Contract No.: K-4279-3-00-80-30
MPR Reference No.: 8140-930

National Job Corps Study: The Impacts of Job Corps on Participants' Literacy Skills

Final Report

July 7, 2000

Steven Glazerman
Peter Z. Schochet
John Burghardt

Submitted to:

U.S. Department of Labor
Employment and Training Administration
Office of Policy and Research
Room N-5637
200 Constitution Ave., NW
Washington, DC 20210

Project Officer:
Daniel Ryan

Project Director:
John Burghardt

Principal Investigators:
Terry Johnson
Charles Metcalf

Submitted by:

Mathematica Policy Research, Inc.
(Prime Contractor)
P.O. Box 2393
Princeton, NJ 08543-2393
(609) 799-3535

In conjunction with:

Battelle Human Affairs Research
Centers (Subcontractor)
4000 N.E. 41st St.
Seattle, WA 98105

Decision Information Resources, Inc.
(Subcontractor)
610 Gray St., Suite 200
Houston, TX 77002

This report has been produced under Contract Number K-4279-3-00-80-30 with the U.S. Department of Labor, Employment and Training Administration. The contents of the report do not necessarily reflect the views or policies of the Department of Labor, nor does mention of trade names, commercial products, or organizations imply endorsement of these by the U.S. Government. Mathematica Policy Research, Inc. assumes full responsibility for the completeness and accuracy of the information presented in this report.

ACKNOWLEDGMENTS

In spring 1994, the Advisory Panel of the National Job Corps Study recommended that the U.S. Department of Labor (DOL) consider funding a substudy of the National Job Corps Study that would examine the impacts of participation in Job Corps on the literacy skills of Job Corps students. DOL concurred with the panel's recommendation, and this report presents the findings of that substudy.

We would like to acknowledge the efforts of several individuals who contributed to the research presented in this report. The design and execution of the study benefited from the guidance of Daniel Ryan, DOL project officer for the National Job Corps Study. We also appreciate the assistance that Irwin Kirsch of the Educational Testing Service (ETS) provided during the design. We would especially like to acknowledge the contributions of Mamoru Ishikawa of DOL's Office of Policy and Research, whose ideas about how to measure the impacts of Job Corps on literacy skills shaped our approach to this study.

Many individuals in three organizations contributed to the data collection and data preparation. The study used the literacy measurement methodology developed by ETS, and several staff at ETS played key roles. Julie Eastland developed field protocols for administering the literacy exercises, trained senior MPR project staff on administration of the literacy exercises, and trained and supervised the individuals who scored the completed literacy exercises. Kentaro Yamamoto performed the statistical analysis for the IRT scoring. At MPR, John Homrighausen, the National Job Corps Study Survey Director, Donna Mikolajewski, and William Beecroft trained all field interviewers in administering the literacy exercises and supervised the MPR field interviewers. At Decision Information Resources, MPR's subcontractor for the National Job Corps Study, Pam Wells and Eleanor Tongee supervised the field interviewers in the South and Southwest regions.

Several people assisted with the preparation of the report. Walter Corson, who provided quality assurance review, offered suggestions that improved the presentation. Laura Berenson and Patricia Ciaccio provided careful editorial support. Cathy Harper and Marjorie Mitchell did an excellent job producing the report.

CONTENTS

Chapter		Page
	EXECUTIVE SUMMARY	xi
I	INTRODUCTION	1
	A. OVERVIEW OF JOB CORPS	2
	1. Outreach and Admissions	3
	2. Job Corps Center Services	3
	3. Placement	6
	B. POLICY ISSUES RELATED TO BASIC SKILLS AND STUDY OBJECTIVES	6
II	SAMPLE DESIGN, DATA, AND METHODS	9
	A. TIMING OF SKILLS MEASUREMENTS	9
	B. SELECTING AN APPROPRIATE TEST	12
	1. Criteria for Selecting a Test	12
	2. Features of Functional Literacy Profiling and the NALS Instrument	16
	C. SAMPLE DESIGN	17
	D. INTERVIEW RESPONSE RATES AND ITEM NONRESPONSE	20
	E. ANALYTIC METHODS	24
III	IMPACTS ON LITERACY SKILLS	27
	A. COMPARING THE LITERACY SKILLS OF ELIGIBLE JOB CORPS APPLICANTS TO THOSE OF OTHER YOUNG ADULTS	27
	B. IMPACTS ON LITERACY SKILLS FOR THE FULL SAMPLE	32
	C. IMPACTS ON LITERACY SKILLS FOR SUBGROUPS	35

CONTENTS *(continued)*

Chapter		Page
IV	INTERPRETATION OF THE IMPACT FINDINGS	41
	A. THEORETICAL FRAMEWORK	42
	B. THE RELATIONSHIP BETWEEN LITERACY SKILLS AND SCHOOLING AND WORK EXPERIENCE	44
	C. THE RELATIONSHIP BETWEEN LITERACY SKILLS AND EARNINGS	53
	REFERENCES	59
	APPENDIX A: ANALYSIS OF NONRESPONSE	
	APPENDIX B: ADMINISTRATION AND SCORING OF THE JOB CORPS ADULT LITERACY ASSESSMENT	
	APPENDIX C: SUPPLEMENTARY TABLES	

TABLES

Table		Page
II.1	DESCRIPTION OF PROSE DOCUMENT, AND QUANTITATIVE LITERACY TASKS AT EACH OF FIVE LEVELS	18
II.2	RESPONSE RATES TO THE LITERACY STUDY, BY RESEARCH STATUS AND KEY SUBGROUP	21
III.1	LITERACY SKILLS OF ELIGIBLE JOB CORPS APPLICANTS COMPARED TO THOSE OF YOUNG ADULTS IN THE NALS	28
III.2	LITERACY SKILLS OF ELIGIBLE JOB CORPS APPLICANTS COMPARED TO THOSE OF ADULTS IN THE NALS, BY EDUCATIONAL ATTAINMENT	31
III.3	IMPACTS ON LITERACY SKILLS	34
III.4	IMPACTS ON AVERAGE LITERACY SCORES, BY SUBGROUP	37
IV.1	PARAMETER ESTIMATES ON SCHOOLING AND WORK EXPERIENCE MEASURES USING PROSE SCORES AS THE DEPENDENT VARIABLE	46
IV.2	PARAMETER ESTIMATES ON SCHOOLING AND WORK EXPERIENCE MEASURES USING DOCUMENT SCORES AS THE DEPENDENT VARIABLE	47
IV.3	PARAMETER ESTIMATES ON SCHOOLING AND WORK EXPERIENCE MEASURES USING QUANTITATIVE SCORES AS THE DEPENDENT VARIABLE	48
IV.4	AVERAGE LITERACY ASSESSMENT SCORES BY EDUCATIONAL ATTAINMENT AND LITERACY DOMAIN	52
IV.5	PARAMETER ESTIMATES ON PROSE, DOCUMENT, AND QUANTITATIVE TEST SCORES USING EARNINGS PER WEEK AND HOURLY WAGES ON THE MOST RECENT JOB IN QUARTER 10 AS THE DEPENDENT VARIABLES	55
A.1	PROCESS FOR OBTAINING THE JOB CORPS LITERACY STUDY SAMPLE	A.4

TABLES *(continued)*

Table		Page
A.2	REASON FOR NONRESPONSE TO LITERACY ASSESSMENT INTERVIEWS, BY RESEARCH STATUS AND WHETHER COMPLETED A 30-MONTH INTERVIEW	A.7
A.3	TIMING OF LITERACY ASSESSMENTS	A.8
A.4	DURATION OF LITERACY ASSESSMENT INTERVIEWS	A.9
A.5	AVERAGE ITEM RESPONSE RATES, BY SUBGROUP AND RESEARCH STATUS	A.11
A.6	CHARACTERISTICS OF RESPONDENTS TO THE LITERACY ASSESSMENT AND THE FULL SAMPLE OF RESPONDENTS AND NONRESPONDENTS, BY RESEARCH STATUS	A.12
A.7	TABE SCORES OF RESPONDENTS AND NONRESPONDENTS TO THE LITERACY ASSESSMENTa	A.16
B.1	BALANCED INCOMPLETE BLOCK DESIGN FOR JOB CORPS LITERACY ASSESSMENT	B.4
C.1	IMPACTS ON LITERACY SKILLS USING THE SIMPLE DIFFERENCES-IN-MEANS APPROACH	C.4
C.2	PARAMETER ESTIMATES FROM MODELS ESTIMATING PROGRAM IMPACTS, BY LITERACY DOMAIN	C.5
C.3	PARAMETER ESTIMATES FROM MODELS ESTIMATING THE ASSOCIATION BETWEEN HUMAN CAPITAL MEASURES AND SKILL SCORES, BY LITERACY DOMAIN	C.8

FIGURES

Figure		Page
I.1	JOB CORPS CENTERS IN PROGRAM YEAR 1995, BY REGION	4
IV.1	THE HYPOTHESIZED RELATIONSHIP BETWEEN FAMILY BACKGROUND, SCHOOLING, LITERACY SKILLS, AND LABOR MARKET BEHAVIOR	43

EXECUTIVE SUMMARY

Job Corps is a major part of federal efforts to provide education and job training to disadvantaged youths. It provides comprehensive services: basic education, vocational skills training, health care and education, counseling, and residential support. More than 60,000 new students between the ages of 16 and 24 enroll in Job Corps each year, at an annual cost to the federal government of more than $1 billion. Currently, the program provides training at 119 Job Corps centers nationwide. The National Job Corps Study is being conducted under contract with the U.S. Department of Labor to provide Congress and program managers with the information they need to assess how well Job Corps attains its goal of helping students become employable, productive citizens.

This report is one of a series presenting findings from the study. It presents estimates of the impacts of Job Corps on participants' literacy and numeracy skills needed to function in the workplace. It builds on the analysis and findings presented in our report on short-term impacts (Schochet et al. 2000). That report relied on interview data collected at baseline, and at 12 and 30 months after random assignment. The current report is based on an in-person literacy skills assessment administered to some sample members in conjunction with the 30-month follow-up interview. These test score data allow us to measure the extent to which Job Corps improves the functional literacy and numeracy skills of Job Corps participants.

STUDY DESIGN, DATA, AND METHODS

The cornerstone of the National Job Corps Study was the random assignment of all youths found eligible for Job Corps to either a program group or a control group. Program group members were permitted to enroll in Job Corps and control group members could not (although they could enroll in other education or training programs). The research sample for the study consists of approximately 9,400 program group members and 6,000 control group members randomly selected from among the nearly 81,000 applicants nationwide who applied to Job Corps for the first time between November 17, 1994 and December 16, 1995 and who were found eligible by February 1996.

The study to measure Job Corps impacts on participants' literacy and numeracy skills was based on a randomly selected subsample of program and control group members. Key features of the study design are as follows:

- ***A single round of skills measurement was conducted in conjunction with the 30-month follow-up interview.*** Because of limited study resources, only a single round of testing could be conducted. We selected the 30-month measurement point under the assumption that the most important study goal is knowing whether and to what extent Job Corps produces differences after most program participants leave the program and spend some time in the workforce.

- ***A total of 3,750 sample members (1,875 program group and 1,875 control group members) was randomly selected for the literacy study.*** The sample was selected from all program and control group members who were eligible for 30-month interviews and who were randomly assigned during the last 7 months of the 16-month sample intake period. The analysis sample contains 1,117 program and 1,156 control group members who completed the literacy test. The overall weighted response rate was 60.2 percent and was similar for program and control group members. On average, respondents attempted 85 percent of the tasks they were asked to perform.

- ***The approach to literacy assessment developed by the Educational Testing Service (ETS) was used for the study.*** Specifically, this study used a version of the assessment instrument that ETS designed for the National Adult Literacy Survey (NALS). We also considered a range of other instruments, including the Test of Adult Basic Education (TABE), which the Job Corps academic education program uses as a diagnostic tool for program participants. We selected the instrument designed by ETS, however, because it focuses on *functional* literacy and numeracy skills rather than academic skills only. Assessing functional skills is meaningful because Job Corps' mission is to prepare its students for a job or for further education that will lead to a job. The ETS approach has been used also in several national studies with populations similar to the population of Job Corps students.

Program impacts were estimated by comparing the average test scores and the test score distributions of program and control group members. We estimated program impacts for the full sample and for key subgroups defined by the following baseline characteristics: gender, presence of children, age, educational attainment, and residential designation status.

