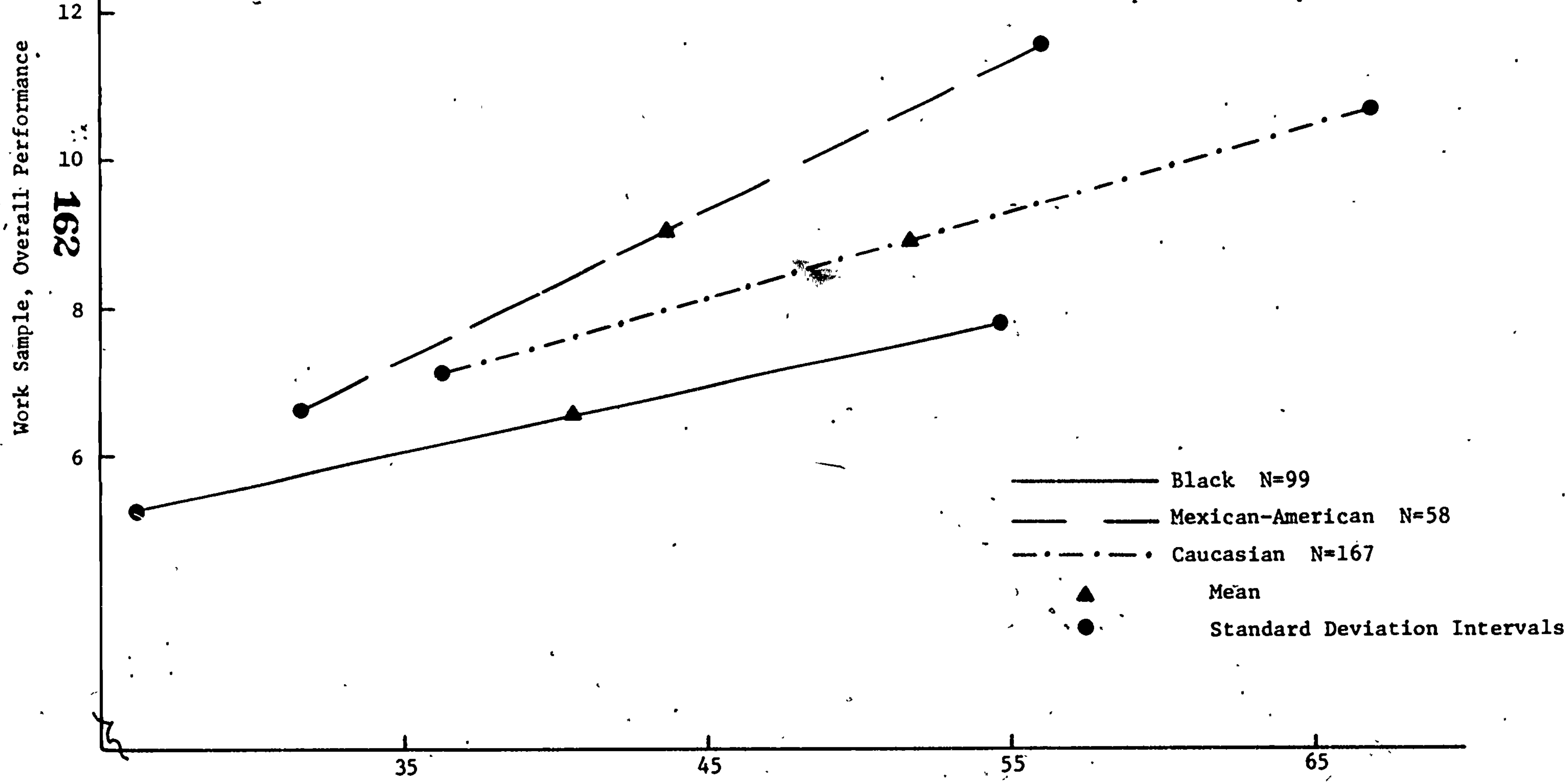
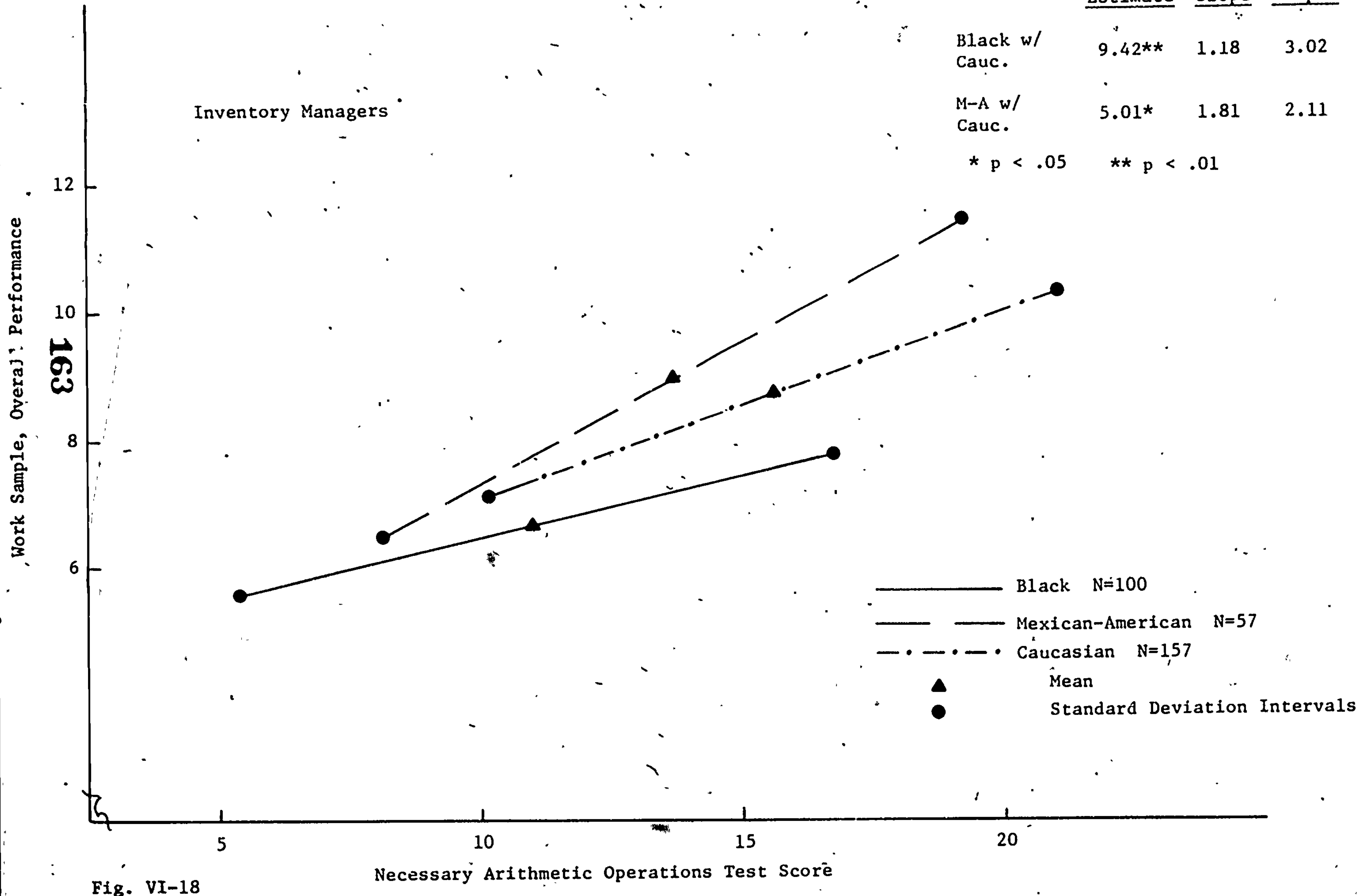


Fig. VI-17



	Error of Estimate	Slope	Intercept
Black w/ Cauc.	9.42**	1.18	3.02
M-A w/ Cauc.	5.01*	1.81	2.11

* p < .05 ** p < .01

Fig. VI-18

	Error of Estimate	Slope	Intercept
Black w/ Cauc.	8.04**	1.01	8.48**
M-A w/ Cauc.	0.60	0.00	2.95

** $p < .01$

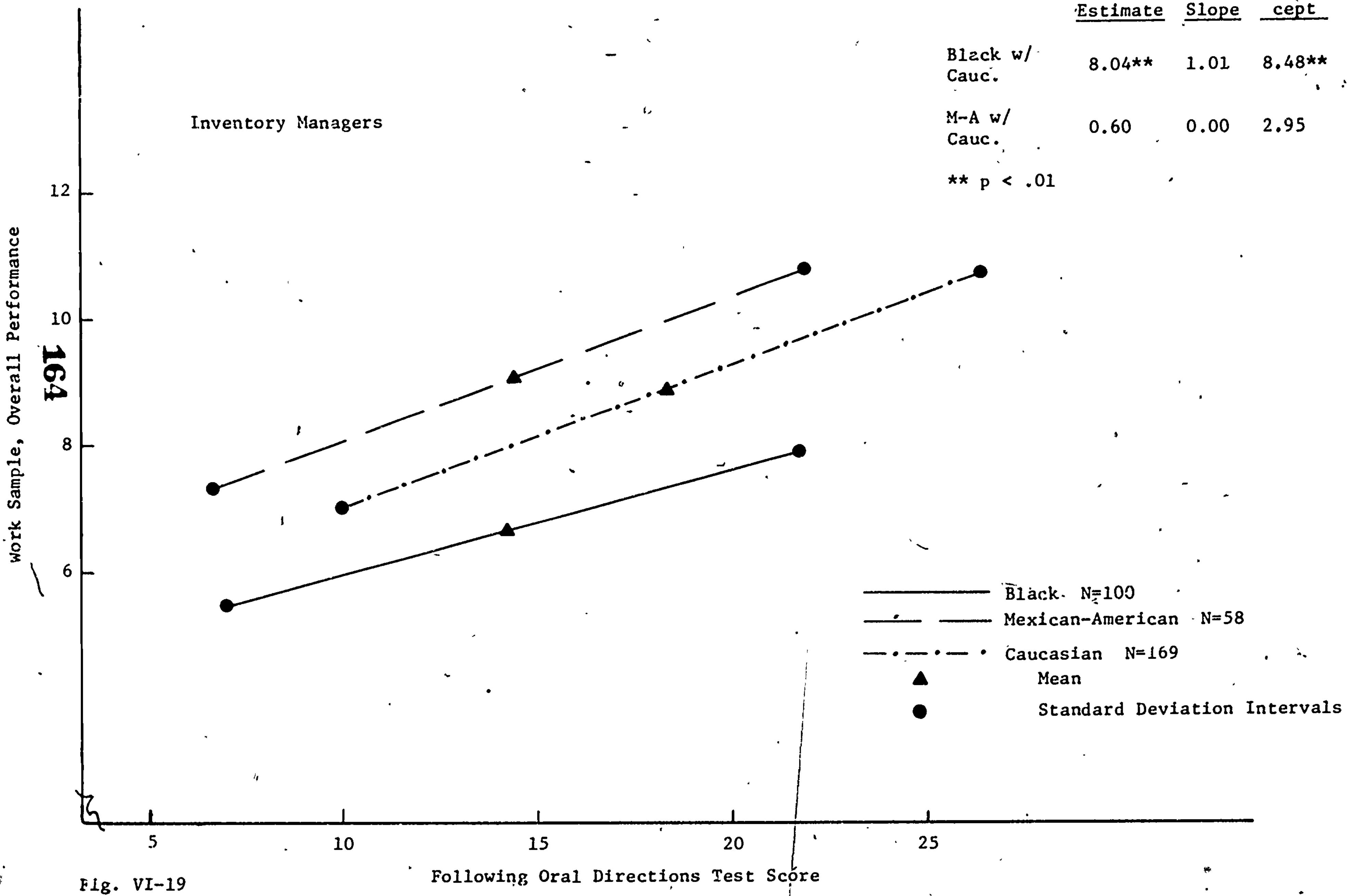


Fig. VI-19

	Error of Estimate	Slope	Intercept
Black w/ Other Cauc.	0.07	0.73	4.23*
M-A w/ S. A. Cauc.	0.01	0.17	0.34

* $p < .05$

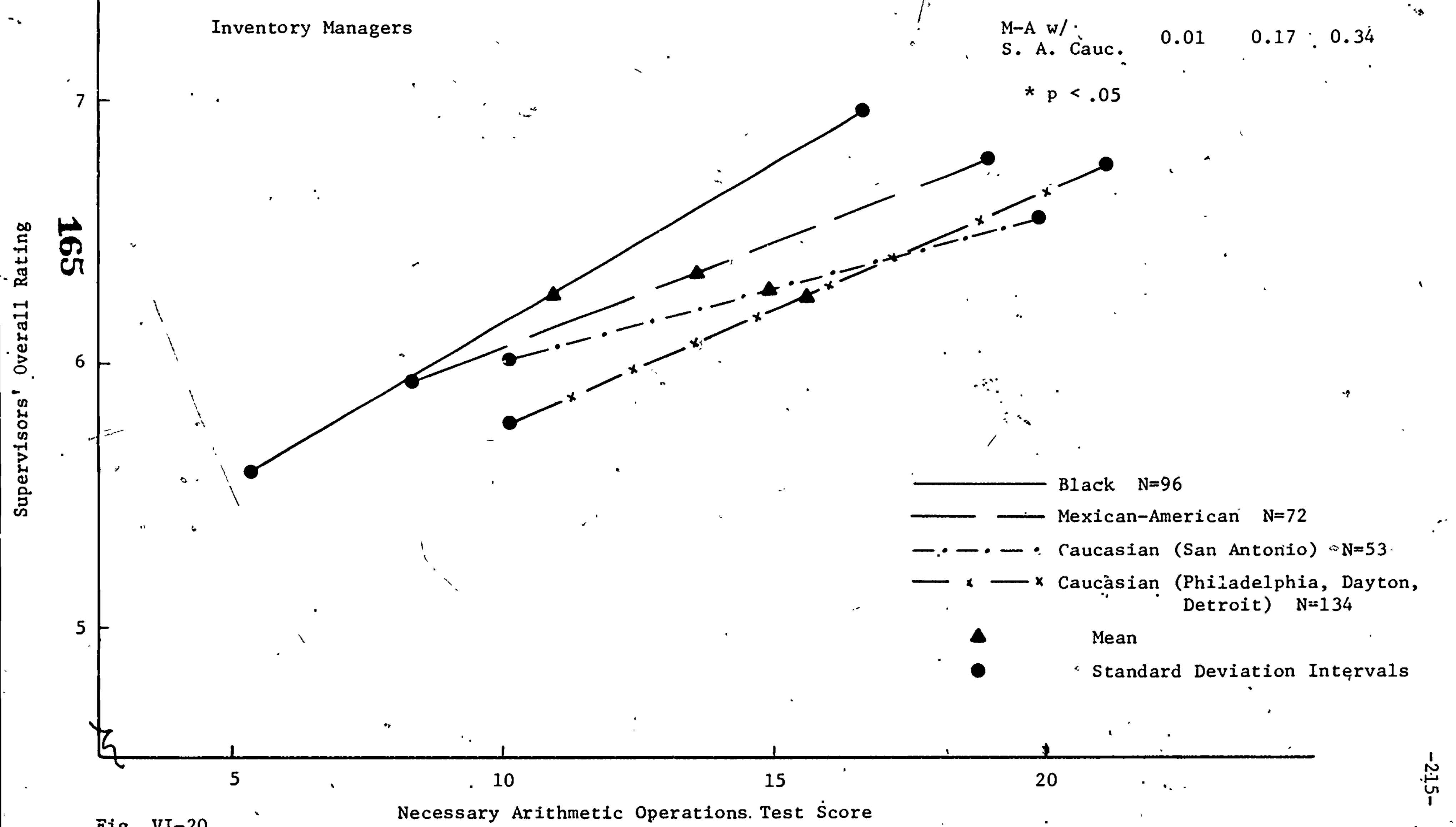


Fig. VI-20

	Error of Estimate	Slope	Intercept
Black w/ Other Cauc.	-0.00	0.82	0.02
M-A w/ S. A. Cauc.	0.12	2.40	0.76

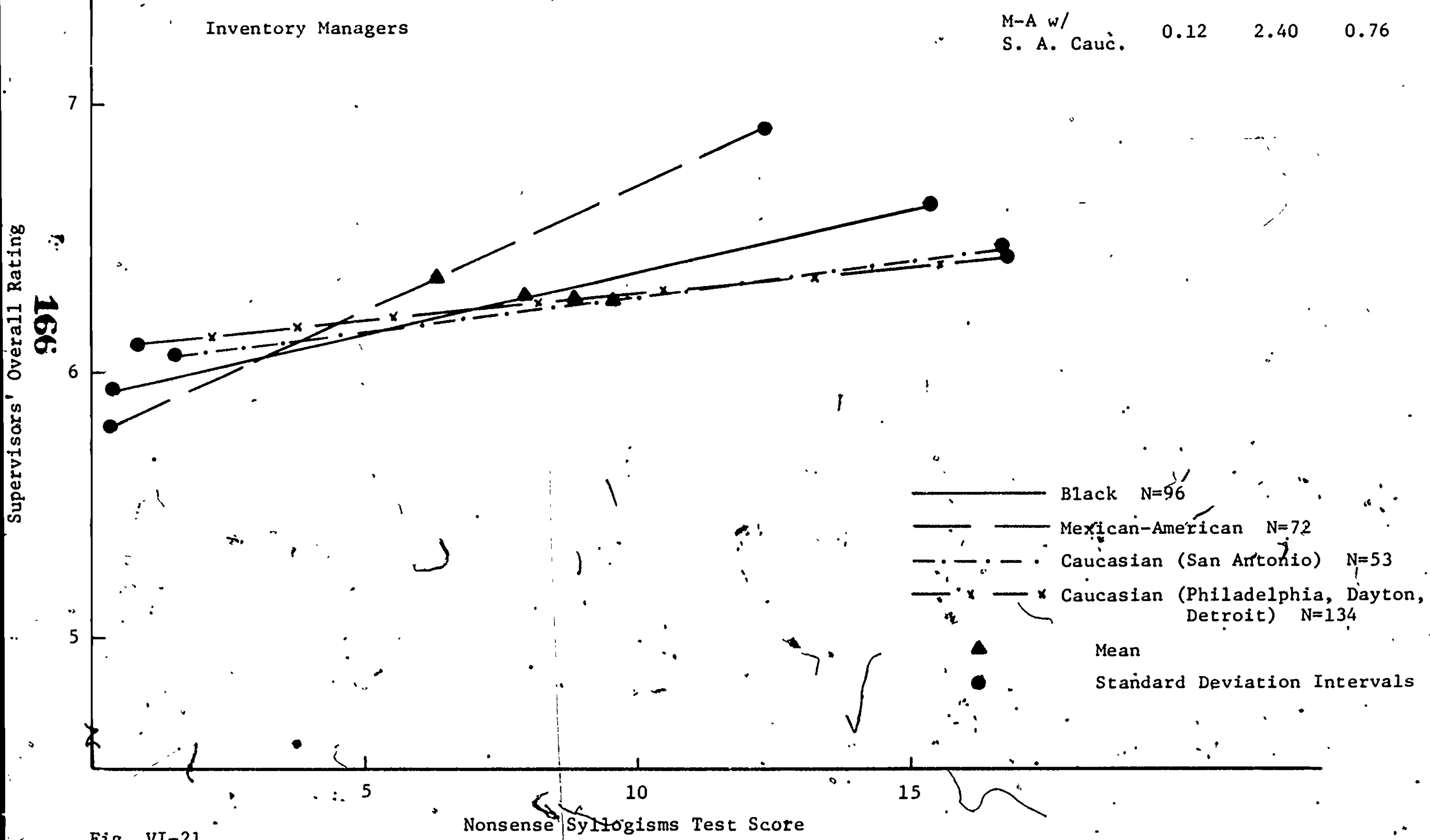
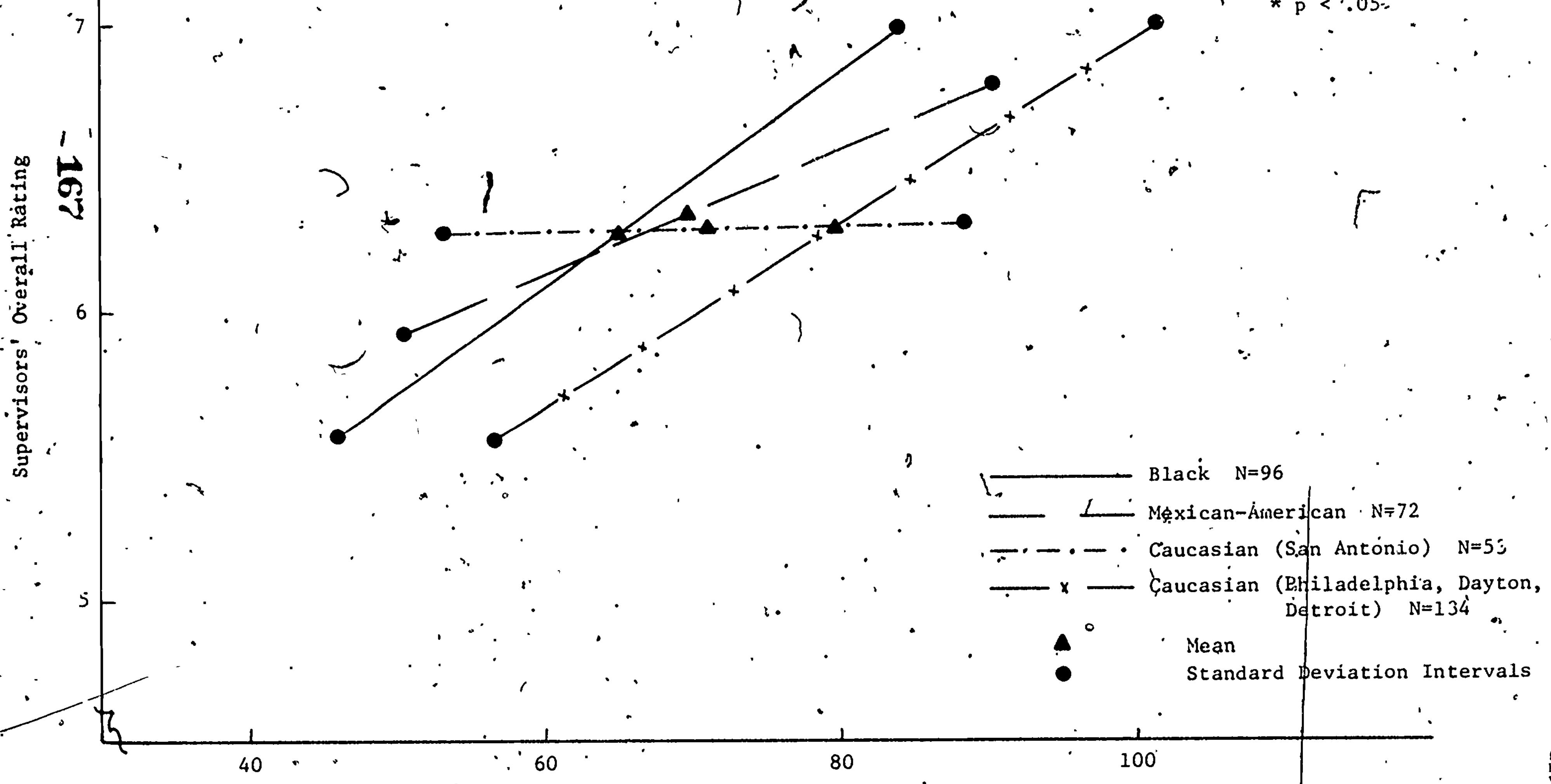


Fig. VI-21

Inventory Managers



	Error of Estimate	Slope	Intercept
Black w/ Other Cauc.	0.04	0.27	4.88*
M-A w/ S. A. Cauc.	0.04	1.44	0.09

* $p < .05$

Fig. VI-22

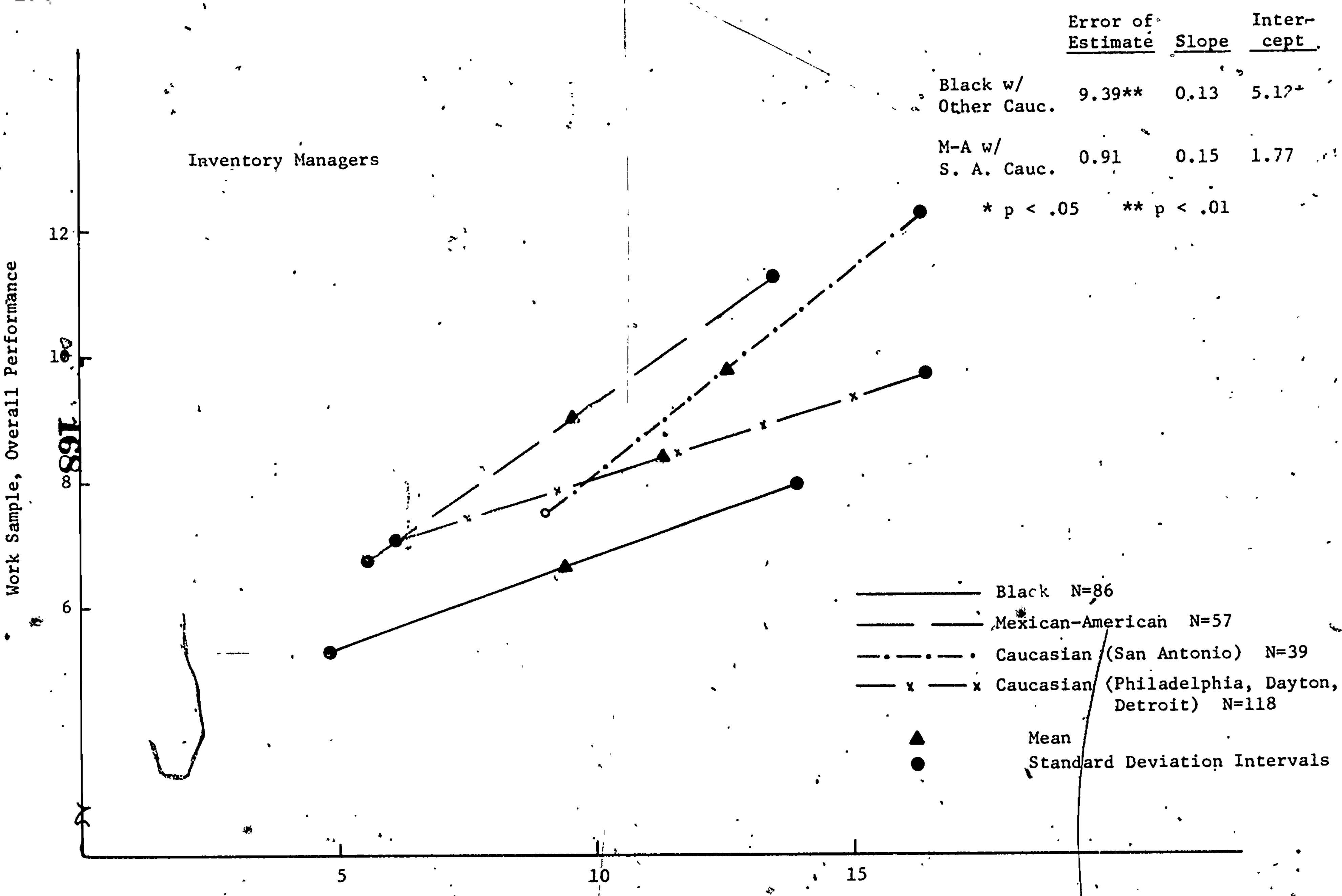

Inventory Managers
Error of Estimate
Slope
Intercept
Black w/ Other Cauc.
9.39**
0.13
5.12*
M-A w/ S. A. Cauc.
0.91
0.15
1.77
* p < .05
** p < .01
Black N=86
Mexican-American N=57
Caucasian (San Antonio) N=39
Caucasian (Philadelphia, Dayton, Detroit) N=118
Mean
Standard Deviation Intervals
Work Sample, Overall Performance
6
8
10
12
Inference Test Score
5
10
15

Fig. VI-23

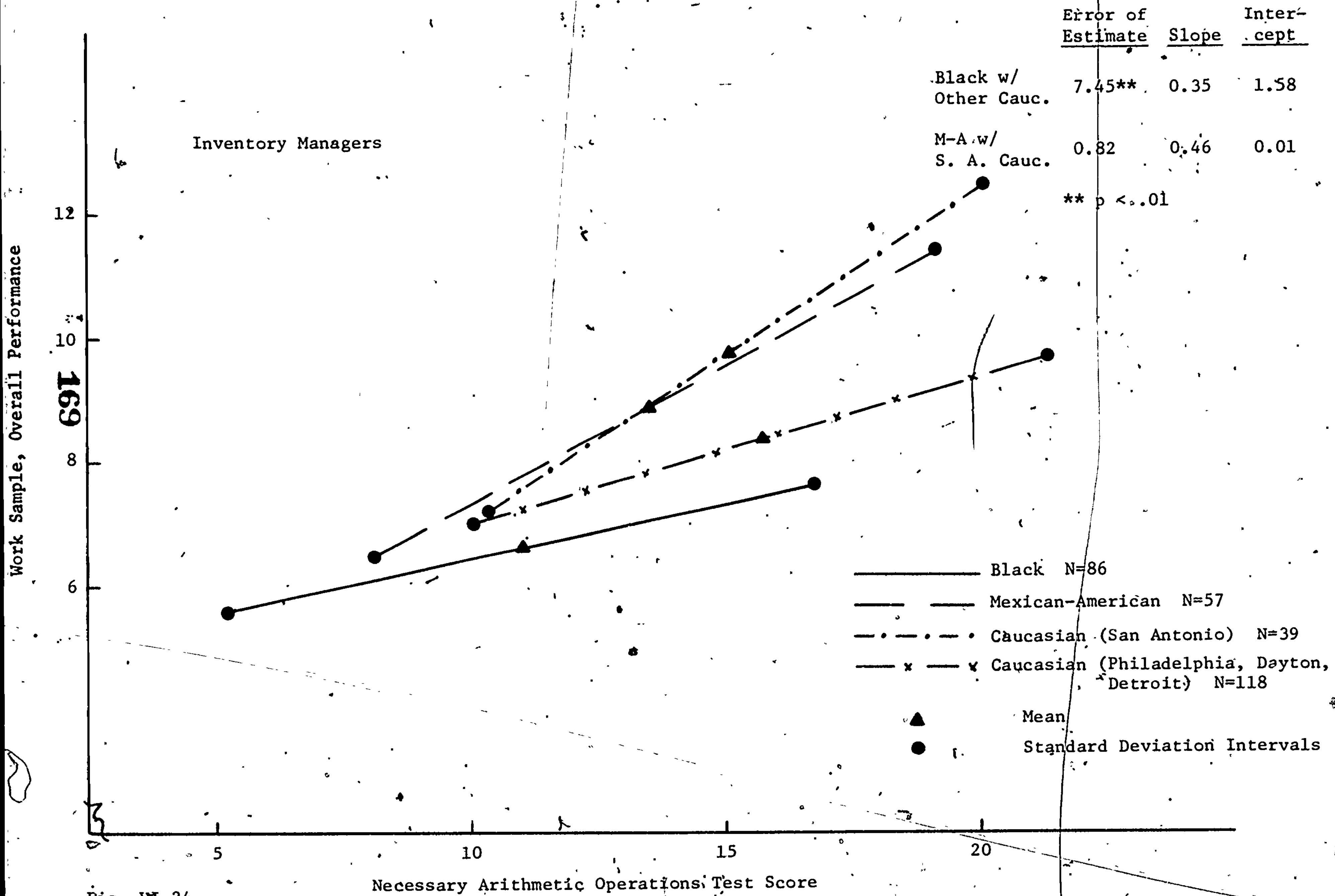

Inventory Managers
Error of Estimate
Slope
Inter-cept
Black w/ Other Cauc.
7.45**
0.35
1.58
M-A w/ S. A. Cauc.
0.82
0.46
0.01
** p < .01
Black N=86
Mexican-American N=57
Caucasian (San Antonio) N=39
Caucasian (Philadelphia, Dayton, Detroit) N=118
Mean
Standard Deviation Intervals
12
10
8
6
5
10
15
20
Work Sample, Overall Performance
Necessary Arithmetic Operations Test Score

Fig. VI-24

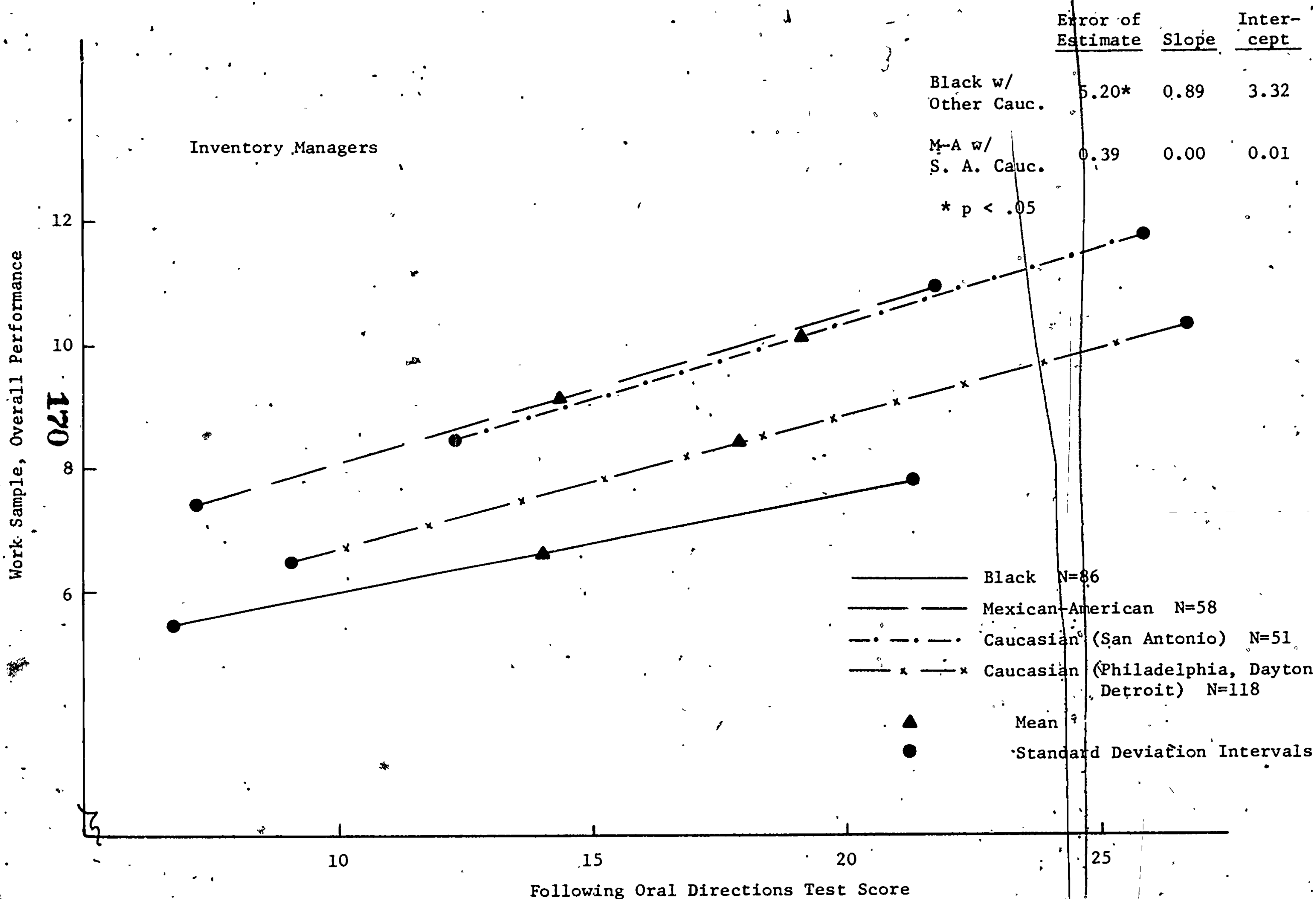

Inventory Managers
Error of Estimate
Slope
Intercept
Black w/ Other Cauc.
5.20*
0.89
3.32
M-A w/ S. A. Cauc.
0.39
0.00
0.01
* p < .05
Work Sample, Overall Performance
12
10
8
6
10
15
20
25
Following Oral Directions Test Score
Black N=86
Mexican-American N=58
Caucasian (San Antonio) N=51
Caucasian (Philadelphia, Dayton, Detroit) N=118
Mean
Standard Deviation Intervals