HOW WE ASSESSED LITERACY SKILLS

The approach to literacy assessment developed by ETS measures the ability to perform a wide range of information-processing tasks that adults encounter in everyday life. The approach posits three dimensions of literacy:

1. ***Prose literacy,*** the knowledge and skills necessary to understand and use information from texts

2. ***Document literacy,*** the knowledge and skills necessary to locate and use information in tables, charts, graphs, and maps

3. ***Quantitative literacy***, the knowledge and skills necessary to perform different arithmetic operations using information embedded in prose and document materials

Proficiency in each of these domains is measured on a scale from 0 to 500. To estimate proficiency, ETS developed a large number of tasks of widely varying difficulty. Test takers are

asked to attempt randomly chosen subsets of the tasks. Test results are then used to estimate proficiency levels for a population group of interest (the program and control groups in our case). A member of a group with a proficiency score of 290 on the prose scale, for example, has an 80 percent probability of correctly completing a prose task at the 290 difficulty level. This same person has a lower probability of completing more difficult tasks and a higher probability of completing less difficult ones.

To facilitate descriptions of the literacy scores of groups and cross-group comparisons, the ETS approach to assessment distinguishes five broad literacy levels. Scores below 225 represent the lowest level of proficiency. Tasks in this range include locating a piece of information in a simple form or document. Scores between 225 and 275 represent Level II proficiency. Level II tasks include locating a piece of information in a more complex document with a distractor or performing a simple calculation with numbers easily found in a document. Scores between 275 and 325 represent Level III, scores between 325 and 375 represent Level IV, and scores between 375 and 500 represent Level V. A score in Level V indicates advanced skills in performing a variety of tasks that involve the use of complex documents.

LITERACY SKILLS OF ELIGIBLE JOB CORPS APPLICANTS COMPARED TO THOSE OF OTHER YOUNG ADULTS

The typical youth served by Job Corps has lower functional literacy scores than the typical young adult in the U.S., especially in the quantitative literacy domain. The average proficiency scores of control group members were 248 for prose, 256 for document, and 231 for quantitative literacy. In comparison, young adults nationally averaged 280 points on the prose and document scales and 277 on the quantitative scale, as reported in the NALS. While 14 percent of young adults nationally performed at the lowest skill level in prose and in document literacy, 28 percent of Job Corps control group members scored at that level on prose and 20 percent on document literacy. For quantitative literacy the gap was considerably wider. About 16 percent of young adults nationally scored in Level I on the quantitative scale, compared to 44 percent of Job Corps control group members.

Part of the reason for these deficits is that Job Corps applicants have considerably lower levels of educational attainment than the general population. Looking within education level, the Job Corps sample more nearly resembled the NALS sample in literacy proficiency. Remaining differences could be explained by a variety of factors. For example, high school dropouts who apply to Job Corps are more skilled in nonquantitative areas than dropouts in the general population, possibly reflecting a greater level of ability among dropouts who apply to Job Corps than dropouts generally.

IMPACTS ON LITERACY SKILLS

The impacts of Job Corps on participants' functional literacy skills were positive in all three domains (see Table 1). Job Corps raised participants' average test scores by about 4 points on the prose scale, 2 points on the document scale, and 5 points on the quantitative scale. The impacts on prose and quantitative literacy are statistically significant (different from zero) at the 10 percent level.

Program impacts on tests scores are often expressed as effect sizes or changes in percentile ranking among the full population. The effect size is the fraction of a standard deviation, which in the case of the literacy assessment used in this study is about 40 points. The effect sizes of the estimated impacts on Job Corps participants are about 0.09 for prose literacy, 0.04 for document literacy, and 0.10 for quantitative literacy. Taking someone who scores at the 50th percentile of a distribution, these effect sizes correspond to increases to the 53rd, 52nd, and 54th percentile of that distribution.

In terms of discrete proficiency levels, Job Corps moved some participants out of Level I. About 3 percent of participants are estimated to have moved out of the lowest proficiency level on the prose scale (2 percent on the document scale and 5 percent on the quantitative scale) and a comparable fraction into Levels II and III. Again, this effect is statistically significant for prose and quantitative literacy, but not for document literacy. Very few members of the program or control groups scored in the top two proficiency levels in any of the three literacy scales.

Positive impacts were found broadly across most key subgroups of students. Nearly all of the impacts estimated at the subgroup level were between 3 and 6 points per Job Corps program participant, although most are not statistically significant because of small sample sizes. Estimated impacts, however, were somewhat larger for older applicants who did not have a GED or high school diploma at random assignment--a group with particularly low skills. Impacts for this group were 6 points on both the prose and document literacy scales, and 11 points on the quantitative scale.

INTERPRETATION OF THE IMPACT FINDINGS

In order to interpret the literacy impact findings, we examined the extent to which the estimated impacts on literacy skills are consistent with our impact findings on other key outcomes that are associated with basic skills. We analyzed the statistical relationships among literacy scores, educational experiences, and labor market experiences for our basic skills sample by combining information on their test scores with information from their baseline and follow-up interviews. We then tried to reconcile the findings.

The analysis relied on a simplified path model of the relationship between family background, schooling, literacy skills, and labor market outcomes. Literacy is considered to be an intermediate (mediating) outcome that is affected by schooling *and* work experience, and that affects later outcomes such as earnings.

TABLE 1

IMPACTS ON AVERAGE LITERACY SCORES

Literacy Domain	Program Group	Control Group	Estimated Impact per Eligible Applicant[a]	Estimated Impact per Participant[b]
Prose	251.0	248.3	2.7*	3.7*
Document	257.6	256.4	1.1	1.6
Quantitative	234.8	231.2	3.6*	4.9*
Sample Size	**1,117**	**1,156**		

SOURCE: Job Corps Literacy Assessment data.

NOTE: All estimates, including program group means, control group means, participant means, and impacts, are regression adjusted.

[a]Estimated impacts for eligible applicants are measured as the difference between the regression-adjusted means for program and control group members.

[b]Estimated impacts for Job Corps participants are measured as the estimated impacts for eligible applicants divided by the proportion of program group members who enrolled in Job Corps. Standard errors for these estimates were inflated to account for the estimation error in the Job Corps participation rate.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

Results of this analysis suggest that the impacts on test scores are broadly consistent with what one might expect on the basis of the schooling and employment experiences of our sample members. Although the estimated impacts on literacy skills appear small relative to the impact on time spent in education and training programs--which was equivalent to about one school year--the two sets of findings are in fact very consistent. The positive impacts on time spent in education and training programs led to gains of about 5 points in the test scores of the program group relative to those of the control group. However, the control group worked more during the 30-month follow-up period. Because work experience appears to improve skills, the greater amount of work by the control group partially offset the gains of the program group due to more hours of schooling. Thus, these two factors combined led to implied program and control group differences on literacy skills that were similar to the observed impacts on literacy skills. The impacts on test scores are also consistent with the large impacts that we found on the attainment of a GED certificate.

We find also that the short-term impacts on earnings were larger than can be explained by the impacts on literacy skills alone, because the association between literacy scores and earnings is very modest within the limited range of literacy skills observed in our sample. Hence, the earnings gains were likely to have been due to other factors influenced by Job Corps that are not captured in the test scores. These factors might include impacts on vocational skills for a specific job that are not captured in the literacy test, improvements in social skills and attitudes about work, and credentialing effects from obtaining a GED or vocational certificate.

I. INTRODUCTION

Job Corps plays a central role in federal efforts to provide employment assistance to disadvantaged youths ages 16 to 24. The program's goal is to help disadvantaged youths become "more responsible, employable, and productive citizens" by providing comprehensive services, including basic education, vocational skills training, counseling, and residential support. Each year, Job Corps serves more than 60,000 new enrollees and costs more than $1 billion.

The National Job Corps Study, funded by the U.S. Department of Labor (DOL), was designed to provide information about the effectiveness of Job Corps in attaining its goal.[1] The cornerstone of the study was the random assignment of all youths found eligible for Job Corps to either a program group or a control group. Program group members were permitted to enroll in Job Corps, and control group members were not (although they could enroll in other training or education programs). The research sample for the study consists of approximately 9,400 program group members and 6,000 control group members randomly selected from among nearly 81,000 eligible applicants nationwide. Sample intake occurred between November 1994 and February 1996.

This report presents estimates of the impacts of Job Corps on participants' literacy and numeracy skills needed to function in the workplace. It builds on the analysis and findings presented in our report on short-term impacts using interview data collected at baseline and at 12 and 30 months after random assignment (Schochet et al. 2000). That report found that Job Corps participation led to beneficial impacts on key outcomes closely related to basic educational skills. Job Corps led to large increases in the receipt of education and training services, as well as in the receipt of General Education Development (GED) and vocational certificates. In addition, the study found modest

[1]The study is being conducted by Mathematica Policy Research, Inc. (MPR) and its subcontractors, Battelle Memorial Institute and Decision Information Resources, Inc.

gains in earnings by the beginning of the third year after a youth was found eligible for Job Corps. Beneficial program impacts were found broadly across most subgroups of students.

In this report, we examine *directly* the extent to which Job Corps improves literacy and numeracy skills. The analysis was conducted using test score data on a randomly selected subsample of program and control group members. A version of the test developed by the Educational Testing Service (ETS) for the National Adult Literacy Survey (NALS) was used for the study. The tests were administered in person in conjunction with the 30-month follow-up interview. The analysis sample for the study contains 2,273 youths (1,117 program and 1,156 control group members). Program impacts were estimated by comparing the distribution of literacy assessment scores of program and control group members, for the full sample and for key subgroups.

In the rest of this chapter, we provide an overview of the Job Corps program, discuss key policy issues related to basic skills, and discuss study objectives.

A. OVERVIEW OF JOB CORPS

The Job Corps program, established by the Economic Opportunity Act of 1964, operated under provisions of the Job Training Partnership Act (JTPA) of 1982.[2] The operational structure of Job Corps is complex, with multiple levels of administrative accountability, several distinct program components, and numerous contractors and subcontractors. DOL administers Job Corps through a national office and nine regional offices. The national office establishes policy and requirements, develops curricula, and oversees major program initiatives. The regional offices procure and administer contracts and perform oversight activities, such as reviews of center performance.

[2]Beginning in July 2000, Job Corps will operate under provisions of the Workforce Investment Act (WIA) of 1998.

Through its regional offices, DOL uses a competitive bidding process to contract out center operations, recruiting and screening of new students, and placement of students in jobs or educational opportunities after they leave the program. At the time of the study, 80 centers were operated under such contracts. In addition, the U.S. Departments of Agriculture and of the Interior operated 30 centers, called Civilian Conservation Centers (CCCs), under interagency agreements with DOL.[3] Figure I.1 shows the location of the 110 Job Corps centers that were in operation at the time our program group members were enrolled and displays the nine Job Corps regions.

Next, we briefly outline the roles of the three main program elements.

1. Outreach and Admissions

Outreach and admissions (OA) agencies conduct recruitment and screening for Job Corps. Private nonprofit firms, private for-profit firms, state employment agencies, and the centers themselves hold contracts with DOL regional offices to perform outreach and admissions work. These agencies provide information to the public through outreach activities (for example, by placing advertisements and making presentations at schools), screen youths to ensure that they meet the eligibility criteria, assign youths to centers (when the regional office delegates this function), and arrange for transportation to centers.

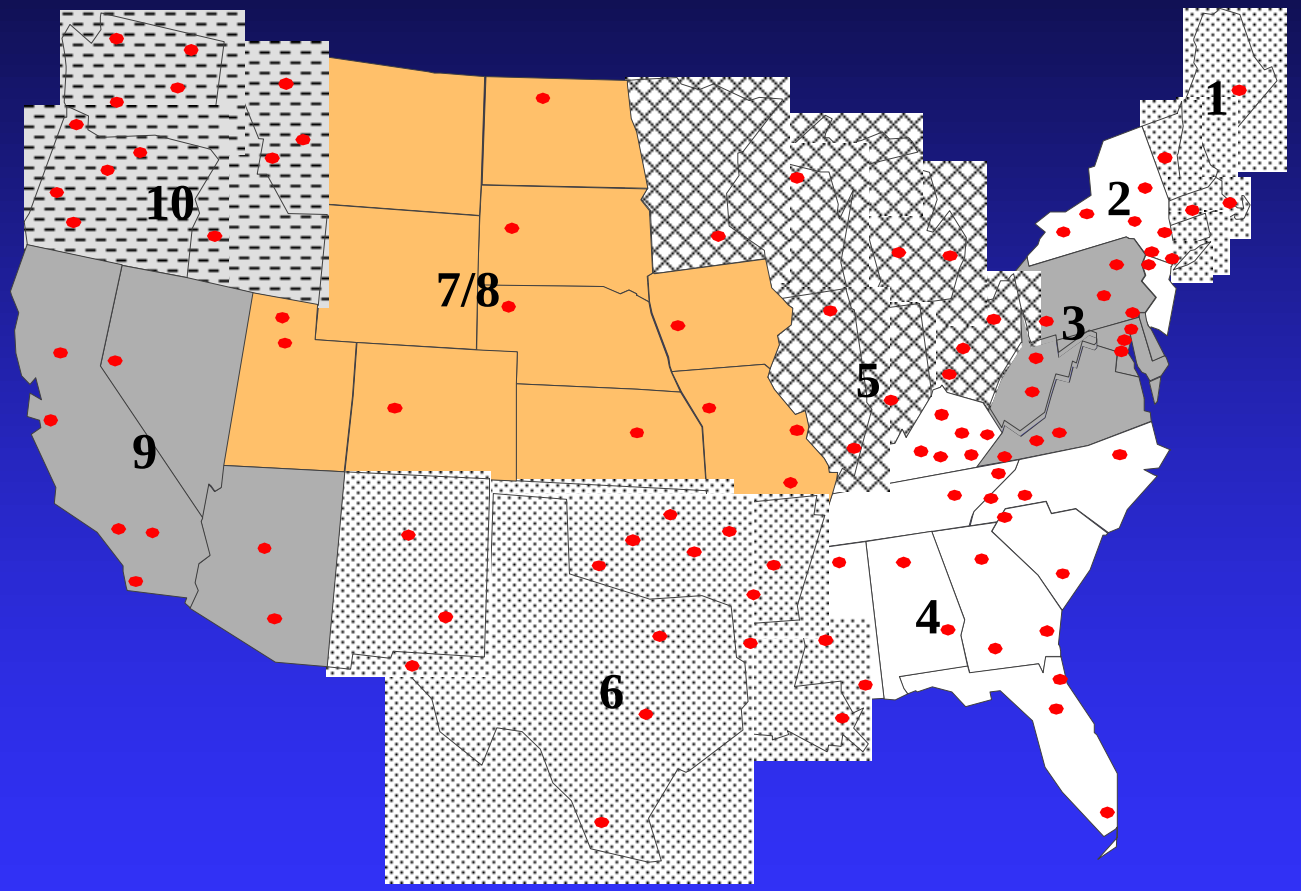
2. Job Corps Center Services

Job Corps is a comprehensive and intensive program. Its major components include basic education, vocational training, residential living (including training in social skills), health care and education, counseling, and job placement assistance. Services in each of these components are tailored to each participant.

[3]Currently, 88 contract centers and 28 CCCs are providing Job Corps training.

FIGURE I.1

JOB CORPS CENTERS IN PROGRAM YEAR 1995, BY REGION



• Indicates one of the 105 Job Corps Centers in the contiguous 48 States and the District of Columbia.

Education. The education component is designed to enable students to learn as fast as their individual abilities permit. Education programs in Job Corps are individualized and self-paced, and they operate on an open-entry and open-exit basis. The programs include remedial education (emphasizing reading and mathematics), world of work (including consumer education), driver education, home and family living, health education, English as a Second Language (ESL) programs, and a GED program of high school equivalency for academically qualified students. About one-fourth of the centers can grant state-recognized high school diplomas.

Vocational Training. The vocational training programs at Job Corps, like the education component, are individualized and self-paced, and operate on an open-entry and open-exit basis. Each Job Corps center offers training in several vocations, typically including business and clerical, health, construction, culinary arts, and building and apartment maintenance. National labor and business organizations provide vocational training at many centers through contracts with the Job Corps national office.

Residential Living. Residential living is the component that distinguishes Job Corps from other publicly funded employment and training programs. The idea behind residential living is that, because participants come from disadvantaged environments, they require new, more supportive surroundings to derive the maximum benefits from education and vocational training. All students must participate in formal social skills training. The residential living component also includes meals, dormitory life, entertainment, sports and recreation, center government, center maintenance, and other related activities. Historically, regulations had limited the number of nonresidential students to 10 percent, but Congress raised that limit to 20 percent in 1993.

Health Care and Education. Job Corps centers offer comprehensive health services to both residential and nonresidential students. Services include medical examinations and treatment;

biochemical tests for drug use, sexually transmitted diseases, and pregnancy; immunizations; dental examinations and treatment; counseling for emotional and other mental health problems; and instruction in basic hygiene, preventive medicine, and self-care.

Counseling and Other Ancillary Services. Job Corps centers provide counselors and residential advisers. These staff help students plan their educational and vocational curricula, offer motivation, and create a supportive environment. Support services are also provided during recruitment, placement, and the transition to regular life and jobs following participation in Job Corps.

3. Placement

The final step in the Job Corps program is placement, which helps students find jobs in training-related occupations with prospects for long-term employment and advancement. Placement contractors may be state employment offices or private contractors, and sometimes the centers themselves perform placement activities. Placement agencies help students find jobs by providing assistance with interviewing and resume writing and services for job development and referral. They are also responsible for distributing the readjustment allowance, a stipend students receive after leaving Job Corps.

B. POLICY ISSUES RELATED TO BASIC SKILLS AND STUDY OBJECTIVES

Policymakers and the general public have expressed mounting concern in recent years about the literacy and numeracy skills of the U.S. workforce. Attention has focused on the effects of declining skill levels on the nation's standard of living and its international competitiveness. Furthering knowledge about the skills required in the modern workplace, proposing acceptable skill levels, and suggesting ways to assess proficiency became explicit goals of DOL with the Secretary's Commission

on Achieving Necessary Skills (SCANS).[4] Improving the skills of youths who move through school into the workforce is an explicit goal of many DOL programs serving youths, including Job Corps.

Many Job Corps participants have poor basic skills. Nearly 80 percent have not completed high school at enrollment, and nearly half read at an eighth-grade level or lower. The Job Corps program takes steps to address the low levels of basic educational skills that participants bring to the program. Specifically, Job Corps provides basic education and job skills training, seeks to place graduates into jobs, and encourages graduates to acquire further education and training after Job Corps. Indeed, the Job Corps academic education component is an important program activity for all students who lack a high school credential and whose literacy or numeracy skills are below an 8.5 grade equivalency level.

At entry, students take the Tests of Adult Basic Education (TABE), including the total reading portion (reading comprehension and vocabulary) and the mathematics computation subtest. These tests are used to determine whether youths require additional academic training, as well as the types of classes they need. Students with low scores are assigned to the Reading Competencies and Math Competencies programs. Those with higher scores may enter classes that prepare them to pass the GED test. Those with high TABE scores who already have a GED certificate or high school diploma may not be required to enroll in the education component but may instead attend full-time vocational training.

Given the policy concerns surrounding educational skill levels and the Job Corps focus on improving skills, an important question is whether Job Corps improves participants' literacy and numeracy skills. The main Job Corps study examined the program's impacts on key areas closely related to basic educational skills using interview data--participants' receipt of education and training services (during both the in-program and postprogram periods), postprogram employment and

[4]U.S. Department of Labor (1991).

earnings, attainment of educational credentials (such as a GED certificate or high school diploma and a vocational certificate), and years of formal schooling. However, the telephone interviews with sample members (and in-person interviews with sample members not interviewed by telephone) did not provide the data necessary for directly measuring program impacts on literacy and numeracy skills. A study of impacts on literacy and numeracy skills requires that these types of skills be measured directly.

This report presents results for a substudy of the National Job Corps Study that directly measures the program's impacts on participants' literacy and numeracy skills. Chapter II describes the sample design, data, and analytic methods used for the analysis. Chapter III presents the impact findings. The final chapter discusses the interpretation of the impact results by comparing the impacts on literacy scores to impacts on other outcomes associated with literacy levels.

II. SAMPLE DESIGN, DATA, AND METHODS

The key features of the design for the study to measure Job Corps impacts on participants' literacy and numeracy skills are as follows:

- ***A single round of skills measurement was conducted in conjunction with the 30-month follow-up interview.*** Because of limited study resources, only a single round of testing could be conducted. We selected the 30-month measurement point under the assumption that the most important study goal is knowing whether and to what extent Job Corps produces differences after most program participants leave the program and spend some time in the workforce.

- ***The approach to literacy assessment developed by ETS was used for the literacy study.*** Specifically, a version of the instruments used by ETS in the NALS were used for the study. We also considered using the TABE, which the Job Corps academic education program uses, as the testing instrument for the study. However, the ETS approach was selected because it focuses on *functional* literacy and numeracy skills and has been used in several national studies with populations similar to that of Job Corps students.

- ***A total of 3,750 sample members (1,875 program group and 1,875 control group members) was selected for the study.*** The sample was selected from all program and control group members who were eligible for 30-month interviews and who were randomly assigned during the last 7 months of the 16-month sample intake period. The analysis sample contains 1,117 program and 1,156 control group members who completed the literacy test.

This chapter discusses key study design issues in more detail: the timing of skills measurement, the selection of an appropriate test, and the sample design. It also discusses interview response rates and the analytic methods that were used to estimate program impacts on literacy scores.

A. TIMING OF SKILLS MEASUREMENTS

The point of follow-up measurement was selected in light of the types of effects that Job Corps is expected to have on participants' literacy and numeracy skills. The most direct effect is likely to be improvements resulting from participation in the Job Corps academic education component.

Attending class and participating in class activities are known to improve students' skills measured by the TABE. On average, TABE reading scores improve by about one grade level, and math scores improve by about two grade levels, for students who remain in Job Corps for at least seven and a half months. These improvements are not measures of impacts, however, because we do not know what would have happened to reading and math skills had students not entered Job Corps. We also do not know whether the improvements are sustained after students leave Job Corps.

In addition to the direct effects of academic classes and other program activities, postprogram activities may affect participants' literacy and numeracy skills after they leave Job Corps. We hypothesized that some Job Corps graduates may undertake additional training or education, which may further improve their literacy skills. In addition, if employees use and improve their reading, writing, and calculating skills in the workplace, moving Job Corps students into good jobs may promote further improvements in these skills. We refer to potential literacy and numeracy skill improvements from these postprogram activities as "indirect effects" of the Job Corps program.

In light of these potential direct and indirect effects of Job Corps on students' literacy skills, we considered two options for a follow-up point to measure impacts on skills. One option was to measure these skills near the time when students left Job Corps. Skill measurement at this point would provide an indication of the direct effects of participating in the Job Corps academic education component. However, because no comparable point would exist for control group members, it would have been necessary to measure these short-term impacts through testing all sample members at some relatively short interval after sample intake (for example, at 9 to 12 months).[1] By itself, this

[1]The average participant in our program group reported staying on center for about eight months, although about one-quarter stayed for more than one year. Thus, conducting basic skills testing at 9 or 12 months would have required testing some sample members on center, which would have created some practical problems.

option would not show whether positive effects were maintained over time. In addition, the fact that Job Corps tests students in the education component periodically would possibly skew the literacy skills results, since the individuals in the program group would have had a recent opportunity to practice their test-taking skills.

The second option was to measure skills at a point after most students had left Job Corps--at 30 or 48 months after sample intake. Measures at one of these points would show the combined impacts of (1) the direct effects of academic education and other Job Corps services, and (2) the indirect effects that might be caused by differences in activities to build literacy skills that program and control groups pursued after Job Corps. Under this design, we would not be able to determine whether skill differences between the program and control groups were the result of academic instruction and other experiences in Job Corps or of differing postprogram employment and education experiences that resulted because of Job Corps.

Our design for the literacy study measured Job Corps' impacts on literacy skills at 30 months after random assignment. We selected this measurement point under the assumption that knowing whether and to what extent Job Corps produces differences in skill levels over the intermediate term is the most important study goal. Although measuring impacts at 30 months may miss some shorter-term impacts that do not persist over time, measuring temporary improvements in skills that do not persist in the intermediate or longer term is of questionable value.

It is important to note that the design--with a single postintervention measurement point and no baseline measurement--limits the questions the study can address. In particular, the design can measure *average* skill gains resulting from Job Corps. Because random assignment produced program and control groups with the same average skills levels at baseline (except for sampling error), program-control differences at followup show the *average* difference in gains. The design,

however, cannot directly address whether differences in literacy skill gains (that is, impacts) were larger for those with low or high skills at baseline (although we can estimate impacts for subgroups of youths defined in terms of baseline characteristics that are correlated with skill levels).