Fig. VI-25

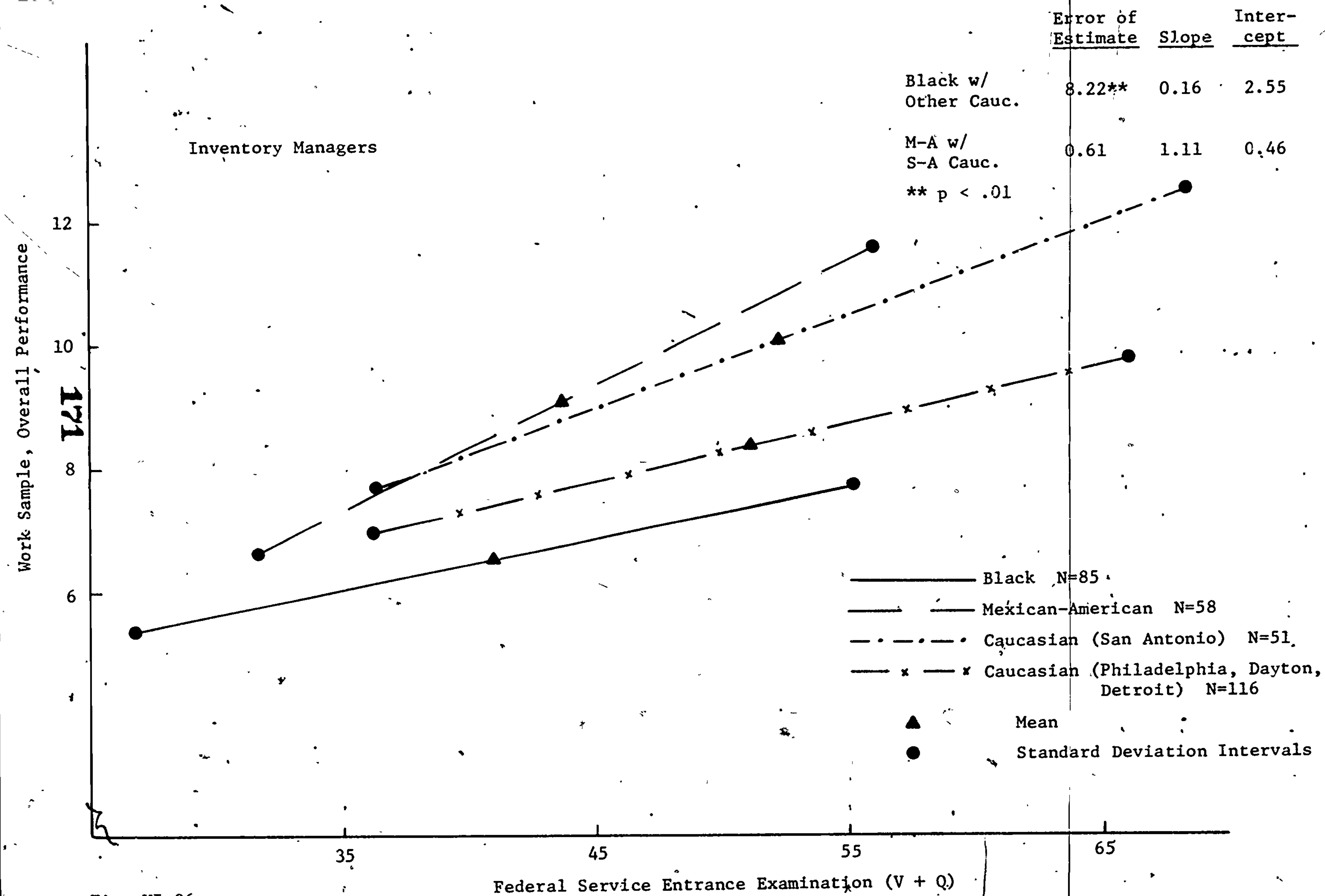

Inventory Managers
Error of Estimate
Slope
Intercept
Black w/ Other Cauc.
8.22**
0.16
2.55
M-A w/ S-A Cauc.
0.61
1.11
0.46
** p < .01
Work Sample, Overall Performance
12
10
8
6
35
45
55
65
Federal Service Entrance Examination (V + Q)
Black N=85
Mexican-American N=58
Caucasian (San Antonio) N=51
Caucasian (Philadelphia, Dayton, Detroit) N=116
Mean
Standard Deviation Intervals

Fig. VI-26

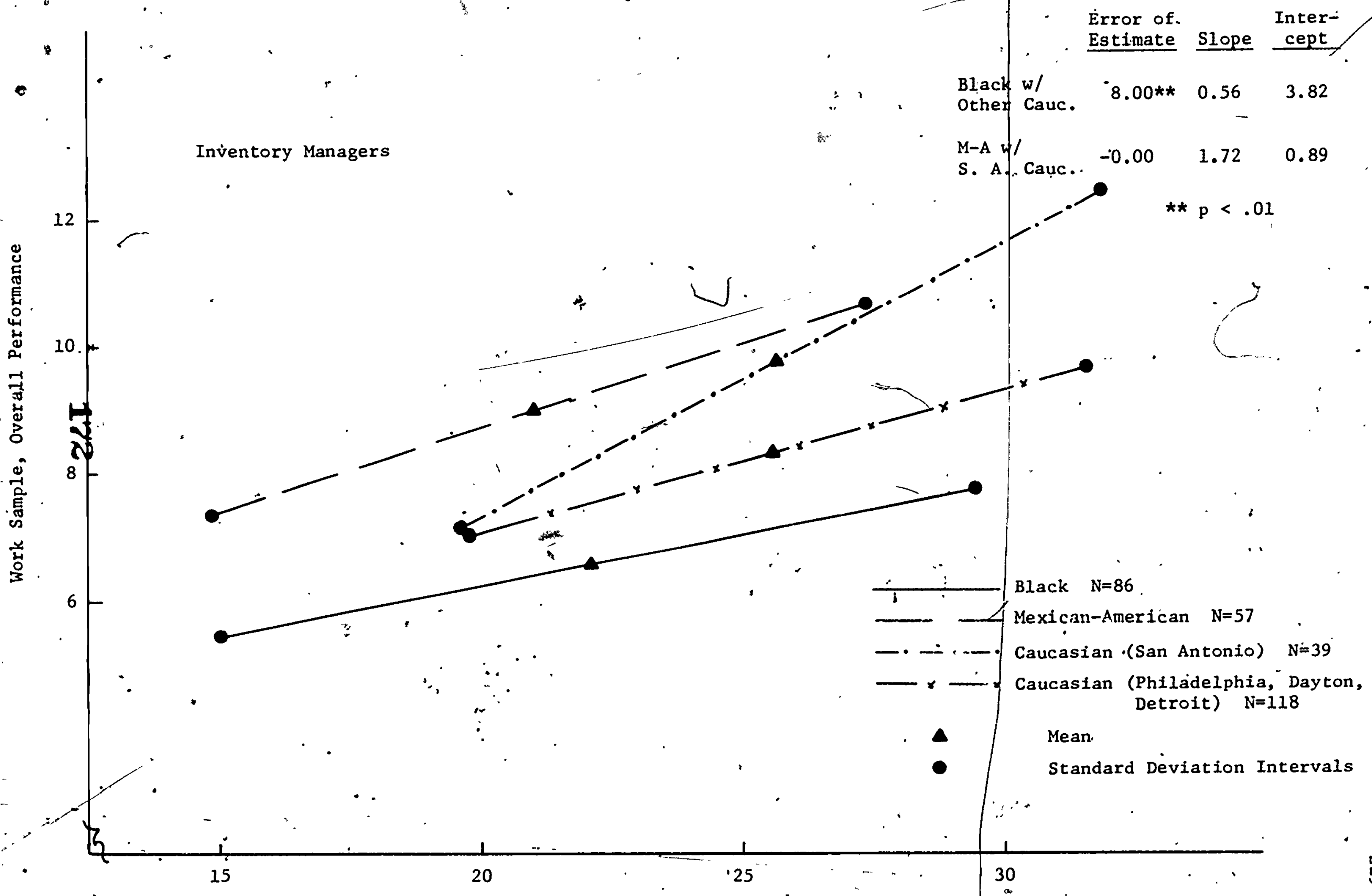

Inventory Managers
Error of Estimate
Slope
Intercept
Black w/ Other Cauc.
8.00**
0.56
3.82
M-A w/ S. A. Cauc.
-0.00
1.72
0.89
** p < .01
Work Sample, Overall Performance
12
10
8
6
15
20
25
30
Vocabulary Test Score
Black N=86
Mexican-American N=57
Caucasian (San Antonio) N=39
Caucasian (Philadelphia, Dayton, Detroit) N=118
Mean
Standard Deviation Intervals

Fig. VI-27

Table VI-3

Criterion Score Means for Medical Technicians

at Different Score Levels on the

Necessary Arithmetic Operations Test

Test Scores	Mean Supervisors' Overall Rating		Mean Job Knowledge Test Scores	
	Black	Caucasian	Black	Caucasian
18 +	7.2 N=5	6.3 N=69	39.4 N=5	40.2 N=69
14 - 17.9	5.9 N=28	6.0 N=81	38.0 N=28	37.3 N=81
10 - 13.9	5.9 N=56	6.3 N=77	32.6 N=56	34.4 N=77
- 9.9	5.7 N=60	5.5 N=55	26.9 N=60	31.3 N=55

Table VI-4

Criterion Score Means for Cartographic Technicians (TOPOCOM) at Different Score Levels on the Map Planning Test

Test Score	Mean Supervisors' Overall Rating			Mean Job Knowledge Test Score			Mean Work Sample Composite		
	Black	Mexican-American	Caucasian	Black	Mexican-American	Caucasian	Black	Mexican-American	Caucasian
24 +	6.0 N=18	6.7 N=16	6.6 N=56	44.3 N=18	38.9 N=16	45.3 N=56	21.9 N=18	24.8 N=16	24.2 N=56
17 - 23.9	5.8 N=28	5.9 N=34	6.0 N=101	34.0 N=28	30.7 N=34	41.4 N=101	19.0 N=28	20.4 N=33	22.0 N=98
10 - 16.9	5.1 N=45	5.5 N=35	5.2 N=67	35.1 N=45	29.9 N=35	37.2 N=67	17.2 N=43	14.4 N=34	20.0 N=65
- 9.9	5.0 N=10	5.2 N=16	5.2 N=17	24.7 N=10	22.4 N=14	29.7 N=17	13.7 N=10	13.1 N=16	16.0 N=17

Table VI-5

Criterion Score Means for

Cartographic Technicians (Coast & Geodetic Survey)

on the Map Planning Test

Test Scores	Mean Supervisors' Overall Rating	
	Black	Caucasian
24 +	3.8 N=2	7.0 N=8
17 - 23.9	7.0 N=10	6.8 N=18
10 - 16.9	5.4 N=16	6.1 N=18
- 9.9	4.2 N=10	5.0 N=6

Table VI-6

Criterion Score Means for Inventory Management Specialists

at Different Score Levels on the

Subtraction and Multiplication Test

Test Scores	Mean Supervisors' Overall Rating			Mean Work Sample Overall Rating		
	Black	Mexican-American	Caucasian	Black	Mexican-American	Caucasian
90 +	7.2 N=14	7.1 N=14	6.9 N=54	6.2 N=14	9.1 N=10	9.4 N=43
70 - 89	6.3 N-28	6.4 N=22	6.4 N=65	6.7 N=24	11.3 N=18	8.4 N=56
50 - 69	6.4 N=48	6.3 N=23	5.8 N=48	7.1 N=44	8.4 N=18	8.6 N=40
- 49	4.8 N=22	5.6 N=13	5.3 N=20	5.7 N=18	5.9 N=11	8.3 N=18

Chapter VII

Multiple Correlations and Cross-Ethnic Cross-Validation Coefficients for Prediction of Criterion Measures

In Chapter V, the relationships of the separate aptitude tests in the predictor batteries with the criterion measures were compared by ethnic group for each occupation studied. Aptitude tests were shown to have substantially equal validity for the different ethnic groups. In Chapter VI, the regression lines of best predictors for each ethnic group in the three occupations were compared with regression lines for these predictors for the other ethnic groups.

In this chapter, the results of stepwise multiple regression analyses are described, in which a best set of predictors was selected by ethnic group for each criterion measure. The prediction equations thus derived for each ethnic group for each criterion measure were then applied to data for the other ethnic samples to provide cross-ethnic cross-validation coefficients in further examination of differential validity.

In each occupation studied, the set of predictors was restricted to no more than three for each criterion measure. Predictors with negative regression weights were not used, so that for some criterion measures, only one or two predictors were included, selected prior to the entry of a negative predictor. In no case was a negatively weighted predictor selected first. Holding the number of predictors selected to three or fewer reduces the possibility of chance or spurious relationships entering into an inflated multiple, and increases the likelihood of relationships holding up in cross-validation. Correction for "shrinkage" in the multiples due to sampling errors is also minimized when the number of variables included is small.[1]

The results for each of the occupations studied will be discussed separately. The research questions to be explored are: 1) How do prediction equations differ (composition, weights, and level of prediction) for different ethnic groups? 2) What is the effect of using prediction equations derived from one ethnic group sample on data for another ethnic group (cross-ethnic cross-validation)? 3) What is the effect of cross-ethnic prediction of criterion scores where test scores of individuals fall one standard deviation below the mean, at the mean, or one standard deviation above the mean of their group? 4) Is the use of differential regression equations for ethnic groups warranted by evidence of bias in prediction?

Medical Technicians

Table VII-1 shows the multiple correlation coefficients (R) and standard regression weights derived for sets of predictors selected for the criterion measures of supervisors' ratings of Learning Ability, Technical Knowledge, and Overall effectiveness, and the Job Knowledge Test, for Medical Technicians. Although multiple correlation coefficients were obtained for all rating scales, only Learning Ability and Technical Knowledge ratings are being reported here in addition to Overall rating, as representing supervisors' judgments of less subjective aspects of job performance. These two scales had as high relationships with predictors as the Overall rating (and in some cases higher) for all occupations studied.

[1] Correction for shrinkage in the multiples for the smallest sample in the present study (N = 57) amounts to no more than .02.

It can be seen in Table VII-1 that multiples and regression weights for the Black sample are higher in every case than for the Caucasian sample. Higher multiples were obtained for the two rating scales than for the Overall rating for the Black sample and on the Learning Ability rating for the Caucasian sample. These differences were to be expected from the differential validities reported in Tables V-1 and V-5 in Chapter V, where the correlations of aptitude test scores with criterion measures for the Black sample were higher in a number of instances than for the Caucasian sample, especially true for the Job Knowledge Test. There is some consistency in the composition of the sets of predictors across criterion measures and ethnic groups. Necessary Arithmetic Operations (a general reasoning test) and the Subtraction & Multiplication Test (number facility) appear selectively as predictors of supervisors' ratings and Job Knowledge Tes[illegible] both ethnic groups.

Table VII-2 is a comparison [illegible] multiple correlation coefficients and cross-ethnic cross-validation coefficients for the same criterion measures shown in Table VII-1, with weights derived for the Black sample used to predict multiples for the Caucasian sample and vice versa. The multiples obtained for the Black sample using Black weights and Caucasian sample using Caucasian weights, shown in Table VII-1, are included again in Table VII-2 for comparison. The multiples for the Black sample, using Black weights, are substantially higher than for the Caucasian sample using Black weights, but the reverse is true when Caucasian weights are used to predict multiples for the Black sample. The multiples for the Black sample thus obtained are higher in every instance than Caucasian sample multiples based on Caucasian weights, although not greatly so, except for the Job Knowledge Test. There is a fair amount of shrinkage in the multiples for the Technical Knowledge rating (from .38 to .21) and Overall rating (from .29 to .17), using cross-ethnic weights, but the multiples for the Learning Ability rating (.42 to .40) and the Job Knowledge Test (.53 to .52) hold up whether Black or Caucasian weights are used.

Table VII-3 shows scores for the same criterion measures predicted from multiple regression equations for ethnic group samples where test scores in the equations are one standard deviation below the mean, at the mean, or one standard deviation above the mean. Results are shown where Black weights were used alternatively on scores from the Black and Caucasian samples and where Caucasian weights were used alternatively on scores for the Black and Caucasian samples. For all four of the criterion measures shown, it can be seen that Blacks with high scores (one standard deviation above the mean) tend to have higher predicted scores when Black weights are used, but those with low scores (one standard deviation below the mean) tend to have slightly higher predicted scores when Caucasian weights are used (except on the Learning Ability rating, which is slightly lower). It can also be seen that in nearly every instance the predicted scores of Caucasians scoring below, at, and above the mean on aptitude tests are higher when Black weights are used in the prediction equation. These results reflect the earlier finding of higher validities (and resulting steeper regression slopes) in the regression equations based on the Black sample, but higher mean scores (and thus a larger intercept constant) in the regression equations based on the Caucasian sample.

In summary, for the Medical Technicians, the multiple correlation coefficients obtained for the Black sample were consistently higher than for the Caucasian sample and, in general, the same aptitude and ability factors were represented and more heavily weighted for the Black sample in the regression equations. In the cross-ethnic cross-validation, the weights derived from the Caucasian sample appeared to be equally valid for the Black sample on the Learning Ability rating and the Job Knowledge Test, although there was some attrition in the other ratings. However, use of the weights derived from the Black sample on the Caucasian sample resulted in attrition in all multiples. Where test scores below the mean, at the mean, and above the mean were used to predict criterion scores by alternative use of ethnic regression weights, use of Black regression equations tended to favor Caucasians with scores at and above the mean, but not those with low scores. Conversely, the use of Caucasian regression equations would benefit Blacks with low test scores but not those with high test scores.

Cartographic Technicians (TOPOCOM)

Table VII-4 shows multiple correlation coefficients and standard regression weights for the sets of predictors selected by stepwise multiple regression for supervisors' ratings of Learning Ability, Technical Knowledge, and Overall effectiveness, Job Knowledge Test, and Work Sample Composite, by ethnic group. It can be seen that while individual tests appear in the sets of predictors for more than one ethnic group, in general the composition of the sets is different for the different ethnic groups for each criterion measure. One exception is noted, where the same set of predictors appears for the

Caucasian sample for both the Job Knowledge Test and the Work Sample Composite. The multiples obtained are fairly consistent in size across ethnic groups, with those for the Black sample as high or higher than for the Mexican-American and Caucasian sample, except for the Work Sample Composite, where the Mexican-American multiple is higher than the other two. It is interesting to note the mix of aptitude and skill factors in the sets of predictors for this occupation by ethnic group. For example, the best predictors for the Black sample on the Job Knowledge Test (for which exceptionally high multiples were obtained for all three groups) were CS Arithmetic, Extended Range Vocabulary, and Necessary Arithmetic Operations (general reasoning), with CS Arithmetic having the largest weight. For the Mexican-American sample, the set includes Surface Development, with the largest weight, Map Planning, and CS Arithmetic. For the Caucasian sample, again the largest weight appears for Surface Development, next highest for Extended Range Vocabulary, and third highest for Following Oral Directions.

Table VII-5 shows multiple correlation coefficients and cross-ethnic cross-validation coefficients obtained where weights derived for each of the three ethnic groups were used alternatively to compute multiples for each sample. It would be expected that some attrition would result when weights derived for one ethnic group are applied to samples for other ethnic groups. This occurs to some degree, but there are exceptions. For example, the use of Caucasian weights on all three ethnic samples for the Learning Ability rating produces multiples of .44, .44, and .43, respectively. Using Black weights alternatively, multiples of .47 (Black sample), .42 (Mexican-American sample), and

.40 (Caucasian sample) are derived. Using Mexican-American weights, multiples of .41 (Black), .48 (Mexican-American), and .40 (Caucasian) are obtained. For the Job Knowledge Test, the results are also remarkably consistent across ethnic groups. For the Black sample, using Black weights, R = .74; using Mexican-American weights, R = .63; and using Caucasian weights, R = .68. For the Mexican-American sample, using Mexican-American weights, R = .61; using Black weights, R = .52; and using Caucasian weights, R = .59. For the Caucasian sample, using Caucasian weights, R = .66; using Mexican-American weights, R = .60; and using Black weights, R = .61.

The results are not quite as consistent for the Technical Knowledge and Overall ratings and the Work Sample Composite. Attrition in the multiples for the Technical Knowledge rating for the Black sample from .47 to .29, and for the Mexican-American sample from .30 to .20, result when Caucasian weights are used. The Caucasian multiple changes from R = .29 using Caucasian weights to .28 (using Mexican-American weights) and .19 (using Black weights).

For the Overall rating, Caucasian weights applied to data for Mexican-American and Black samples produces multiples that are as high or almost as high as when respective weights for these groups are applied to their own data. Conversely, when Black weights are applied to data for Mexican-American and Caucasian samples, attrition is fairly great: Mexican-American (Mexican-American) = .28, Black (Mexican-American) = .16; Caucasian (Caucasian) = .35, Black (Caucasian) = .17. The attrition when using Mexican-American weights is also fairly great: Black (Black) = .36, Mexican-American (Black) =

.24; Caucasian (Caucasian) = .35, Mexican-American (Caucasian) = .20.

For the Work Sample Composite, applying Caucasian weights to data for Black and Mexican-American samples results in some attrition in the multiples obtained when using the regression weights obtained for their own samples: Black (Black) = .39, Caucasian (Black) = .30; Mexican-American (Mexican-American) = .53, Caucasian (Mexican-American) = .33. Using Black and Mexican-American weights alternatively on the other two ethnic groups results in less attrition.

Some insight into the fluctuations as the result of cross-ethnic cross-validation may be obtained by examining the sets of predictors for the various criterion measures by ethnic group, shown in Table VII-4.

At the bottom of Table VII-5, the multiples and cross-ethnic multiples are shown for supervisors' Overall rating for the Coast & Geodetic Survey sample, using weights derived from the Black and Caucasian TOPOCOM samples. Multiple regression analysis was not carried out for the Coast & Geodetic Survey sample because of the small size of the ethnic group samples. Supervisors' ratings were the only criterion measures available for this sample. In general, both the Caucasian and the Black TOPOCOM weights produced higher multiples for prediction of supervisors' Overall rating for the Coast & Geodetic Survey sample than for the TOPOCOM sample, suggesting that the TOPOCOM regression equations are equally valid for this group.

Table VII-6 shows predicted criterion scores using ethnic and cross-ethnic regression equations where test scores of the respective ethnic samples are one standard deviation below the mean, at the mean, and one standard deviation above the mean.

It can be seen that the Black group with scores below the mean tend to have higher predicted criterion scores when Caucasian weights are used for all criterion measures shown. For example, on the Job Knowledge Test, the Black equation used on Black scores predicts a score of 24.7; the Caucasian equation predicts 26.0 This is also true at the mean. For ratings, predicted criterion scores are not as consistent for high scorers, but differences are small. For the Job Knowledge Test and the Work Sample Composite, again Blacks scoring high tend to have slightly higher predicted scores when the Caucasian regression equations are used. For the Mexican-Americans, the results are not as clear-cut. Ratings predicted from low scores tend to be lower using Caucasian weights, but Job Knowledge Test and Work Sample Composite predicted scores are slightly higher. This is generally true for scores at the mean, although use of Caucasian weights result in a slightly lower criterion score for the Work Sample Composite than the Mexican-American weights (17.0 from 17.7). For high-scoring Mexican-Americans, prediction of Overall rating is the same using either Mexican-American or Caucasian weights; prediction of Job Knowledge Test scores is higher using Caucasian weights (38.9 - 41.4), and lower for the Work Sample Composite (25.1 - 22.0).

For the Caucasian sample, Black regression equations predict ratings slightly lower for low-scoring Caucasians than Caucasian equations, and Mexican-American weights predict slightly higher ratings. For high-scoring Caucasians, both Black and Mexican-American weights predict higher ratings than those predicted by Caucasian weights. On the Job Knowledge Test, Black weights tend to predict higher scores for

Caucasians who score in both the low and high groups, with more of a difference for the high scorers. On the Work Sample Composite, Black weights produce lower predicted scores for both low and high Caucasian scorers than Caucasian weights.

In summary, for Cartographic Technicians, while sets of predictors differ for the respective ethnic groups, multiple correlation coefficients were fairly consistent in size across ethnic groups. Those obtained for the Black and Mexican-American samples were as high as, and in some cases higher than, those obtained for the Caucasian sample.

In general, the use of Caucasian regression equations on data for the other two samples predicted performance of Blacks and Mexican-Americans with less attrition in multiples than might be expected in cross-ethnic cross-validation. In cross-ethnic comparisons of predicted criterion scores where test scores were below the mean, at the mean, or above the mean, Caucasian regression equations tended to favor Blacks by predicting higher criterion scores for both the low-scoring and high-scoring Black groups. This was not as consistently so for the Mexican-Americans, but the over- and under-predictions were not great.

It may be concluded that any bias resulting from use of Caucasian regression weights tended to be in favor of the other ethnic groups, and that the use of differential prediction equations for the separate ethnic groups would not, in general, be warranted for Cartographic Technicians.

Inventory Management Specialists

Table VII-7 shows the multiple correlation coefficients and standard regression weights for supervisors' ratings of Learning Ability, Technical Knowledge, and Overall effectiveness and the Work Sample Composite by ethnic group, for Inventory Management Specialists. The level of prediction is consistent across criterion measures and ethnic groups, with R ranging from .34 to .51 for the Black sample, from .38 to .70 for the Mexican-American sample, and from .35 to .51 for the Caucasian sample. It is interesting that every aptitude test except Object-Number (associative [rote] memory)[1] appears as a predictor for one or more criterion measures for one or more ethnic groups. This confirms, to some extent, observations of the research team regarding the complexity of the Inventory Manager's job, in that a fairly wide range of aptitude factors appears to contribute unique variance to prediction of performance. (See validity coefficients for individual tests in Table V-4 and V-8 in Chapter V.)

There is little overlapping in sets of predictors across ethnic groups, except for the Subtraction & Multiplication Test (number facility) which appears as a predictor for ratings in all three groups, but in only two sets are there as many as two predictors in common (Learning Ability rating for Blacks and Mexican-Americans).

Table VII-8 shows the multiple correlation coefficients and cross-ethnic cross-validation coefficients by criterion measure, where multiple regression equations computed for each ethnic group were alternatively applied to the data for each other ethnic group. Multiples for within-group prediction (shown in Table VII-7) are again included for comparison.

As seen for the previous occupations, the multiples resulting from use of Caucasian regression equations do not differ to any marked degree from those resulting from use of Black weights for Black sample data and use of Mexican-American weights for Mexican-American sample data. There

[1] The Object-Number Test was the only test not correlated significantly with rating or work sample criteria. See Tables V-4 and V-8, Chapter V.

are greater differences (multiples are lower), especially in ratings, when Mexican-American weights are applied to Black and Caucasian data, and when Black weights are applied to Mexican-American data, but differences are less when Black weights are applied to Caucasian data. It is apparent that the most consistent results are obtained when Caucasian weights are used for all samples.

Table VII-9 shows criterion scores predicted when regression equations for each ethnic group were used alternatively on test scores one standard deviation below the mean, at the mean, and one standard deviation above the mean for the ethnic samples.

Caucasian regression equations used with Black sample data tend to underpredict ratings across the scoring range, and to overpredict Work Sample Composite scores when compared to those computed by Black regression equations. Caucasian regression equations used with Mexican-American data tend to predict lower ratings and lower Work Sample Composite scores across the scoring range, when compared with criterion scores predicted by Mexican-American regression equations.

Black regression equations tend to predict higher Caucasian ratings and lower Work Sample Composite scores across the range than those predicted by Caucasian equations. Black regression equations tend to predict about the same criterion scores for low-scoring Mexican-Americans as Mexican-American equations, but lower Work Sample Composite scores.

Mexican-American equations tend to predict higher criterion scores for both Black and Caucasian samples across the range, when compared to scores predicted by each group's own respective equations, with few exceptions.

In general, the use of Caucasian regression equations favors Blacks in the upper scoring range in predicting Work Sample criterion scores, but not in predicting ratings. In the lower scoring range, differences in prediction are trivial. For Mexican-Americans in both the lower and upper scoring ranges, lower Work Sample Composite scores are predicted by Caucasian regression weights than by Mexican-American equations.

In summary, for Inventory Management Specialists, the level of prediction was consistent for criterion measures across ethnic groups. All but one aptitude test in the predictor battery was represented at least once in the sets of predictors, suggesting the range of aptitude factors related to job performance in this occupation. Cross-ethnic cross-validation coefficients showed some attrition in multiples, particularly when Mexican-American weights were applied to Black and Caucasian data, but results were more consistent, with less attrition in multiples, when Caucasian weights were used for Blacks and Mexican-Americans. When cross-ethnic regression equations were used alternatively to predict criterion scores for ethnic subjects scoring below, at, and above the mean, Caucasian weights favored Blacks in the upper scoring range. Differences were trivial in the lower scoring range. However, somewhat lower criterion scores were predicted by Caucasian equations for Mexican-Americans in the upper scoring range on both ratings and Work Sample Composite. Differences in the lower scoring range were again trivial.

Prediction and Cross-Validation Across Occupations

The comparisons for the separate occupations do not take into consideration the relative level and accuracy of predictions attained by cross-validation. The multiple correlation coefficients and cross-ethnic cross-validation coefficients have been plotted from Tables VII-2, VII-5, and VII-8 for the Black - Caucasian samples (Figure VII-1) and for the Mexican-American - Caucasian samples (Figure VII-2).

In each figure, the distance between each point and the diagonal line represents the loss in prediction when regression weights from a different ethnic group (i.e., Caucasian weights for a Black sample, or vice versa) are used. It can be seen that, in general, very similar multiples are obtained and that there are no striking discrepancies.

Summary

For all three occupations, level of prediction achieved was fairly consistent across ethnic groups and criterion measures, although in some instances multiples for Black and Mexican-American samples were higher than for Caucasian samples. Learning Ability ratings were, in general, predicted at a higher level than Overall ratings, and multiples for these ratings were consistently as high as for the Job Knowledge Tests and Work Samples. The number of factors represented in the sets of tests selected by regression analysis as predictors was fairly narrow for the Medical Technicians, with only four out of nine aptitude tests represented in the sets of predictors for both ethnic groups. For the other two occupations, all of the aptitude tests except one were represented in at least one set of predictors for at least one ethnic group, demonstrating the validity of the researchers' observations of unique job performance factors.

In cross-ethnic cross-validation of findings, where regression equations for each ethnic group were alternatively applied to data for each other ethnic group, less differential attrition in multiples resulted for all three occupations when Caucasian weights were used for prediction. For the Medical

Technicians, Black regression weights produced higher multiples for the Caucasian sample because the same aptitude and ability factors were represented in, and more heavily weighted for, the Black sample. However, for all three occupations, Caucasian regression equations appeared to be about equally valid predictors for both Blacks and Mexican-Americans. This holds up even where test scores are one standard deviation below, at, or above the mean. Use of Caucasian weights resulted in about the same or slightly higher predicted criterion scores for Blacks and Mexican-Americans across the range, with some exceptions. It may be concluded that differential regression equations for separate ethnic groups would not be warranted by the evidence of these findings.