B. SELECTING AN APPROPRIATE TEST

This section first discusses the criteria that were used to select an appropriate testing instrument for measuring the impacts of Job Corps on participants' literacy skills. It then describes the features of functional literacy profiling and the instrument used in the NALS study, which forms the basis for the instrument used in our literacy study.

1. Criteria for Selecting a Test

The following criteria were used to select an appropriate testing instrument for measuring the impacts of Job Corps on participants' literacy skills:

The test should focus on functional workplace skills rather than on academic skills. Job Corps' mission is to prepare its students for a job or for further education that will lead to a job. Thus, it is appropriate to measure the extent to which Job Corps improves basic skills by comparing how well the program and control group perform literacy and numeracy tasks called for in the workplace. Although Job Corps offers an academic education component whose main goal is to improve academic skills, the ability to perform job-related and everyday tasks is a more meaningful measure of the program's long-term effects on students' skills.

A related concern is that participation in academic classes tends to improve performance on tests of basic academic skills and that some degradation in performance can be expected after students stop attending classes. This tendency creates the potential for differences in test performance between the program and control groups solely because program group members will have been more likely than

control group members to have recently attended academic classes. Short-term improvements in academic skills will generate few benefits if they are not accompanied by longer-term improvements in the performance of reading, writing, and mathematics tasks in the workplace and daily life.

The test should accommodate diverse skill levels. Job Corps serves youths with a wide range of skills. For example, more than 18 percent of sample members reported having a high school diploma at the baseline interview, and more than 16 percent reported that they had not completed the ninth grade. Similarly, the median entry-level total reading score on the TABE is an eighth-grade equivalent level. However, about one-fourth of students read at the 5th-grade level or lower, and one-fourth read at the 11th-grade level or higher. To provide a valid measure of average skill levels, the test selected must provide valid results for students with high and low skill levels.

The test should be administered by field interviewers in one 60-minute session to minimize nonresponse. The National Job Corps Study sample is dispersed across the country. With small numbers of sample members in any location, it was not feasible to administer the tests in a group setting. Accordingly, tests needed to be administered by trained interviewers in one-on-one sessions with individual sample members. The amount of time required to take the test was a major consideration in selecting a test. Longer tests are likely to have higher refusal rates and higher rates of noncompletion or partial completion. In our judgment, a test that required more than about one hour to administer would create difficulties in securing and maintaining participants' cooperation.

The test should support comparisons between the National Job Corps Study sample and other well-defined population groups. Use of a basic skills test that has been administered to a nationally representative sample of young adults allows us to compare the skills of former Job Corps students and control group members with the skills of other young adults.

We considered the merits of several candidate tests for the literacy skills study using these criteria and concluded that the instrument used in the NALS study would best meet study objectives. The main alternative to the NALS that we considered was the TABE, which the Job Corps academic education program uses. We rejected the TABE, however, for three main reasons.

First, unlike the NALS, the TABE does not focus on functional and workplace literacy skills. It focuses instead on achievement in reading, mathematics, language, and spelling.

Second, the TABE takes too long to administer. The entire TABE battery, including the locator, requires nearly six hours of testing time (excluding the time required to score the locator and determine the appropriate test level). Testing time can be reduced by using a subset of the tests or by using only the locator and the TABE survey form. However, the locator plus the survey form still requires approximately 90 minutes of testing time. In contrast, the NALS requires about 45 minutes for the testing portion plus 15 minutes for the background information portion.

Finally, the TABE does not support comparisons with nationally representative samples of youths similar to the Job Corps population. The TABE was calibrated with a sample of individuals drawn from adult basic education programs, vocational-technical schools, juvenile correctional institutions, and adult correctional institutions. Scores show achievement of the examinee relative to the norming group. However, the norming sample was not selected scientifically, and the norming data are not representative of any population beyond the norming sample. Consequently, comparisons with nationally representative samples are not possible. In contrast, the NALS, conducted in 1992, included a nationally representative sample of 13,600 adults plus a sample of

1,100 inmates in federal and state prisons. In addition, 14 states provided funding to increase the sample size in their states to 1,000 people. About 26,000 people completed the NALS assessment.[2]

We also considered and eliminated several other tests. We eliminated the Adult Basic Learning Examination (ABLE), Basic English Skills Test (BEST), and the California Student Assessment System (CASAS) from further consideration because they do not support comparisons with other well-defined population groups. We eliminated the General Aptitude Test Battery (GATB) and Armed Forces Qualifying Tests (AFQT) because they are designed to predict how well a person will perform in specific jobs and in the military, respectively. We eliminated the National Educational Longitudinal Study (NELS) because it is designed to test the academic skills of students and is not appropriate for young adults who are out of school. Finally, we eliminated two other ETS literacy tests similar to the NALS instrument: the Test of Applied Literacy Skills (TALS) because it requires two and a half hours to administer, and the test used in the DOL Workplace Literacy Study (WLT) because it has never been used in a field setting.

In sum, we selected the instrument used in the NALS study because it focuses on functional literacy skills, rather than academic skills, and provides the capability to test persons with different skill levels within a field interview setting. In addition, the use of the NALS-based assessment allows us to compare the skill levels of youths served by Job Corps with those of the broader youth population.

---

[2]In addition, other functional literacy tests developed by ETS and that are similar to NALS have been used in several large studies that include disadvantaged young adults. For example, the National Assessment of Educational Progress (1986) assessed a nationally representative sample of people ages 21 to 25 in 1985. The DOL Workplace Literacy Study (1990), conducted from November 1989 to June 1990, included a nationally representative sample of first-time JTPA applicants determined eligible for JTPA (2,500 applicants selected from 14 states) and a separate nationally representative sample of people who were served by state employment service (ES) offices or had filed an unemployment insurance (UI) claim (3,277 individuals selected from the same 14 states).

2. Features of Functional Literacy Profiling and the NALS Instrument

Literacy assessments based on the TABE and other standardized tests assume that people who read below a particular grade level lack the literacy and numeracy skills necessary to function in mainstream adult life. A problem with this approach is that it is based on performance on multiple-choice tests in a school setting. Yet, the types of literacy materials that adults encounter in daily life are often very different from those found in a school setting. Competency-based literacy tests use tasks that are more like the tasks adults perform at home and on the job.

The profile approach to literacy assessment developed by ETS measures the ability to perform a wide range of information-processing tasks that adults encounter in everyday life. The profile approach posits three dimensions of literacy:

1. ***Prose literacy***--the knowledge and skills necessary to understand and use information from texts

2. ***Document literacy***--the knowledge and skills necessary to locate and use information in tables, charts, graphs, and maps

3. ***Quantitative literacy***--the knowledge and skills necessary to perform different arithmetic operations using information embedded in prose and document materials

Proficiency in each of these domains is measured on a scale from 0 to 500. To estimate proficiency, ETS developed a large number of tasks of widely varying difficulty. Test takers are asked to attempt randomly chosen subsets of the tasks. Test results are then used to estimate proficiency levels for a population group of interest (the program and control groups in our case) using item response theory (IRT; see Appendix B for further details). A member of a group with a proficiency score of 290 on the prose scale, for example, has an 80 percent probability of correctly

completing a prose task at the 290 difficulty level. This same person has a lower probability of completing more difficult tasks and a higher probability of completing less difficult ones.[3]

To facilitate descriptions of the literacy levels of groups and cross-group comparisons, the profiling approach to assessment distinguishes five broad literacy levels. Table II.1 shows the levels, the proficiency score that places an individual at each level, and a description of the knowledge and skills demonstrated at each level.

ETS developed the profiling approach as a method for assessing the average skill levels *of groups* within the population. A balanced incomplete block (BIB) spiraling sample design is used, where each respondent completes a randomly selected subset of a large set of tasks. The testing is structured to provide accurate estimates of average skills within a group, but it does not provide reliable and valid information about the skills of *individuals*. The NAEP (1986), the DOL Workplace Literacy Study (1990), and the NALS (1992) have all used this approach.

C. SAMPLE DESIGN

The sample selected for the literacy skills study consists of 1,875 program group and 1,875 control group members. These youths were randomly selected from 7,348 sample members (4,478 program group and 2,870 control group members), satisfying two criteria. First, we selected youths who were eligible for 30-month follow-up interviews, because skills testing was conducted in conjunction with the 30-month follow-up interviews.[4] Second, we selected youths who were randomly assigned in August 1995 or later (and, thus, who became eligible for 30-month interviews

[3]ETS designed the scale so that when a test taker's proficiency equals the item's difficulty, the item will be completed successfully 80 percent of the time. This RP80 criterion is discussed more fully in Kirsch et al. 1993.

[4]Youths were eligible for 30-month interviews if they completed either the baseline interview or the 12-month follow-up interview (see Schochet 2000 for more details).

TABLE II.1

DESCRIPTION OF PROSE, DOCUMENT, AND QUANTITATIVE LITERACY TASKS
AT EACH OF FIVE LEVELS

Level (Score)	Prose Tasks	Document Tasks	Quantitative Tasks
Level 1 (0 to 225)	Tasks at this level are the least demanding in terms of what the reader must do to produce a correct response. Typically, these tasks require the reader to locate one piece of information through a literal match between the question and the stimulus material. If a distractor or plausible right answer is present, it tends to be located away from where the correct information is found.	Tasks at this level are the least demanding. They require the reader to either locate a piece of information based on a literal match or to enter information from personal knowledge.	Tasks at this level require a single, relatively simple operation for which the numbers are given and the arithmetic operation specified.
Level 2 (226 to 275)	Some tasks at this level require the reader to locate a single piece of information; however, these tasks tend to involve several distractors or a match based on low-level inferences. Tasks at this level also begin to require the reader to integrate information by pulling together two or more pieces or by comparing and contrasting information.	Tasks at this level begin to become more varied. Some still require the reader to match a single piece of information; however, some tasks involve several distractors or a match based on low-level inferences. Tasks at this level also begin to require the reader to integrate information.	Tasks at this level require the use of a single operation involving numbers that are either stated in the question or easily located in the material. In addition, the operation needed is either stated in the question or easily determined on the basis of the format of the problem--for example, entries on a bank deposit slip or order form.
Level 3 (276 to 325)	Tasks at this level require the reader to search fairly dense text for literal or synonymous matches on the basis of more than one feature of information, or to integrate information from a relatively long text passage that does not contain organizational aids such as headings.	Tasks at this level require the reader to either integrate three pieces of information or to find relevant information in rather complex tables or graphs in which distractors are present.	In tasks at this level, two or more numbers needed to solve the problem must be found in the stimulus material. Also, the operations needed can be determined from arithmetic relation terms.
Level 4 (326 to 375)	Tasks at this level continue to demand more from the reader. Not only are multiple-feature matching and integration of information from complex displays included, but the degree to which readers must draw inferences also increases. Conditional information that must be considered is frequently present in tasks at this level.	Tasks at this level continue to demand more from the reader. Multiple-feature matching and integration of information are included, and the degree to which readers must draw inferences increases. Tasks often require the reader to make five or more responses with no designation of the correct number of responses. Conditional information is also present and must be taken into account.	Tasks at this level require two or more sequential operations or the application of a single operation in which the quantities must be located in complex displays and/or the operations must be inferred from verbal information given or prior knowledge.
Level 5 (376 to 500)	At this level, tasks require the reader to search for information in dense text containing plausible distractors, to make high-level text-based inferences or use specialized background knowledge, and to compare and contrast sometimes complex information.	Tasks at this level require the most from the reader. The reader must search through complex displays containing multiple distractors, make high-level text-based inferences, or use specialized knowledge.	Quantitative tasks at this level are the most demanding. They require the reader to perform multiple operations and to extract salient features of a problem from stimulus material, or to rely on background knowledge to determine the quantities or operations needed.

SOURCE: Descriptions are adapted from *Beyond the School Doors: The Literacy Needs of Job Seekers Served by the U.S. Department of Labor*.

starting in February 1998). Thus, our impact estimates pertain *only* to the population of eligible program applicants who were randomly assigned during the last 7 months of the 16-month sample intake period for the main study. We judged that a shorter, more concentrated fielding period would reduce data collection costs and that the later cohort would be sufficiently representative of those in the full study population to meet the objectives of the basic skills study.