Table VII-1

MULTIPLE CORRELATION COEFFICIENTS (R) AND STANDARD REGRESSION WEIGHTS FOR PREDICTING CRITERION MEASURES FROM SELECTED APTITUDE TEST SCORES, BY ETHNIC GROUP SAMPLE

Medical Technicians

Aptitude Tests	Supervisors' Ratings						Job Knowledge Test	
	Learning Ability		Technical Knowledge		Overall Rating			
	Black R = .42	Caucasian R = .36	Black R = .38	Caucasian R = .16	Black R = .29	Caucasian R = .16	Black R = .53	Caucasian R = .40
	Standard Regression Weights							
Subtraction & Multiplication	.26	.14	.22		.29			.14
Vocabulary				.13			.22	.19
Hidden Figures								
Necessary Arithmetic Operations	.23	.18	.23			.16	.35	.23
Pin-Dexterity		.16		.10			.18	
Number Comparison								
Gestalt Completion								
Picture-Number								
Paper Folding								

Table VII-2

MULTIPLE CORRELATION COEFFICIENTS AND CROSS-ETHNIC CROSS-VALIDATION COEFFICIENTS FOR PREDICTION OF CRITERION MEASURES FROM APTITUDE TEST SCORES

Medical Technicians

Criterion Measures	Black Weights		Caucasian Weights	
	Black Sample (N=166)	Caucasian Sample (N=297)	Black Sample (N=166)	Caucasian Sample (N=297)
Supervisors' Ratings				
Learning Ability	.42	.32	.40	.36
Technical Knowledge	.38	.13	.21	.16
Overall Rating	.29	.13	.17	.16
Job Knowledge Test	.53	.24	.52	.40

Table VII-3

CRITERION SCORES PREDICTED BY ETHNIC AND CROSS-ETHNIC MULTIPLE REGRESSION EQUATIONS WHERE APTITUDE TEST SCORES ARE BELOW THE MEAN, AT THE MEAN, AND ABOVE THE MEAN IN THE ETHNIC SAMPLES

Medical Technicians

Criterion Measures		Predicted Criterion Scores: Test Scores One Standard Deviation Below the Mean: Using Black Weights	Using Caucasian Weights	Test Scores At the Mean: Using Black Weights	Using Caucasian Weights	Test Scores One Standard Deviation Above the Mean: Using Black Weights	Using Caucasian Weights
Rating Scales							
Learning Ability	Black	4.9	4.6	5.8	5.4	6.7	6.2
	Caucasian	5.4	5.0	6.5	6.0	7.6	6.9
Technical Knowledge	Black	4.5	4.7	5.2	5.0	6.0	5.4
	Caucasian	5.0	4.9	6.0	5.3	7.0	5.7
Overall	Black	5.1	5.4	5.7	5.6	6.3	5.9
	Caucasian	5.4	5.5	6.1	5.9	6.7	6.2
Job Knowledge Test	Black	23.6	27.9	31.4	32.2	39.2	36.4
	Caucasian	28.4	30.6	37.6	35.7	46.9	40.8

Table VII-4

MULTIPLE CORRELATION COEFFICIENTS AND STANDARD REGRESSION WEIGHTS

FOR PREDICTORS BY ETHNIC GROUP SAMPLE

Cartographic Technicians (TOPOCOM)

Standard Regression Weights

Aptitude Tests	Learning Ability: Black R = .47	Learning Ability: Mexican-American R = .48	Learning Ability: Caucasian R = .43	Supervisors' Ratings, Technical Knowledge: Black R = .47	Supervisors' Ratings, Technical Knowledge: Mexican-American R = .30	Supervisors' Ratings, Technical Knowledge: Caucasian R = .29	Supervisors' Ratings, Overall Rating: Black R = .36	Supervisors' Ratings, Overall Rating: Mexican-American R = .28	Supervisors' Ratings, Overall Rating: Caucasian R = .35	Job Knowledge Test: Black R = .74	Job Knowledge Test: Mexican-American R = .61	Job Knowledge Test: Caucasian R = .66	Work Sample Composite: Black R = .39	Work Sample Composite: Mexican-American R = .53	Work Sample Composite: Caucasian R = .49
Coordination	.09														
Hidden Figures		.25			.15	.10		.21					.20	.23	
Vocabulary				.19			.13						.14		
Object-Number				.10			.12								
Card Rotations														.28	
CS Arithmetic	.29	.16		.36	.11		.26		.12	.36	.21				
Map Planning		.23	.31		.14	.16		.12	.18		.21			.19	
Surface Development	.23		.20						.16		.38	.36			.29
Maze Tracing Speed													.22		
Following Oral Directions						.12						.26			.21
Identical Pictures															
Extended Range Vocabulary										.26		.27			.13
Necessary Arithmetic Operations										.31					

Table VII-5

MULTIPLE CORRELATION COEFFICIENTS AND CROSS-ETHNIC CROSS-VALIDATION COEFFICIENTS FOR PREDICTION OF CRITERION MEASURES FROM APTITUDE TEST SCORES

Cartographic Technicians (TOPOCOM)

Criterion Measures	Black Weights			Mexican-American Weights			Caucasian Weights		
	Black Sample (N=99)	Mexican-American Sample (N=99)	Caucasian Sample (N=240)	Black Sample (N=99)	Mexican-American Sample (N=99)	Caucasian Sample (N=240)	Black Sample (N=99)	Mexican-American Sample (N=99)	Caucasian Sample (N=240)
Supervisors' Ratings									
Learning Ability	.47	.42	.40	.41	.48	.40	.44	.44	.43
Technical Knowledge	.47	.17	.19	.36	.30	.28	.29	.20	.29
Overall Rating	.36	.16	.17	.24	.28	.20	.32	.28	.35
Job Knowledge Test	.74	.52	.61	.63	.61	.60	.68	.59	.66
Work Sample Composite	.39	.47	.42	.38	.53	.40	.30	.33	.49

Cartographic Technicians (Coast & Geodetic Survey)

	Black Weights *		Caucasian Weights *	
	Black Sample (N=38)	Caucasian Sample (N=50)	Black Sample (N=38)	Caucasian Sample (N=50)
Supervisors' Overall Rating	.47	.36	.48	.44

* Derived from TOPOCOM multiple regressions

Table VII-6

CRITERION SCORES PREDICTED BY ETHNIC AND CROSS-ETHNIC MULTIPLE REGRESSION EQUATIONS WHERE APTITUDE TEST SCORES ARE BELOW THE MEAN, AT THE MEAN, AND ABOVE THE MEAN IN THE ETHNIC SAMPLES

Cartographic Technicians (TOPOCOM)

		Predicted Criterion Scores								
		Test Scores One Standard Deviation Below the Mean			Test Scores At the Mean			Test Scores One Standard Deviation Above the Mean		
Criterion Measures		Using Black Weights	Using Mexican-American Weights	Using Caucasian Weights	Using Black Weights	Using Mexican-American Weights	Using Caucasian Weights	Using Black Weights	Using Mexican-American Weights	Using Caucasian Weights
Supervisors' Ratings										
Learning Ability	Black	4.4	4.6	4.6	5.4	5.7	5.5	6.5	6.8	6.4
	Mexican-American	4.5	4.6	4.6	5.5	5.7	5.5	6.5	6.7	6.3
	Caucasian	4.9	5.0	5.0	6.0	6.1	5.8	7.1	7.2	6.7
Technical Knowledge	Black	4.1	5.0	4.8	5.3	5.7	5.5	6.4	6.4	6.1
	Mexican-American	4.2	5.0	4.7	5.3	5.7	5.3	6.4	6.4	5.9
	Caucasian	4.6	5.2	5.0	5.8	5.9	5.7	7.0	6.7	6.3
Overall	Black	4.5	5.5	4.8	5.4	5.9	5.5	6.4	6.3	6.3
	Mexican-American	4.6	5.4	4.8	5.5	5.8	5.5	6.5	6.3	6.3
	Caucasian	4.9	5.6	5.1	5.9	6.0	5.9	6.9	6.5	6.7
Job Knowledge Test	Black	24.7	22.3	26.0	35.3	30.7	36.2	46.0	39.2	46.5
	Mexican-American	24.7	22.7	23.1	34.7	30.8	32.3	44.6	38.9	41.4
	Caucasian	30.6	26.4	30.1	41.4	34.8	40.4	52.2	43.2	50.6
Work Sample Composite	Black	13.2	11.2	13.4	18.3	18.6	19.0	23.4	25.9	24.6
	Mexican-American	12.5	10.3	11.9	17.5	17.7	17.0	22.4	25.1	22.0
	Caucasian	14.5	12.8	15.8	20.0	20.6	21.4	25.4	28.3	27.1

Table VII-7

MULTIPLE CORRELATION COEFFICIENTS AND STANDARD REGRESSION WEIGHTS

FOR PREDICTORS BY ETHNIC GROUP SAMPLE

Inventory Management Specialists

Aptitude Tests	Supervisors' Ratings: Learning Ability			Supervisors' Ratings: Technical Knowledge			Supervisors' Ratings: Overall Rating			Work Sample Composite		
	Black R = .51	Mexican-American R = .53	Caucasian R = .41	Black R = .34	Mexican-American R = .42	Caucasian R = .36	Black R = .46	Mexican-American R = .38	Caucasian R = .35	Black R = .43	Mexican-American R = .70	Caucasian R = .51
	Standard Regression Weights											
Number Comparison			.20			.27						.21
Hidden Figures	.27	.19			.23							
Vocabulary												.24
Object-Number												
Letter Sets				.22			.25				.21	
Nonsense Syllogisms		.33			.32			.29				
Subtraction & Multiplication	.27	.25	.16	.22			.33	.20	.28			
Extended Range Vocabulary											.34	
Necessary Arithmetic Operations											.31	
Following Oral Directions										.20		.25
Inference	.20									.28		
Civil Service No. 170 (V+Q)			.18			.17			.15			

Table VII-8

MULTIPLE CORRELATION COEFFICIENTS AND CROSS-ETHNIC

CROSS-VALIDATION COEFFICIENTS FOR PREDICTION OF CRITERION

MEASURES FROM APTITUDE TEST SCORES

Inventory Management Specialists

Criterion Measures	Black Weights			Mexican-American Weights			Caucasian Weights		
	Black Sample (N=112)	Mexican-American Sample (N=72)	Caucasian Sample (N=190)	Black Sample (N=112)	Mexican-American Sample (N=72)	Caucasian Sample (N=190)	Black Sample (N=112)	Mexican-American Sample (N=72)	Caucasian Sample (N=190)
Supervisors' Ratings									
Learning Ability	.51	.41	.29	.45	.53	.26	.44	.43	.41
Technical Knowledge	.34	.16	.29	.16	.42	.13	.30	.23	.36
Overall Rating	.46	.18	.33	.32	.38	.22	.41	.28	.35
Work Sample Overall Performance	.43	.59	.44	.35	.70	.39	.39	.54	.51

Table VII-9

CRITERION SCORES PREDICTED BY ETHNIC AND CROSS-ETHNIC MULTIPLE REGRESSION EQUATIONS WHERE APTITUDE TEST SCORES ARE BELOW THE MEAN, AT THE MEAN, AND ABOVE THE MEAN IN THE ETHNIC SAMPLES

Inventory Management Specialists

Criterion Measures		Predicted Criterion Scores								
		Test Scores One Standard Deviation Below the Mean			Test Scores At the Mean			Test Scores One Standard Deviation Above the Mean		
		Using Black Weights	Using Mexican-American Weights	Using Caucasian Weights	Using Black Weights	Using Mexican-American Weights	Using Caucasian Weights	Using Black Weights	Using Mexican-American Weights	Using Caucasian Weights
Rating Scales										
Learning Ability	Black	4.6	4.5	4.8	6.0	6.0	5.7	7.4	7.5	6.7
	Mexican-American	5.0	4.7	5.1	6.3	6.1	6.0	7.7	7.5	6.[illegible]
	Caucasian	5.2	4.9	5.1	6.8	6.5	6.2	8.3	8.2	7.2
Technical Knowledge	Black	5.0	5.2	5.3	5.8	6.3	6.0	6.7	7.3	6.8
	Mexican-American	5.3	5.4	5.6	6.1	6.3	6.2	6.9	7.2	6.9
	Caucasian	5.4	5.4	5.5	6.3	6.5	6.3	7.2	7.7	7.0
Overall	Black	5.2	5.4	5.1	6.2	6.4	5.8	7.2	7.3	6.5
	Mexican-American	5.5	5.5	5.3	6.5	6.4	6.0	7.5	7.2	6.7
	Caucasian	5.7	5.6	5.5	6.8	6.7	6.3	7.9	7.7	7.0
Work Sample Overall Performance	Black	5.0	4.8	4.8	6.6	8.7	7.8	8.2	12.6	10.8
	Mexican-American	5.1	5.4	5.2	6.6	9.0	8.0	8.1	12.6	10.7
	Caucasian	5.6	6.8	5.8	7.4	10.8	8.9	9.2	14.7	12.0

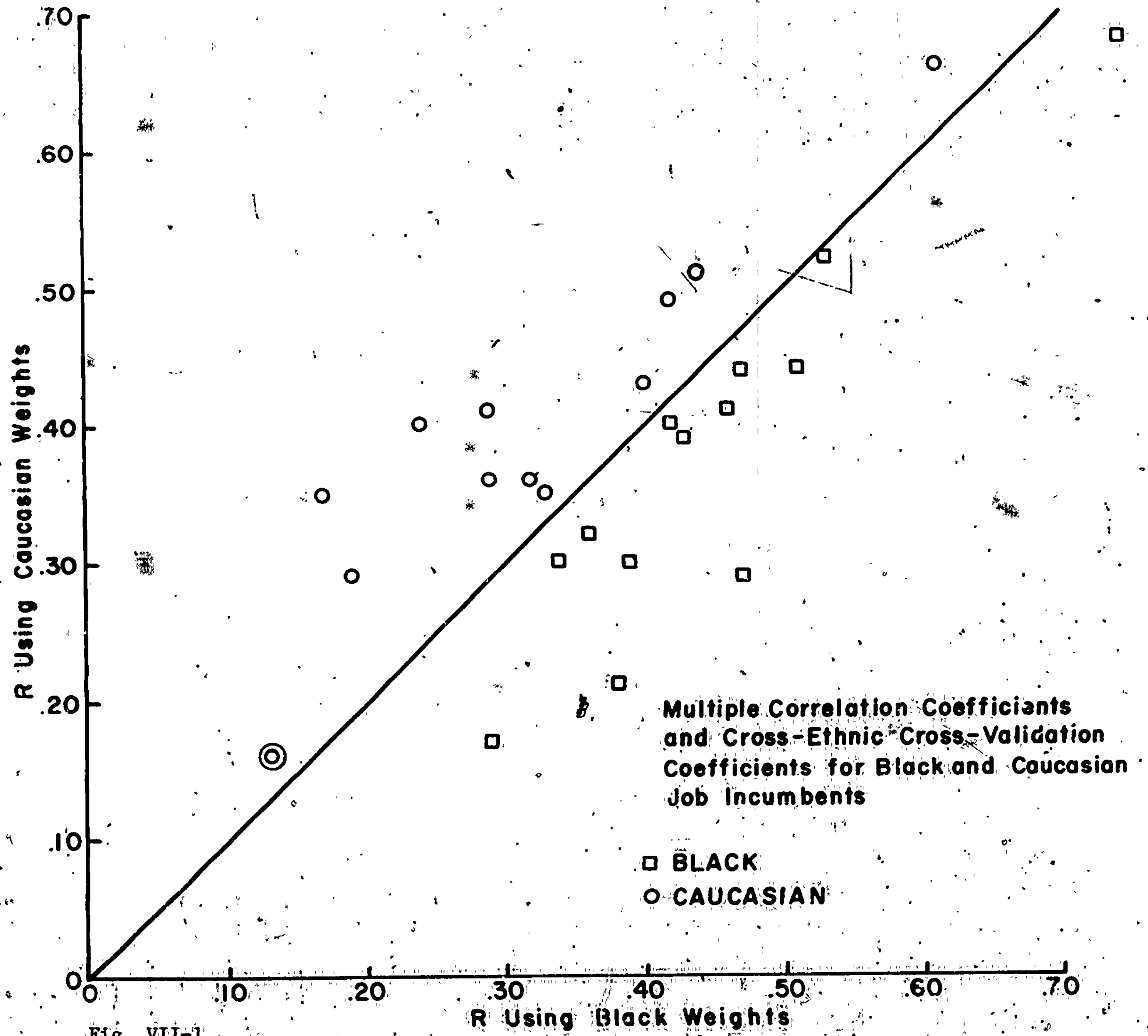

.70
.60
.50
.40
.30
.20
.10
0
R Using Caucasian Weights
.10
.20
.30
.40
.50
.60
.70
R Using Black Weights
Multiple Correlation Coefficiants
and Cross-Ethnic Cross-Validation
Coefficients for Black and Caucasian
Job Incumbents
BLACK
CAUCASIAN

Fig. VII-1

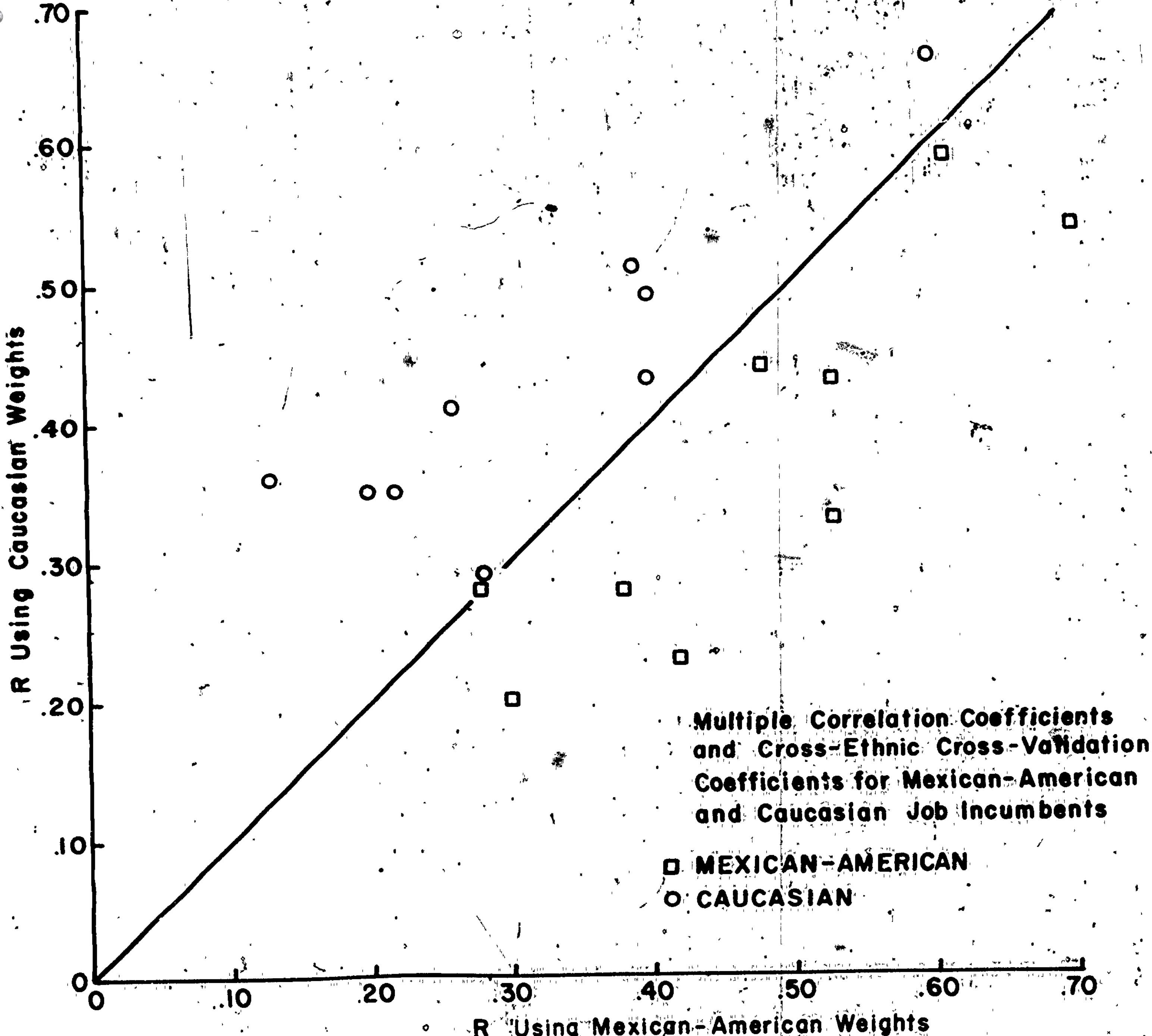

Multiple Correlation Coefficients and Cross-Ethnic Cross-Validation Coefficients for Mexican-American and Caucasian Job Incumbents
MEXICAN-AMERICAN
CAUCASIAN
R Using Caucasian Weights
R Using Mexican-American Weights
0
.10
.20
.30
.40
.50
.60
.70

Fig. VII-2

Chapter VIII

Biasing Effects of Ethnic Rater-Ratee Interaction

For each of the occupations studied, ethnic identification of supervisors who completed the rating forms was obtained, as well as for the job incumbent being rated. Thus it was possible to study the data by ethnic group of both rater and ratee to see whether there were any biasing effects from ethnic rater-ratee interaction.

Differences in Mean Ratings

Table VIII-1 shows the mean rating on each rating scale for the different ethnic rater-ratee combinations in the Medical Technician sample. Black raters assigned higher mean ratings to Black technicians than to Caucasian technicians on eight out of nine scales. Caucasian raters assigned higher mean ratings to Caucasian technicians on all nine scales.

Table VIII-2 shows the mean ratings by ethnic rater-ratee combinations for Cartographic Technicians at TOPOCOM installations. Again, Black raters assigned higher mean ratings to Black technicians than to Caucasian technicians on seven out of eight rating scales. Mexican-American raters assigned higher mean ratings to Mexican-American technicians than to Caucasian technicians on all eight scales. Caucasian raters assigned higher mean ratings to Caucasian technicians than to Mexican-American technicians on four out of eight scales, and higher mean ratings to Caucasian technicians than to Black technicians on all eight scales.

A somewhat different pattern emerges in Table VIII-3, which shows the mean ratings by ethnic rater-ratee combinations for Cartographic Technicians in the Coast and Geodetic Survey Sample. Here, the Black for these two variables by ethnic group rater-ratee combinations. As can be seen from the table, Black raters rating Black Medical Technicians assigned considerably higher mean ratings on the Job Knowledge scale than did Black raters rating Caucasian technicians. The Job Knowledge Test scores for the two rating combinations varied in the opposite direction. Neither of these differences in means, however, reached the level of statistical significance. The correlation between rating score and test was .50 for Black raters rating Black technicians, and .09 for Black raters rating Caucasian technicians. These coefficients are significantly different at the .01 level. It appears that the Black raters were more lenient when rating Black technicians than when rating Caucasian technicians, or that they considered different aspects of performance in rating the two groups.

The regression lines for predicting the Job Knowledge ratings from the Job Knowledge Test scores for the four groups of Medical Technicians are shown in Figure VIII-1. The regression line for Black technicians rated by Black supervisors lies above the regression line for Caucasian technicians rated by Black supervisors except at the lower end of the test score range. Similarly, the regression line for Black technicians rated by Caucasian supervisors lies above the regression line for Caucasian technicians rated by Caucasian supervisors except at the lower end. In both instances, Black technicians with moderate to high Job Knowledge Test scores received higher ratings on the Job Knowledge scale than did Caucasians with equivalent scores. It is particularly noticeable that ratings assigned to Caucasian technicians by Black supervisors on the Job Knowledge scale bear little relationship to the Job Knowledge Test

Caucasian supervisors rating Caucasian technicians rated them slightly higher on the average than they rated Black technicians, although the mean differences are not statistically significant. Job Knowledge Test scores vary in the same direction, but here the difference is significant, again suggesting that the Caucasian raters are possibly more lenient toward the Black technicians. There is a higher correlation between test scores and ratings for the Black technicians than for the Caucasian technicians, but the coefficients are not significantly different.

Regression lines for predicting Job Knowledge ratings from Job Knowledge Test scores for Cartographic Technicians at TOPOCOM installations are shown in Figures VIII-2, VIII-3, and VIII-4.

Figure VIII-2 shows the regression lines for Black and Caucasian technicians rated by Black supervisors. The two lines are approximately parallel, but the regression line for Black technicians lies above the regression line for Caucasian technicians, indicating that for equivalent test scores, Black technicians receive higher ratings than do Caucasian technicians.

Figure VIII-3 shows the regression lines for Mexican-American and Caucasian technicians rate[illegible] by Mexican-American supervisors. Except at the upp[illegible] end of the scale, the regression line for Mexican-American technicians lies above the line for Caucasian technicians. Thus, Mexican-American technicians with <u>low</u> test scores receive higher ratings than do Caucasian technicians with equivalent scores.

Figure VIII-4 shows the regression lines for Black, Mexican-American, and Caucasian technicians rated by Caucasian supervisors. The regression line for Caucasian technicians lies above the line for Black technicians, although the two lines are fairly close together. The regression line for Mexican-American technicians lies above the other two regression lines. Thus, Mexican-American technicians appear to receive higher ratings than those of the other two ethnic groups with equivalent test scores.

Relationship of Learning Ability Ratings and Objective Aptitude and Job Performance Measures

It may be useful to consider the relationship between ratings on another rating scale for the various rater-ratee combinations and objective measures of aptitude and job performance. The Learning Ability rating was selected on the logical ground that this is the aspect of job performance which aptitude tests usually can best be expected to predict. This expectation is confirmed by the fact that ratings on this scale have higher correlations with the aptitude tests than do the other ratings for all three occupations studied.

Table VIII-7 shows, for the Medical Technicians, the correlation between the Learning Ability rating on the one hand and aptitude test scores and Job Knowledge Test scores on the other, by ethnic group rater-ratee combinations.

For the Black raters, higher validity coefficients are shown for Black technicians than for Caucasian technicians on five of ten measures. However, the level of validity coefficients appears to be higher in general for the Black technicians. The average correlation coefficient (by r to z transformation) is .27 for Black technicians rated by a Black supervisor, and .17 for Caucasian technicians rated by a Black supervisor.

For Caucasian raters, similarly, the validities for Black technicians were higher than for Caucasian technicians on five out of ten tests. The average coefficient for the Black technicians is .25 and for Caucasian technicians .23.

Table VIII-8 shows the correlations between the Learning Ability rating and aptitude tests, Job Knowledge Test, and Work Sample tasks for Cartographic Technicians at TOPOCOM installations, again by ethnic rater-ratee combinations.

For those technicians rated by Black supervisors, the correlation coefficients are higher for Black technicians than for Caucasian technicians on 15 out of the 17 measures. The average correlation coefficient for Black technicians is .44 and for Caucasian technicians .22.

For those rated by a Mexican-American supervisor, the coefficients are higher for Mexican-American technicians than for Caucasian technicians on only three out of the 17 measures. The average correlation coefficient for Mexican-American technicians is .26 and for Caucasian technicians .42.

For those rated by Caucasian supervisors, higher validities for the Caucasian technicians than for the Black technicians are shown on nine out of the 17 measures, and higher coefficients for the Caucasian technicians than for the Mexican-American technicians appear for only six out of the 17 measures. The average coefficient is .22 for Black technicians, .24 for Mexican-American technicians, and .22 for Caucasian technicians.

Table VIII-9 shows the correlation between the Learning Ability rating and the aptitude tests for Cartographic Technicians in the Coast and Geodetic Survey sample. For those technicians rated by a Black supervisor, there were higher validities for Black technicians than for Caucasian technicians on seven out of 13 measures. The average coefficient is .59 for Black technicians and .47 for Caucasian technicians.

For those rated by a Caucasian supervisor, there are higher coefficients for the Caucasian technicians than for Black technicians on seven out of 13 measures. The average coefficient is .27 for both groups of ratees.

Table VIII-10 shows the correlation coefficients between Learning Ability rating and aptitude test scores and Work Sample Overall Performance scores by ethnic group rater-ratee combinations for Inventory Management Specialists.

For those managers rated by a Black supervisor, higher coefficients appear for Black managers than for Caucasian managers on 11 out of 15 measures. The average coefficient is .40 for Black managers and .28 for Caucasian managers. For those managers rated by a Caucasian supervisor, higher coefficients appear for Caucasian managers than for Black managers on only two out of 15 measures, and higher coefficients appear for Caucasian managers than for Mexican-American managers on only one out of 15 measures. The average correlation coefficient is .26 for Black managers, .28 for Mexican-American managers, and .20 for Caucasian managers.

The comparison of relative size of correlation coefficient is summarized in Table VIII-11. Black raters appear to base their ratings on the qualities measured by the objective instruments to a greater degree when rating members of their own ethnic group than when rating Caucasians. Other raters appear to have the opposite tendency, that is to base their ratings on the qualities measured by the objective instruments to a greater degree when rating members of other ethnic groups than when rating members of their own group.

Table VIII-12 shows the mean correlation coefficient (computed by r to z transformation) for each of the ethnic group rater-ratee combinations by occupation and for the overall total. Within each of the occupational groups, as well as for the total, the highest average coefficient is for the group of Black ratees rated by Black raters.

Mexican-American raters were available in only one study. The average correlation for Caucasian technicians rated by Mexican-American supervisors is quite high. The average coefficients for the remaining groups, Caucasians rated by Black supervisors, Mexican-Americans rated by Mexican-American supervisors, and all three ethnic groups rated by Caucasian supervisors, are considerably lower and approximately equal.

Summary

The evidence in differences in mean ratings, differing relationships between Job Knowledge Test scores and Job Knowledge ratings, and differing relationships between Learning Ability and objective measures for the various ethnic group rater-ratee combinations suggests the following conclusions:

1. Raters appear to be more lenient toward members of their own ethnic groups.

2. Raters of different ethnic groups appear to base their ratings on different aspects of performance, and these may also vary according to the ethnic group being rated.

Table VIII-1

Means of Supervisors' Ratings

by Ethnic Group Rater-Ratee Combination

Medical Technicians

ETHNIC GROUP OF TECHNICIAN: OF RATER:	Black Black	Caucasian Black	Black Caucasian	Caucasian Caucasian
	N = 43-50	N = 146-148	N = 38-41	N = 272-276
Flexibility	6.2	5.5	5.3	5.4
Organization	6.3	6.1	5.6	6.0
Interest	6.1	6.2	5.4	5.8
Learning Ability	6.6	5.8	5.6	6.0
Job Knowledge	6.2	5.5	5.1	5.3
Technique	6.5	5.8	5.8	6.0
Need for Supervision	6.5	5.9	5.6	6.1
Communication	6.3	6.0	5.3	5.7
Overall Rating	6.3	5.8	5.6	5.9

Table VIII-2

Means of Supervisors' Ratings

by Ethnic Group Rater-Ratee Combination

Cartographic Technicians (TOPOCOM)

ETHNIC GROUP OF TECHNICIAN: OF RATER:	Black Black	Caucasian Black	Mexican-American Mexican-American	Caucasian Mexican-American	Black Caucasian	Mexican-American Caucasian	Caucasian Caucasian
	N=20	N=53	N=99	N=26	N=99	N=100	N=240
Accuracy	5.9	5.4	6.0	5.7	5.4	5.8	5.9
Interest	5.3	5.6	5.9	5.7	5.3	5.9	5.9
Learning Ability	6.5	5.4	5.8	5.4	5.3	5.6	5.9
Job Knowledge	6.2	5.3	5.7	5.3	5.2	5.7	5.7
Dexterity	6.2	5.9	6.4	6.1	5.8	6.2	6.1
Need for Supervision	6.0	5.2	5.6	4.9	5.0	5.5	5.6
Perseverance	6.4	6.0	7.2	6.9	6.1	6.2	6.2
Overall Rating	6.4	5.9	6.1	5.5	5.4	5.5	5.9

Table VIII-3

Means of Supervisors' Ratings

by Ethnic Group Rater-Ratee Combination

Cartographic Technicians (Coast and Geodetic Survey)

ETHNIC GROUP OF TECHNICIAN: OF RATER:	Black Black	Caucasian Black	Black Caucasian	Caucasian Caucasian
	N=16	N=21	N=28	N=42
Accuracy	5.0	5.0	5.8	5.4
Interest	4.9	5.4	5.9	6.0
Learning Ability	4.5	5.4	5.5	5.9
Job Knowledge	4.8	5.3	5.2	5.6
Dexterity	5.9	5.7	6.3	6.
Need for Supervision	5.2	5.1	5.6	5.8
Perseverance	5.4	6.0	6.4	6.8
Overall Rating	4.8	5.8	5.9	6.0

Table VIII-4

Means of Supervisors' Ratings

by Ethnic Group Rater-Ratee Combination

Inventory Management Specialists

ETHNIC GROUP OF MANAGER: OF RATER:	Black Black	Caucasian Black	Black Caucasian	Mexican-American Caucasian	Caucasian Caucasian
	N=19	N=26	N=106	N=68	N=191
Organization	5.5	5.3	6.1	6.3	6.3
Interest	6.3	5.7	6.0	6.2	6.1
Learning Ability	5.7	5.3	6.1	6.1	6.2
Communication	5.3	5.4	5.8	5.5	6.0
Technical Knowledge	5.7	5.3	5.9	6.2	6.3
Stability	6.0	5.7	6.0	6.0	6.0
Dependability	5.7	6.0	6.4	6.5	6.6
Practical Judgment	5.6	5.6	6.2	6.3	6.4
Cooperation	6.8	7.5	7.3	6.9	7.0
Overall Rating	6.3	5.7	6.2	6.3	6.2

Table VIII-5

Mean Ratings by Ethnic Group of Rater

	Black raters			Caucasian raters		
	Higher mean ratings to Blacks	Equal	Higher mean ratings to Caucasians	Higher mean ratings to Caucasians	Equal	Higher mean ratings to Blacks
	Number of scales			Number of scales		
Medical Technicians	8	0	1	9	0	0
Cartographic Technicians (TOPOCOM)	7	0	1	8	0	0
Cartographic Technicians (C & G)	2	1	5	6	0	2
Inventory Management Specialists	6	1	3	7	2	1
Total	23	2	10	30	2	3

	Mexican-American raters			Caucasian raters		
	Higher mean ratings to Mexican-Americans	Equal	Higher mean ratings to Caucasians	Higher mean ratings to Caucasians	Equal	Higher mean ratings to Mexican-Americans
	Number of scales			Number of scales		
Cartographic Technicians (TOPOCOM)	8	0	0	4	3	1
Inventory Management Specialists				6	2	2
Total	8	0	0	10	5	3

Table VIII-6

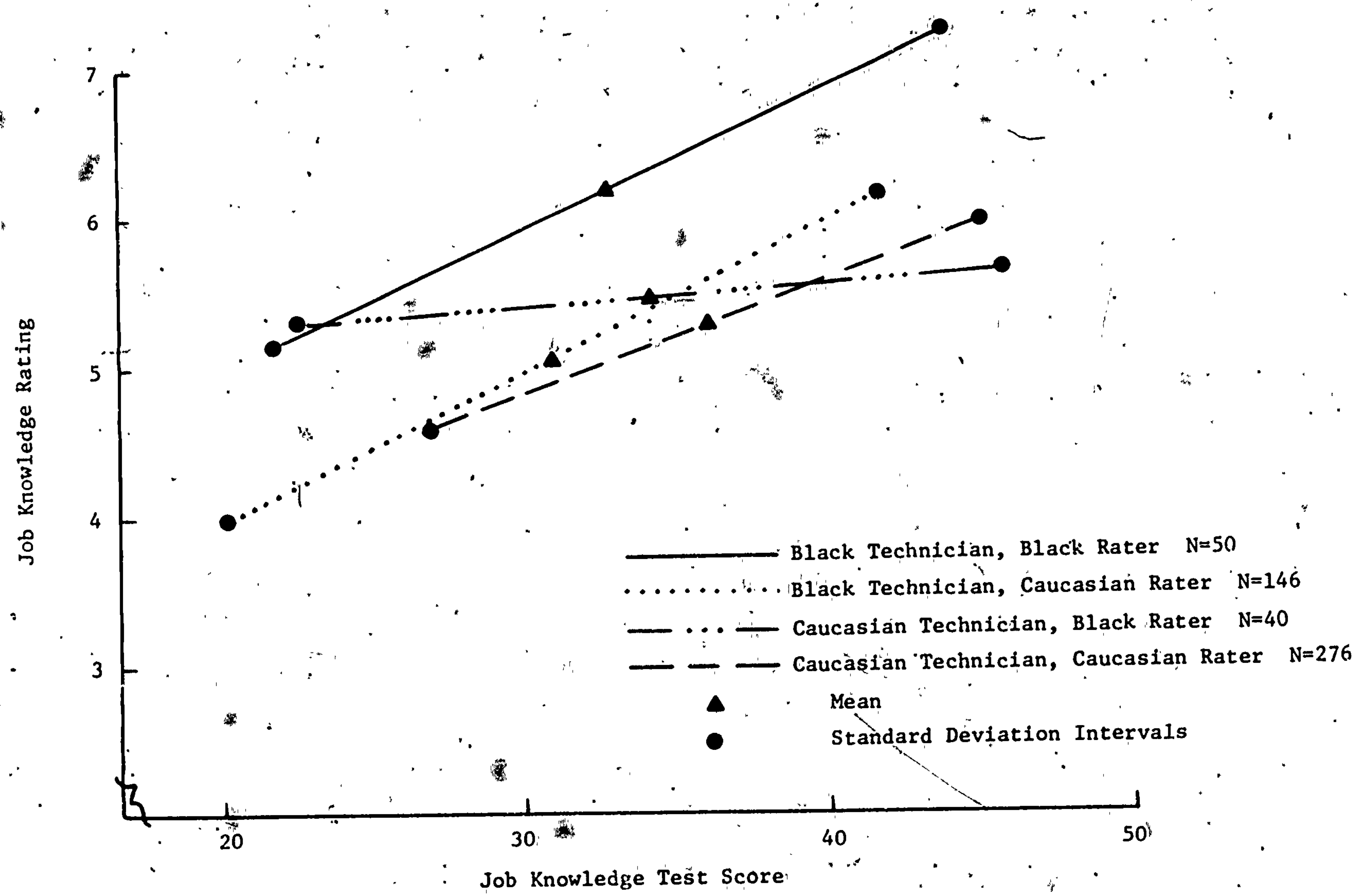
Mean Job Knowledge Ratings and Job Knowledge Test Scores and Correlations of Ratings with Test Scores by Ethnic Group Rater-Ratee Combination

Medical Technicians

ETHNIC GROUP OF TECHNICIAN: OF RATER:	Black Black	Caucasian Black	Black Caucasian	Caucasian Caucasian
	N=50	N=40	N=148	N=275
Job Knowledge Rating	6.2	5.5	5.1	5.3
Job Knowledge Test Score	32.6	34.2	31.0	36.0
Correlation	.50	.09	.56	.39

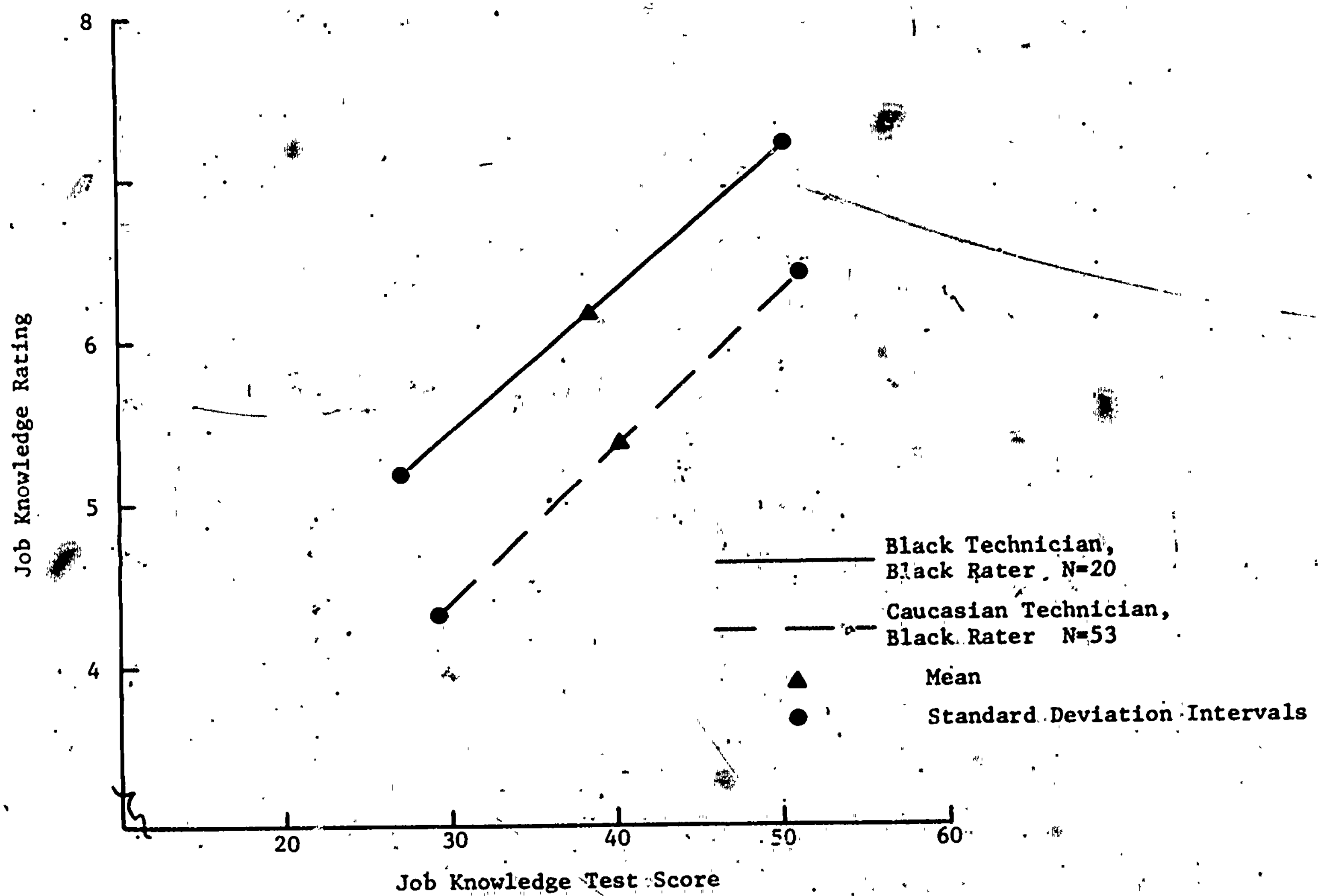
Cartographic Technicians (TOPOCOM)

ETHNIC GROUP OF TECHNICIAN: OF RATER:	Black Black	Caucasian Black	Mexican-American Mexican-American	Caucasian Mexican-American	Black Caucasian	Mexican-American Caucasian	Caucasian Caucasian
	N=20	N=53	N=97	N=26	N=99	N=99	N=240
Job Knowledge Rating	6.2	5.3	5.7	5.3	5.2	5.7	5.7
Job Knowledge Test Score	38.6	40.1	30.5	33.0	34.9	30.6	40.4
Correlation	.44	.42	.29	.61	.35	.37	.36



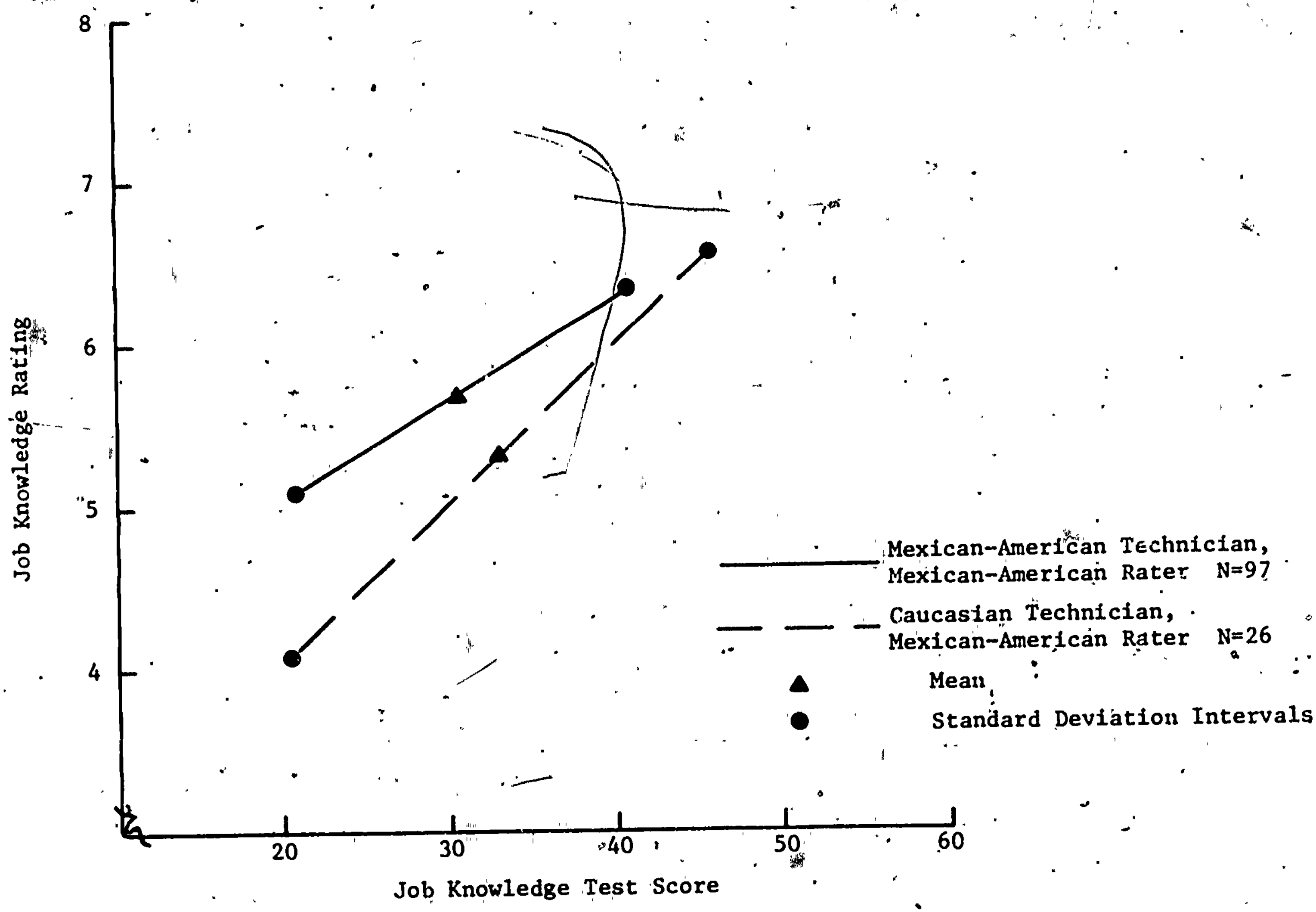
Prediction of Job Knowledge Rating from Job Knowledge Test Score, using data from four groups of Medical Technicians

Fig. VIII-1



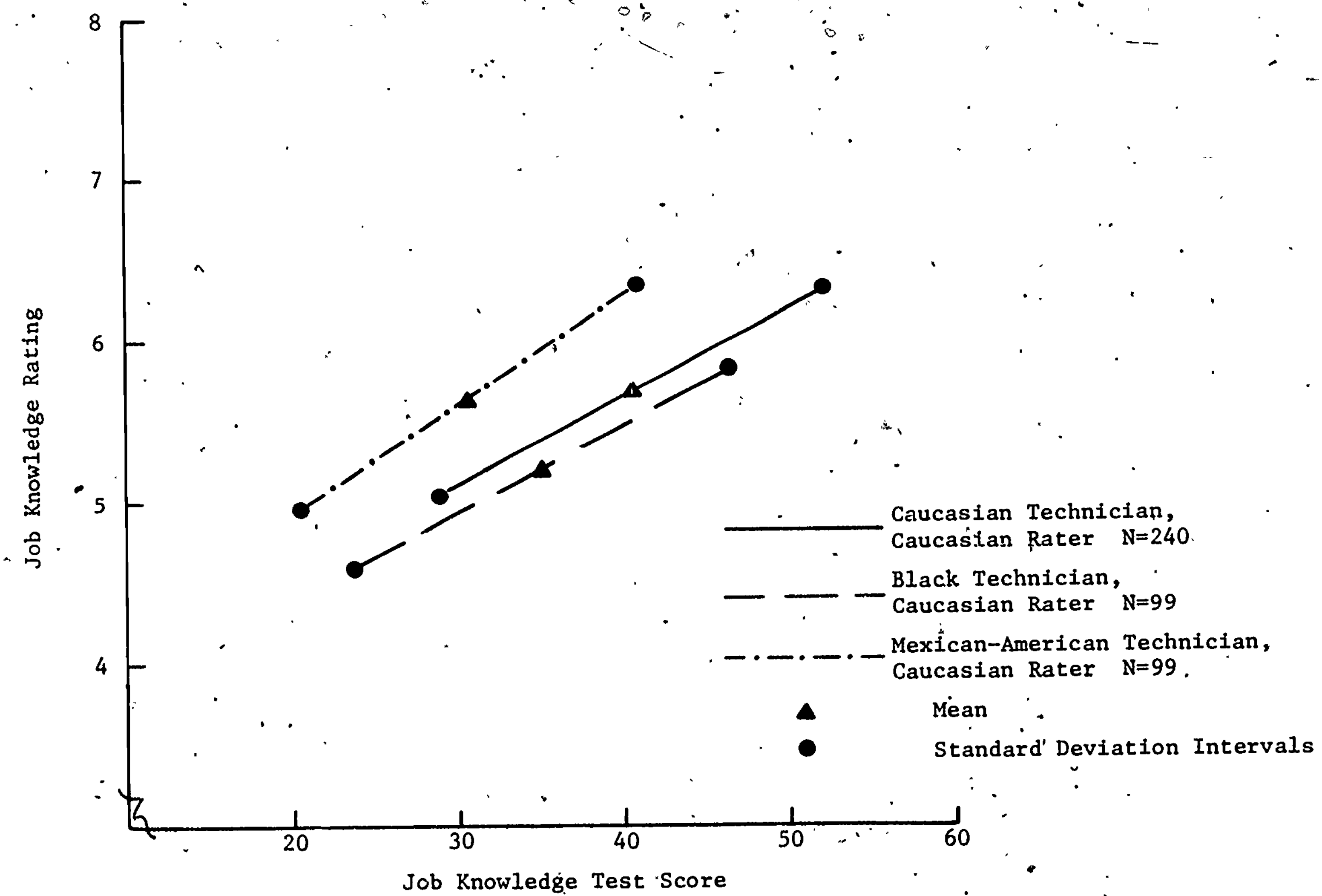
Prediction of Job Knowledge Rating from Job Knowledge Test Score, using data from Cartographic Technicians rated by Black supervisors.

Fig. VIII-2



Prediction of Job Knowledge Rating from Job Knowledge Test Score, using data from Cartographic Technicians rated by Mexican-American supervisors.

Fig. VIII-3



Prediction of Job Knowledge Rating from Job Knowledge Test Score, using data from Cartographic Technicians rated by Caucasian supervisors.

Fig. VIII-4

Table VIII-7

Correlation of Learning Ability Rating with Aptitude Tests and Job Knowledge Test

by Ethnic Group Rater-Ratee Combination

Medical Technicians

ETHNIC GROUP OF TECHNICIAN: OF RATER:	Black Black	Caucasian Black	Black Caucasian	Caucasian Caucasian
	N=50	N=40	N=148	N=275
Subtraction & Multiplication	.43	.17	.39	.27
Vocabulary	.44	.10	.18	.08
Hidden Figures	-.12	.01	-.01	.13
Necessary Arithmetic Operations	.34	.34	.34	.29
Pin-Dexterity	.07	.22	.21	.25
Number Comparison	.29	.09	.27	.16
Gestalt Completion	.19	.07	.17	.20
Picture-Number	.03	.10	.07	.23
Paper Folding	.24	.25	.23	.28
Job Knowledge	.59	.28	.53	.36

Table VIII-8

Correlation of Learning Ability Rating with Aptitude Tests, Job Knowledge Test, and Work Samples

by Ethnic Group Rater-Ratee Combination

Cartographic Technicians (TOPOCOM)

ETHNIC GROUP OF TECHNICIAN: OF RATER:	Black Black	Caucasian Black	Mexican-American Mexican-American	Caucasian Mexican-American	Black Caucasian	Mexican-American Caucasian	Caucasian Caucasian
	N=20	N=53	N=97	N=26	N=99	N=99	N=240
Coordination	.47	-.04	.15	.07	.11	.14	.21
Hidden Figures	.33	.32	.37	.36	.22	.35	.21
Vocabulary	.48	.06	.03	.21	.13	.03	.02
Object-Number	.42	.08	.03	.46	.14	.14	.04
Card Rotations	.42	.35	.27	.44	.21	.08	.29
Arithmetic	.59	.20	.31	.36	.37	.26	.22
Map Planning	.51	.33	.35	.53	.28	.29	.37
Surface Development	.36	.26	.31	.61	.36	.27	.32
Maze Tracing Speed	.51	.32	.33	.57	.16	.22	.29
Following Oral Directions	.44	.19	.24	.51	.28	.31	.32
Identical Pictures	.34	.26	.26	.40	.31	.19	.17
Extended Range Vocabulary	.22	-.09	.05	.09	.12	.13	-.08
Necessary Arithmetic Operations	.55	.23	.33	.46	.26	.31	.29
Job Knowledge	.48	.55	.42	.66	.30	.46	.38
Geometric Restitution	.11	.12	.19	.48	.11	.22	.20
Logical Contouring	.43	-.06	.33	.38	.13	.24	.11
Pull-Up	.67	.41	.44	.41	.15	.33	.24

Table VIII-9

Correlation of Learning Ability Rating with Aptitude Tests

by Ethnic Group Rater-Ratee Combination

Cartographic Technicians (Coast & Geodetic Survey)

ETHNIC GROUP OF TECHNICIAN: OF RATER:	Black Black	Caucasian Black	Black Caucasian	Caucasian Caucasian
	N=16	N=21	N=28	N=42
Coordination	.55	.19	.15	.04
Hidden Figures	.42	.58	.33	.47
Vocabulary	.72	.14	.04	-.11
Object-Number	.50	.02	.29	.12
Card Rotations	.79	.64	.23	.26
Arithmetic	.55	.60	.60	.40
Map Planning	.35	.47	.35	.39
Surface Development	.72	.60	.38	.38
Maze Tracing Speed	.43	.64	.00	.40
Following Oral Directions	.61	.64	.28	.42
Identical Pictures	.68	.59	.38	.41
Extended Range Vocabulary	.55	-.04	.18	-.24
Necessary Arithmetic Operations	.61	.64	.25	.43

Table VIII-10

Correlation of Learning Ability Rating with Aptitude Tests and Work Sample

by Ethnic Group Rater-Ratee Combination

Inventory Management Specialists

ETHNIC GROUP OF TECHNICIAN: OF RATER:	Black Black	Caucasian Black	Black Caucasian	Mexican-American Caucasian	Caucasian Caucasian
	N=19	N=26	N=104	N=67	N=178
Number Comparison	.48	.25	.32	.35	.36
Hidden Figures	.49	-.06	.37	.30	.08
Vocabulary	.10	.13	.12	.16	.12
Object-Number	.21	.11	-.06	.12	.07
Letter Sets	.45	.33	.28	.25	.23
Nonsense Syllogisms	.45	.40	.23	.45	.08
Subtraction & Multiplication	.53	.54	.33	.34	.28
Extended Range Vocabulary	-.04	.36	.06	.17	.17
Necessary Arithmetic Operations	.46	.24	.34	.42	.28
Following Oral Directions	.46	.07	.31	.33	.20
Inference	.54	.32	.28	.20	.18
FSEE - Verbal	.09	.50	.25	.26	.22
- Abstract	.43	.27	.36	.23	.13
- Quantitative	.64	.36	.39	.35	.26
Work Sample	.47	.28	.35	.31	.28

Table VIII-11

Correlation of Objective Measures with Learning Ability Ratings

	Black raters			Caucasian raters		
	Higher validities for Blacks	Equal	Higher validities for Caucasians	Higher validities for Caucasians	Equal	Higher validities for Blacks
Medical Technicians	5	1	4	5	0	5
Cartographic Technicians (TOPOCOM)	15	0	2	9	0	8
Cartographic Technicians (C & G)	7	0	6	7	1	5
Inventory Management Specialists	11	0	4	2	1	12
Total	38	1	16	23	2	30

	Mexican-American raters			Caucasian raters		
	Higher validities for Mexican-Americans	Equal	Higher validities for Caucasians	Higher validities for Caucasians	Equal	Higher validities for Mexican-Americans
Cartographic Technicians (TOPOCOM)	3	0	14	6	0	11
Inventory Management Specialists				1	1	13
Total	3	0	14	7	1	24

Table VIII-12

Average Correlation of Learning Ability Rating with Objective Measures

by Ethnic Group Rater-Ratee Combination

Occupation	Ethnic group of ratee	Ethnic group of rater		
		Black	Mexican-American	Caucasian
Medical Technician	Black	.27		.25
	Caucasian	.17		.23
Cartographic Technician (TOPOCOM)	Black	.44		.22
	Mexican-American		.26	.24
	Caucasian	.22	.42	.22
Cartographic Technician (C & G)	Black	.59		.27
	Caucasian	.47		.27
Inventory Management Specialist	Black	.40		.26
	Mexican-American			.28
	Caucasian	.28		.20
Total	Black	.44		.25
	Mexican-American		.26	.26
	Caucasian	.28	.42	.25

Chapter IX

Factor Analysis of Predictor and Criterion Variables by Ethnic Group

The relationships of single measurements of aptitude or ability to a single criterion of job performance, or of a combination of several measures of aptitude to a single criterion, are important in assessing the validity of selection programs. However, these kinds of analyses would not necessarily show the existence of different patterns of abilities from one ethnic group to another, if these patterns did in fact exist. Previous work by Lesser, Fifer, and Clark (1965), Stodolsky and Lesser (1967), and Flaugher (1971), found different patterns of mean scores from one cultural group to another.

To explore the possibility that such differential patterns would be found in the present data, test, work sample, and supervisors' rating data were factor analyzed separately by ethnic group for each occupation. Selected background variables were added by extension. Principal factor solutions were obtained, and for each sample five factors were retained and rotated orthogonally by the varimax procedure. Adherence to a strict numerical criterion (eigenvalue of 1.00 or greater) would have called for retaining only four factors in several instances, but it was decided to retain a uniform number for all analyses for ease of comparison.

Medical Technicians

These analyses, done separately for Black and Caucasian technicians, contained 28 variables: supervisors' ratings on nine scales, peer ratings on nine scales, nine aptitude tests, and the Job Knowledge Test. Age, Sex, and Civil Service Salary Grade were added by extension. For both supervisor and peer ratings, the value used for each technician was the average rating; that is, if four co-workers rated an individual on Flexibility, the score used for that scale was the average of these four ratings.

The results for Black technicians are shown in Table IX-1. (Loadings of .25 or higher are underlined; loadings of .23 or .24 are in parentheses.) Factor I appears to be primarily a supervisors' general opinion factor, shared somewhat by the opinions of co-workers, as shown by the loadings on peer ratings. The loadings on the Job Knowledge Test and the Subtraction and Multiplication Test indicate that the abilities measured by these two tests were apparently taken into account by the supervisor and, to some extent, by the co-workers in forming opinions reached in their ratings. Salary Grade also receives a positive loading on Factor I, indicating that salary grade of an individual may influence supervisors' ratings and, to a lesser extent, ratings of co-workers.

Factor II appears to represent the co-workers' general opinion. This factor also has substantial loadings on most of the scales for supervisors' ratings, indicating a further sharing of opinion not covered in Factor I. Although, again, the Job Knowledge Test has a loading on this factor, neither the Subtraction and Multiplication Test nor any of the other aptitude tests apparently relate to this portion of co-workers' evaluations. Salary Grade also receives a positive loading on this factor.

Factor III has its highest loading on the Pin-Dexterity Test, and the next highest on the Gestalt Completion and Number Comparison tests. In fact, Factor III picks up some loading from all of the aptitude tests other than Vocabulary. This seems to be an analytical function factor which has been found in other studies. The high loading on the Pin-Dexterity Test probably is a reflection of the rather complex directions for this test. It should be pointed out that neither Factor III nor either of the remaining two factors picks up any loadings on supervisors' ratings or co-workers' ratings.

Factor IV has its highest loading on the Necessary Arithmetic Operations Test, and its next highest loading on the Vocabulary Test, with substantial loadings on Paper Folding, Picture-Number, Subtraction and Multiplication, and Job Knowledge tests. This factor is probably measuring verbal ability primarily, including the ability needed to understand the verbal directions on otherwise nonverbal tests. The loadings of this factor on supervisors' ratings of Job Knowledge and Learning Ability may indicate some real relationship other than chance variance, but such small loadings should be interpreted with caution. The negative loading on Age perhaps reflects a speededness in this factor, in that with increasing age, individuals tend to score lower on speeded tests.

Factor V, with its highest loadings on the Subtraction and Multiplication Test, and the only other loadings on Number Comparison and on Age (negative loading), may probably best be interpreted as speed of numerical operations.

For the Caucasian sample, the loadings shown in Table IX-2 reveal a similar pattern. Factor I represents predominantly supervisors' overall opinion, with all nine rating scales receiving quite high loadings. This opinion is also reflected in part in ratings by co-workers, with four of the rating scales receiving loadings over .25 and none with a loading lower than .19. The Job Knowledge Test has a positive loading, but none of the aptitude tests has any sizeable loading. Salary Grade again receives a positive loading on this factor.

Factor II reflects predominantly co-worker overall opinion, as for the Black technicians, also shared in part by supervisors, with loadings of .25 or higher on six of the nine supervisors' rating scales. Salary Grade has a loading of .26, but neither the Job Knowledge Test nor any of the aptitude tests show loadings on this factor.

Factor III, with highest loadings on the Paper Folding, Gestalt Completion, and Pin-Dexterity tests, and substantial loadings on the Hidden Figures, Number Comparison, and Picture-Number tests, probably is a measure of analytical functioning. This factor also has a positive loading on supervisors' rating of Flexibility, which indicates that the abilities measured by tests of spatial visualization, speed of closure, and finger dexterity were being considered when supervisors evaluated technicians on their ability to shift readily from one activity to another. Age receives a negative loading, again probably reflecting speededness of the tests, or perhaps less flexibility in shifting from one activity to another.

Factor IV appears to be a measure of ability to work with numbers. The only substantial loadings are those on the Subtraction and Multiplication, Number Comparison, Necessary Arithmetic Operations, and Picture-Number tests. None of the ratings has a substantial loading on this factor.

Factor V has its highest loadings on the Vocabulary Test, substantial loadings on the Necessary Arithmetic Operations and Job Knowledge Tests, and loading to a lesser extent on the Paper Folding Test. Co-workers' ratings on Communication (ability to communicate effectively) and supervisors' ratings on the same scale have positive loadings. Apparently, Factor V is a verbal ability factor.

The ratings of Black technicians by both supervisors and co-workers are related to the abilities measured by the Job Knowledge Test, and the supervisors' evaluations are related to some extent to the abilities measured by the Subtraction and Multiplication Test; but other than that, there seems little overlap between the abilities measured by the aptitude tests and supervisors' and co-workers' evaluation of job performance.

The ratings of Caucasian technicians by supervisors and co-workers are related to some degree to the abilities measured by the Job Knowledge Test. Two rating scales, Flexibility for supervisors' ratings and Communication for co-workers' ratings, showed some overlap with aptitude test measurement.

On the whole, the structure of the five factors for the two groups is quite similar. There appears to be little overlap of supervisors' or co-workers' ratings with aptitude measures, although for both groups the Job Knowledge Test, as a performance measure, loads on the rating factors.

Cartographic Technicians

The next three analyses show the results for Cartographic Technicians at the Army Topographic Command (TOPOCOM). Twenty-five variables were included in these analyses: supervisors' ratings on eight scales, 13 aptitude tests, the Job Knowledge Test, and three work sample tasks. In addition, Age, Sex, Salary Grade, and Years of Experience were added by extension.

Table IX-3 shows the rotated factor loadings for Black Cartographic Technicians at TOPOCOM installations.

Factor I reflects supervisors' general opinion, with high loadings on all rating scales, evidence of a general "halo" effect in the ratings. Only one of the tests, Civil Service Arithmetic, has a loading which may possibly be meaningful (.23). Salary Grade has a substantial loading (.35).

Factor II has its highest loadings on the Surface Development, Civil Service Arithmetic, and Following Oral Directions tests. The Necessary Arithmetic Operations, Identical Pictures, Maze Tracing Speed, Card Rotations, and Hidden Figures tests also have high loadings. The Object-Number Test has a substantial loading as well. The Job Knowledge Test has a high loading, and the Pull-up work sample task has a substantial loading. Age has a substantial negative loading and there is also a negative loading on Years of Experience, which may be meaningful. The Learning Ability rating also has a substantial loading. Again, this may be an analytical functioning factor, but in view of the high negative loading on Age, it may be more accurate to call this factor speed of functioning at perceptual and physical tasks.

Factor III has its highest loadings on the Vocabulary and Extended Range Vocabulary tests. There are also loadings on the Necessary Arithmetic Operations and Job Knowledge tests, as well as on Age and Sex, and to some degree on Years of Experience. This appears to be a verbal ability factor. The positive loading on Sex indicates a slight difference on this factor in favor of women.

Factor IV has its highest loadings on the Logical Contouring and

Pull-up work sample tasks. There are also loadings on the Job Knowledge Test and the Hidden Figures Test, with a possibly meaningful loading on the Maze Tracing Speed Test. Salary Grade and Years of Experience also receive positive loadings. This may be a job knowledge or experience factor, with no apparent differentiation for age.

Factor V has its highest loadings on the Coordination, Maze Tracing Speed, and Identical Pictures tests, with a possibly meaningful loading on Map Planning. There are negative loadings on the Job Knowledge Test and Years of Experience. It is possible that this is a spatial ability factor, although the absence of loadings on the Card Rotations and Surface Development tests argues against that, leading to the tentative conclusion that it is an error factor.

Table IX-4 shows the loadings for Mexican-American technicians at TOPOCOM installations.

Factor I is again the rater "halo" factor, with substantial loadings on all eight rating scales. It also receives substantial loadings on the Job Knowledge Test and on the Pull-up work sample task, while the loading on the Hidden Figures Test is possibly meaningful.

Factor II has its highest loadings on the Necessary Arithmetic Operations and Surface Development tests, and other substantial loadings on the Following Oral Directions, Hidden Figures, Job Knowledge, Map Planning, Maze Tracing Speed, Card Rotations, and Civil Service Arithmetic tests. It also has a possibly meaningful loading on the Logical Contouring work sample task, and positive loadings on the Learning Ability and Interest ratings. There are negative loadings on Age, Salary Grade, and Years of Experience. This can probably be correctly described as an analytical functioning factor.

Factor III has its highest loadings on the Maze Tracing Speed and Identical Pictures tests. There are also substantial loadings on Civil Service Arithmetic, the Pull-up and Logical Contouring work sample tasks, and the Coordination, Hidden Figures, Surface Development, Map Planning, and Job Knowledge tests. There are negative loadings on Age and Years of Experience. Perhaps this can be called ability at visualizing spatial relationships.

Factor IV is a verbal ability factor. The two vocabulary tests receive very high loadings. Following Oral Directions also receives a positive loading. It should be noted that in contrast to earlier analyses, Age did not receive a positive loading here. Apparently, the older Mexican-American technicians had no advantage over the younger ones on tests of English vocabulary.

Factor V has its highest loadings on the Restitution and Logical Contouring work sample tasks, and other positive loadings on Card Rotations, the Pull-up work sample task, and Salary Grade. There is a negative loading on the Object-Number Test. Since all three work sample tasks had loadings on this factor, perhaps it could be considered a work sample factor.

Table IX-5 shows the loadings for Caucasian technicians at TOPOCOM installations.

Factor I reflects the supervisors' general opinion, with substantial loadings on all rating scales, indicating a substantial halo effect. The Job Knowledge Test also receives a substantial loading, as does Salary Grade.

Factor II has its highest loadings on the Maze Tracing Speed, Identical Pictures, Map Planning, and Card Rotations tests. There are also substantial loadings on the Hidden Figures, Civil Service Arithmetic, Following Oral Directions, Coordination, Necessary Arithmetic Operations, and Job Knowledge tests. The supervisors' ratings on Dexterity also receive a substantial loading and there are negative loadings on Age and Years of Experience. This may be an analytical functioning factor, but it appears more accurate to call it speed of functioning at perceptual and physical tas

Factor III has its highest loadings on the Pull-up work sample task, and the Job Knowledge and Surface Development tests. There are substantial loadings on the Logical Contouring and Restitution work sample tasks, and the Following Oral Directions, Maze Tracing Speed, Card Rotations, Hidden Figures, and Necessary Arithmetic Operations tests. There is a positive loading on Learning Ability rating. A negative loading on Sex indicates that men do better on this factor than do women. This factor perhaps can be described as ability at spatial visualization, although the loadings on the Following Oral Directions and Necessary Arithmetic Operations tests, and to some extent the loading on the Job Knowledge Test, are not consistent with this description.

Factor IV is again the verbal ability factor, with very high loadings on both vocabulary tests and other loadings on the Job Knowledge and Necessary Arithmetic Operations tests, as well as on Age and Sex, a pattern previously found.

Factor V has its highest loadings on the Necessary Arithmetic Operations and Civil Service Arithmetic tests, with other loadings on the Following Oral Directions, Job Knowledge, and Map Planning tests. There are negative loadings on Age and Years of Experience. Probably this is a numerical ability factor, with an element of speededness.