We adopted several design features to maximize the precision of the impact estimates given our sample size of 3,750 youths (which we judged was adequate to detect relatively modest impacts on average literacy skills).[5] First, equal numbers of program and control group members were selected for the study even though the research sample for the main study contains more program group members. Second, we set sampling rates to the basic skills sample that "undid" the oversampling of certain youths to the initial research sample (and, hence, to the 30-month follow-up sample). Thus, the literacy skills sample is nearly a simple random sample (that is, a self-weighting sample) of eligible applicants in the sample frame.[6]

Youths were selected for testing *before* the basic 30-month follow-up interviewing effort began. Youths in the basic skills sample who completed the 30-month follow-up interview (either by telephone or in person) were asked to participate in the literacy skills study. A testing date was scheduled for those who agreed to participate. The test was administered in person, and test respondents received a $25 incentive payment in addition to the $10 incentive payment for the basic 30-month follow-up interview. Testing started in February 1998 and ended in February 1999.

---

[5]See the discussion of sample precision in Section D of this chapter.

[6]The sample is not completely self-weighting because only youths in selected areas were eligible for baseline interviews after 45 days after random assignment; youths in the nonselected areas who did not complete baseline interviews within 45 days after random assignment were not eligible for follow-up interviews (see Schochet 2000 for further details). It was not possible to "undo" this clustering when selecting the sample for the literacy skills study. Thus, we constructed sample weights that account for the small clustered portion of the sample.

Youths were randomly administered one of seven tests. Each test contained 37 to 40 tasks across the three literacy domains, and these tasks were selected from the full set of 82 tasks. There were six core tasks that were contained in all tests. All youths were administered a baseline form that was used for scoring.

ETS scored the tests using IRT scaling (see Appendix B). Because each respondent completed only a randomly selected subset of tasks, it was not possible to obtain reliable literacy scores for each youth. Instead, the scaling procedure estimated the distribution (mean and variance) of each individual's ability level in each of the three literacy domains (prose, document, and quantitative). The procedure took into account an individual's raw score, question difficulty, and the scores of those with similar characteristics to the individual. ETS simulated five plausible values from each youth's ability distribution in each domain. These plausible values were used in the impact analysis.

D. INTERVIEW RESPONSE RATES AND ITEM NONRESPONSE

The overall (weighted) proportion of those in the study sample who completed at least some portion of the literacy test was 60.2 percent, and was similar for the program group (59.6 percent) and the control group (60.8 percent). Interviews were completed with 2,273 of the 3,750 youths eligible for skills testing. The response rate was 73.3 percent among those who completed the main 30-month follow-up interview.[7] The average interview was completed about 36 months after random assignment.

The response rates differed somewhat across some key subgroups (Table II.2). For example, the response rate was higher for females than males (65 percent, compared to 57 percent) and for younger sample members than older ones (62 percent for those 16 and 17 years old,

[7]There are 34 youths who completed a literacy skills test but who did not complete the 30-month follow-up interview, because the testing period lasted slightly longer than the main survey in order to increase response rates for the literacy study.

TABLE II.2

RESPONSE RATES TO THE LITERACY STUDY, BY RESEARCH STATUS AND KEY SUBGROUP
(Percentages)

Subgroup	Program Group	Control Group	Combined Sample
Full Sample	59.6	60.8	60.2
Gender			
Male	56.5	57.0	56.8
Female	64.1	66.6	65.3
Age at Application			
16 to 17	60.5	62.6	61.6
18 to 19	60.0	61.6	60.8
20 to 21	56.2	59.7	58.0
22 to 24	59.7	52.3	56.2
Region			
1	52.8	53.5	53.1
2	38.6	50.7	44.9
3	57.4	59.5	58.5
4	62.9	62.6	62.7
5	64.9	65.3	65.1
6	59.0	63.3	61.1
7/8	67.3	64.0	65.7
9	55.2	51.5	53.3
10	67.1	75.2	71.4
Race/Ethnicity			
White, non-Hispanic	57.3	59.9	58.6
Black, non-Hispanic	62.6	64.4	63.5
Hispanic	55.8	52.8	54.3
Other	52.5	53.0	52.8
Education			
Completed 12th grade	62.0	63.8	62.9
Did not complete 12th grade	58.7	60.1	59.4

TABLE II.2 *(continued)*

Subgroup	Program Group	Control Group	Combined Sample
Convictions			
Ever convicted or adjudged delinquent	59.3	60.6	60.0
Never convicted or adjudged delinquent	58.9	60.6	59.8
Residential Designation Status			
Resident	59.0	60.1	59.6
Nonresident	63.0	64.7	63.9
Sample Size[a]	**1,875**	**1,875**	**3,750**

SOURCE: Literacy assessment test score data and program intake (ETA-652) data.

[a]Sample includes those selected for the basic skills study.

compared to 57 percent for those 20 and older). The response rates across regions ranged from 45 percent in Region 2 to 71 percent in Region 10 and were larger in rural areas than in urban areas. Because of these subgroup differences, the sample weights were adjusted to help reduce the potential bias in the impact estimates due to interview nonresponse using the methods described in Appendix A.

For the overall sample of 1,117 program group and 1,156 control group members, the minimum detectable impact on scaled literacy scores is 5.3 scaled score points.[8] This means that there is a high probability that we would find a statistically significant impact on document, prose, and quantitative scores if the actual program impact per participant were 5.3 points or more. Because the range of scores within a NALS level is typically 50 points (for example, Level 2 scores range from 225 to 275), the minimum detectable impact for the total sample corresponds to roughly one-tenth of a NALS level increase. These detectable impacts are reasonable to expect on the basis of typical TABE reading and math score gains for Job Corps participants while in the program and on observed correlations between literacy test scores and earnings (DOL 1993).

Respondents attempted most of the literacy and numeracy tasks that they were asked to perform. On average, respondents attempted 85 percent of the tasks. About one-quarter of respondents attempted at least 97 percent of the tasks, and only about 5 percent of the sample attempted less than one-half of the tasks. These figures are similar for the program and control groups.

[8]This figure is calculated using a one-tailed test with a confidence level of 95 percent at 80 percent power and a standard deviation of test scores of 40 as determined from our data. The variance of the impact estimates is assumed to be reduced by 7.8 percent owing to the use of regression models (that is, assuming an R^2 value of .15). Because 73 percent of program group members enrolled in Job Corps, detectable impacts were inflated by (1/.73) to obtain impacts for program participants.

E. ANALYTIC METHODS

The analytic methods used to estimate program impacts on literacy assessment test scores for the full sample and for key subgroups are similar to those described in our main impact report. In this section, we summarize the main features of our approach.

The random assignment design ensures that no systematic observable or unobservable differences between program and control group members existed at the point of random assignment, except for the opportunity to enroll in Job Corps. Thus, unbiased estimates of program impacts for *eligible applicants* were estimated as the simple differences in the distributions of test scores between program and control group members. The statistical significance of the impact estimates was gauged using t-tests (for differences in average scores) and chi-squared tests (for differences in the distribution of scores across the five levels of proficiency discussed earlier). Sample weights were used in the calculations to account for the sample design and interview nonresponse.

We estimated impacts in each literacy domain using the first of the five plausible values provided by ETS. We repeated the analysis using the mean of the five values and each of the other four plausible values separately to test the robustness of study findings. The variance of each estimate took into account the variance of test scores across individuals, as well as the variance of the mean scores across the five plausible values (see Appendix A).

Impacts per *participant* were estimated by dividing the impacts per eligible applicant by the proportion of program group members who enrolled in Job Corps (73 percent). This approach yields unbiased estimates of program impacts under the assumption that Job Corps has no effect on those who do not enroll in the program. The impact estimates per participant should be interpreted as the distribution of test scores that Job Corps participants would have had if they had not enrolled in the program. We present impacts per eligible applicant and impacts per participant.

We also estimated "regression-adjusted" impact estimates using multivariate models that control for other factors that affect test scores. The multivariate regression approach increases the precision of the estimated program impacts and the power of significance tests relative to the differences-in-means approach. In addition, the use of multivariate models adjusts for any random residual differences in the observable baseline characteristics of program and control group members. The control variables used in the regression models were constructed using baseline interview data. We selected control variables that had predictive power in the regression models and that were correlated with research status.

We used different estimation techniques depending on the nature of the skills measured. We estimated impacts on average test scores using ordinary least squares (OLS) procedures and estimated impacts on the overall distribution of test scores across the five NALS proficiency levels using ordered logit regression procedures.[9] We used a two-stage instrumental variables procedure to estimate regression-adjusted impacts for *participants*.[10]

[9]Our main approach was to estimate regression models without using sample weights because there is no theoretical reason to estimate weighted regressions when unequal weighting of the sample is due to exogenous factors. Instead, we included as control variables in the regression models variables that were used to construct the weights. We estimated weighted regression models to check the robustness of study findings and found that the unweighted and weighted results were nearly identical.

[10]In the first stage, we estimated a logit model where the probability that a program group member enrolled in Job Corps was regressed on an indicator variable equal to 1 if the youth was in the program group and zero otherwise, as well as on other control variables. In the second stage, we regressed literacy scores on the predicted values from this logit model and other control variables. The parameter estimates on these predicted values represent impacts per participant on literacy scores. The standard errors of the estimates were inflated for the first-stage estimation error.

The impact estimates using the differences-in-means and regression approaches are very similar. We present the regression-adjusted results in the main body of this report and present the differences-in-means results in Appendix Table C.2.[11]

The impact findings are also very similar using scores measured as the mean of the five plausible values and using scores measured as any of the single plausible values. We present the impact findings using the first plausible value in each domain.

We present impacts for the overall sample, as well as for the following key subgroups of youths defined by their characteristics at baseline: gender and the presence of children (males, females with children, and females without children); age and educational attainment (16 to 17, 18 to 24 without a high school credential, and 18 to 24 with a high school credential); and residential designation status (residents and nonresidents). We selected these subgroups on the basis of the impact findings presented in our main report on short term impacts. Regression-adjusted impact estimates for a subgroup were obtained by including as control variables terms formed by interacting a binary indicator of program group research status with subgroup indicator variables. We estimated a separate regression model for each of the three categories of subgroups. We conducted statistical tests to gauge the statistical significance of the impacts for each subgroup and whether impacts differed across levels of a subgroup (for example, for males and females).

[11]In the main impact report, we focused on impacts estimated using the differences-in-means approach because of the large number of weighting cells, which complicated the estimation of the regression-adjusted impacts. However, we focus on the regression-adjusted estimates for the literacy skills study because the literacy skills sample is nearly self-weighting and because it is important to increase the precision of the impact estimates because of relatively small sample sizes.

III. IMPACTS ON LITERACY SKILLS

This chapter presents the impacts of Job Corps on each of the three domains of literacy skills discussed in the previous chapter: prose literacy, document literacy, and quantitative literacy. In the first section, we present a descriptive analysis, comparing the literacy levels of our sample to a nationally representative sample. The next section presents impact findings by comparing the mean test scores and distribution of test scores of program group and control group members. The final section presents estimates of impacts on literacy for key subgroups.

A. COMPARING THE LITERACY SKILLS OF ELIGIBLE JOB CORPS APPLICANTS TO THOSE OF OTHER YOUNG ADULTS

The test used in this study supports comparisons with several other recent adult literacy surveys, as discussed in Chapter II. Therefore, we are able to compare the proficiency levels of the National Job Corps Study sample with those of well defined population groups. Table III.1 compares the test score results from the Job Corps sample with those from the National Adult Literacy Survey (NALS), a nationally representative sample of adults, in each of the three literacy domains. The NALS test was nearly the same as the one used in the Job Corps study. For the Job Corps sample, we use only control group members. For the NALS sample, we present results for young adults aged 19 to 26, who represent roughly the same age group as the Job Corps sample when the testing occurred (about 36 months after random assignment).

To compare the two samples in each literacy domain, we tabulated the mean proficiency level, and the proportion of each group whose skills placed them in each of the five proficiency levels defined by ETS. As discussed in Chapter II, the literacy assessments were scored on a

TABLE III.1

LITERACY SKILLS OF ELIGIBLE JOB CORPS APPLICANTS COMPARED TO THOSE OF YOUNG ADULTS IN THE NALS

Proficiency Level	Job Corps[a]	NALS Young Adults[b]
Prose (Percentages)		
Level I	28	14
Level II	46	29
Level III	24	37
Level IV	2	18
Level V	0	2
Average Prose Proficiency	248	280
Document (Percentages)		
Level I	20	14
Level II	47	29
Level III	29	37
Level IV	3	18
Level V	0	2
Average Document Proficiency	256	280
Quantitative (Percentages)		
Level I	44	16
Level II	37	28
Level III	17	37
Level IV	2	16
Level V	0	2
Average Quantitative Proficiency	231	277
Sample Size	**1,156**	**3,344**

SOURCE: Kirsch et al. 1992; Job Corps literacy assessment test data.

NOTE: Level I scores are 225 or less, Level II scores are between 225 and 275, Level III scores are between 275 and 325, Level IV scores are between 325 and 375, and Level V scores are between 375 and 500.

[a]Job Corps averages pertain to the control group only and are weighted for nonresponse and the sample design.

[b]NALS young adults are ages 19 to 24.

scale from 0 to 500, representing degrees of proficiency along each dimension of literacy (prose, document, and quantitative). For example, a low score (below 225) on the document scale indicates that an individual has very limited skills in processing information from tables, charts, graphs, and maps, including those that are short and uncomplicated. On the other hand, a high score (above 375) indicates advanced skills in performing a variety of tasks that involve the use of complex documents (Kirsch et al 1993; p. 8). Level I on any of the scales is defined as having a score less than 225 points. Level II is defined as any score between 225 and 275. Level III is defined as any score between 275 and 325. Level IV is any score between 325 and 375. Scores higher than 375 are considered Level V proficiency.

Control group members in the Job Corps sample had literacy skills that averaged in the middle to low end of Level II (see Table III.1). On average, they scored highest in document literacy (256 points), somewhat lower in prose literacy (248 points), and lowest in quantitative literacy (231 points). On the prose literacy scale, 28 percent of Job Corps applicants scored at Level I and nearly half at Level II. The results were similar for the document scale, with 20 percent scoring at Level I and again, nearly half, at Level II. Job Corps applicants performed much lower, however, on the quantitative scale, where about 44 percent scored at Level I and 37 percent scored at Level II. This left only 24 percent, 29 percent, and 17 percent of Job Corps applicants scoring in Level III on the prose, document, and quantitative scales, respectively. Very few Job Corps applicants scored above Level III on any of the scales.

When compared to young adults in the NALS sample, these proficiency levels of Job Corps applicants were quite low in all three areas. The average proficiency score for young adults nationally was much higher than for Job Corps applicants and was nearly the same across the three domains--280 scale points in prose and document literacy and 277 points in quantitative literacy.

Thus, the typical young adult in the U.S. scored about 32, 24, and 46 scale points higher in the prose, document, and quantitative domains, respectively, than did the typical Job Corps applicant.

These results reflect the disadvantaged background of Job Corps applicants relative to the broader population of young adults. Over three quarters of NALS sample members had at least a high school diploma or GED on the day they were tested, compared to fewer than half of Job Corps control group members. Job Corps sample members were more than three times more likely to have a learning disability than NALS sample members, and were much more likely to be members of a minority racial or ethnic group. These findings are consistent with our earlier report on baseline characteristics (Schochet 1998) that showed Job Corps serves a disadvantaged population.

Since Job Corps applicants are educationally disadvantaged, we also compared them to NALS participants with similar educational attainment (Table III.2).[1] This shows, for example, how Job Corps applicants without a high school or a GED certificate ("dropouts") compare with dropouts nationally.

For the prose and document literacy scales, dropouts who applied to Job Corps had higher skills than dropouts nationally, a difference of 12 points on the prose scale and 27 points on the document scale. This may reflect a greater level of ability among dropouts who apply to Job Corps compared to all dropouts. On the other hand, high school graduates in the two samples had similar skill levels, but Job Corps applicants who went on to trade schools or two-year colleges had lower skills, 19 points lower on the prose scale and 12 points lower on the document scale, than their NALS counterparts. This indicates that Job Corps control group members who completed some

[1]These comparisons are based on the full sample of NALS participants, not just young adults, because the NALS report did not differentiate age categories within literacy profiles by educational attainment. However, comparisons of literacy skills between young adults and the full sample in the NALS report (Kirsch et al. 1993) suggest that skill differences by age group are small and likely due to differences in educational attainment.

TABLE III.2

LITERACY SKILLS OF ELIGIBLE JOB CORPS APPLICANTS COMPARED TO THOSE OF ADULTS IN THE NALS, BY EDUCATIONAL ATTAINMENT

	Educational Attainment[a]					
	No GED or High School Diploma[b]		GED or High School Diploma Only[b]		More Than High School or GED[b]	
Proficiency Level	Job Corps[c]	NALS	Job Corps[c]	NALS	Job Corps[c]	NALS
Prose (Percentages)						
Level I	36	46	17	15	16	8
Level II	49	36	47	37	25	23
Level III	15	15	32	38	49	45
Level IV	0	2	3	9	10	22
Level V	0	0	0	1	0	3
Average Prose Proficiency	237	225	260	269	275	294
Document (Percentages)						
Level I	26	50	14	19	10	9
Level II	52	35	42	40	40	27
Level III	22	14	38	33	37	42
Level IV	1	2	6	8	12	20
Level V	0	0	0	1	2	2
Average Document Proficiency	247	220	266	264	278	290
Quantitative (Percentages)						
Level I	57	49	29	17	16	8
Level II	33	32	44	35	41	23
Level III	10	16	23	36	34	42
Level IV	0	3	3	11	7	23
Level V	0	0	0	1	1	4
Average Quantitative Proficiency	217	220	246	269	266	295
Sample Size	**654**	**5,478**	**400**	**7,169**	**102**	**7,620**

SOURCE: Kirsch et al. 1993; Job Corps literacy assessment data.

NOTE: Level I scores are 225 or less, Level II scores are between 225 and 275, Level III scores are between 275 and 325, Level IV scores are between 325 and 375, and Level V scores are between 375 and 500.

[a] Educational attainment is based on self-reports on the day of the test.

[b] NALS data are adjusted so both samples have equal proportions completing grades 0 to 8, grades 9 to 11, grade 12, a GED, some college, and a two-year degree. Graduates of four-year colleges are excluded from both samples.

[c] Job Corps averages pertain to the control group only and are weighted for nonresponse and the sample design.

postsecondary school had somewhat lower skills than the average person with this level of schooling.[2]

Comparing the quantitative skills of Job Corps applicants with those of similarly educated adults in the NALS shows that the Job Corps population had a considerable disadvantage in this area. For all three education groups, the Job Corps control group members scored lower than their NALS counterparts. Consistent with the differences in prose and document literacy skills, the disadvantage was greater for those who attended school beyond high school and lesser for dropouts.

In sum, we find that Job Corps serves youths whose literacy and numeracy skills are lower than those of the average young adult in the U.S., because Job Corps serves youth with less education than the typical youth. The skills of Job Corps students are more similar when compared to the skills of young adults with similar educational levels. Even so, Job Corps applicants who are dropouts appear to have somewhat better skills than dropouts nationally, and Job Corps applicants who have a high school credential have somewhat lower quantitative skills than young adults with similar educational backgrounds nationally.

B. IMPACTS ON LITERACY SKILLS FOR THE FULL SAMPLE

The impacts of Job Corps on participants' literacy skills were estimated by comparing the test scores of program and control group members. For each of the three literacy scales, we compared the *average* skill levels and the *distribution* of skill levels for the two research groups. The skill distributions are described as the proportion of program and control group members whose scores placed them in each of the five discrete skill levels described in Chapter II.

---

[2]NALS data were adjusted so both samples have equal proportions completing grades 1 to 8, grades 9 to 11, grade 12, a GED certificate, some college, and a two-year degree. There were no graduates of four-year colleges in the Job Corps sample, so these individuals were excluded from the NALS summary presented here.

The impact findings are presented in Table III.3. All figures were regression-adjusted using the methods described in Chapter II. The full regression results are displayed in Appendix Table C.1, and results using the simple differences-in-means approach are presented in Table C.2. We present impacts per eligible applicant as well as impacts per Job Corps participant.

Job Corps raised participants' average test scores by about 4 points on the prose scale, 2 points on the document scale, and 5 points on the quantitative scale. The impacts on prose and quantitative literacy are statistically significant (different from zero) at the 10 percent level.[3]

Program impacts on tests scores are often expressed as effect sizes or changes in percentile ranking among the full population. The effect size is the fraction of a standard deviation, which in the case of the literacy assessment used in this study is about 40 points.[4] The effect sizes of the estimated impacts on Job Corps participants are about 0.09 for prose literacy, 0.04 for document literacy, and 0.10 for quantitative literacy. Taking someone who scores at the 50th percentile of a distribution, these effect sizes correspond to increases to the 53rd, 52nd, and 54th percentile of that distribution.

[3]Unfortunately, very little experimental evidence from similar programs exists that could serve as a benchmark for helping to assess the magnitude of the impacts and place them in perspective. The evaluation of California's Greater Avenues to Independence (GAIN) program (Martinson and Friedlander 1994) used a similar assessment to measure two-year impacts of adult basic education for mostly female single parents on welfare. The authors reported average impacts of -0.6 points on the document scale and 2.6 points on the quantitative scale, both of which were smaller than their Job Corps counterparts and neither of which was statistically significant. Despite the methodological similarities with this study, the GAIN evaluation examined a different intervention aimed at a very different population.

[4]The actual standard deviations were calculated separately for each literacy domain, using Job Corps control group members. They were 41.3, 40.1, and 47.1 points for prose, document, and quantitative literacy, respectively.

TABLE III.3

IMPACTS ON LITERACY SKILLS

Proficiency Level	Program Group	Control Group	Estimated Impact Per Eligible Applicants	Program Group Job Corps Participants	Estimated Impact Per Participants[b]
Prose (Percentages)					
Level I	24.3	26.8	-2.5*[c]	24.7	-3.4*[c]
Level II	48.2	48.2	0.0	48.4	0.0
Level III	24.2	22.2	2.1	23.8	2.8
Level IV	3.2	2.8	0.4	3.1	0.6
Level V	0.0	0.0	0.0	0.0	0.0
Average Prose Proficiency	251.0	248.3	2.7*	250.3	3.7*
Document (Percentages)					
Level I	18.9	20.0	-1.1	19.2	-1.5
Level II	47.4	47.8	-0.4	47.8	-0.5
Level III	29.3	28.1	1.2	28.9	1.7
Level IV	4.2	3.9	0.3	4.0	0.4
Level V	0.1	0.1	0.0	0.1	0.0
Average Prose Proficiency	257.6	256.4	1.1	256.9	1.6
Quantitative (Percentages)					
Level I	40.7	44.4	-3.7**[c]	41.2	-5.0**[c]
Level II	38.6	37.3	1.3	38.6	1.8
Level III	18.5	16.5	2.1	18.1	2.8
Level IV	1.9	1.6	0.3	1.8	0.4
Level V	0.2	0.2	0.0	0.2	0.1
Average Quantitative Proficiency	234.8	231.2	3.6*	234.0	4.9*
Sample Size	**1,117**	**1,156**		**818**	

SOURCE: Job Corps Literacy Assessment data.

NOTES: 1. All estimates, including program group means, control group means, participant means, and impacts, are regression adjusted.
2. Level I scores are 225 or less, Level II scores are between 225 and 275, Level III scores are between 275 and 325, Level IV scores are between 325 and 375, and Level V scores are between 375 and 500.

[a]Estimated impacts for eligible applicants are measured as the difference between the regression-adjusted means for program and control group members.

[b]Estimated impacts for Job Corps participants are measured as the estimated impacts for eligible applicants divided by the proportion of program group members who enrolled in Job Corps. Standard errors for these estimates were inflated to account for the estimation error in the Job Corps participation rate.

[c]The significance levels pertain to statistical tests for differences in the distribution of the outcome measure for the program and control group members.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

In terms of proficiency, Job Corps moved some participants out of Level I. About 3 percent of participants are estimated to have moved out of the lowest proficiency level on the prose scale (2 percent on the document scale and 5 percent on the quantitative scale) and a comparable fraction into Levels II and III. Again, this effect is statistically significant for prose and quantitative literacy, but not for document literacy. Very few members of the program or control groups scored in the top two proficiency levels in any of the three literacy scales.

The magnitude of these impacts is small in terms of the types of tasks that the average program and control group member could do. However, as discussed in the next chapter, the impact findings on literacy skills are consistent with the short-term impact findings on key outcomes associated with literacy levels that are reported in our main impact report.