In summary, for Cartographic Technicians (TOPOCOM), Factor I, representing supervisors' general opinion, and Factor II, an analytical functioning factor with a speededness element, are like those found for Medical Technicians, and loading patterns are similar for the three ethnic groups. Factor III is a verbal ability factor for the Black sample and a spatial visualization factor for the Mexican-American and Caucasian samples, while Factor IV appears as the verbal ability factor for Mexican-Americans and Caucasians and a combination of spatial visualization, job knowledge, and experience for Blacks. Factor V, which accounts for a very small part of the total variance, differs somewhat for the three groups, in that for Blacks the factor represents primarily spatial visualization ability, for Mexican-Americans performance on the work sample tasks, and for Caucasians numerical ability, because of high loadings on the Necessary Arithmetic Operations and Civil Service Arithmetic tests. There were also loadings for knowledge and experience on this factor for Caucasians.

While some differences are seen in the factor patterns for the separate ethnic groups, the differences are not substantial. The loadings are quite similar within factors, and reflect essentially the same relationships between predictor and criterion variables described in earlier chapters.

The next two analyses show the results for Cartographic Technicians at the Coast and Geodetic Survey. There are 21 variables in these analyses, including supervisors' ratings on eight scales and 13 aptitude tests. Four variables, Age, Sex, Salary Grade, and Years of Experience were added by extension.

Table IX-6 shows the rotated factor loadings for Black Technicians at the Coast and Geodetic Survey. Factor I is the rater halo or supervisors' opinion factor, with high loadings on all eight rating scales. This halo, or general opinion, appears to be related to performance on several of the tests. The Civil Service Arithmetic, Identical Pictures, and Object-Number tests have substantial loadings, and the Card Rotations and Surface Development tests have loadings which may be meaningful. Salary Grade also has a substantial loading on this factor.

Factor II has substantial loadings on all of the tests other than the two vocabulary tests and the Object-Number Test. The Vocabulary Test, which was administered with speeded conditions, has a loading which may be meaningful. The rating on Learning Ability also has a loading which may be meaningful. This appears to be a general, other than verbal, ability factor, although perhaps it could be labeled ability at speeded tasks.

Factor III has its highest loadings on the Hidden Figures and Object-Number tests, and substantial loadings on the Card Rotations, Map Planning, Surface Development, Maze Tracing Speed, and Identical Pictures tests. In addition, the Civil Service Arithmetic Test has a loading which may be meaningful. The Learning Ability and Dexterity ratings have substantial loadings, and the Job Knowledge rating has a loading which also may be meaningful. Age and Sex have substantial negative loadings. This appears to be an analytical functioning factor.

Factor IV has its highest loadings on the two vocabulary tests, and other substantial loadings on the Civil Service Arithmetic, Surface Development, and Following Oral Directions tests. Age and Salary Grade also have substantial loadings. This appears to be a verbal ability factor, but it is not clear why the Surface Development Test should load on a verbal factor.

Factor V has its highest loading on the Coordination Test, and other substantial loadings on the Card Rotations, Map Planning, Maze Tracing Speed, Following Oral Directions, and Identical Pictures tests. It also receives a substantial loading on the Need for Supervision rating and a substantial negative loading on Age. This appears to be primarily a perceptual and spatial relations factor, with an element of speededness negatively related to age.

Table IX-7 shows the factor loadings for Caucasian technicians at the Coast and Geodetic Survey. Factor I again is the rater halo or supervisors' opinion factor, with high loadings on all eight rating scales. This opinion is apparently influenced by the abilities measured by several of the tests. The Hidden Figures, Civil Service Arithmetic, Following Oral Directions, and Necessary Arithmetic Operations tests all have substantial loadings, and the loading on the Surface Development Test may also be meaningful. Salary Grade also has a loading which may be meaningful.

Factor II has loadings on all of the tests except the two vocabulary tests. It also has loadings on several of the rating scales. Ratings of Learning Ability, Job Knowledge, Dexterity, and Need for Supervision all have substantial loadings, and the loadings on the Overall scale may also be meaningful. Age, Salary Grade, and Years of Experience all have substantial negative loadings, so apparently the younger technicians are better at what this factor is measuring. This probably is the analytical functioning factor.

Factor III is obviously a verbal ability factor, with very high loadings on the two vocabulary tests and a possibly meaningful loading on the Necessary Arithmetic Operations Test. Age and Salary Grade also have substantial loadings.

Factor IV has its highest loadings on the Coordination and Identical Pictures tests, with substantial negative loadings on the Dexterity rating, Civil Service Arithmetic, and Necessary Arithmetic Operations tests. There is some possibility that this is an error factor, as it is difficult to explain why those who do best on the Coordination Test should be rated poorly in Dexterity and also do poorly on an arithmetic test.

Factor V has substantial positive loadings on the Object-Number, Civil Service Arithmetic, Map Planning, and Following Oral Directions tests, and substantial negative loadings on the Hidden Figures Test, Age, and Sex, and a possibly meaningful negative loading on Salary Grade. This is most likely an error factor.

Again, there are some differences in factor order between the Black and Caucasian samples. Several factors appear to be error factors, in that some of the loadings are difficult to rationalize. The smallness of the Coast and Geodetic Survey samples may explain some anomalies.

The last three analyses show the results for Inventory Management Specialists. There are 32 variables in these analyses, including supervisors' ratings on ten scales, 12 aptitude tests, and ten scores on the work sample. Again, Age, Sex, Salary Grade, and Years of Experience were added by extension.

Table IX-8 shows the rotated factor loadings for Black Inventory Management Specialists.

Factor I is clearly rater's overall halo, with all rating scales receiving high loadings. The Subtraction and Multiplication Test also receives a substantial loading on this factor.

Factor II is performance on the Work Sample measure. To some extent, this can be interpreted as a scorer's halo factor, since all but one of the Work Sample scores received a substantial loading on this factor. The factor loadings on several of the Work Sample scales run considerably lower than the supervisors' rater halo shown in Factor I. Thus, it can reasonably be argued that the scorer's halo effect is minimal. One test, the Inference Test, received a substantial loading on this factor, and Salary Grade similarly has a substantial loading.

Factor III has loadings on a number of the aptitude tests. The highest loadings are on the two vocabulary tests, with other high loadings on the Federal Service Entrance Examination, the Inference Test, and the Necessary Arithmetic Operations Test. Other tests with substantial loadings include Letter Sets, Nonsense Syllogisms, and Following Oral Directions. One of the Work Sample scores, "Organizes systematically," receives a substantial loading on this factor as well. This probably can best be interpreted as a verbal reasoning factor.

Factor IV has its highest loadings on the Hidden Figures and Number Comparison tests, with other substantial loadings on the Following Oral Directions, Letter Sets, Subtraction and Multiplication, and Necessary Arithmetic Operations tests, and Federal Service Entrance Examination. The Work Sample score, "Productivity," also has a substantial loading, as does the rating on Learning Ability. Age has a substantial negative loading. This factor probably can be considered as ability at speeded operations.

Factor V has its highest loadings on two of the Work Sample scores, "Maintains controls" and "Shows Inventory Manager knowledge." It also has a substantial loading on "Takes problem solving action." The Object-Number Test and the Subtraction and Multiplication Test also have substantial loadings, while Age and Years of Experience have substantial negative loadings. This may be an error factor or perhaps might be considered memory for details.

Table IX-9 shows the rotated factor loadings for Mexican-American Inventory Management Specialists. Factor I represents the rater's overall judgment, or halo, with very substantial loadings on all ten rating scales. The Nonsense Syllogisms Test also has a substantial loading, as does the "Analyzes problems" score from the Work Sample. Sex has a negative loading, indicating a slight tendency to rate Mexican-American males higher.

Factor II reflects primarily performance on the Work Sample, with substantial loadings on all of the Work Sample scores except "Productivity." There are also substantial loadings on several of the aptitude tests, including the Extended Range Vocabulary, Inference, Federal Service Entrance Examination, and Number Comparison tests.

Factor III has its highest loadings on the two vocabulary tests and the Federal Service Entrance Examination. There are also substantial loadings on several other tests, including Inference, Necessary Arithmetic Operations, Hidden Figures, Following Oral Directions, Letter Sets, Subtraction and Multiplication, and Object-Number. Several of the Work Sample scores also have substantial loadings on this factor. These include "Organizes systematically," "Quality of action," "Takes problem solving action," "Analyzes problems," "Follows directions," and scorer's rating of overall Work Sample performance. The supervisor's rating on effective communication may also be meaningful. This factor appears to represent verbal reasoning ability, primarily.

Factor IV has its highest loadings on the Number Comparison and Subtraction and Multiplication tests. There are also substantial loadings on the Letter Sets, Necessary Arithmetic Operations, Following Oral Directions, and Object-Number tests, and the Federal Service Entrance Examination. There are also substantial loadings on supervisors' ratings of Learning Ability and effective Communication, and on the Work Sample rating of overall performance. There is a negative loading on the supervisors' rating of Cooperation and on Age. Sex has a positive loading, indicating better performance by women. This factor appears to represent ability at speeded operations.

Factor V has its highest loadings on the Nonsense Syllogisms and Following Oral Directions tests, and on the Work Sample "Productivity" score. The Letter Sets and Necessary Arithmetic Operations tests have substantial loadings, as do two additional scores from the work sample: "Analyzes problems" and "Maintains controls." Sex has a negative loading. This may be an error factor, or it may possibly represent ability to work effectively with somewhat unfamiliar types of materials.

Table IX-10 shows the rotated factor loadings for Caucasian Inventory Management Specialists. Factor I is clearly a rater halo factor, with loadings on each of the ten rating scales. There is also a substantial loading on the Subtraction and Multiplication Test.

Factor II is a work sample performance factor, with very substantial loadings on all ten scores. There are also substantial loadings on three of the aptitude tests: Number Comparison, Following Oral Directions, and Vocabulary.

Factor III has its highest loadings on the two vocabulary tests and the Federal Service Entrance Examination. There are also very substantial loadings on a number of the other aptitude tests. These are Letter Sets, Inference, Necessary Arithmetic Operations, Nonsense Syllogisms, and Hidden Figures. None of the rating scales or work sample scores has any substantial loading on this factor, but Years of Experience has a substantial negative loading. This factor apparently represents primarily verbal reasoning ability.

Factor IV has its highest loading on the Number Comparison Test, and the next highest loading on the Subtraction and Multiplication Test. There were also substantial loadings on the Letter Sets, Hidden Figures, Necessary Arithmetic Operations, and Following Oral Directions tests. Age has a substantial negative loading. Apparently this factor represents ability at speeded operations.

Factor V has its highest loading on the Object-Number Test, and a substantial loading on the Subtraction and Multiplication Test. There is a negative loading on the Hidden Figures Test and on the "Productivity" Work Sample score. Quite possibly this is an error factor, since there does not appear to be a logical explanation for these relationships.

For Inventory Management Specialists, the first four factors show quite uniform patterns for the three ethnic groups. The fifth factor appears to be an error factor, since relationships of variables loading on this factor are not readily interpretable for any of the three ethnic groups. On the whole, however, the factor patterns are quite similar across ethnic groups.

Summary and Conclusions

In general, differing patterns of abilities are not readily discernible from one ethnic group to another in the three occupations studied. Factor structures appear to be quite similar across occupations and ethnic groups. With a few exceptions, abilities measured by the aptitude tests do not appear to have been taken into account in ratings by supervisors of various aspects of job performance. However, supervisors do appear to be reflecting in their ratings to some degree aspects of job performance measured by the Job Knowledge Test and Work Sample. Various aptitude tests are related to the latter criterion measures to a much greater degree than to ratings.

Table IX-1

Factor Analysis of Predictor and Criterion Variables

Medical Technicians

Black Sample

Variables Included in Factor Analysis	Rotated Loadings				
	I	II	III	IV	V
Supervisors' Ratings					
Flexibility	80	27	10	06	11
Organization	82	29	02	-03	13
Interest	80	19	-03	-01	10
Learning Ability	83	33	17	19	06
Job Knowledge	85	22	06	22	01
Technique	84	30	04	06	04
Need for Supervision	80	31	00	07	02
Communication	78	22	03	16	-06
Overall	85	27	05	-03	06
Peer Ratings					
Flexibility	21	76	07	-02	04
Organization	(24)	79	-05	05	12
Interest	18	76	00	-02	08
Learning Ability	31	78	09	05	02
Job Knowledge	32	77	02	13	-06
Technique	(24)	83	03	04	-04
Need for Supervision	26	79	04	03	01
Communication	27	67	03	14	06
Overall	21	84	-01	-02	-02
Aptitude Tests					
Subtraction & Multiplication	26	08	(24)	28	71
Vocabulary	13	-03	10	46	06
Hidden Figures	04	-11	35	12	-03
Necessary Arithmetic Operations	19	03	40	60	16
Pin-Dexterity	08	-03	87	-13	16
Number Completion	09	10	62	10	31
Gestalt Completion	03	07	64	22	-03
Picture-Number	-15	17	30	33	09
Paper Folding	00	12	48	34	-07
Job Knowledge Test	45	28	26	41	04
Sum of Squared Loadings	6.98	6.26	2.36	1.30	.76
Variables Added by Extension					
Salary Grade	35	31	02	19	-05
Sex	-07	05	-16	13	08
Age	-04	-10	(-24)	-28	29

Sample size = 168

Table IX-2

Factor Analysis of Predictor and Criterion Variables

Medical Technicians

Caucasian Sample

Variables Included in Factor Analysis	Rotated Loadings I	II	III	IV	V
Supervisors' Ratings					
Flexibility	72	(23)	25	17	-03
Organization	79	32	07	07	-06
Interest	79	17	03	09	00
Learning Ability	80	25	22	15	06
Job Knowledge	75	31	00	-05	19
Technique	77	34	11	02	00
Need for Supervision	85	25	-03	-03	03
Communication	73	(23)	00	-04	22
Overall	82	32	05	04	01
Peer Ratings					
Flexibility	19	77	07	15	-01
Organization	21	81	-02	08	-03
Interest	31	69	-03	09	-01
Learning Ability	27	75	17	13	07
Job Knowledge	(23)	71	08	00	11
Technique	(23)	80	-03	04	-02
Need for Supervision	26	80	02	-09	05
Communication	(24)	67	01	07	25
Overall	26	82	03	02	-06
Aptitude Tests					
Subtraction & Multiplication	08	08	-11	72	11
Vocabulary	02	06	09	02	67
Hidden Figures	03	-04	55	10	11
Necessary Arithmetic Operations	11	03	60	33	39
Pin-Dexterity	09	07	68	17	-12
Number Comparison	-02	05	49	50	-06
Gestalt Completion	06	01	72	-01	-03
Picture-Number	08	15	28	31	03
Paper Folding	09	03	73	-05	(23)
Job Knowledge Test	29	06	19	19	33
Sum of Squared Loadings	6.17	5.92	2.73	1.19	.99
Variables Added by Extension					
Salary Grade	32	26	-22	-11	16
Sex	-01	00	04	08	-03
Age	-06	03	-34	(24)	-20

Sample size = 297

Table IX-3

Factor Analysis of Predictor and Criterion Variables

Cartographic Technicians (TOPOCOM)

Black Sample

Variables Included in Factor Analysis	Rotated Loadings I	II	III	IV	V
Supervisors' Ratings					
Accuracy	91	18	-01	04	08
Interest	84	05	11	-07	-06
Learning Ability	85	31	00	12	04
Job Knowledge	89	21	07	16	-06
Dexterity	82	09	-06	-01	19
Need for Supervision	91	17	07	11	00
Perseverance	82	-02	22	-06	00
Overall	92	15	05	08	00
Aptitude Tests					
Coordination	06	18	-12	05	57
Hidden Figures	13	53	19	26	03
Vocabulary	13	16	87	14	10
Object-Number	16	43	02	-15	10
Card Rotations	09	58	-05	19	14
CS Arithmetic	(23)	73	07	17	-15
Map Planning	09	68	09	16	(23)
Surface Development	15	75	10	12	-03
Maze Tracing Speed	01	55	-03	(23)	38
Following Oral Directions	09	71	20	-06	02
Identical Pictures	10	66	-05	17	35
Extended Range Vocabulary	09	22	88	08	-10
Necessary Arithmetic Operations	13	69	40	09	-03
Work Sample					
Restitution	03	18	13	21	03
Logical Contouring	03	03	14	65	10
Pull-up	06	31	-06	55	-01
Job Knowledge Test	15	61	31	39	-33
Sum of Squared Loadings	6.29	4.92	2.02	1.31	.89
Variables Added by Extension					
Age	-04	-50	53	01	-08
Sex	04	-01	27	-19	-02
Salary Grade	50	-11	18	33	09
Years of Experience	07	(-24)	(24)	44	-34

Sample size = 100

Table IX-4

Factor Analysis of Predictor and Criterion Variables

Cartographic Technicians (TOPOCOM)

Mexican-American Sample

Variables Included in Factor Analysis	Rotated Loadings I	II	III	IV	V
Supervisors' Ratings					
Accuracy	92	19	03	-05	09
Interest	91	(23)	05	06	-08
Learning Ability	90	30	17	00	02
Job Knowledge	93	12	06	-04	06
Dexterity	84	06	22	-07	09
Need for Supervision	92	11	10	03	14
Perseverance	83	03	12	11	-13
Overall	94	10	06	-01	09
Aptitude Tests					
Coordination	06	10	46	03	-03
Hidden Figures	(23)	60	36	-06	02
Vocabulary	04	12	05	84	00
Object-Number	03	17	13	-05	-31
Card Rotations	-03	34	47	08	31
CS Arithmetic	14	38	51	08	-01
Map Planning	20	44	26	-05	13
Surface Development	09	70	31	07	05
Maze Tracing Speed	04	43	75	-05	03
Following Oral Directions	12	68	15	25	-13
Identical Pictures	10	28	72	01	-07
Extended Range Vocabulary	04	15	-02	82	04
Necessary Arithmetic Operations	14	71	22	17	05
Work Sample					
Restitution	09	35	08	03	43
Logical Contouring	19	(23)	33	-12	46
Pull-up	35	10	49	-05	28
Job Knowledge Test	35	71	22	17	05
Sum of Squared Loadings	6.94	3.31	2.71	1.58	.82
Variables Added by Extension					
Age	-13	-39	-40	14	08
Sex	-07	-21	-14	11	-07
Salary Grade	.22	-32	05	-01	33
Years of Experience	-10	-36	(-24)	-08	16

Sample size = 100

Table IX-5

Factor Analysis of Predictor and Criterion Variables

Cartographic Technicians (TOPOCOM)

Caucasian Sample

Variables Included in Factor Analysis	Rotated Loadings I	II	III	IV	V
Supervisors' Ratings					
Accuracy	92	14	06	07	03
Interest	85	04	14	05	09
Learning Ability	84	19	25	-09	17
Job Knowledge	92	06	17	05	04
Dexterity	79	28	05	-05	-05
Need for Supervision	89	13	20	-02	09
Perseverance	81	00	-01	11	03
Overall	92	17	08	-06	09
Aptitude Tests					
Coordination	13	39	-02	-03	08
Hidden Figures	11	48	27	17	07
Vocabulary	04	04	16	84	13
Object-Number	01	30	-01	19	09
Card Rotations	12	56	39	-15	03
CS Arithmetic	14	45	-01	14	53
Map Planning	19	62	15	00	(24)
Surface Development	11	53	50	03	20
Maze Tracing Speed	12	68	32	-12	04
Following Oral Directions	13	42	42	19	40
Identical Pictures	01	66	20	00	03
Extended Range Vocabulary	-04	-05	21	85	10
Necessary Arithmetic Operations	09	30	31	(23)	61
Work Sample					
Restitution	11	14	38	13	10
Logical Contouring	03	10	42	06	00
Pull-up	14	10	58	06	01
Job Knowledge Test	26	27	54	30	36
Sum of Squared Loadings	6.29	3.12	2.10	1.79	1.16
Variables Added by Extension					
Age	05	-52	-09	41	-39
Sex	-17	-11	-37	34	-11
Salary Grade	49	-16	06	08	-15
Years of Experience	17	-42	10	19	-27

Sample size = 240

Table IX-6

Factor Analysis of Predictor and Criterion Variables

Cartographic Technicians (Coast & Geodetic Survey)

Black Sample

Variables Included in Factor Analysis	Rotated Loadings I	II	III	IV	V
Supervisors' Ratings					
Accuracy	91	13	18	05	08
Interest	93	15	11	00	-03
Learning Ability	83	(23)	28	19	14
Job Knowledge	92	14	(24)	04	08
Dexterity	86	02	29	12	03
Need for Supervision	88	07	10	13	25
Perseverance	90	16	-02	04	-05
Overall	93	21	12	02	-01
Aptitude Tests					
Coordination	08	33	-04	10	83
Hidden Figures	14	31	71	22	-13
Vocabulary	02	(24)	10	87	09
Object-Number	29	02	73	06	05
Card Rotations	(23)	53	45	15	31
CS Arithmetic	37	63	(24)	29	11
Map Planning	22	51	32	03	26
Surface Development	(24)	43	58	34	-02
Maze Tracing Speed	-12	64	41	-05	35
Following Oral Directions	20	70	04	44	26
Identical Pictures	31	38	54	-08	44
Extended Range Vocabulary	13	15	17	91	-02
Necessary Arithmetic Operations	17	80	06	21	02
Sum of Squared Loadings	7.02	3.28	2.53	2.19	1.38
Variables Added by Extension					
Age	18	-08	-56	35	-31
Sex	01	-10	-48	10	12
Salary Grade	52	-02	06	27	-16
Years of Experience	06	-17	-22	-15	-06

Sample size = 38

Table IX-7

Factor Analysis of Predictor and Criterion Variables

Cartographic Technicians (Coast & Geodetic Survey)

Caucasian Sample

Variables Included in Factor Analysis	Rotated Loadings I	II	III	IV	V
Supervisors' Ratings					
Accuracy	90	18	01	-14	-12
Interest	86	12	04	20	02
Learning Ability	83	40	-08	-03	-03
Job Knowledge	83	26	12	02	02
Dexterity	69	40	00	-28	-14
Need for Supervision	79	41	-04	-13	-01
Perseverance	78	01	01	18	28
Overall	93	(23)	01	-09	02
Aptitude Tests					
Coordination	-01	32	07	50	-03
Hidden Figures	27	73	-09	01	-40
Vocabulary	09	00	88	08	07
Object-Number	-01	25	10	-04	47
Card Rotations	18	68	03	03	05
CS Arithmetic	28	65	19	-43	30
Map Planning	16	75	05	07	28
Surface Development	(23)	79	07	20	00
Maze Tracing Speed	18	74	-11	15	12
Following Oral Directions	29	76	12	-05	28
Identical Pictures	22	75	-22	(24)	03
Extended Range Vocabulary	-06	04	88	-04	07
Necessary Arithmetic Operations	31	73	(23)	(-24)	05
Sum of Squared Loadings	6.04	5.65	1.77	.83	.77
Variables Added by Extension					
Age	-16	-66	27	02	-26
Sex	-11	-10	22	11	-27
Salary Grade	(24)	-39	26	10	(-24)
Years of Experience	-03	-38	13	06	-02

Sample size = 51

Table IX-8

Factor Analysis of Predictor and Criterion Variables

Inventory Management Specialists

Black Sample

Variables Included in Factor Analysis	Rotated Loadings I	II	III	IV	V
Supervisors' Ratings					
Organization	87	12	05	21	-02
Interest	88	07	07	11	-06
Learning Ability	81	19	12	25	-00
Communication	82	04	06	08	01
Technical Knowledge	87	10	08	06	-01
Stability	85	10	04	12	08
Dependability	85	10	13	07	13
Practical Judgment	89	15	11	08	-00
Cooperation	66	15	-02	-03	(24)
Overall	90	08	10	16	08
Aptitude Tests					
Number Comparison	21	09	07	69	11
Hidden Figures	17	13	22	70	-10
Vocabulary	03	17	87	04	05
Object-Number	-04	-08	10	(23)	50
Letter Sets	20	11	60	49	13
Nonsense Syllogisms	06	21	65	08	-01
Subtraction & Multiplication	29	-03	13	46	25
Extended Range Vocabulary	-09	16	91	-03	-06
Necessary Arithmetic Operations	21	15	71	32	08
Following Oral Directions	08	(24)	53	54	15
Inference	13	26	81	01	15
FSEE	19	17	83	26	11
Work Sample					
Takes Problem Solving Action	11	73	21	01	27
Uses New Procedures	10	88	20	12	21
Shows I. M. Knowledge	19	56	04	00	63
Analyzes Problems	10	71	22	-00	-12
Organizes Systematically	10	73	31	15	-09
Maintains Controls	14	(23)	09	03	74
Follows Directions	13	83	05	(23)	-04
Productivity	07	41	06	43	14
Quality of Actions	17	87	22	02	10
Overall Performance	16	90	17	06	10
Sum of Squared Loadings	7.59	5.59	4.97	2.40	1.64
Variables Added by Extension					
Sex	-09	-02	-13	16	13
Age	-19	08	13	-33	-26
Salary Grade	-03	25	-09	-16	-12
Years of Experience	05	07	-14	-06	-29

Sample size = 114

Table IX-9

Factor Analysis of Predictor and Criterion Variables

Inventory Management Specialists

Mexican-American Sample

Variables Included in Factor Analysis	Rotated Loadings				
	I	II	III	IV	V
Supervisors' Ratings					
Organization	86	-02	10	07	02
Interest	89	-00	-01	-03	-01
Learning Ability	81	09	15	37	05
Communication	72	11	(24)	34	05
Technical Knowledge	75	01	17	13	05
Stability	89	03	04	-02	08
Dependability	84	09	-15	-10	06
Practical Judgment	85	14	06	11	05
Cooperation	63	21	10	-28	12
Overall	90	-04	10	02	04
Aptitude Tests					
Number Comparison	10	26	03	67	11
Hidden Figures	21	03	58	18	01
Vocabulary	-05	19	80	04	10
Object-Number	-06	12	26	29	-32
Letter Sets	-09	14	38	58	37
Nonsense Syllogisms	34	21	40	05	60
Subtraction & Multiplication	20	19	31	63	-05
Extended Range Vocabulary	01	40	76	01	08
Necessary Arithmetic Operations	16	29	59	38	26
Following Oral Directions	20	13	45	28	52
Inference	06	35	65	12	02
FSEE	14	32	76	25	21
Work Sample					
Takes Problem Solving Action	15	68	34	05	-09
Uses New Procedures	06	88	22	13	08
Shows I. M. Knowledge	-06	85	03	11	-02
Analyzes Problems	33	40	30	18	33
Organizes Systematically	22	57	48	-14	03
Maintains Control	-03	68	-00	05	33
Follows Directions	13	80	27	16	07
Productivity	02	17	11	14	86
Quality of Actions	04	86	35	18	08
Overall Performance	08	81	32	29	15
Sum of Squared Loadings	7.22	5.77	4.54	2.23	2.06
Variables Added by Extension					
Sex	-25	18	-04	29	-25
Age	-22	-08	(-23)	-42	-21
Salary Grade	14	(23)	22	(24)	13
Years of Experience	04	12	15	04	13

Sample size = 72

Table IX-10

Factor Analysis of Predictor and Criterion Variables

Inventory Management Specialists

Caucasian Sample

Variables Included in Factor Analysis	Rotated Loadings I	II	III	IV	V
Supervisors' Ratings					
Organization	86	13	-01	08	10
Interest	85	09	04	06	-02
Learning Ability	85	19	09	09	10
Communication	81	09	15	18	-01
Technical Knowledge	86	12	07	11	01
Stability	88	09	07	08	02
Dependability	86	07	06	01	03
Practical Judgment	90	09	19	02	02
Cooperation	66	-11	09	01	-20
Overall	91	14	07	06	04
Aptitude Tests					
Number Comparison	19	28	17	74	03
Hidden Figures	03	15	50	39	-41
Vocabulary	04	26	79	-21	00
Object-Number	03	04	18	20	69
Letter Sets	13	14	70	45	02
Nonsense Syllogisms	05	-02	61	04	11
Subtraction & Multiplication	26	12	07	69	26
Extended Range Vocabulary	08	19	84	-22	02
Necessary Arithmetic Operations	17	16	71	39	01
Following Oral Directions	06	28	58	35	-11
Inference	13	17	78	07	06
FSEE	14	22	87	15	07
Work Sample					
Takes Problem Solving Action	-01	75	20	05	(24)
Uses New Procedures	10	89	18	11	13
Shows I. M. Knowledge	11	78	01	05	13
Analyzes Problems	04	76	20	-07	-11
Organizes Systematically	18	73	19	16	-02
Maintains Controls	14	71	-03	04	-01
Follows Directions	11	83	15	20	06
Productivity	02	66	10	17	-28
Quality of Actions	08	89	(23)	06	-02
Overall Performance	11	88	(23)	04	-12
Sum of Squared Loadings	7.48	6.83	5.10	2.01	1.01
Variables Added by Extension					
Sex	09	17	-09	(24)	20
Age	-17	-16	-22	-35	-19
Salary Grade	21	-02	-15	-04	-02
Years of Experience	02	04	-39	-07	04

Sample size ≐ 200

Chapter X

Comparisons of Background and Task Variables by Occupation by Ethnic Group

As described in Chapter II, a comprehensive personal history questionnaire was included in the test battery for each occupation studied. The questionnaire for each occupation also contained a detailed task list. The purpose was to explore possible differences in composition of the ethnic samples, to relate background variables to measures of aptitude and job performance by ethnic group, and to determine whether there were variations in the type of work performed by members of different ethnic groups.

Composition of Ethnic Samples

Tables X-1 (Medical Technicians), X-2 (Cartographic Technicians), and X-3 (Inventory Management Specialists) show the distribution of selected background variables expressed as percent of total sample, by ethnic group. As will be seen, the variables selected for comparison are those considered most likely to be moderators of job performance: age, sex, education, source and length of training on and off the job, source and amount of experience on and off the job, and present GS-level.

For Medical Technicians, computations of chi-square on Black and Caucasian distributions of selected variables show significant differences at the .01 level for Age (Caucasian technicians tended to be older on the average), Source of training (more Blacks had training in an accredited school and in government hospitals while more Caucasians had training in military service and in civilian hospitals), Length of time employed at present installation (71 percent of Blacks employed four years or longer versus 55 percent of Caucasians), and Total years of experience as a Medical Technician (57 percent of Caucasian technicians had 12 or more years of experience versus 43 percent of Black technicians). Chi-square for Total years in Civil Service as a Medical Technician was significant at the .05 level (23 percent of Caucasian technicians had 16 or more years of service versus 14 percent of the Black technicians). Chi-squares for Sex, Amount of education, Length of training, and Salary grade level were not significant.

For Cartographic Technicians, F-ratios at the .01 level of significance were found for Amount of education (Blacks tended to have more education than Mexican-Americans and Caucasians: 61 percent of Blacks had 1 or 2 years college or more versus 28 percent of Mexican-Americans and 31 percent of Caucasians); Amount of training prior to becoming Cartographic Technician (54 percent of Mexican-Americans had a moderate amount of training or more versus 25 percent of Blacks and 22 percent of Caucasians); Years of experience as Cartographic Technician prior to 1371 series (31 percent of Mexican-Americans had 12 or more years experience versus 17 percent of Blacks and 16 percent of Caucasians). F-ratios significant at the .05 level were found for Sex (62 percent Black males and 79 percent Mexican-American males versus 34 percent Caucasian males); and Total years of experience as Cartographic Technician in 1371 series (71 percent of Mexican-Americans had 8 or more years in 1371 series, 58 percent of Blacks, and 46 percent of Caucasians). F-ratios for Age, Source of formal training, Length of training on the job, and Present salary grade were not significant.

For Inventory Management Specialists, chi-square values were computed. Significant differences at the .01 level were found for Sex (62 percent male Mexican-Americans, 52 percent male Caucasians, and 37 percent male Blacks); Age (30 percent Black and 39 percent Caucasian inventory managers were 50 years old or older versus 11 percent of Mexican-Americans); Education (51 percent of the Blacks had 1 - 2 years of college or more versus 44 percent Mexican-Americans and 33 percent Caucasians); Length of time at present installation (85 percent of Mexican-Americans had been employed at present installation 8 years or longer versus 68 percent Blacks and 45 percent Caucasians); and Present salary grade (8 percent Mexican-Americans in Grade 11 versus 21 percent of Blacks and 20 percent of Caucasians, and 19 percent Mexican-Americans in Grade 7 versus 6 percent Blacks and 4 percent Caucasians). Chi-squares for Amount and Source of training, Length of training, and Total years of experience were not significant.

Differences in Distribution of Salary Grades Across Ethnic Samples

Some differences in the distribution of the ethnic samples by Present salary grade (GS-level) were found. To determine whether these distributions are significantly different, chi-square values were computed separately by ethnic group versus each other ethnic group for this variable for all three occupations. For Medical Technicians, chi-square was not significant, as noted above. For Cartographic Technicians, chi-square was not significant for Blacks versus Caucasians, but was significant at the .01 level for Mexican-Americans versus the total Caucasian sample. However, when the sample of Caucasians was subdivided into those at San Antonio and those at all other installations (see Chapter VIII for other such comparisons) chi-square was not significant for differences between the samples of Mexican-Americans and Caucasians at the San Antonio installation. For the Inventory Management Specialist samples, chi-square was not significant for Blacks versus Caucasians. Chi-square was significant, at the .01 level, for Mexican-Americans versus all Caucasians, and when the distribution for Caucasians at San Antonio was compared with the Mexican-American sample (all from San Antonio), the difference was significant at the .05 level. The difference is at the GS-7 level. A few more Mexican-American Inventory Managers at the GS-7 level had been included in the study than Caucasian GS-7's at San Antonio, to increase the ethnic sample size. The effect of this difference on aptitude and performance measures may be observed in later comparisons.

Relationship of Selected Background Variables to Aptitude and Job Performance Measures

In this section, relationships of selected background variables to aptitude and job performance measures will be discussed where they are significantly different from zero, by ethnic group and by occupation. More detail will be provided in table form for correlations of present salary grade (GS-level) with aptitude tests and criterion measures because of differences described in the previous section and because factor analyses presented in Chapter IX showed loadings of salary grade on factors on which aptitude and job performance measures also loaded.

For Medical Technicians correlations of selected background variables with aptitude and job performance measures for the two ethnic groups had no notable patterns of difference. Table X-4 shows the correlations of <u>Present salary grade (GS-level)</u> with aptitude test scores, selected supervisors' ratings, and Job Knowledge Test score. Significant correlations are seen for present GS-level for both Blacks and Caucasians with supervisors' Learning Ability and Job Knowledge ratings and Overall Rating, and with Job Knowledge Test (in each case higher for Blacks). A few positive and significant correlations of present GS-level with aptitude tests appear for Blacks. The correlations of present GS-level with aptitude tests for Caucasians are in general negative and low.