C. IMPACTS ON LITERACY SKILLS FOR SUBGROUPS

The sample used in this literacy study is considerably smaller than that used for the overall Job Corps impact study, which limits our ability to make inferences about impacts for subgroups. It is useful, however, to present impacts for the main subgroups. From the 30-month follow-up of the full Job Corps study sample, we found that impacts on hours of education and training and on high school credential receipt were large across almost all subgroups. Similarly, impacts on earnings near the end of the observation period were positive for most subgroups. As expected, then, we found that the positive impacts on literacy test scores for the full sample were present for most subgroups as well. As for the overall sample, impacts were generally smallest within each subgroup on the document scale and largest on the quantitative scale (the domain in which sample members needed the most improvement). Consistent with this, the impacts tended to be larger for subgroups with lower proficiency levels, such as older applicants without a high school credential at baseline.

Table III.4 presents subgroup findings. It shows average scores for the program and control groups, the impacts per eligible applicant, and the impacts per participant for key subgroups defined by baseline characteristics: gender and children (males, females without children, and females with children), age and education (age 16 to 17, age 18 to 24 without a high school credential, and age 18 to 24 with a high school credential)[5], and residential status (residential and nonresidential program slot designees). All figures are regression adjusted. We also estimated the impacts on the distribution of test scores across proficiency levels. Results from these analyses are largely consistent with the impact findings on average test scores, so they are not presented here.

Control group members had fairly similar average literacy scores across subgroups defined by baseline characteristics. The only substantial differences are those between control group members with and without a high school credential at baseline. Older graduates (those with a high school diploma or GED certificate at baseline) scored an average of 264 points on the prose scale, 271 on document literacy, and 252 on quantitative literacy, which is roughly 20 points higher than the corresponding scores for the oldest nongraduates and the youngest sample members, almost all of whom were also nongraduates.

Although many of the impacts estimated at the subgroup level are not statistically significant, nearly all were between 3 and 6 points per Job Corps program participant. One exception was the older applicants who already had a GED or high school diploma when they were randomly assigned. Impacts for this group were 2 points on the prose scale and less than a point on the other two scales. On the other hand, the program had larger impacts on older dropouts, where impacts for participants were 6 points on both the prose and document literacy scales, and 11 points on the quantitative scale.

[5]Nearly all applicants age 16 or 17 at baseline had no high school credential at baseline time, so this age group is not disaggregated by educational attainment.

TABLE III.4

IMPACTS ON AVERAGE LITERACY SCORES, BY SUBGROUP

Subgroup	Program Group	Control Group	Impact Per Eligible Applicant[a]	Program Group Job Corps Participants	Impact Per Participant[b]
Gender and Children					
Males					
Prose	247.7	245.1	2.7	247.3	3.4
Document	255.0	253.0	2.0	254.8	2.5
Quantitative	233.3	228.9	4.5	233.0	5.7*
Females with Children					
Prose	253.4	250.8	2.6	252.4	4.4
Document	261.3	258.8	2.5	260.6	4.3
Quantitative	235.5	232.0	3.4	234.1	5.8
Females with No Children					
Prose	256.6	253.1	3.6	255.6	4.8
Document	261.7	261.6	0.0	260.6	0.1
Quantitative	238.4	235.0	3.4	237.0	4.6
Age Group and Education at Random Assignment					
Age 16-17					
Prose	245.2	242.3	3.0	244.8	3.8
Document	252.2	252.3	-0.1	251.8	-0.1
Quantitative	227.2	224.3	2.9	226.4	3.6
Age 18-24, No HS Credential					
Prose	249.5	245.3	4.2	249.4	6.0
Document	255.9	251.7	4.2	255.9	6.1
Quantitative	233.3	225.6	7.7**	233.4	11.1**
Age 18-24, with HS Credential					
Prose	265.7	264.3	1.4	264.3	2.0
Document	271.6	271.6	0.1	270.5	0.1
Quantitative	253.3	252.8	0.5	252.4	0.7
Residential Status					
Residential Designee					
Prose	250.2	248.1	2.1	249.4	2.8
Document*[c]	256.6	256.3	0.3	255.9	0.4
Quantitative*[c]	233.8	231.2	2.6	232.9	3.5

TABLE III.4 *(continued)*

Subgroup	Program Group	Control Group	Impact Per Eligible Applicant[a]	Program Group Job Corps Participants	Impact Per Participant[b]
Nonresidential Designee					
Prose	255.6	249.5	6.1*	255.6	9.7*
Document*[c]	262.8	257.1	5.8*	262.9	9.2*
Quantitative*[c]	240.1	231.0	9.2**	240.0	14.5**

SOURCE: Baseline interview data and Job Corps Literacy Assessment data.

NOTE: All estimates, including program and control group means, participant means, and impacts, are regression-adjusted.

[a] Estimated impacts for eligible applicants are measured as the difference between the regression-adjusted means for program and control group members.

[b] Estimated impacts for Job Corps participants are measured as the estimated impacts for eligible applicants divided by the proportion of program group members who enrolled in Job Corps. Standard errors for these estimates were inflated to account for the estimation error in the Job Corps participation rate.

[c] Asterisks next to an outcome indicate significance level of the statistical test for differences in impacts across levels of that subgroup.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

Impacts for nonresidential designees were also somewhat larger than average--10 points on the prose scale, 9 points on the document scale, and 15 points on the quantitative scale.

We conducted additional hypothesis tests to examine whether impacts differed across levels of the same subgroup, for example impacts on males compared to impacts on females.[6] We did not find evidence that any of the subgroup impact differences were statistically significant except for the impacts on residential designees and nonresidential designees, which were significant for document and quantitative literacy, but not for prose.

In sum, the limited evidence available for subgroups is consistent with findings reported elsewhere that program impacts were broadly distributed across the key subgroups, with some slightly higher or lower than others.

[6]These were F-tests performed on the constraint that all program group indicator by subgroup level interaction terms were equal.

IV. INTERPRETATION OF THE IMPACT FINDINGS

This section discusses the extent to which the estimated impacts on literacy skills are consistent with our impact findings on key outcomes associated with literacy skills. As described in our main impact report, the average Job Corps participant receives an additional year of education and training relative to what would have been received had the youth not enrolled in Job Corps. Job Corps participation also leads to very large gains in the attainment of GED and vocational certificates, and to modest gains in short-term postprogram earnings. There is evidence that higher educational levels are associated with higher literacy levels, and that higher literacy levels are in turn associated with better labor market outcomes in the full adult population (Kirsch et al. 1993) and in the Job Training Partnership Act (JTPA) and Unemployment Insurance (UI)/Employment Service (ES) populations (Kirsch et al. 1992). Thus, in this section, we treat literacy assessment scores as an intermediate outcome, and examine the relationship between the impact findings on literacy scores and the impact findings on education-related and labor market outcomes. The goal of this analysis is to provide information that can be used to help interpret our literacy skills impacts.

Our results suggest that the impacts on test scores are broadly consistent with what one might expect on the basis of the schooling and employment experiences of our sample members. Although the estimated impacts on literacy skills appear small relative to the impacts on the education-related outcomes, the two sets of findings are in fact very consistent because of the modest association between our human capital measures and test scores. The positive impacts on hours spent in academic and vocational training classes led to gains of about 5 points in the test scores of the program group relative to the control group. However, the control group worked more during the 30-month follow-up period, which partially offset these skills gains. Thus, these two factors combined led to implied impacts on literacy skills that were similar to the observed impacts on literacy skills.

We find also that the short-term earnings impacts were larger than can be explained by the impacts on literacy skills alone. This is because the association between literacy scores and earnings is very modest within the limited range of literacy skills observed in our sample. Hence, the earnings gains were likely to have been due to other factors influenced by Job Corps that are not captured in the test scores. These factors might include impacts on vocational skills for a specific job that are not captured in the literacy test, improvements in social skills and attitudes about work, and credentialing effects from obtaining a GED or vocational certificate.

In the remainder of this section, we first provide a theoretical framework that we use to examine the link between educational attainment, labor market outcomes, and literacy skill scores. In the second and third sections, we present empirical results on these relationships.

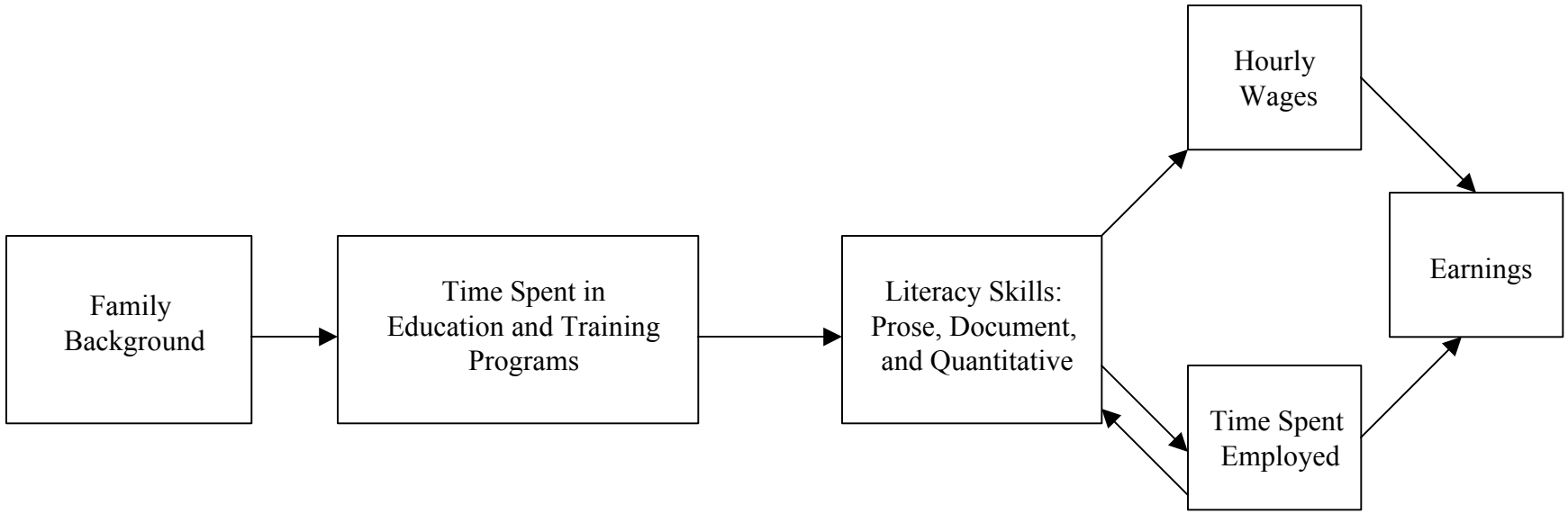
A. THEORETICAL FRAMEWORK

Figure IV.1 displays a simplified path model of the relationship between family background, schooling, literacy skills, and labor market outcomes. The arrows indicate the hypothesized direction of the relationships. Family background and demographic variables (such as parent's education) influence educational levels (measured by the amount of education and training received). Education-related outcomes are in turn expected to influence functional literacy skills in all three domains. Literacy skills are likely to correlate with one's productivity in the workplace, and hence, with hourly wage rates and the time spent employed. Finally, we expect that the time spent employed should in turn influence literacy skills as workers obtain skills on the job.[1]

[1]For simplicity, some arrows are omitted from the figure. For example, family background measures may have a direct effect on literacy skills even after controlling for the effects of schooling. Similarly, schooling measures may have a direct effect on labor market outcomes, even after controlling for the effects of skill levels. In addition, we omit from the figure the effect of innate ability on schooling, literacy levels, and labor market outcomes.

FIGURE IV.1

THE HYPOTHESIZED RELATIONSHIP BETWEEN FAMILY BACKGROUND, SCHOOLING, LITERACY SKILLS, AND LABOR MARKET BEHAVIOR



NOTE: For simplicity, some arrows are omitted from the figure. For example, family background measures may have a direct effect on literacy skills even after controlling for the effects of schooling. Similarly, schooling measures may have a direct effect on labor market outcomes, even after controlling for the effects of skill levels. In addition, we omit from the figure the effect of innate ability on schooling, literacy levels, and labor market outcomes.

In the path model, literacy is considered to be an *intermediate* (mediating) outcome that is affected by schooling and work experience, and that affects later outcomes (such as earnings). Thus, information on the empirical relationship between these measures can be used to help interpret our impact findings on literacy skills. Our goal is not to deduce causal relations, or to test formally the path model. This analysis is very difficult because of the complex relationship between these outcomes. Instead, our goal is to examine the broad relationships between schooling, literacy skills, and labor market outcomes to provide explanations for the impacts on literacy skills that we observe.