For Cartographic Technicians, some trends may be noted for certain other background variables. In general, <u>Age</u> correlates positively with Vocabulary test scores, but negatively with some aptitude test scores for all ethnic groups. <u>Time in grade level</u> also tends to correlate negatively and significantly with measures of ability (both ratings and certain aptitude test scores) for all ethnic groups. For Mexican-Americans, a significant and positive relationship is shown for speaking knowledge of English with those aptitude tests where reading or language comprehension is an element. This effect was discussed in a previous chapter, where Mexican-Americans were found to score lower on vocabulary tests. (Almost all Mexican-Americans indicated that a language other than English was spoken in their parents' homes.) Table X-5 shows the correlations of <u>Present salary grade (GS-level)</u> with selected aptitude test scores and job performance measures for Cartographic Technicians by ethnic group. For the Black and Caucasian samples, as also shown for Medical Technicians, GS-level relates positively and significantly (at the .01 level) with selected supervisors' ratings, indicating that those in the higher GS-levels were rated higher on Learning Ability, Technical Knowledge, and Overall job performance. The correlation of present salary grade with Work Sample Composite was significant for the Black sample only (at the .05 level). The only significant relationship of present salary grade with any aptitude test score was found with the Vocabulary Test for the Black sample. Present salary grade correlated significantly only with supervisors' rating of Technical Knowledge for Mexican-Americans.

For Inventory Management Specialists, as found for the other two occupations, Age tended to correlate negatively with measures of learning ability (both aptitude test scores and supervisors' rating of Learning Ability) for all three ethnic groups. Level of education attained, on the other hand, was correlated positively and significantly with scores on the Federal Service Entrance Examination for all three ethnic groups. In addition, for Mexican-Americans, Standing in high school graduating class was correlated positively and significantly with Vocabulary Test scores and Highest level of math used on job with Subtraction and Multiplication Test scores.

Table X-6 shows the correlations of Present salary grade (GS-level) with selected aptitude test scores and job performance measures for Inventory Management Specialists. The pattern here is quite different from that shown for Cartographic Technicians. The only significant relationship of GS-level for Blacks is with Work Sample Overall Performance (at the .05 level), and for the Caucasian sample, with supervisors' Technical Knowledge and Overall ratings (at the .05 level). Correlations of GS-level with aptitude test scores for both Blacks and

Caucasians are low, usually negative, and non-significant. For Mexican-Americans, those at higher GS-levels tended to receive higher ratings on Learning Ability and Technical Knowledge and to perform at a higher level on the Work Sample. They also tended to score higher on most aptitude tests, including the FSEE, implying a significant relationship between abilities and present level on the job.

Comparisons of Non-Test Selected and Test-Selected Inventory Managers

As described in Chapter II, test scores were seldom or never included as part of the process for selecting Medical Technicians or Cartographic Technicians. However, Inventory Management Specialists had been required in many instances to take a written test (usually the Federal Service Entrance Examination), for entry into a career development training program at the GS-5 level, leading to the journeyman level tested in this study. In the Personal History Questionnaire, inventory managers were asked to respond to the question:

> "What kind of test were you required to take when you were hired or reclassified as an inventory manager?"
>
> 1. None
> 2. Federal Service Entrance Examination
> 3. Other

Responses expressed in percent of ethnic group sample were as follows:

	No test	FSEE	Other test
Black	50	40	10
Mexican-American	64	20	16
Caucasian	41	53	6

Table X-7 shows means and standard deviations of selected predictor and criterion scores by ethnic group for those who were not tested, those who were screened on the FSEE, and those who were screened on some other test prior to entry into an inventory management job. It will be noted that without exception those who were screened on the FSEE received higher scores on the average on every measure than those who were not required to take this test prior to entry. In general, those required to take some other test also scored higher on the average than those not screened by any test. Differences in supervisors' ratings for those who had taken a test as a pre-selection requirement are not as striking as those on aptitude, job knowledge, and work sample measures. However, supervisors' ratings are higher to some degree in every instance as well.

Background Variables as Differential Predictors Across Ethnic Groups

Cartographic Technicians

An exploratory analysis of data from Cartographic Technicians was undertaken in an effort to determine whether certain background variables might affect the level of accuracy of prediction across ethnic groups. That is, was it possible that some aptitude tests might have varying degrees of predictive accuracy for individuals from different backgrounds. To reduce the number of available background variables, a factor analysis of the Personal History Questionnaire was performed. The first five factors were selected and rotated using the variance procedure.

The five factors were called: (1) length of time as cartographer, but not 1371 series; (2) length of time in lower grades; (3) length of time in 1371 series; (4) socioeconomic standing of parents; and (5) educational and training level. The Cartographic Technicians were then subdivided according to their scores on each of these factors through the use of the moderated regression technique (Rock, et al., 1968). Of particular interest was the possibility of finding differential prediction for members of the two extreme groups on each factor. For example, different kinds of tests might predict more accurately for individuals from a high socioeconomic background than for those from a low socioeconomic background.

The possibility of differential prediction was evaluated for five tests against two criterion measures. The tests selected were Necessary Arithmetic Operations, Surface Development, Hidden Figures, Extended Range Vocabulary, and Map Planning. The two criterion measures were the "pull-up" task from the Work Sample and the Supervisors' Overall Rating.

Only one of the five factors led to groupings of individuals who appeared to be characterized by differential predictive accuracy. That is, the two extreme groups on the fifth factor, educational and training level, differed with respect to accuracy of prediction on both criterion measures. Table X-8 presents the multiple correlations and zero-order validity coefficients against both criterion measures for the "low education and training," "high education and training," and "total" groups. For both criterion measures, the multiple correlation is higher for the "high" groups than for the "low" and/or "total" group. Also, as might be expected, the "high education and training" group had higher mean scores on all the predictors. It was somewhat surprising, however, that the best predictor for the "high education and training" group for both criterion measures was the Surface Development Test, while the least accurate predictor was the Extended Range Vocabulary Test.

Little if anything of a conclusive nature can be drawn from this analysis, however, since the samples were too small for replication. It does suggest, however, that background factors which do in fact cut across racial groups might also have an impact on differential predictive accuracy.

Comparison of Tasks Performed by Ethnic Group Members

A detailed task list was included as part of the Personal History Questionnaire for each of the three occupations. Response patterns were analyzed by ethnic group to determine whether, for a given occupation, the different ethnic groups were performing essentially the same tasks.

Job assignments of Medical Technicians were examined by GS-level to find out what their tasks were and how often they were performed. Chi-square values were computed to determine whether response patterns to the task list items showed significant differences within grade level between ethnic groups. No major differences were found.

Job assignments of Cartographic Technicians were examined by installation by ethnic group. Chi-square values were computed. No major differences were disclosed among ethnic groups. A total of 24 out of 130 tasks showed a difference in at least one installation. However, no one task was significantly different at all installations. Six tasks were different at two installations. This is fewer than would be expected by chance.

Inventory Management Specialists' job assignments were first examined by ethnic group. There were few differences between Blacks and Caucasians, but for about 40 items, the tasks performed by Mexican-

Americans appeared to differ from the job activities of the other two groups. The data were further analyzed by comparing the responses of San Antonio Caucasians (who worked with the Mexican-Americans there) with those of all Caucasians. Table X-9 shows a selection of tasks from these analyses. As may be seen, significant differences between the two Caucasian groups are indicated, but there were no significant differences between the Mexican-Americans and the Caucasians working at the same installation in San Antonio.

It may be concluded that no significant differences in the patterns of tasks performed by members of respective ethnic groups were found to exist in any of the occupations studied.

Summary and Conclusions

A number of different comparisons of background and task variables were made for the three occupations studied, to examine possible ethnic group differences which might affect other results obtained. The findings are summarized as follows:

1. Differences in composition of ethnic samples on selected background variables were negligible for all three occupations.

2. Distributions of salary grades by ethnic group by occupation were not significantly different with one exception, where a few more Inventory Management Specialists from one ethnic group were added at the lowest salary grade level tested in order to increase the ethnic sample.

3. When relationships of selected background variables to aptitude and job performance measures were examined, Present salary grade was significantly related to measures of job performance but only in a few

instances to aptitude test scores for Blacks and Caucasians in all three occupations. This is also true for Mexican-American Cartographic Technicians. However, for Mexican-American Inventory Management Specialists, significant relationships were found of Present salary grade to both aptitude and performance measures. Among other background variables, Age tended to correlate positively with vocabulary measures, but negatively with speeded tests in general.

4. Inventory-Management Specialists screened on the FSEE prior to job entry received higher average scores on every measure than those who were not tested as a pre-selection requirement.

5. Exploratory analyses of Cartographic Technician data were generally inconclusive as to the moderating effects of selected background variables on level of predictive accuracy across ethnic groups.

6. For a given occupation, no major differences were found in the patterns of tasks performed by members of respective ethnic groups. Apparent differences between Mexican-American Inventory Managers and members of the other two ethnic groups disappeared when Caucasians working at San Antonio were compared with their Mexican-American counterparts. The task differences were by installation rather than by ethnic group.

Table X-1

Distributions of Selected Background Variables

Expressed as Percent of Total Sample, by Ethnic Group

Medical Technicians

		Percent	
		Black	Caucasian
Sex	Male	46	47
	Female	54	53
Age	60 +	2	2
	50 - 59	8	19
	40 - 49	29	31
	30 - 39	43	22
	20 - 29	18	25
	Less than 20	0	1
Education	Advanced study	5	2
	College degree	8	7
	College, more than 2 years	21	18
	College, 2 year terminal	7	5
	College, less than 2 years	32	31
	High school graduate	20	31
	Some high school	4	4
	8th grade or less	0	1
Source of training as Medical Technician	Accredited school	40	31
	Military service	17	28
	Government hospital	23	11
	Civilian hospital	7	13
	Civilian laboratory	5	6
	Other	6	10
Length of training as Medical Technician	24 months or longer	16	11
	18 - 23 months	11	6
	12 - 17 months	49	52
	Less than 12 months	23	31

Table X-1 (Continued)

		Percent	
		Black	Caucasian
Total years of experience as a Medical Technician	20 or more	8	25
	16 - 19	14	12
	12 - 15	21	16
	8 - 11	21	16
	4 - 7	18	17
	2 - 3	5	6
	Less than 2	10	8
Total years in Civil Service as a Medical Technician	20 or more	5	14
	16 - 19	9	9
	12 - 15	16	12
	8 - 11	16	14
	4 - 7	28	20
	2 - 3	14	14
	Less than 2	12	16
Length of time employed at present installation	8 years or longer	48	38
	4 - 7 years	23	17
	1 - 3 years	14	23
	6 - 11 months	9	7
	Less than 6 months	5	14
Present salary grade (GS) level	8 or higher	4	5
	7	21	20
	6	36	41
	5	27	24
	4 or lower	12	10

Table X-2

Distributions of Selected Background Variables

Expressed as Percent of Total Sample, by Ethnic Group

Cartographic Technicians

		Percent		
		Black	Mexican-American	Caucasian
Sex	Male	62	79	34
	Female	38	21	66
Age	60 +	2	0	2
	50 - 59	6	9	13
	40 - 49	39	27	23
	30 - 39	36	62	26
	20 - 29	18	2	36
Education	1 or more years graduate	0	0	1
	3 or 4 years college	20	1	5
	1 or 2 years college	41	27	25
	Tech or Voc institute	15	13	18
	11th or 12th grade	24	56	50
	9th or 10th grade	0	2	2
	8th grade or less	0	1	0
Amount of training before entering GS-1371	Quite a lot	7	13	10
	A moderate amount	18	41	12
	None, or hardly any	75	46	77
Source of formal training as Carto Tech before entering GS-1371	High school	4	13	12
	Junior or 4-year college	4	1	4
	Tech or Voc institute	9	9	2
	In the military	6	6	5
	Civilian government mapping organization	53	38	48
	Commercial mapping organization	0	25	1
	Other location	1	3	0
	No prior formal training	23	5	28

Table X-2 (Continued)

		Percent Black	Percent Mexican-American	Percent Caucasian
Source of on-the-job training as Carto Tech before entering 1371 series	At a Tech or Voc institute	1	2	0
	In the military	3	3	4
	At a commercial map-making organization	0	24	2
	At a governmental map-making organization	70	56	56
	Other	1	2	1
	No prior on-the-job training	25	13	37
Length of training after entering 1371 series	More than 8 months	35	32	32
	7 - 8 months	10	1	4
	5 - 6 months	13	18	16
	3 - 4 months	21	23	22
	1 - 2 months	13	13	11
	Less than 1 month	8	13	15
Total years of experience as Carto Tech in 1371 series	20 or more	4	3	4
	16 - 19	20	7	15
	12 - 15	21	27	13
	8 - 11	13	34	14
	4 - 7	26	27	37
	2 - 3	14	1	14
	Less than 2	1	0	3
Total years of experience as Carto Tech outside 1371 series	20 or more	1	7	4
	16 - 19	10	8	6
	12 - 15	6	16	6
	8 - 11	12	12	12
	4 - 7	14	27	23
	2 - 3	13	12	11
	Less than 2	44	17	38

Table X-2 (Continued)

		Percent		
		Black	Mexican-American	Caucasian
Length of time employed at present installation	8 years or longer	69	90	57
	4 - 7 years	23	10	35
	1 - 3 years	5	0	6
	6 - 11 months	3	0	1
	Less than 6 months	0	0	1
Present salary grade (GS) level	12	1	0	0
	11	5	0	8
	10	0	0	0
	9	52	83	55
	8	10	0	8
	7	32	17	23
	6	0	0	0
	5	0	0	5

Table X-3

Distributions of Selected Background Variables

Expressed as Percent of Total Sample, by Ethnic Group

Inventory Management Specialists

		Percent		
		Black	Mexican-American	Caucasian
Sex	Male	37	62	52
	Female	58	32	38
	No response	5	5	10
Age	60 +	5	0	6
	50 - 59	25	11	33
	40 - 49	42	50	29
	30 - 39	18	31	13
	20 - 29	4	3	9
	No response	6	5	10
Education	Graduate school	3	3	2
	3 or 4 years college	20	15	22
	1 or 2 years college	28	26	9
	1 or 2 years tech or business institute	17	9	19
	11th - 12th grade or GED diploma	27	41	34
	9th - 10th grade	0	3	4
	8th grade or less	0	1	0
	No response	5	3	10
Amount of training before entering GS-2010	Quite a lot	13	5	6
	A moderate amount	10	16	8
	None, or hardly any	62	73	73
	No response	15	5	13

Table X-3 (Continued)

		Percent		
		Black	Mexican-American	Caucasian
Source of formal training in inventory management before entering 2010 series	High school	8	11	6
	Junior college	3	3	1
	Four-year college	2	7	3
	Tech or business institute	6	7	3
	In the military	6	8	6
	Civilian governmental activity	23	18	21
	Non-governmental commercial organization	0	1	1
	Other location	1	0	0
	No prior formal training	37	40	45
	No response	15	5	14
Source of on-the-job training before entering 2010 series	Work-study program	0	1	0
	In the military	4	7	6
	Civilian governmental activity	44	46	34
	Non-governmental commercial organization	1	5	3
	Other	3	3	2
	No prior on-the-job training	36	34	43
	No response	12	4	12
Length of training after entering 2010 series	More than 8 months	27	15	18
	7 or 8 months	4	0	2
	5 or 6 months	11	9	8
	3 or 4 months	19	12	16
	1 or 2 months	9	23	20
	Less than 1 month	18	34	22
	No response	12	7	13

Table X-3 (Continued)

		Percent		
		Black	Mexican-American	Caucasian
Total years of experience as Inventory Management Specialist in 2010 series	20 or more	4	0	4
	16 - 19	5	3	6
	12 - 15	18	15	10
	8 - 11	23	26	19
	4 - 7	28	34	38
	2 - 3	11	12	8
	Less than 2	5	5	5
	No response	7	5	10
Total years of experience in inventory management outside 2010 series	20 or more	4	1	6
	16 - 19	4	4	2
	12 - 15	6	7	4
	8 - 11	10	7	4
	4 - 7	11	18	12
	2 - 3	7	4	3
	Less than 2	35	35	45
	No response	24	24	23
Length of time employed at present installation	8 years or longer	68	85	45
	4 - 7 years	18	9	36
	1 - 3 years	6	3	8
	6 - 11 months	2	0	1
	Less than 6 months	0	0	0
	No response	5	3	10
Present salary grade (GS) level	11	21	8	20
	9	67	68	66
	7	6	19	4
	No response	6	5	10

Table X-4

Correlations Between Present Salary Grade and

Selected Predictor and Criterion Variables

Medical Technicians

Aptitude Tests	Black N=168	Caucasian N=297
Subtraction & Multiplication	.15*	-.01
Vocabulary	.02	.09
Hidden Figures	.13	-.02
Necessary Arithmetic Operations	.17*	-.15**
Pin-Dexterity	-.06	-.07
Number Comparison	.04	-.13*
Gestalt Completion	.03	-.13*
Picture-Number	.15*	-.05
Paper Folding	.00	-.12*
Criterion Variables		
Learning Ability Rating	.37**	.24**
Job Knowledge Rating	.50**	.44**
Supervisors' Overall Rating	.45**	.29**
Job Knowledge Test Score	.48**	.21**

* significant at .05 level
** significant at .01 level

Table X-5

Correlations Between Present Salary Grade and

Selected Predictor and Criterion Variables

Cartographic Technicians (TOPOCOM)

Aptitude Tests	Black N≈101	Mexican-American N≈101	Caucasian N≈240
Hidden Figures	.13	-.15	.10
Vocabulary	.24*	-.08	.08
CS Arithmetic	.04	-.06	-.04
Map Planning	.02	-.14	-.06
Surface Development	.05	-.14	-.02
Necessary Arithmetic Operations	.09	-.13	-.06
Criterion Variables			
Work Sample Composite	.21*	.15	.08
Learning Ability Rating	.42**		.33**
Technical Knowledge Rating	.54**	.20*	.48**
Supervisors' Overall Rating	.50**	.19	.40**

* significant at .05 level
** significant at .01 level

Table X-6

Correlations Between Present Salary Grade and

Selected Predictor and Criterion Variables

Inventory Management Specialists

	Black	Mexican-American	Caucasian
Aptitude Tests	N=112	N=72	N=194
Number Comparison	-.09	.15	-.04
Hidden Figures	-.16	.23*	-.07
Vocabulary	-.08	.23*	-.10
Nonsense Syllogisms	.06	.24*	-.11
Subtraction & Multiplication	.00	.27*	.02
Necessary Arithmetic Operations	-.17	.39**	-.06
FSEE	.02	.38**	-.13
Criterion Variables			
Work Sample Overall Performance	.22*	.40**	-.08
Learning Ability Rating	.03	.24*	.10
Technical Knowledge Rating	.06	.36**	.22*
Supervisors' Overall Rating	-.05	.17	.18*

* significant at .05 level
** significant at .01 level

Table X-7

Means and Standard Deviations of Selected Predictor and Criterion Scores--Those Not Tested vs. Those Tested Before Hiring or Reclassification

Inventory Management Specialists

		Black			Mexican-American			Caucasian		
		N=50	N=40	N=10	N=45	N=14	N=11	N=73	N=95	N=12
		No Test	FSEE	Other Test	No Test	FSEE	Other Test	No Test	FSEE	Other Test
Aptitude Tests										
Number Comparison	M	38.5	42.8	42.9	41.3	46.7	48.6	37.8	43.1	45.0
	SD	11.4	10.2	10.4	10.3	9.5	9.9	13.0	9.7	11.2
Hidden Figures	M	4.5	6.9	6.6	6.2	9.5	7.1	5.5	9.2	8.8
	SD	4.3	6.6	4.8	4.6	5.0	5.2	5.2	5.6	4.5
Vocabulary	M	20.2	25.9	21.7	20.0	24.9	22.2	22.9	28.5	24.8
	SD	5.8	6.9	7.3	6.0	4.2	6.4	6.2	4.3	6.2
Nonsense Syllogisms	M	5.4	11.3	6.6	5.1	10.6	6.4	5.6	11.7	10.3
	SD	6.4	7.7	5.0	5.0	6.6	7.0	5.9	8.5	9.1
Subtraction & Multiplication	M	62.7	67.9	67.7	65.7	77.6	84.7	71.8	78.8	88.9
	SD	19.4	17.2	18.5	20.3	12.6	19.9	21.8	20.9	18.6
Necessary Arithmetic Operations	M	9.8	13.7	12.4	12.3	17.5	15.0	12.5	18.0	17.5
	SD	5.1	5.2	4.5	5.0	4.6	5.5	5.2	3.9	4.9
FSEE	M	35.5	52.0	40.6	38.8	52.9	51.0	40.5	60.8	49.2
	SD	11.5	12.7	10.4	9.9	11.1	10.0	12.3	10.6	10.1
Work Sample										
Scorer's Rating of Overall Performance	M	6.0	7.7	7.1	7.7	11.8	10.4	7.3	9.7	10.1
	SD	3.0	3.6	2.6	3.8	3.8	3.7	4.1	4.6	3.6
Rating Scales										
Learning Ability	M	5.9	6.4	6.3	5.9	6.8	6.0	5.7	6.5	5.8
	SD	1.9	1.9	1.3	1.7	1.3	2.2	1.8	1.8	1.8
Technical Knowledge	M	5.6	6.2	5.9	6.2	6.8	6.2	5.9	6.5	6.0
	SD	2.1	1.7	1.5	1.6	1.3	2.0	1.8	1.6	1.9
Overall	M	6.2	6.4	6.4	6.3	7.3	6.0	6.0	6.6	5.7
	SD	1.8	1.7	1.3	1.8	1.2	1.7	1.7	1.5	2.0

Table X-8

Differential Prediction of the "Pull-Up" and Overall Rating Criteria -

Pull-Up Criterion

Moderator = Education and Training

Predictors	Total Group N=418 r_{xy}	Total Group R=34 $\overline{X}$	Low Group N=91 r_{xy}	Low Group R=24 $\overline{X}$	High Group N=54 r_{xy}	High Group R=51 $\overline{X}$
Necessary Arithmetic Operations	26	12.77	20	11.33	38	14.52
Surface Development	30	28.20	08	22.43	44	35.99
Hidden Figures	26	11.10	18	9.36	33	13.75
Extended Range Vocabulary	08	18.00	06	15.34	19	17.17
Map Planning	24	18.00	12	15.71	37	20.24
Moderator $\overline{X}$	0.00		-1.22		1.71	

Overall Rating Criterion

Moderator = Education and Training

Predictors	Total Group N=434 r_{xy}	Total Group R=31 $\overline{X}$	Low Group N=97 r_{xy}	Low Group R=39 $\overline{X}$	High Group N=79 r_{xy}	High Group R=47 $\overline{X}$
Necessary Arithmetic Operations	22	12.71	29	11.03	31	14.89
Surface Development	27	27.74	34	21.67	37	35.21
Hidden Figures	22	10.99	28	9.26	19	12.91
Extended Range Vocabulary	01	17.98	26	15.09	-10	18.96
Map Planning	22	-	29	-	33	-
Moderator $\overline{X}$	0.00		-1.22		1.55	

Table X-9

Comparisons of Selected Tasks for Inventory Management Specialists

Within and Across Locations Studied

	Percent reporting "significant part of job every day" or "substantial part of job, at least several times a week"			
	Black	Mexican-American (San Antonio)	Caucasian (San Antonio)	Caucasian (All locations)
Authorize purchase of additional quantity of materiel over amount normally used	25	5	4	19
Work with Procurement in cancelling contracts	15	3	1	10
Make manual buy for urgent request	54	3	7	37
Use mathematical formulas to figure net assets	40	18	19	29
Request expediting action to assure on-time delivery of item	61	42	30	47
Arrange for transfer of stock from one depot to another	35	18	18	26
Request Cataloging Branch to assign Federal Stock Numbers to items and place in FSN Catalog	8	22	22	9
Submit back order status report	33	15	19	20
Cancel purchase orders	42	15	13	23

Chapter XI.

An Alternative Approach to Culture Fairness when Tests are Used in Selection

Recently a number of definitions of "test bias" have been proposed. The one definition which has been most often applied is the Cleary (1968) definition of "test bias." Simply stated, Cleary defines a test as being "culture-fair" for populations A and B when the regression equation based on population A neither systematically over- nor under-predicts level of performance for members of population B. Thorndike (1971) has proposed an alternative definition which can lead to entirely opposite conclusions with respect to whether or not a test is "culture-fair" in a prediction situation.

Thorndike suggests that a test may be judged "culture-fair" if the overlap on the criterion scores between groups A and B is essentially equivalent to their overlap on the predictors. Assuming normally distributed criterion and predictor scores within each group, and equivalent standard deviations across groups, the measurement of overlap reduces to comparing the differences between group means scaled by their standard deviations on the criterion and on the predictor. For example, if the means for group B are one standard deviation apart on the criterion, they should also be one standard deviation apart on the predictor. If the two groups' means are only one-half standard deviation apart on the criterion, but a full standard deviation apart on the predictor, the test would be judged unfair in the sense that for any one specified criterion cutting score, the members of each group would not have the same opportunity for being selected as would be represented by the proportion of their group well as the reliability of such criteria. Past research suggests that more often than not, criteria such as ratings tend to have lower reliabilities (i.e., inter-rater reliabilities) than the objective tests used to predict them (Conklin, 1923; Stead and Shartle, 1940; Ghiselli and Brown, 1955). Thus, before any conclusions can be drawn with respect to culture-fair test usage based on the Thorndike approach, it would seem reasonable to correct the criterion scores for unreliability.

The fact that the subjective nature of ratings almost rules out the possibility of "color-blindness" can lead to even more serious problems." The possibility of bias in ratings is documented in other reports forthcoming from the present study. Of course, the possibility of bias in a subjective criterion tends to reduce the usefulness, if not appropriateness, of the Thorndike criterion, or for that matter, any definition regarding "culture-fair" usage of tests.

Results of a modified Thorndike analysis of culture-fair usage of tests against (1) a subjective rating criterion corrected for unreliability, (2) an objective work sample criterion corrected for estimated unreliability, and (3) a criterion based on an objective test purporting to measure job knowledge are presented here. The analyses were performed for the ethnic samples in the three occupations in the present study: Black and Caucasian Medical Technicians; Black, Mexican-American, and Caucasian Cartographic Technicians; and Black, Mexican-American, and Caucasian Inventory Management Specialists.

The correction for unreliability in the subjective ratings was made in an effort to see if the "true" mean difference in criterion rating means more closely approximated the difference in predictor score means.

The hypothetical "true" mean difference may be defined as that difference between minority and majority group means which one would expect to find if the criterion rating was as reliable as the predictor test. The assumption here, of course, is that bias in the ratings is at a minimum. However, additional insight into the possibility of bias in the ratings may be inferred from the comparison of relative differences between group means on the subjective rating criterion and the group mean differences which were found when an objective job knowledge test and/or work sample was used as a criterion. Admittedly, a paper-and-pencil test has its drawbacks due to its somewhat theoretical nature. It would be fairer to point out here that a possible drawback of a job knowledge test as a criterion is that some of the same skills which play a role in the predictor also affect the criterion scores (e.g., "test taking" skills, if in fact these exist other than in a hypothetical sense). However, it does have the following advantages: (1) in general, its reliability approximates that of the test used to predict it, (2) it is "color-blind," and (3) in this particular situation, it was deemed to be an appropriate measure of job knowledge by supervisory staff members who participated in developing the test.

The corrections for unreliability of ratings were originally suggested by Angoff (in a personal communication) and were estimated in the following manner. Two critical ratios were formed, one for the criterion means, CR'_{yy}, and one for the predictor means, CR_{xx}.

$$CR'_{yy} = \frac{\overline{Y}_C - \overline{Y}_B}{\overline{\sigma}_{yt}} \quad (1)$$

Where: $\overline{Y}_C$ = mean score on the criterion for the Caucasian group

$\overline{Y}_B$ = mean score on the criterion for the Black group

and:

$$\overline{\sigma}_{yt} = \sqrt{\frac{\sigma_C^2 + \sigma_B^2}{2} \cdot \frac{r_{yy}}{r_{xx}}}$$

where: σ_C^2 = variance of the Caucasian criterion scores

σ_B^2 = variance of the Black criterion scores

r_{yy} = reliability of the criterion rating

r_{xx} = reliability of the predictor test

while CR_{xx} is simply:

$$\frac{\overline{X}_C - \overline{X}_B}{\overline{\sigma}_x}$$

Equation (1) above simply expresses the expected difference between the group means on the criterion (in terms of the standard deviations of their true scores) if the criterion ratings had the same reliability as the predictor. The prime in the ratio CR'_{yy} indicates that this correction for unreliability has been made. The reliabilities, based on inter-rater agreement, were .46 for the Black Medical Technicians and .66 for the Caucasian Medical Technicians. The average reliability was then used in equation (1). In short, if the predictor and the criterion were equally reliable, would CR'_{yy} equal CR_{xx}? If the answer is yes, then one could conclude by this modified Thorndike definition that there was no unfairness in selection against either group and that the "unfairness" found before the correction was simply due to differential reliability between predictor and criterion. However, if CR_{xx} is still significantly larger than CR'_{yy}, one may conclude that the test is biased against the Black group. The assumption here is that the Caucasian means are higher than the Black means and that there is little or no racial bias in the rating procedure.

Results and Discussion

Medical Technicians

Table XI-1 presents the critical ratios for each of the predictor tests against each of two criteria for Medical Technicians. In all cases, the critical ratios, CR'_{yy} and CR_{xx}, are both positive, indicating that the Caucasian means were higher on both the predictor and the criterion. For example, in the case of the Subtraction and Multiplication Test predicting the Job Knowledge Test, the CR_{yy} is equal to .44, indicating that the Caucasian group was approximately one-half standard deviation above the Black group on the criterion, while the CR_{xx} ratio indicates that the Caucasian group was also about one-half standard deviation above the Black group on the Subtraction and Multiplication Test. When tests were used for criteria, no correction for unreliability was made since r_{xx} then approximated r_{yy}.

Using the Thorndike definition, one would conclude that there is little or no unfairness against the Black group when the Job Knowledge Test is used as the criterion and the Subtraction and Multiplication Test is the predictor. Similar findings apply to the remaining tests, with the exception of Vocabulary and Necessary Arithmetic Operations. Out of the eight possible predictors, only these two tests show a significant bias against the Black group when the Job Knowledge Test is the criterion (i.e., CR_{yy} is significantly less than CR_{xx}), where significance is arbitrarily defined as a critical ratio difference of at least one-quarter standard deviation.

An inspection of the critical ratios for the predictor tests against the Overall rating would seem to lead to the opposite conclusion. Closer inspection of the Black and Caucasian validity coefficients for each of the tests suggests that they are almost all near zero and, as a result, the application of any definition of culture fairness would be meaningless. Either the tests are inappropriate and/or the criterion ratings include a reliable but large proportion of invalid variance. Results from earlier studies concerning supervisors' ratings among Medical Technicians lend some support to a hypothesis of bias in the supervisors' ratings. The finding that certain ethnic combinations of rater and ratee lead to systematic but not necessarily valid sources of variance in the overall ratings may help to explain why objective ("color-blind") predictors have little variance in common with such a criterion.

Cartographic Technicians

Table XI-2 shows similar data for Black, Mexican-American, and Caucasian Cartographic Technicians. More specifically, Table XI-2 presents the critical ratios for Caucasians and Blacks and Caucasians and Mexican-Americans when the Supervisors' Overall Rating was used as the criterion. Of the four predictors (Hidden Figures, Map Planning, Surface Development, and Necessary Arithmetic Operations) having similar non-zero validity coefficients, no one test was consistently unbiased against all groups. Hidden Figures and Map Planning could be used as "culture-fair" selection instruments (against the rating criterion) if there was only a Caucasian and Black applicant pool. However, if the applicant pool also included

Mexican-Americans, these two tests would no longer be "culture-fair" instruments, since they would tend to discriminate against the Mexican-Americans and in favor of the Caucasians.

It should be noted here that a test may discriminate in favor of the majority group as well as certain minority groups, but not necessarily all minority groups. However, before any conclusions can be drawn from the above analysis concerning the culture fairness of tests against supervisory ratings, additional evidence concerning the validity of the supervisory ratings should be presented. The Cartographic Technicians' study was especially appropriate for such an analysis since three types of criteria were available for comparison. Two of these criteria, a paper-and-pencil Job Knowledge Test and a composite Work Sample, were objective "color-blind" measures, and the third was the Supervisors' Rating.

As pointed out before, a paper-and-pencil job knowledge test may be unsuitable as a criterion for identifying unfair test selection procedures since it may be argued that the theoretical responses to an "artificial" job situation may not reflect the skills necessary for the successful performance of the actual job tasks. Figure XI-1 presents a paradigm indicating the relationships between the three criterion measures.

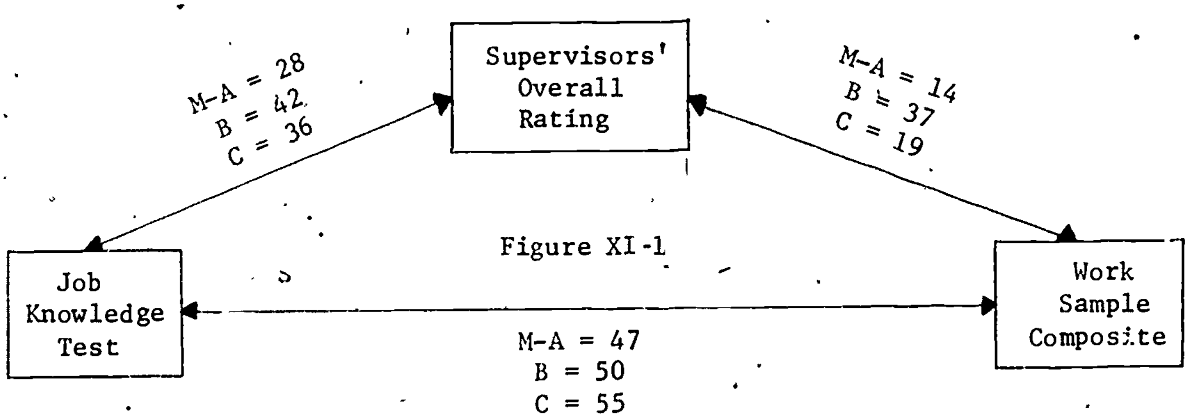


Figure XI-1

Inspection of Figure XI-1 suggests that the Job Knowledge Test is more highly related to the actual tasks performed on the job than is the subjective Supervisors' Rating. Since ratings are not "color-blind," it is possible that they may include nonrelevant variance due to racial bias. It is particularly interesting to note that two of the three correlations between the Supervisors' Overall Rating and the Work Sample Composite are quite low, while the correlations between the two objective measures, one "theoretical" and one operational, are all close to .50. It would seem that the "theoretical" job knowledge measure has much in common with the objective work sample, yet neither have much in common with the subjective ratings.

Table XI-3 presents the critical ratios for the various group comparisons when the Work Sample Composite is used as a criterion. Of the six tests which have similar non-zero correlations with the Work Sample criterion--Hidden Figures, Card Rotations, Map Planning, Surface Development, Maze Tracing Speed, and Necessary Arithmetic Operations--Card Rotations, Map Planning, and Surface Development appear to be relatively culture-fair regardless of the groups to be compared. Map Planning, in fact, has a slight tendency to favor the Blacks and Mexican-Americans when these minority groups are compared with the Caucasian majority. It is also interesting to note that, besides being relatively culture-fair, these tests have relatively high face validity for this particular job.

Table XI-4 presents the critical ratios for the various group comparisons when the Job Knowledge Test is used as a criterion. Hidden Figures, Vocabulary, Card Rotations, Map Planning, Surface Development, Maze Tracing Speed, Extended Range Vocabulary, and Necessary Arithmetic Operations all had non-zero validity coefficients across all groups. In general, when the Caucasians are compared with the Blacks, only one test out of eight appears to be biased against selecting the appropriate proportion of Blacks, and that test is Necessary Arithmetic Operations. In fact, five out of the eight tests are slightly biased ($CR_{xx} < CR'_{yy}$) in favor of the Blacks. These tests are Hidden Figures, Vocabulary, Card Rotations, Map Planning, and Extended Range Vocabulary. When Caucasians are compared with Mexican-Americans with respect to these eight predictors, the bias, if any, is always in favor of the Mexican-Americans.

As in the case of the Medical Technicians, when tests were evaluated against an objective criterion, little or no predictive bias was found.

Inventory Management Specialists

Table XI-5 presents the critical ratios for the predictor tests against Supervisors' Overall Rating for Inventory Management Specialists. Only two predictors, Subtraction and Multiplication and Necessary Arithmetic Operations, out of ten predictors show similar non-zero relationships with the rating criterion. In both instances, the critical ratios indicate that these two tests are biased against both the Black and the Mexican-American minority groups. However, Table XI-6 shows that when the Work Sample overall scores are used as the criterion, six predictors--Hidden Figures, Vocabulary, Letter Sets, Extended Range Vocabulary, Necessary Arithmetic Operations, and Inference--all have non-zero validity coefficients across all groups. When the Caucasian and Black Inventory Management Specialists are compared, the critical ratios CR_{xx} and CR_{yy} are approximately equal for five of these six tests, indicating littl or no predictive bias. Necessary Arithmetic Operations tends to be biased against the Blacks ($CR_{yy} - CR_{xx} \approx \frac{\sigma}{4}$). For the remaining five tests there is a slight tendency for the selection procedure to be in favor of the Blacks ($CR_{yy} > CR_{xx}$). However, when the Caucasian group is compared with the Mexican-American group, the "bias" is reversed. That is, with the possible exception of the Hidden Figures Test, the remaining five predictors appear to be biased against the Mexican-Americans. Further inspection of the types of predictors which lead to the greatest, as well as the least, bias against the Mexican-Americans suggests that possible language problems hampered their performance to a greater degree on the predictors than on the Work Sample criterion. For example, the two most biased tests against the Mexican-Americans were the two vocabulary tests, while non-verbal tests such as Number Comparison and Hidden Figures showed little or no bias.

Conclusions

With few exceptions, the direction of differences between majority and minority group means on the predictors was also reflected in the criterion means corrected for unreliability.

However, when Supervisors' Ratings were used as the criterion, the difference between minority and majority group means on this criterion were much less pronounced than their corresponding difference on the predictor means. On the surface, this finding suggests that tests were biased, in the Thorndike sense, against the minority group when subjective ratings were used. However, other analyses of the rating data which are reported elsewhere in this report (Chapter VIII) seem to indicate a potential ethnic bias in the ratings. The finding of a possible racial contamination of the Supervisors' Ratings was particularly noticeable for the Medical Technicians. Additional evidence against using subjective ratings as the criterion in a "test bias" study was also found in the analysis of the Cartographic Technicians data.

When the Job Knowledge Test and/or the Work Samples were used as the criterion, much of the evidence for "test bias" against Blacks and Mexican-Americans disappeared. One exception to this was the Mexican-American Inventory Management Specialist group. Compared to the Caucasian group, they tended to do better, on the average, on the Work Sample than their predictor scores would suggest. It is quite possible that a language problem did hamper their performance on particular predictor tests used in this study. The fact that the two verbal predictors showed the greatest amount of "bias" against the Mexican-American group lends additional support to a possible language problem.

The results from this study suggest that if we accept an objective test or work sample as being more desirable measures of job performance than supervisors' ratings, and carefully select predictors which indeed do measure aptitudes and skills relevant to successful performance on the job, tests will lead to fair selection procedures for the Black minority group. It should, however, be pointed out that one might reach different conclusions if other than the Thorndike definition were used. Findings with respect to culture-fair selection of Mexican-Americans were less clear.

Based on what we have learned from the present study, as well as from earlier studies, concerning the problems inherent in achieving unbiased supervisors' ratings, it may well be that the use of an objective job knowledge test in lieu of the more subjective supervisors' ratings is the only reasonable answer. This need for a "color-blind" criterion would seem to be particularly germane in those cases where the potential bias in the ratings may have an ethnic basis.

This possibility of criterion bias raises interesting questions with respect to the nature of test bias. That is, all approaches have assumed that mean differences in criterion measures between groups are in fact true differences between the performance levels of the group members. Any predictor that does not accurately reflect these criterion differences is rejected as a "biased" predictor. If, however, some or all of these criterion differences have been due to bias and not to differences in "true" performance, the use of differential prediction has acted to maintain the bias in selection either for or against the minority group. Future "test bias" research should settle on a subjective criterion measure such as ratings only if the ratings can be shown to reflect to a high degree the same rank ordering obtained on "hard" criterion measures, such as objectively scores job samples, which can be considered reasonably "color-blind."

Table XI-1

Critical Ratios

Medical Technicians

Criteria

	Supervisors' Overall Rating				Job Knowledge Test			
			Black	Caucasian			Black	Caucasian
Predictor	CR_{yy}	CR_{xx}	r_{xy}		CR_{yy}	CR_{xx}	r_{xy}	
Subtraction & Multiplication	.12	.57	.29	.13	.44	.56	.34	.23
Vocabulary	.12	.77	.09	.06	.44	.78	.32	.27
Hidden Figures	.11	.47	.02	.05	.44	.46	.15	.14
Necessary Arithmetic Operations	.11	.84	.17	.16	.44	.86	.46	.34
Number Comparison	.12	.61	.17	.04	.44	.60	.23	.14
Gestalt Completion	.12	.33	.05	.11	.44	.33	.25	.17
Picture-Number	.11	.36	-.02	.15	.44	.37	.21	.16
Paper Folding	.11	.65	.08	.11	.44	.60	.22	.21

Table XI-2

Correlations and Critical Ratios--

Aptitude Tests with Supervisors' Overall Rating

Cartographic Technicians (TOPOCOM)

Aptitude Tests	Correlations xy: Black	Correlations xy: Mexican-American	Correlations xy: Caucasian	Critical ratios: Caucasian/Black CR_{yy}	Critical ratios: Caucasian/Black CR_{xx}	Critical ratios: Caucasian/Mexican-American CR_{yy}	Critical ratios: Caucasian/Mexican-American CR_{xx}
Hidden Figures	.21	.29	.21	.24	.22	.07	.39
Vocabulary	.19	-.02	.01	.27	.12	.08	.66
Object-Number	.19	.01	.02	.24	.36	.08	-.04
Card Rotations	.16	.03	.26	.27	.28	.08	.34
Map Planning	.24	.23	.30	.26	.29	.07	.35
Surface Development	.28	.22	.28	.27	.49	.08	.60
Maze Tracing Speed	.14	.15	.27	.28	.53	.08	.44
Extended Range Vocabulary	.17	.03	-.07	.27	.14	.08	.83
Necessary Arithmetic Operations	.25	.22	.19	.25	.82	.08	.76

Table XI-3

Correlations and Critical Ratios--

Aptitude Tests with Work Sample Composite

Cartographic Technicians (TOPOCOM)

Aptitude Tests	Correlations xy: Black	Correlations xy: Mexican-American	Correlations xy: Caucasian	Critical ratios: Caucasian/Black CR'_{yy}	Critical ratios: Caucasian/Black CR_{xx}	Critical ratios: Caucasian/Mexican-American CR'_{yy}	Critical ratios: Caucasian/Mexican-American CR_{xx}
Hidden Figures	.35	.43	.29	.20	.23	.41	.43
Vocabulary	.22	-.02	.22	.40	.08	.46	.66
Object-Number	.05	.03	.11	.39	.36	.44	-.03
Card Rotations	.26	.43	.34	.40	.29	.47	.36
Map Planning	.29	.39	.24	.38	.29	.44	.36
Surface Development	.31	.38	.41	.41	.50	.45	.61
Maze Tracing Speed	.30	.45	.34	.41	.54	.48	.46
Extended Range Vocabulary	.18	.05	.23	.41	.11	.47	.84
Necessary Arithmetic Operations	.28	.39	.35	.38	.82	.44	.77

Table XI-4

Correlations and Critical Ratios--

Aptitude Tests with Job Knowledge Test

Cartographic Technicians (TOPOCOM)

Aptitude Tests	Correlations xy			Critical ratios			
	Black	Mexican-American	Caucasian	Caucasian/Black		Caucasian/Mexican-American	
				CR_{yy}	CR_{xx}	CR_{yy}	CR_{xx}
Hidden Figures	.40	.45	.40	.43	.23	.92	.42
Vocabulary	.46	.22	.39	.43	.12	.90	.67
Object-Number	.28	.10	.17	.43	.36	.90	-.03
Card Rotations	.34	.36	.32	.43	.28	.90	.34
Map Planning	.35	.42	.36	.43	.29	.90	.35
Surface Development	.58	.55	.55	.43	.50	.89	.59
Maze Tracing Speed	.36	.46	.37	.43	.53	.90	.44
Extended Range Vocabulary	.50	.23	.39	.43	.14	.90	.84
Necessary Arithmetic Operations	.66	.44	.54	.43	.83	.90	.77

Table XI-5

Correlations and Critical Ratios--

Aptitude Tests with Supervisors' Overall Rating

Inventory Management Specialists

Aptitude Tests	Correlations xy: Black	Correlations xy: Mexican-American	Correlations xy: Caucasian	Critical ratios: Caucasian/Black CR'_{yy}	Critical ratios: Caucasian/Black CR_{xx}	Critical ratios: Caucasian/Mexican-American CR'_{yy}	Critical ratios: Caucasian/Mexican-American CR_{xx}
Number Comparison	.29	.10	.28	.06	.12	-.06	-.21
Hidden Figures	.30	.25	.10	.07	.43	.04	.12
Vocabulary	.14	.04	.12	.06	.50	-.06	.71
Object-Number	.04	.00	.07	.06	.07	-.06	.50
Letter Sets	.32	-.04	.21	.06	.49	.08	.28
Nonsense Syllogisms	.17	.33	.10	.06	.17	-.05	.38
Subtraction & Multiplication	.39	.25	.32	.06	.64	-.07	.30
Extended Range Vocabulary	.03	.10	.14	.06	.37	-.06	.72
Necessary Arithmetic Operations	.34	.25	.25	.06	.80	-.06	.34
Inference	.23	.13	.20	.05	.49	-.07	.57

Table XI-6

Correlations and Critical Ratios--

Aptitude Tests with Work Sample, Overall Performance

Inventory Management Specialists

Aptitude Tests*	Correlations xy: Black	Correlations xy: Mexican-American	Correlations xy: Caucasian	Critical ratios: Caucasian/Black CR_{yy}	Critical ratios: Caucasian/Black CR_{xx}	Critical ratios: Caucasian/Mexican-American CR_{yy}	Critical ratios: Caucasian/Mexican-American CR_{xx}
Number Comparison	.17	.36	.34	.64	.13	-.08	-.15
Hidden Figures	.21	.29	.30	.56	.50	-.05	.12
Vocabulary	.32	.41	.37	.61	.53	-.07	.74
Object-Number	.04	.20	.06	.61	.01	-.07	-.56
Letter Sets	.28	.49	.29	.62	.50	-.07	.24
Nonsense Syllogisms	.29	.38	.13	.56	.23	-.06	.38
Subtraction & Multiplication	.08	.37	.13	.65	.59	-.08	.29
Extended Range Vocabulary	.28	.58	.32	.64	.36	-.08	.63
Necessary Arithmetic Operations	.33	.60	.35	.61	.83	-.07	.35
Inference	.39	.56	.34	.59	.50	-.07	.47

Chapter XII

Conclusions and Inferences

The preceding chapters have described in detail the statistical treatment of the data obtained in this study and the conclusions derived. This chapter will recapitulate the more important findings, and draw some of the inferences that follow from them.

Conclusions

1. There is little in the data to support the hypotheses of differential validity for the wide variety of tests studied for the ethnic groups included in this study. Tests which were valid for one ethnic group were also valid for the other ethnic group(s). This held true when the tests were used to predict all three kinds of criterion measures.
2. Tests valid against one kind of criterion were generally valid against other criteria also.
3. When supervisors' ratings are used as the criterion, there is little difference in the regression lines for different ethnic groups; i.e., a particular test score predicts the same level of job performance for all ethnic groups.
4. When work samples or job knowledge tests are used as criteria, there usually are differences in the regression lines between majority and minority ethnic groups. In these instances, a given test score is associated with higher job performance for the Caucasian group than for the other two groups.
5. When test scores are combined by a multiple regression

equation, there is no practical loss in predictability when the equation developed for one ethnic group is used for prediction of criterion scores for the other ethnic groups.

6. Supervisors' ratings were affected by interaction of ethnic group membership of the rater with the ethnic group membership of the ratee. Raters tended to give higher ratings to ratees of their own ethnic group. Ratings of Black job incumbents by Black supervisors had higher correlation with other measures than did ratings of Caucasian job incumbents by Black supervisors. In contrast, ratings of Caucasian job incumbents by Mexican-American supervisors had higher correlations with other measures than did their ratings of Mexican-Americans. Ratings by Caucasian raters of all three ethnic groups correlated about equally well with other measures.

7. There is no substantial difference in background or experience variables for the different ethnic groups. Possibly for this reason, use of moderator variables such as length of experience or amount of education did not produce significant improvement in predictability.

8. Mean scores for minority groups on aptitude tests are generally about one-half standard deviation below the mean scores for Caucasians. There are, however, a few instances where the mean for Mexican-Americans is above the mean for Caucasians.

9. Mean scores for minority groups on job knowledge tests and

work samples are similarly about one-half standard deviation below the Caucasian mean score. There is one instance where the mean score for Mexican-Americans is above the mean score for Caucasians.

10. In contrast to the differences on tests and work samples, the means of supervisors' ratings for minority groups are very close to those for Caucasians.

11. Factor analyses of the test and criterion measures show very similar patterns for all groups.

Inferences and Implications

It is perhaps appropriate to comment on some of the inferences or implications which follow from the data which have been presented. Some of these comments reflect or overlap those made by speakers at the invitational conference previously mentioned (Crooks, Ed. 1972), although no effort will be made to recapitulate all of these comments here.

In view of the evidence from this study and other evidence presented by Boehm and Ruch, it appears that differential validity, if not entirely a statistical artifact where it does appear, is at best an isolated phenomenon. Furthermore, the use of different regression lines for different ethnic groups usually will operate to reduce rather than enhance employment opportunities for minority groups.

Evidence of the effect of rater-ratee ethnic group interaction on ratings makes the use of supervisors' ratings (or similar judgmental variables such as grades) as criteria of job performance a somewhat dubious proceeding where different ethnic groups are involved.

However, it is reassuring to find, in this study at least, that tests selected to predict supervisors' ratings for one group also made valid predictions of more "solid" criteria such as job knowledge tests and work samples for all groups.

It appears evident that the recognized problem of minority under-representation in technical, professional, and managerial jobs--or, in many instances, outright unemployment--will not be solved by the differential prediction approach, since this approach more often than not will reduce rather than increase the proportion of minority applicants employed. The question still remains, "What can be done to solve this problem?"

Answers--or hypotheses about the answers--to that question go beyond the scope of this research. Yet the question is too important to ignore or evade. One suggestion (Brown, 1972) is that employers use a low cutting score and provide additional training to those who score low. This is a reasonable suggestion where employers can afford to invest substantial training and are not bound by legal requirements to hire from the top of a qualifying list. If test scores are used to predict subsequent job performance, why should they not also be used to predict how much or what kind of training the new employee needs?

There is reason for some cautious optimism in the fact that background variables and test scores <u>do</u> seem to have the same meaning for individuals from different ethnic groups. It appears possible that if all groups have equal opportunity in <u>all</u> aspects of their lives--and it is evident that so far minority groups have <u>not</u> had equal opportunity--then eventually the differences in test scores between groups will become inconsequential.

References

Boehm, V. R. Negro-white differences in validity of employment and training selection procedures: Summary of research evidence. Journal of Applied Psychology, 1972, 56(1), 33-39.

Boehm, V. R. Test bias: Issues and research. Paper presented at symposium, Public Personnel Association, St. Louis, October 1972.

Brown, R. C., Jr. In Crooks, L. A. (Ed.), An investigation of sources of bias in the prediction of job performance: A six-year study. Princeton, New Jersey: Educational Testing Service, 1972.

Campbell, J. T., Pike, L. W., & Flaugher, R. L. Prediction of job performance for Negro and white medical technicians--A regression analysis of potential test bias: Predicting job knowledge scores from an aptitude battery. PR-69-6. Princeton, New Jersey: Educational Testing Service, 1969.

Campbell, J. T., Pike, L. W., Flaugher, R. L., & Mahoney, M. H. Prediction of job performance for Negro and white medical technicians: The prediction of supervisors' ratings from aptitude tests, using a cross-ethnic cross-validation procedure. PR-70-18. Princeton, New Jersey: Educational Testing Service, 1970.

Cleary, T. A. Test bias: Prediction of grades of Negro and white students in integrated colleges. Journal of Educational Measurement, 1968, 5, 115-124.

Cole, N. S. Bias in selection. ACT Research Report No. 51. Iowa City, Iowa: American College Testing Program, 1972.

Conklin, E. S. The scale of values method for studies in genetic psychology. University of Oregon Publications, 1923, 3, No. 1.

Crooks, L. A. Issues in the development and validation of In-basket exercises for specific objectives. RM-68-23. Princeton, New Jersey: Educational Testing Service, 1968.

Crooks, L. A. The In-basket study. PR-72-1. Princeton, New Jersey: Educational Testing Service, 1972.

Crooks, L. A. (Ed.) An investigation of sources of bias in the prediction of job performance: A six-year study. Princeton, New Jersey: Educational Testing Service, 1972.

Crooks, L. A., & Mahoney, M. H. Prediction of job performance for Black, Mexican-American, and Caucasian inventory management specialists: Instrumentation development and description of sample. PR-71-23. Princeton, New Jersey: Educational Testing Service, 1971.

Darlington, R. D. Another look at "culture fairness." Journal of Educational Measurement, 1971, 8, 71-82.

Farr, J. L., O'Leary, B. S., Pfeiffer, C. M., Goldstein, I. L., & Bartlett, C. J. Ethnic group membership as a moderator in the prediction of job performance: An examination of some less traditional predictors. Technical Report No. 2. September 1971. Silver Spring, Maryland: American Institutes for Research, 1971.

Flanagan, J. C. The American high-school student: Project TALENT. U. S. Office of Education Cooperative Research Project No.635. Pittsburgh: University of Pittsburgh, 1964.

Flanagan, J. C. Flanagan industrial tests manual. Chicago: Science Research Associates, Inc., 1965.

Flaugher, R. L. Patterns of test performance by high school students of four ethnic identities. Project Access Research Report No. 2. RB-71-25. Princeton, New Jersey: Educational Testing Service, 1971.

Flaugher, R. L., Campbell, J. T., & Pike, L. W. Prediction of job performance for Negro and white medical technicians--Ethnic group membership as a moderator of supervisor's ratings. PR-69-5. Princeton, New Jersey: Educational Testing Service, 1969.

Frederiksen, N., Saunders, D. R., & Wand, B. The In-basket test. Psychological Monographs: General and Applied, 1957, 71, 9 (Whole No. 438).

French, J. W., Ekstrom, R. B., & Price, L. A. Kit of reference tests for cognitive factors. Princeton, New Jersey: Educational Testing Service, 1963.

Ghiselli, E. E., & Brown, C. W. Personnel and industrial psychology. (2nd edition.) New York: McGraw-Hill, 1955.

Glennon, J. R., & Albright, L. W. A catalog of life history items. Washington, D. C.: Scientific Affairs Committee, American Psychological Association, Division 14, 1966.

Griggs vs. Duke Power Company, 401 U. S. 424.

Guinn, N., Tupes, E. C., & Alley, W. E. Cultural subgroup differences in the relationships between Air Force aptitude composites and training criteria. Technical Report 70-35. Lackland Air Force Base, Texas: Human Resources Research Center, 1970.

Hemphill, J. K., Griffiths, D. E., & Frederiksen, N. Administrative performance and personality: A study of the principal in a simulated elementary school. New York: Teachers College Bureau of Publications, Columbia University, 1962.

Kirkpatrick, J. J., Ewen, R. B., Barrett, R. S., & Katzell, R. A.

Testing and fair employment. New York: New York University Press, 1968.

Lesser, G. S., Fifer, F., & Clark, H. Mental abilities of children from different social class and cultural groups. Monographs of the Society for Research in Child Development, 1965, 30 (Whole Issue No. 4).

Linn, R. L. Fair test use in selection. Review of Educational Research, Vol. 43, No. 2, Spring 1973, 139-161.

Lopez, F. M., Jr. Current problems in test performance of job applicants: I. Personnel Psychology, 1966, 19, 10-18.

Morsh, J. E., & Christal, R. E. Impact of the computer on job analysis in the United States Air Force. PRL-TR-66-19, AD-656 304 Lackland Air Force Base, Texas: Personnel Research Laboratory, Aerospace Medical Division, October 1966.

O'Leary, B. S., Farr, J. L., & Bartlett, C. J. Ethnic group membership as a moderator of job performance. Technical Report No. 1. April 1970. Washington, D. C.: American Institutes for Research, 1970.

Parry, M. E. Prediction of job performance for Black, Mexican-American, and Caucasian cartographic technicians: Instrumentation development and description of the sample. PR-71-22. Princeton, New Jersey: Educational Testing Service, 1971.

Parry, M. E., & Mahoney, M. H. Prediction of job performance for Negro and white medical technicians: Otto Analyzer: An anchored rating scale's "main anchor." PR-70-16. Princeton, New Jersey: Educational Testing Service, 1970.

Pike, L. W. Prediction of job performance for Negro and white medical technicians: Dev opment of the instrumentation. PR-69-5. Princeton, New Jersey: Educational Testing Service, 1969.

Rock, D. A., Campbell, J. T., & Evans, F. R. Prediction of job performance for Negro and white medical technicians: Aptitude and rating factors. PR-70-17. Princeton, New Jersey: Educational Testing Service, 1970.

Ruch, W. W. A re-analysis of published differential validity studies. Paper presented at the symposium on differential validation under EEOC and OFCC testing and selection regulations. 80th Annual Convention of the American Psychological Association, Honolulu, September, 1972.

Shuey, A. M. The testing of Negro intelligence. (2nd edition.) New York: Social Science Press, 1966.

Smith, P. C., & Kendall, L. M. Retranslation of expectations: An approach to the construction of unambiguous anchors for rating scales. Journal of Applied Psychology, 1965, 47(2), 149-155.

Stanley, J. C. Predicting college success of educationally disadvantaged students. Science, 1971, Vol. 171, 640-647.

Stead, W. H., & Shartle, C. L. Occupational counseling techniques. New York: American Book, 1940.

Stodolsky, S. S., & Lesser, G. Learning patterns in the disadvantaged. Harvard Educational Review, 1967, 37, 546-593.

Thorndike, R. L. Concepts of culture-fairness. Journal of Educational Measurement, 1971, 8, 63-70.

Appendices

Appendix to Chapter II

An Attempt to Minimize Sources of Bias in Supervisors' Ratings by Use of a Standard Stimulus

In an effort to minimize known sources of bias in supervisors' ratings such as halo effect, rater leniency, and social desirability set, and to focus on the behavior to be rated on each scale beyond the definitions and anchors given, a fictitious incumbent was created for each occupation studied. On the facing page for each scale, the typical behavior of this incumbent was described as it related to the aspect of performance defined by the scale. On the basis of this behavioral description, the supervisor was asked to rate the fictitious person on the scale before rating the real incumbents he or she supervised. It was theorized that ratings made on such a standard stimulus would not only provide a benchmark of the supervisor's rating behavior as a basis for adjusting for possible bias in other ratings, but in addition might influence the rater to consider typical job behavior more carefully in rating those actually supervised.

A detailed description of the procedure and the findings in the Medical Technician study may be found in Parry and Mahoney (1970), which will only be summarized here. A similar analysis for Cartographic Technicians was also made. Because the findings in these two studies as to the desirability of adjusting ratings on the basis of benchmark rating behavior were somewhat inconclusive, no such detailed analysis was made for Inventory Management Specialists. However, the fictitious incumbent was again used in the rating procedure as a focus and further guide for the supervisors in rating their "real" incumbents.

The descriptions of the fictitious persons were made as realistic as possible by referring to them by name (Otto Analyzer, Medical Technician;

Mark Bench, Cartographic Technician; and Mel Stripp, Inventory Management Specialist). Although there was no effort to make the character of Otto Analyzer consistent from scale to scale in the Medical Technician study, an attempt was made in the other two occupations to describe Mark Bench and Mel Stripp across scales with consistent and believeable attributes, as possibly having more face validity. An overall rating was not asked for on the Medical Technician study, but was elicited for the other two. Examples of Mark Bench's descriptions for three scales are given in Appendix Figure II-1.

Analysis of the Data

A mean rating for the imaginary person was obtained for each of the scales where he was described, based on the number of supervisors who rated him. These mean ratings were assumed to be his "true" ratings. The means and standard deviations for these ratings for each scale for the Medical Technicians are given in Appendix Table II-A, and for the Cartographic Technicians in Appendix Table II-B. The possible ratings for each scale ranged from 1 (lowest) to 9 (highest). As may be seen from the variations in means across scales in these two tables, the supervisors did seem to be attending to the behavioral descriptions given. The ratings that supervisors had assigned to their "real" people were then adjusted, by scale, according to how each supervisor had rated the standard stimulus in relation to his mean ratings. For example, if the supervisor rated Mark Bench 8 and his mean rating for a given scale was 6.4, 1.6 was subtracted from the rating the supervisor gave each of his technicians on this scale. For any scale, if the imaginary person had not been rated by the supervisor, no adjustment could be made, and the supervisor's original ratings for that scale were used. In cases where individuals had been rated by more than one supervisor, an average of the adjusted ratings was obtained. Appendix Table II-C shows the means and standard deviations, by scale, for the ratings received by Black and Caucasian Medical Technicians. For the Black technicians, the adjusted mean ratings were higher than the originals on five scales and lower on two. For the Caucasian technicians, the adjusted mean ratings were higher than the originals on four scales, lower on two, and equal on one. Appendix Table II-D shows the means and standard deviations, by scale, for the ratings received by Black, Mexican-American, and Caucasian Cartographic Technicians. For the Black technicians, the adjusted mean ratings were higher than the originals on one scale and lower on six. For the Mexican-American technicians, the adjusted mean ratings were higher than the originals on two scales and lower on five. For the Caucasian technicians, the adjusted mean ratings were higher on five scales and lower on two. For all three ethnic groups, across both occupations, the variance of the adjusted ratings was greater than that of the unadjusted ratings for every scale but two. This may partly be due to the fact that the method of adjustment used increased the range of the scales. For example, if 3 were being added to all ratings given by a supervisor on a particular scale, and a technician had originally been given a 9, he would now receive a 12.

The adjusted ratings for each technician were then used as new criterion measures. The scores from the aptitude test battery administered to the technicians were correlated with the adjusted ratings and compared with the correlations obtained between the aptitude tests and the unadjusted ratings. Appendix Table II-E presents the correlations for Black

Medical Technicians, Appendix Table II-F for Caucasian Medical Technicians, Appendix Table II-G for Black Cartographic Technicians, Appendix Table II-H for Mexican-American Cartographic Technicians, and Appendix Table II-I for Caucasian Cartographic Technicians. For the Black Medical Technicians the correlations based on the adjusted ratings were lower than those based on the unadjusted ratings in 65 percent of the cases, unchanged in five percent, and higher in 30 percent. For the Caucasian technicians they were lower in 70 percent, unchanged in eight percent, and higher in 22 percent. The differences in either direction were small, ranging from .00 to .11. No pattern of change was evident, either by rating scale, aptitude test, or race. For the Black Cartographic Technicians the correlations based on the adjusted ratings were lower than those based on the unadjusted ratings in 44 percent of the cases, unchanged in eight percent, and higher in 48 percent. For the Mexican-American technicians they were lower in 57 percent of the cases, unchanged in 13 percent, and higher in 30 percent. For the Caucasian technicians they were lower in 53 percent of the cases, unchanged in 12 percent, and higher in 35 percent. Again, the differences in either direction were small, ranging from .00 to .15.

Job Knowledge Test scores were also correlated with the adjusted supervisors' ratings (Appendix Tables II-E to II-I) for both occupations. The correlations were lower than had been obtained with the unadjusted ratings with only one exception (Caucasian) for the Medical Technicians and five exceptions for the Cartographic Technicians (four Mexican-American and one Black). If one considers the Job Knowledge Test as another criterion measure, it is less related to those aspects of job performance which the adjusted ratings measure than it is to those which the unadjusted ratings measure.

When considering the correlations between the Job Knowledge Test and the job knowledge ratings, the difference between the correlations based on adjusted and unadjusted ratings is greater for the Black Medical Technicians than for the Caucasian Medical Technicians (a difference of .11 versus .02). For both ethnic groups, the correlation of Job Knowledge Test scores with unadjusted job knowledge ratings is higher than with adjusted ratings. For the Cartographic Technicians the differences are less marked (.06 for the Caucasians versus .05 for the Mexican-Americans and .04 for the Blacks). For the Blacks, the correlation of Job Knowledge Test scores with unadjusted job knowledge ratings is lower than with adjusted ratings, and for the Mexican-Americans and Caucasians it is higher with unadjusted than with adjusted.

Discussion

Why did adjusting the ratings in general fail to improve their correlations with the aptitude tests and the Job Knowledge Test? The reasons are not entirely clear. Apparently adjusting the technicians' ratings on the basis of pooling the judgments of a large number of supervisors resulted in the addition of more random variation instead of correcting for differential preferences in rating behavior.

One possible explanation is that the descriptions of the various aspects of the fictitious person's job performance were too specific and limited. If too little information was given on which to base the ratings of his performance, then adjusting all the ratings given by the supervisors according to their ratings of him could not be expected to improve the correlations between test scores and ratings. It may be easier for a supervisor to provide ratings for his technicians based on many relevant behaviors and incidents than it is to rate a simulated individual on the basis of only a paragraph description. The ratings of this person may, therefore, include more error variance than the ratings of the "real" people.

Another possible explanation is that the supervisors did not take the task of rating the fictive incumbent as seriously as they did rating their own personnel. If the raters exercised less care in rating him, this could explain the lower correlations found when adjusted ratings were used.

For the seven scales used from each of the two studies, adjusting the supervisors' ratings on the basis of their ratings of an imaginary subordinate did not generally improve the correlations between the aptitude measures and the ratings of job performance. It is possible that another type of analysis, such as a correction for each supervisor across scales rather than for each scale, would produce different results, although this would seem unlikely in view of the rather limited findings of the present analysis.

It is felt that the standard stimulus may still have served the other purpose for which he was included in the rating scales, that of helping to orient the supervisor toward the qualities of each job dimension to be rated. Also, it may be that having to rate a fictitious person, where the supervisor knew that his responses could be compared with those of other supervisors, tended to make him more attentive to the standards he was using. While taking account of the definition given for the scale and the description given of the fictitious individual, he may have been led to exercise greater care in rating the actual incumbents.

Mark Bench really loves maps. He collects old maps and books about cartography and likes to show each new find to his less enthusiastic co-workers. He spends a lot of time reading and studying about maps and mapping, but it is not always relevant to the work he is doing. He is inclined to take more interest in his work if he can spend time figuring out how to use a new piece of equipment, assembling source material, making elaborate plans, and looking up specifications. He would like to leave the routine compiling to others, although it is part of his job.

(Interest)

When faced with an assignment requiring fine detail, crowded features, or the use of small equipment such as gravers, Mark Bench seems to be "all thumbs." His contouring is sometimes so messy that it looks like a pile of wet spaghetti and he has to re-do it several times before it is acceptable. His eye-hand coordination is off just enough so that he has difficulty putting stick up just where he wants it. On tasks where less precision is required, he produces quite acceptable work.

(Dexterity)

Using the descriptions of Mark Bench and his work given for the previous scales, please give him an overall rating. This may not be a true evaluation of him as a cartographic technician, because you have never actually seen any of his work or have never met him in person. From your impression of him, rate him on this scale along with the other people in your group.

(Overall Performance)

Appendix Figure II-1 Examples of Behavioral Description for Fictitious Cartographic Technician

Appendix Table II-A

Means and Standard Deviations for the

Ratings Given to Fictitious Medical Technician by All Supervisors

N=200

Rating Scales	Mean	S.D.
Organization	6.4	2.0
Interest	6.6	2.3
Learning Ability	4.4	1.3
Job Knowledge	4.3	1.5
Technique	1.9	1.2
Need for Supervision	5.3	1.4
Communication	4.3	1.8

Appendix Table II-B

Means and Standard Deviations for the

Ratings Given to Fictitious Cartographic Technician by All Supervisors

(TOPOCOM and Coast & Geodetic Survey)

N=130

Rating Scales	Mean	S.D.
Interest	6.0	1.9
Learning Ability	4.6	1.4
Job Knowledge	5.2	1.3
Dexterity	1.7	1.1
Need for Supervision	3.3	1.7
Perseverance	3.8	1.4
Overall Rating	4.2	1.5

Appendix Table II-C

Means and Standard Deviations for Original and

Adjusted Supervisors' Ratings of

Black and Caucasian Medical Technicians

Rating Scales	Black (N=166)				Caucasian (N=285)			
	Mean		S.D.		Mean		S.D.	
	Orig.	Adj.	Orig.	Adj.	Orig.	Adj.	Orig.	Adj.
Organization	5.75	5.73	1.82	2.26	5.99	6.12	1.79	2.33
Interest	5.57	5.88	1.78	2.43	5.90	6.08	1.74	2.46
Learning Ability	5.79	5.82	1.88	2.07	5.96	5.82	1.77	2.05
Job Knowledge	5.23	5.09	1.88	2.01	5.30	5.35	1.74	2.08
Technique	5.90	5.97	1.73	1.73	6.01	6.06	1.73	1.86
Need for Supervision	5.72	5.82	1.97	2.08	6.10	5.98	1.86	2.07
Communication	5.49	5.85	1.81	2.13	5.75	5.75	1.75	2.06

Appendix Table II-D

Means and Standard Deviations for

Original and Adjusted Supervisors' Ratings of TOPOCOM

Black, Mexican-American, and Caucasian Cartographic Technicians

Rating Scales	Black (N=101)				Mexican-American (N=101)				Caucasian (N=241)			
	Mean		S.D.		Mean		S.D.		Mean		S.D.	
	Orig.	Adj.	Orig.	Adj.	Orig.	Adj.	Orig.	Adj.	Orig.	Adj.	Orig.	Adj.
Interest	5.34	5.28	1.68	2.23	5.86	5.57	1.70	1.78	5.87	5.95	1.61	1.96
Learning Ability	5.44	5.28	1.74	1.88	5.63	5.46	1.71	1.75	5.84	5.67	1.72	1.93
Job Knowledge	5.30	5.07	1.73	1.93	5.63	5.51	1.76	1.96	5.69	5.59	1.78	1.88
Dexterity	5.86	5.74	1.58	1.71	6.26	6.31	1.67	1.66	6.10	6.17	1.60	1.66
Need for Supervision	5.11	5.00	1.75	2.02	5.47	5.40	1.76	2.01	5.57	5.74	1.79	1.98
Perseverance	6.05	6.20	1.64	1.89	6.65	6.83	1.53	1.61	6.19	6.27	1.68	1.97
Overall Rating	5.46	5.39	1.86	2.16	5.78	5.75	1.63	1.75	5.88	5.94	1.74	2.07

Appendix Table II-E

Correlations Between Tests and Supervisors' Ratings

for Black Medical Technicians

Original Above - Adjusted Below

(N=166)

Test / Scale: Supervisors' Ratings	Subtraction & Multiplication	Vocabulary	Hidden Figures	Necessary Arithmetic	Finger Dexterity	Number Comparison	Gestalt Completion	Picture Number	Paper Folding	Job Knowledge Test
Organization	.32	.10	.03	.19	.09	.15	.06	-.07	.01	.45
	.27	.08	.04	.23	-.01	.08	.02	-.12	-.01	.39
Interest	.29	.10	.05	.17	.04	.07	.03	-.07	.00	.38
	.20	.11	.12	.20	.02	.06	.04	-.09	-.02	.20
Learning Ability	.36	.20	.02	.34	.20	.26	.18	.06	.21	.58
	.37	.14	.00	.30	.13	.26	.11	.07	.18	.53
Job Knowledge	.32	.20	.07	.33	.09	.18	.14	-.02	.09	.58
	.25	.18	.13	.25	.04	.16	.06	-.07	.06	.47
Technique	.28	.15	.04	.23	.08	.17	.08	-.06	.07	.49
	.25	.08	.02	.21	.04	.11	.03	-.02	.09	.44
Need for Supervision	.27	.10	.03	.25	.05	.10	.04	-.00	.09	.43
	.27	.07	.01	.21	.02	.08	.06	.01	.08	.38
Communication	.23	.21	-.02	.22	.06	.14	.12	-.05	.11	.50
	.25	.19	.03	.19	.05	.15	.09	-.03	.06	.35

Appendix Table II-F

Correlations Between Tests and Supervisors' Ratings for Caucasian Medical Technicians

Original Above - Adjusted Below

(N=285)

Test / Scale — Supervisors' Ratings	Subtraction & Multiplication	Vocabulary	Hidden Figures	Necessary Arithmetic	Finger Dexterity	Number Comparison	Gestalt Completion	Picture Number	Paper Folding	Job Knowledge Test
Organization	.15 .14	.01 -.02	.03 .06	.16 .15	.16 .08	.05 .03	.14 .10	.14 .09	.10 .09	.27 .25
Interest	.12 .06	.06 .01	.04 -.04	.15 .08	.12 .01	.07 .02	.06 -.02	.12 .04	.10 .08	.22 .17
Learning Ability	.24 .19	.07 .09	.12 .09	.30 .26	.26 .18	.16 .07	.19 .18	.20 .22	.28 .23	.35 .29
Job Knowledge	.09 .06	.13 .07	-.01 -.02	.12 .09	.10 .04	-.01 -.07	.03 -.02	.06 .06	.12 .06	.35 .33
Technique	.11 .08	.06 .03	.06 .05	.16 .17	.17 .15	.08 .06	.14 .13	.19 .22	.15 .14	.23 .18
Need for Supervision	.05 -.02	.05 .05	.03 .03	.09 .07	.07 -.01	-.01 -.08	.03 .00	.09 .01	.06 .04	.27 .30
Communication	.06 .02	.17 .18	.08 .01	.13 .14	.06 .03	.01 -.01	.03 .06	.05 .10	.10 .11	.27 .26

Appendix Table II-G

Correlations Between Tests and Supervisors' Ratings

for Black Cartographic Technicians (TOPOCOM Sample)

Original Above - Adjusted Below

(N=101)

Test / Scale — Supervisors' Ratings	Coordination	Hidden Figures	Vocabulary	Object-Number	Card Rotation	CS Arithmetic	Map Planning	Surface Development	Maze Tracing	Following Oral Directions	Identical Pictures	Extended Range Vocabulary	Necessary Arithmetic Operations	Job Knowledge Test
Interest	.02 .07	.17 .18	.21 .19	.14 .06	.14 .18	.23 .15	.08 .07	.21 .14	-.03 -.03	.16 .12	.08 .02	.18 .20	.16 .13	.15 .08
Learning Ability	.15 .14	.29 .29	.17 .20	.21 .14	.28 .29	.42 .36	.33 .34	.41 .37	.20 .21	.32 .25	.33 .33	.16 .20	.32 .31	.35 .31
Job Knowledge	.04 .00	.23 .28	.26 .38	.20 .24	.19 .21	.41 .41	.26 .27	.30 .33	.17 .18	.21 .25	.25 .28	.21 .36	.30 .37	.39 .43
Dexterity	.19 .16	.14 .12	.03 .01	.16 .10	.15 .12	.21 .20	.16 .19	.21 .15	.13 .11	.09 .09	.24 .21	.04 .02	.12 .11	.12 .09
Need for Supervision	.09 .07	.27 .31	.21 .21	.21 .23	.14 .18	.35 .35	.21 .23	.29 .33	.17 .18	.22 .25	.17 .18	.18 .21	.27 .31	.31 .29
Perseverance	.04 .07	.13 .20	.27 .22	.15 .11	.07 .03	.25 .14	.10 .06	.08 .06	-.07 -.02	.08 .07	.02 -.03	.24 .21	.20 .15	.13 .05
Overall Rating	.04 .00	.21 .30	.19 .26	.19 .17	.16 .21	.31 .33	.24 .26	.28 .31	.14 .13	.18 .24	.21 .23	.17 .26	.25 .29	.28 .26

Appendix Table II-H

Correlations Between Tests and Supervisors' Ratings

for Mexican-American Cartographic Technicians (TOPOCOM Sample)

Original Above - Adjusted Below

(N=99)

Test / Scale: Supervisors' Ratings	Coordination	Hidden Figures	Vocabulary	Object-Number	Card Rotation	CS Arithmetic	Map Planning	Surface Development	Maze Tracing	Following Oral Directions	Identical Pictures	Extended Range Vocabulary	Necessary Arithmetic Operations	Job Knowledge Test
Interest	.11	.34	.01	.03	.07	.26	.29	.26	.17	.33	.19	.10	.32	.47
	.08	.32	.03	.02	.06	.28	.23	.22	.16	.25	.18	.09	.30	.44
Learning Ability	.17	.41	.01	.12	.19	.34	.39	.35	.33	.32	.26	.07	.36	.51
	.13	.39	-.01	.10	.21	.28	.34	.32	.28	.31	.20	.05	.35	.47
Job Knowledge	.10	.27	-.06	.05	.07	.23	.24	.24	.18	.17	.15	.03	.25	.37
	.11	.21	-.05	.04	.03	.24	.20	.18	.20	.06	.14	-.01	.18	.32
Dexterity	.17	.40	-.07	.11	.06	.15	.23	.23	.17	.13	.28	.02	.22	.38
	.19	.43	-.10	.10	.11	.17	.28	.23	.23	.13	.29	-.02	.24	.41
Need for Supervision	.11	.32	-.00	-.01	.13	.19	.29	.26	.17	.17	.18	.08	.25	.44
	.14	.32	-.04	.01	.08	.26	.26	.19	.17	.17	.15	.02	.24	.48
Perseverance	.15	.27	.05	.04	.06	.20	.23	.13	.10	.23	.17	.11	.10	.38
	.11	.32	.02	.03	.09	.17	.20	.10	.08	.20	.14	.09	.10	.43
Overall Rating	.05	.29	-.02	.01	.04	.21	.23	.22	.15	.15	.18	.03	.22	.42
	.05	.31	.00	-.01	.06	.27	.25	.21	.19	.18	.18	.01	.25	.45

Appendix Table II-I

Correlations Between Tests and Supervisors' Ratings

for Caucasian Cartographic Technicians (TOPOCOM Sample)

Original Above - Adjusted Below

(N=240)

Test / Scale Supervisors' Ratings	Coordination	Hidden Figures	Vocabulary	Object-Number	Card Rotation	CS Arithmetic	Map Planning	Surface Development	Maze Tracing	Following Oral Directions	Identical Pictures	Extended Range Vocabulary	Necessary Arithmetic Operations	Job Knowledge Test
Interest	.10 .08	.18 .20	.11 .06	.04 .06	.17 .26	.22 .20	.20 .27	.19 .29	.19 .27	.24 .27	.11 .14	.03 -.02	.20 .19	.34 .29
Learning Ability	.21 .17	.25 .25	.03 .07	.04 .04	.31 .32	.25 .23	.40 .38	.34 .32	.32 .31	.33 .29	.20 .20	-.05 -.01	.29 .24	.45 .42
Job Knowledge	.16 .12	.21 .13	.11 .16	.03 .03	.20 .21	.17 .14	.25 .22	.22 .21	.18 .14	.24 .22	.09 .03	.04 .10	.18 .17	.38 .32
Dexterity	.26 .24	.26 .24	.00 -.02	.08 .10	.25 .22	.22 .21	.27 .25	.27 .23	.28 .24	.22 .19	.19 .16	-.09 -.10	.11 .10	.27 .22
Need for Supervision	.18 .21	.20 .17	.07 .08	.07 .07	.28 .27	.22 .23	.30 .29	.26 .26	.28 .30	.30 .31	.14 .13	.00 .00	.24 .29	.39 .33
Perseverance	.07 .10	.11 .11	.13 .11	.01 .06	.08 .14	.15 .11	.20 .21	.09 .14	.11 .15	.13 .12	.01 -.00	.06 .03	.09 .09	.25 .21
Overall Rating	.18 .17	.21 .19	.01 .05	.02 .05	.26 .27	.24 .21	.30 .27	.28 .26	.27 .26	.25 .25	.14 .15	-.07 -.01	.19 .19	.36 .31

Appendix to Chapter III

Appendix Table III-A

INTERCORRELATIONS OF PREDICTORS WITHIN ETHNIC GROUPS

Medical Technicians

(Black - first line; Caucasian - second line)
N=168 N=297

Test	1.	2.	3.	4.	5.	6.	7.	8.	9.
1. Subtraction &	1.00	.22	.08	.43	.13	.44	.22	.22	.16
Multiplication	1.00	.08	.11	.37	.19	.44	.08	.24	.08
2. Extended Range		--	.10	.29	-.08	.18	.18	.15	.20
Vocabulary			.15	.33	-.01	.03	.02	.08	.22
3. Hidden			--	.26	.26	.16	.30	.14	.15
Figures				.40	.37	.33	.38	.22	.42
4. Necessary				--	.10	.39	.36	.34	.43
Arithmetic Operations					.39	.40	.42	.36	.56
5. Pin-					--	.45	.21	.10	.19
Dexterity						.48	.50	.26	.47
6. Number						--	.42	.18	.33
Comparison							.35	.29	.32
7. Gestalt							--	.25	.44
Completion								.20	.52
8. Picture-								--	.30
Number									.13
9. Paper Folding									--

Appendix Table III-B

INTERCORRELATIONS OF PREDICTORS WITHIN ETHNIC GROUPS

Cartographic Technicians - TOPOCOM Sample

(Black - first line; Mexican-American - second line; Caucasian - third line)
N=101 N=99 N=241

Test	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.
1. Coordination	1.00	.06	-.15	.21	.14	.08	.17	.12	.33	.20	.29	-.12	.18
	1.00	.25	.05	.12	.26	.18	.10	.22	.40	.20	.36	.03	.12
	1.00	.21	-.03	.13	.24	.16	.25	.23	.33	.23	.20	-.01	.20
2. Hidden Figures		--	.32	.12	.43	.47	.48	.53	.39	.40	.45	.34	.44
			.03	.16	.35	.34	.40	.59	.49	.48	.51	.08	.53
			.20	.26	.31	.28	.34	.50	.41	.38	.34	.15	.31
3. Vocabulary			--	.13	.08	.22	.19	.24	.06	.26	.11	.88	.46
				-.02	.13	.14	.03	.16	.04	.30	-.10	.77	.24
				.13	-.01	.21	.11	.15	-.02	.31	.09	.82	.34
4. Object-Number				--	.25	.27	.25	.32	.24	.29	.30	.12	.28
					-.07	.12	.17	.12	.08	.09	.17	.04	.09
					.03	.20	.23	.16	.14	.16	.23	.14	.19
5. Card Rotations					--	.42	.50	.48	.40	.35	.53	.11	.38
						.36	.33	.40	.49	.31	.39	.05	.32
						.26	.46	.54	.55	.43	.48	-.06	.26
6. CS Arithmetic						--	.53	.54	.34	.57	.53	.26	.63
							.31	.40	.60	.32	.51	.09	.49
							.42	.29	.31	.44	.41	.11	.53
7. Map Planning							--	.56	.53	.51	.56	.24	.50
								.37	.43	.34	.28	.03	.31
								.44	.54	.45	.44	.01	.41
8. Surface Development								--	.44	.57	.46	.27	.54
									.55	.54	.43	.16	.62
									.49	.56	.44	.11	.49
9. Maze Tracing Speed									--	.33	.54	.10	.34
										.40	.68	.01	.46
										.42	.55	-.07	.30
10. Following Oral Directions										--	.48	.34	.66
											.26	.29	.61
											.37	.29	.59
11. Identical Pictures											--	.08	.41
												.04	.40
												.01	.23
12. Extended Range Vocabulary												--	.49
													.27
													.30
13. Necessary hmetic ations													--

Appendix Table III-C

INTERCORRELATIONS OF PREDICTORS WITHIN ETHNIC GROUPS

Cartographic Technicians - Coast & Geodetic Survey Sample

(Black - first line; Caucasian - second line)
N=38 N=51

Test	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.
1. Coordination	1.00	.02	.21	.02	.46	.33	.36	.19	.47	.52	.49	.14	.29
	1.00	.29	.13	.10	.18	.00	.25	.29	.29	.26	.37	.02	.10
2. Hidden Figures		--	.30	.53	.54	.45	.31	.70	.48	.35	.42	.41	.32
			-.10	.02	.41	.38	.44	.63	.58	.60	.56	-.09	.58
3. Vocabulary			--	.12	.34	.46	.23	.45	.21	.54	.14	.86	.39
				.09	.06	.21	.08	.08	-.08	.13	-.14	.79	.20
4. Object-Number				--	.37	.40	.38	.46	.24	.18	.57	.25	.09
					.10	.29	.33	.16	.29	.36	.09	.16	.20
5. Card Rotations					--	.52	.51	.65	.62	.58	.58	.29	.56
						.51	.57	.67	.48	.52	.60	.03	.60
6. CS Arithmetic						--	.54	.49	.50	.70	.53	.47	.61
							.56	.46	.49	.73	.45	.17	.77
7. Map Planning							--	.44	.48	.48	.53	.16	.50
								.65	.61	.69	.64	.12	.56
8. Surface Development								--	.40	.52	.51	.48	.55
									.67	.66	.62	.08	.63
9. Maze Tracing Speed									--	.45	.56	.10	.48
										.55	.68	-.04	.55
10. Following Oral Directions										--	.43	.54	.68
											.58	.14	.69
11. Identical Pictures											--	.10	.40
												-.19	.50
12. Extended Range Vocabulary												--	.33
													.22
13. Necessary Arithmetic Operations													--

Appendix Table III-D

INTERCORRELATIONS OF PREDICTORS WITHIN ETHNIC GROUPS

Inventory Management Specialists

(Black - first line; Mexican-American - second line; Caucasian - third line)
N=112 N=72 N=200

	Test	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.
1.	Number Comparison	1.00	.38	.21	.07	.41	.12	.41	.07	.25	.35	.11	.36
		1.00	.17	.23	.08	.31	.13	.45	.23	.33	.29	.13	.25
		1.00	.34	.17	.15	.47	.10	.54	.13	.41	.38	.27	.33
2.	Hidden Figures		--	.23	.16	.50	.25	.19	.21	.35	.47	.26	.38
				.42	-.02	.34	.17	.30	.36	.45	.42	.36	.47
				.33	.12	.52	.27	.11	.36	.47	.42	.34	.46
3.	Vocabulary			--	.09	.55	.46	.22	.88	.63	.43	.69	.76
					.16	.34	.36	.27	.73	.50	.36	.45	.65
					.09	.47	.31	.06	.85	.50	.38	.57	.69
4.	Object-Number				--	.19	.02	.05	.03	.13	.16	.20	.18
						.22	.08	.20	.22	.18	.07	.18	.26
						.22	.10	.17	.11	.17	.18	.13	.22
5.	Letter Sets					--	.38	.21	.50	.64	.63	.53	.66
							.33	.35	.35	.51	.46	.38	.60
							.45	.34	.50	.62	.56	.60	.71
6.	Nonsense Syllogisms						--	.16	.53	.50	.47	.56	.55
								.16	.46	.45	.61	.30	.54
								.08	.36	.44	.32	.43	.50
7.	Subtraction & Multiplication							--	.12	.32	.33	.11	.32
									.36	.42	.42	.33	.47
									.04	.40	.20	.21	.27
8.	Extended Range Vocabulary								--	.56	.43	.69	.76
										.54	.33	.52	.71
										.49	.43	.62	.73
9.	Necessary Arithmetic Operations									--	.61	.63	.68
											.59	.59	.72
											.59	.60	.76
10.	Following Oral Directions										--	.56	.59
												.39	.53
												.50	.60
11.	Inference											--	.74
													.69
													.74
12.	FSEE Total												--

Appendix to Chapter V

Appendix Table V-A

Correlations of Aptitude Tests with Supervisors' Ratings*

and Job Knowledge Test by Ethnic Group

Medical Technicians

(Black - first line; Caucasian - second line)
N=166 N=285

Test	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	R_9	JKT
Subtraction &	.35	.32	.29	.36	.32	.28	.27	.23	.29	.34
Multiplication	.24	.15	.12	.24	.09	.11	.05	.06	.13	.23
Extended Range	.11	.10	.10	.20	.20	.15	.10	.21	.09	.32
Vocabulary	-.00	.01	.06	.07	.13	.06	.05	.17	.06	.27
Hidden Figures	.04	.03	.05	.02	.07	.04	.03	-.02	.02	.15
	.16	.03	.04	.12	-.01	.06	.03	.08	.05	.14
Necessary Arithmetic	.29	.19	.17	.34	.33	.23	.25	.22	.17	.46
Operations	.29	.16	.15	.30	.12	.16	.09	.13	.16	.34
Pin-Dexterity	.14	.09	.04	.20	.09	.08	.05	.06	.10	.28
	.26	.16	.12	.26	.10	.17	.07	.06	.12	.17
Number	.16	.14	.07	.26	.18	.17	.10	.14	.17	.23
Comparison	.16	.05	.07	.16	-.01	.08	-.01	.01	.04	.14
Gestalt	.14	.06	.03	.18	.14	.08	.04	.12	.05	.25
Completion	.23	.14	.06	.19	.03	.14	.03	.03	.11	.17
Picture-	-.03	-.07	-.07	.06	-.02	-.06	-.00	-.05	-.02	.21
Number	.22	.14	.12	.20	.06	.19	.09	.05	.15	.16
Paper	.13	.00	.00	.21	.09	.07	.09	.11	.08	.22
Folding	.26	.10	.10	.28	.12	.15	.06	.10	.11	.21

*Rating Scales

R_1 = Flexibility
R_2 = Organization
R_3 = Interest
R_4 = Learning Ability
R_5 = Job Knowledge
R_6 = Technique
R_7 = Low Need for Supervision
R_8 = Communication
R_9 = Overall Performance Rating

JKT = Job Knowledge Test

Appendix Table V-B

Correlations of Aptitude Tests with Supervisors' Ratings*,

Job Knowledge Test and Work Sample** by Ethnic Group

Cartographic Technicians (TOPOCOM)

(Black - first line; Mexican-American - second line; Caucasian - third line)
N=100 N=99 N=240

Tests	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	JKT	WS_1	WS_2	WS_3	WS_4
	.12	.02	.15	.04	.19	.09	.04	.04	-.06	-.03	.18	.11	.13
Coordination	.10	.11	.17	.10	.17	.11	.15	.05	.18	.16	.15	.23	.25
	.21	.10	.21	.16	.26	.18	.07	.18	.21	.12	.06	-.08	.05
	.25	.17	.29	.23	.14	.27	.13	.21	.41	.26	.17	.29	.35
Hidden Figures	.36	.34	.41	.27	.40	.32	.27	.29	.45	.29	.34	.30	.43
	.21	.18	.25	.21	.26	.20	.11	.21	.40	.19	.19	.23	.29
	.14	.21	.17	.26	.03	.21	.27	.19	.46	.16	.20	.11	.22
Vocabulary	-.05	.01	.01	-.06	-.07	-.00	.05	-.02	.22	.06	-.08	-.02	-.02
	.11	.11	.03	.11	.00	.07	.13	.01	.39	.18	.13	.14	.22
	.29	.14	.21	.20	.16	.21	.15	.19	.28	.02	-.16	.24	.05
Object-Number	.03	.03	.12	.05	.11	-.01	.04	.01	.10	-.03	.04	.05	.03
	.08	.04	.04	.03	.08	.07	.01	.02	.17	.08	.06	.08	.11
	.20	.14	.28	.19	.15	.14	.07	.16	.34	.06	.15	.32	.26
Card Rotations	.02	.07	.19	.07	.06	.13	.06	.04	.34	.26	.31	.35	.43
	.20	.17	.31	.20	.25	.28	.08	.26	.33	.23	.24	.25	.34
	.34	.23	.42	.41	.21	.35	.25	.31	.62	.12	.15	.34	.30
CS Arithmetic	.24	.26	.34	.23	.15	.19	.20	.21	.44	.18	.25	.32	.34
	.20	.22	.25	.17	.22	.22	.15	.24	.43	.13	.01	.07	.11
	.18	.08	.33	.26	.16	.21	.10	.24	.35	.12	.20	.27	.29
Map Planning	.26	.29	.39	.25	.23	.29	.23	.23	.42	.25	.31	.27	.39
	.36	.29	.20	.40	.25	.27	.30	.20	.30	.16	.15	.18	.24

(continued on next page)

Appendix Table V-C

Correlations of Aptitude Tests with Supervisors' Ratings* by Ethnic Group

Cartographic Technicians (Coast & Geodetic Survey)

(Black - first line; Caucasian - second line)
N=38 N=50

Test	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8
Coordination	.18 -.03	.13 .09	.26 .12	.15 .15	.08 -.02	.29 .07	.10 .05	.13 .00
Hidden Figures	.27 .41	.29 .34	.40 .51	.34 .37	.32 .60	.24 .52	.20 .11	.30 .42
Vocabulary	.12 .09	.06 .12	.28 .02	.11 .19	.18 .01	.19 .01	.09 .11	.09 .07
Object-Number	.43 -.06	.36 -.01	.44 .09	.42 .10	.41 .10	.31 .11	.28 .09	.35 .07
Card Rotations	.42 .30	.32 .22	.51 .39	.45 .31	.42 .37	.41 .42	.23 .22	.37 .30
CS Arithmetic	.48 .40	.49 .23	.61 .48	.50 .47	.37 .50	.40 .53	.46 .24	.51 .44
Map Planning	.39 .26	.25 .27	.44 .45	.35 .28	.36 .39	.29 .46	.25 .19	.31 .32
Surface Development	.39 .31	.35 .34	.53 .47	.43 .40	.48 .40	.33 .48	.27 .28	.38 .40
Maze Tracing Speed	.07 .26	.03 .28	.21 .45	.08 .36	-.00 .36	.09 .42	-.06 .20	.11 .33
Following Oral Directions	.32 .36	.29 .36	.43 .52	.36 .42	.23 .49	.32 .49	.33 .34	.32 .46
Identical Pictures	.39 .29	.41 .32	.51 .52	.48 .38	.42 .37	.42 .43	.36 .20	.44 .36
Extended Range Vocabulary	.21 -.06	.18 -.01	.35 -.11	.19 .05	.25 .02	.24 -.06	.19 -.04	.21 -.01
Necessary Arithmetic Operations	.27 .44	.25 .36	.39 .51	.28 .49	.25 .48	.30 .59	.24 .24	.33 .48

*Rating Scales

R_1 = Accuracy
R_2 = Interest
R_3 = Learning Ability
R_4 = Job Knowledge
R_5 = Dexterity
R_6 = Low Need for Supervision
R_7 = Perseverance
R_8 = Overall Rating

Appendix Table V-D

Correlations of Aptitude Tests with Supervisors' Ratings*

and Work Sample** by Ethnic Group

Inventory Management Specialists

(Black - first line; Mexican-American - second line; Caucasian - third line)

N(R)=112 N(WS)=100 | N(R)=72 N(WS)=57 | N(R)=191 N(WS)=157

Test	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	R_9	R_{10}	WS_1	WS_2	WS_3	WS_4	WS_5	WS_6	WS_7	WS_8	WS_9	WS_{10}
Number Comparison	.34	.24	.34	.21	.23	.26	.25	.31	.21	.29	.15	.21	.17	.09	.24	.20	.29	.15	.14	.17
	.12	.15	.33	.29	.18	.05	-.04	.17	.06	.10	.23	.36	.16	.28	.24	.14	.43	.27	.35	.36
	.26	.20	.34	.31	.32	.26	.23	.26	.02	.28	.28	.35	.27	.26	.35	.20	.41	.28	.36	.34
Hidden Figures	.30	.27	.37	.25	.22	.20	.23	.25	.10	.30	.18	.19	.13	.12	.33	.09	.27	.28	.19	.21
	.28	.15	.32	.34	.28	.20	.10	.20	.20	.25	.25	.18	.19	.15	.32	.10	.23	.13	.33	.29
	.01	.10	.08	.18	.11	.11	.11	.15	.14	.10	.17	.26	.12	.19	.28	.14	.28	.25	.28	.30
Vocabulary	.11	.07	.10	.13	.12	.08	.16	.12	.12	.14	.30	.34	.16	.31	.40	.15	.26	.15	.35	.32
	.05	-.04	.15	.23	.12	-.02	-.15	.06	.06	.04	.40	.33	.22	.40	.43	.22	.35	.23	.43	.41
	.05	.07	.13	.17	.08	.07	.15	.19	.04	.12	.32	.33	.20	.36	.29	.14	.31	.27	.38	.37
Object-Number	.01	.03	.03	.01	.04	.10	.06	-.07	.01	.04	.16	.04	.14	-.10	.10	.12	.02	.14	-.02	.04
	.05	-.05	.06	.06	-.04	-.09	-.03	-.07	-.04	.00	.12	.07	.21	.06	.12	.02	.23	-.09	.19	.20
	.06	.07	.06	.15	.10	.09	.07	.07	.06	.07	.11	.15	.11	.07	.04	.05	.14	.05	.07	.06
Letter Sets	.32	.26	.31	.28	.27	.27	.33	.32	.21	.32	.32	.35	.17	.20	.30	.23	.25	.32	.28	.28
	-.04	-.10	.24	.29	.06	-.03	-.07	.09	-.06	-.04	.19	.28	.26	.31	.14	.25	.36	.40	.38	.49
	.15	.18	.25	.32	.22	.20	.18	.26	.13	.21	.29	.32	.19	.17	.37	.18	.33	.23	.31	.29
Nonsense Syllogisms	.10	.12	.26	.00	.07	.11	.22	.23	.05	.17	.32	.37	.14	.28	.33	.13	.17	.20	.36	.29
	.30	.32	.40	.40	.36	.40	.28	.36	.30	.33	.22	.36	.20	.46	.45	.21	.42	.51	.38	.38
	.07	.07	.13	.09	.10	.06	.11	.17	.05	.10	.15	.10	.05	.11	.11	-.01	.09	.05	.15	.13
Subtraction & Multiplication	.36	.24	.35	.24	.27	.32	.30	.28	.17	.39	.07	.17	.21	.07	.07	.19	.14	.19	.13	.08
	.24	.17	.36	.40	.20	.21	.07	.24	.12	.25	.31	.34	.19	.44	.25	.11	.36	.14	.39	.37
	.33	.29	.31	.34	.26	.32	.27	.25	.13	.32	.18	.22	.20	.11	.21	.13	.29	.18	.19	.13

(continued on next page)

Appendix Table V-D (Continued)

Test	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	R_9	R_{10}	WS_1	WS_2	WS_3	WS_4	WS_5	WS_6	WS_7	WS_8	WS_9	WS_{10}
Extended Range Vocabulary	.00	-.01	-.04	.02	.02	-.04	.03	.03	-.06	.03	.26	.29	.08	.31	.38	.11	.23	.06	.33	.28
	.09	.00	.15	.27	.14	.03	-.04	.10	.20	.10	.45	.53	.35	.53	.53	.25	.52	.24	.66	.58
	.07	.11	.18	.19	.13	.13	.13	.22	.03	.14	.26	.29	.14	.34	.26	.12	.27	.18	.34	.32
Necessary Arithmetic Operations	.29	.29	.36	.21	.25	.29	.35	.34	.11	.34	.29	.37	.20	.27	.33	.18	.26	.35	.37	.33
	.22	.10	.39	.39	.33	.18	.09	.30	.12	.25	.28	.47	.34	.44	.36	.34	.47	.34	.54	.60
	.21	.21	.28	.31	.23	.25	.20	.29	.17	.25	.31	.35	.17	.25	.36	.12	.31	.21	.37	.35
Following Oral Directions	.24	.20	.35	.14	.13	.21	.20	.18	.15	.22	.36	.46	.26	.28	.36	.21	.29	.34	.36	.36
	.25	.18	.34	.23	.30	.24	.14	.28	.14	.25	.31	.31	.17	.35	.41	.22	.35	.46	.36	.41
	.11	.13	.18	.22	.18	.13	.06	.21	.12	.13	.33	.39	.23	.31	.39	.21	.36	.30	.29	.42
Inference	.18	.19	.30	.15	.20	.18	.24	.24	.14	.23	.37	.40	.31	.37	.48	.23	.27	.24	.43	.39
	.12	.05	.22	.22	.21	.10	-.04	.21	.07	.13	.54	.50	.26	.35	.43	.37	.41	.16	.49	.56
	.13	.17	.21	.26	.21	.20	.15	.30	.09	.20	.32	.30	.18	.30	.29	.17	.26	.19	.33	.34
FSEE (VA)	.12	.16	.19	.13	.17	.11	.14	.15	.03	.13	.28	.34	.22	.30	.42	.16	.22	.14	.35	.34
	.19	.16	.24	.36	.09	.16	.06	.22	.23	.19	.42	.48	.31	.43	.48	.33	.48	.27	.53	.51
	.14	.16	.24	.24	.19	.15	.13	.28	.09	.16	.34	.35	.17	.30	.31	.10	.29	.21	.36	.35
FSEE (AR)	.31	.31	.36	.27	.36	.32	.27	.30	.10	.37	.30	.28	.19	.14	.32	.23	.17	.16	.21	.25
	.08	.00	.28	.36	.08	.05	-.01	.17	.02	-.04	.24	.23	.19	.24	.18	.22	.22	.07	.35	.41
	.09	.10	.15	.17	.15	.10	.10	.18	.14	.10	.23	.30	.22	.19	.28	.24	.27	.25	.30	.34
FSEE (QR)	.35	.34	.42	.30	.33	.33	.33	.40	.27	.38	.34	.42	.26	.30	.41	.17	.31	.33	.36	.34
	.21	.16	.36	.39	.30	.24	.18	.34	.22	.24	.34	.50	.33	.44	.42	.40	.45	.42	.55	.58
	.18	.16	.28	.26	.27	.24	.18	.31	.22	.25	.26	.39	.26	.22	.34	.15	.33	.19	.35	.36
FSEE (VA +QR)	.22	.25	.30	.21	.25	.21	.23	.26	.13	.24	.33	.40	.25	.32	.45	.18	.28	.23	.29	.37
	.22	.18	.32	.41	.19	.21	.08	.30	.25	.24	.43	.54	.36	.48	.51	.40	.52	.36	.60	.60
	.17	.18	.28	.27	.25	.21	.17	.32	.16	.22	.34	.40	.23	.30	.36	.14	.34	.22	.40	.40

*Rating Scales

R_1 = Organization
R_2 = Interest
R_3 = Learning Ability
R_4 = Communication
R_5 = Technical Knowledge
R_6 = Stability/ Adaptability
R_7 = Dependability
R_8 = Judgment
R_9 = Cooperation
R_{10} = Overall Rating

**Work Sample

WS_1 = Takes Leading Action
WS_2 = Uses New Procedures
WS_3 = Shows I. M. Knowledge
WS_4 = Analyzes Problems
WS_5 = Organizes Systematically
WS_6 = Maintains Controls
WS_7 = Follows Directions
WS_8 = Productivity
WS_9 = Quality of Actions
WS_{10} = Rating of Overall Performance