B. THE RELATIONSHIP BETWEEN LITERACY SKILLS AND SCHOOLING AND WORK EXPERIENCE

Our model posits that literacy skills are influenced by schooling and work experience because both of these activities are associated with an individual's ability to read, write, and speak in English, and to compute and solve problems at levels necessary to function on the job and in society. To examine the relationship between human capital measures and literacy skills, we estimated multivariate regression models where literacy scores were regressed on measures of participation in education and training programs and work experience. The parameters of the models were estimated using ordinary least squares (OLS) methods, and separate regressions were estimated for prose, document, and quantitative scores.

The regression models included control variables pertaining to the period before random assignment (baseline variables) as well as those pertaining to the period between random assignment and the 30 months after random assignment (since the basic skills tests were all administered after the 30 month point). The baseline variables were similar to those that were used to obtain the regression-adjusted impacts on literacy skills. They were constructed using baseline interview data and include demographic and family background variables as well as education-related and labor

market variables pertaining to the period prior to random assignment.[2] The control variables pertaining to the post-random assignment period included the following schooling and work experience measures: (1) total hours spent in education and training programs (including Job Corps and other programs) and in some specifications, total hours of academic classroom training and total hours of vocational training; and (2) total hours employed over the 30-month period. We focus on the parameter estimates for the post-random assignment measures because as discussed below, these parameter estimates can be used, along with the study's short-term impact estimates on time spent employed and in school to help explain the impacts on literacy skills.

It is important to note that for logistical and cost reasons we did not conduct skills testing at baseline. Thus, we cannot fully control for differences in skill levels between sample members at baseline. The control variables measured at baseline serve only as proxies for baseline differences in ability levels. Thus, as stated above, we do not view the estimated relationships between the human capital measures and test scores as true structural parameters because of likely omitted variable (sample selection) biases.[3]

Tables IV.1 to IV.3 display the parameter estimates for the post-random assignment schooling and work experience variables in each literacy domain. We indicate the statistical significance of the parameter estimates by asterisks. We present results for various model specifications defined by

[2]The baseline interview obtained detailed information only about jobs in the prior year. Thus, we have only limited information about lifetime work experience. However, this is not likely to be a serious problem because Job Corps typically serves youths with limited labor market experience.

[3]These biases could in principle be eliminated by using instrumental variables that are correlated with schooling and employment decisions, but are uncorrelated with unobservable factors that affect literacy levels. However, we were unable to specify measures from our survey data that met these stringent conditions.

TABLE IV.1

PARAMETER ESTIMATES ON SCHOOLING AND WORK EXPERIENCE MEASURES USING PROSE SCORES AS THE DEPENDENT VARIABLE

	Schooling Measures Included in the Model	
Control Variables	Hours in Education and Training	Hours in Academic Classes and in Vocational Training
Total Hours Employed During the 30 Months After Random Assignment	0.002***	0.002***
Total Hours Ever in an Education or Training Program During the 30-Month Period	0.002***	
Total Hours in Academic Classes During the 30-Month Period		-0.001
Total Hours in Vocational Training During the 30-Month Period		0.005***
Ratio of Implied Impacts from the Model to Actual Impacts on Prose Scores[a]	0.34	0.76
Sample Size	**2,273**	**2,273**

SOURCE: Literacy assessment test score data on 1,117 program and 1,156 control group members.

NOTE: The regression models also included control variables pertaining to the period before random assignment. These included demographic and family background variables as well as education-related and labor market variables.

[a]The implied impacts on literacy skills were calculated as a weighted average of the impacts on the schooling and work experience measures, where the weights were the parameter estimates on these measures.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

TABLE IV.2

PARAMETER ESTIMATES ON SCHOOLING AND WORK EXPERIENCE MEASURES USING DOCUMENT SCORES AS THE DEPENDENT VARIABLE

	Schooling Measures Included in the Model	
Control Variables	Hours in Education and Training	Hours in Academic Classes and in Vocational Training
Total Hours Employed During the 30 Months After Random Assignment	0.002***	0.002***
Total Hours Ever in an Education or Training Program During the 30-Month Period	0.002*	
Total Hours in Academic Classes During the 30-Month Period		-0.001
Total Hours in Vocational Training During the 30-Month Period		0.005***
Ratio of Implied Impacts from the Model to Actual Impacts on Document Scores[a]	0.78	0.69
Sample Size	**2,273**	**2,273**

SOURCE: Literacy assessment test score data on 1,117 program and 1,156 control group members.

NOTE: The regression models also included control variables pertaining to the period before random assignment. These included demographic and family background variables as well as education-related and labor market variables.

[a]The implied impacts on literacy skills were calculated as a weighted average of the impacts on the schooling and work experience measures, where the weights were the parameter estimates on these measures.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

TABLE IV.3

PARAMETER ESTIMATES ON SCHOOLING AND WORK EXPERIENCE MEASURES USING QUANTITATIVE SCORES AS THE DEPENDENT VARIABLE

	Schooling Measures Included in the Model	
Control Variables	Hours in Education and Training	Hours in Academic Classes and in Vocational Training
Total Hours Employed During the 30 Months After Random Assignment	0.003***	0.003***
Total Hours Ever in an Education or Training Program During the 30-Month Period	0.003*	
Total Hours in Academic Classes During the 30-Month Period		0.0
Total Hours in Vocational Training During the 30-Month Period		0.006***
Ratio of Implied Impacts from the Model to Actual Impacts on Quantitative Scores[a]	0.47	0.94
Sample Size	**2,273**	**2,273**

SOURCE: Literacy assessment test score data on 1,117 program and 1,156 control group members.

NOTE: The regression models also included control variables pertaining to the period before random assignment. These included demographic and family background variables as well as education-related and labor market variables.

[a]The implied impacts on literacy skills were calculated as a weighted average of the impacts on the schooling and work experience measures, where the weights were the parameter estimates on these measures.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

which education-related variables are included in the model. Appendix Table C.3 displays the full regression results for the second model specification.[4]

For each specification, the tables also present the ratio of literacy skill impacts that are *implied* by the regression model to the *actual* literacy skill impacts. These ratios were calculated by decomposing the impact on literacy skills into a weighted average of the impacts on the schooling and work experience measures, and into residual factors. The weights were the parameter estimates from the regression models. For example, for the first model specification, we decomposed the impact on literacy skills (3_{NALS}) as follows:

$$(1)\quad I_{NALS} = \$_W I_W + \$_S I_S + residual,$$

where $\$_W$ and $\$_S$ are the parameter estimates on hours of work and hours of school, respectively, and 3_W and 3_S are the estimated impacts of Job Corps on these measures. The residual term is the parameter estimate on the indicator variable signifying that the youth is in the program group that is included in the regression models. The ratio of implied to actual impacts on literacy skills for this specification was calculated as follows:

$$(2)\quad Ratio = \frac{\$_W I_W + \$_S I_S}{I_{NALS}} .$$

[4]Statistical tests indicate that the parameter estimates on the schooling and work experience variables are similar for program and control group members. Thus, we present findings for the pooled sample.

Ratios near 1 suggest that the observed impacts on literacy skills are consistent with what one might expect on the basis of impacts on the human capital measures. Ratios substantially smaller or larger than 1 suggest that there are significant residual differences between program and control group members that account for the observed impacts on literacy skills.

Our results indicate that increases in hours of schooling and hours of work are associated with increases in literacy skills in all three domains (see specification 1). We find that an additional 1,000 hours of schooling (which is equivalent to about one school year) increased document and prose scores by about 2 scale points and quantitative scores by about 3 scale points, all else equal. An additional 1,000 hours of work had a similar effect. These parameter estimates are statistically significant at the 1 percent level, but are smaller than expected.

One possible explanation for the small effects is that time spent employed and in school during the 30-month follow-up period varied considerably across sample members (because, for example, program group members spent very different amounts of time in Job Corps and control group members had very different education and labor market experiences), while variation in literacy scores was relatively small (because most of the sample had test scores in Levels I and II).

The implied impacts on literacy skills from specification 1 are about one-third of the actual impacts on prose scores, about half of the actual impacts on quantitative scores, and nearly 80 percent of the actual impacts on document scores. For example, the 1,000 hour impact per participant on hours spent in education and training programs led to about a 3 scale point increase in quantitative scores, whereas the -275 hour impact on hours of employment led to about a 0.8 scale point decrease in quantitative scores. Thus, these two factors combined led to a 2.2 point increase in quantitative scores, which is about half the actual impact per participant on these scores.

Interestingly, the positive association between hours of education and literacy levels came from hours spent in vocational training and not from hours spent in academic classes (see specification 2). A 1,000 hour increase in the amount of vocational training led to a 5 point gain in test scores in each domain, and these effects are statistically significant at the 1 percent level. Furthermore, the implied impacts on test scores from this specification are very similar to the actual impacts on test scores (that is, the reported ratios are near 1). This finding strongly suggests that vocational skills training plays a strong role in improving the functional literacy skills of Job Corps students.

To further investigate the association between schooling and literacy skills, we calculated average test scores in each domain for those at different educational levels at the time the test was administered (Table IV.4).[5] In particular, we examined the differences in test scores between those with and without a high school credential because we found large impacts of Job Corps on the receipt of GED certificates (although we found no program impacts on college attendance or on highest grade completed).

We find that those with GED certificates or high school diplomas scored about 25 points higher in all three domains than those without a high school credential, and those who completed some college scored about 15 points higher than those with only a high school credential.[6] Furthermore, the 16 percentage point impact on GED attainment per participant implies a 4 scale point impact on literacy test scores, which is similar to the actual impacts. These findings provide further evidence

[5]The figures are not regression-adjusted because it is difficult to interpret the effect of a high school credential on test scores after controlling for the effects of other schooling measures and demographic characteristics that are intended to capture differences in the ability levels of sample members.

[6]Interestingly, these differences in scores by educational level are similar to those found in the nationally representative NALS study (Kirsch et al. 1993).

TABLE IV.4

AVERAGE LITERACY ASSESSMENT SCORES
BY EDUCATIONAL ATTAINMENT AND
LITERACY DOMAIN

	Literacy Domain		
Educational Attainment	Prose	Document	Quantitative
Completed Grades 0 to 8 and Did Not Receive a GED Certificate	237.3	248.2	218.5
Completed Grades 9 to 11 and Did Not Receive a GED Certificate	237.1	245.1	217.5
Received a GED Certificate	264.9	270.3	250.7
Received a High School Diploma	254.9	264.0	241.6
Completed Some College	276.6	281.1	268.5
Sample Size	**2,273**	**2,273**	**2,273**

SOURCE: Literacy assessment test score data on 1,117 program and 1,156 control group members.

NOTE: All figures were calculated using sample weights to adjust for nonresponse and the sample design. The figures are not regression-adjusted because it is difficult to interpret the effect of a high school credential on test scores after controlling for the effects of other schooling measures and demographic characteristics that are intended to capture differences in the ability levels of sample members.

that the observed impacts on test scores are consistent with the impacts on education-related measures.

In sum, we find that literacy levels are influenced by schooling and employment measures for youths served by Job Corps. Furthermore, the impacts on functional literacy scores in all three literacy domains appear to be consistent with what one would expect on the basis of our impact findings on schooling and work experience, and on the empirical relationships between these human capital measures and literacy.

C. THE RELATIONSHIP BETWEEN LITERACY SKILLS AND EARNINGS

The approach to literacy assessment developed by ETS gauges one's ability to perform literacy and numeracy tasks called for in the workplace. Thus, literacy scores are expected to be positively associated with one's productivity, and hence, with labor market outcomes such as hourly wages and earnings. In this section, we address the following question: Are the impacts of Job Corps on participants' literacy skills consistent with the program impacts on key labor market outcomes?

To address this issue, we estimated regression models in which labor market outcomes were regressed on test scores in all three domains.[7] We then used the methods described in the previous section to compare the implied impacts on labor market outcomes based on these relationships to the actual labor market impacts. The labor market outcomes that we examined were earnings per week in the tenth quarter after random assignment (that is, in the last quarter of the 30-month follow-up period), and hourly wages in the most recent job in quarter 10 (for the two-thirds of sample members who were employed in quarter 10).

[7]We did not include other control variables in these regression models because it is not clear how to interpret the parameter estimate on literacy scores after controlling for the effects of schooling and other variables that the literacy scores are intended to capture. For example, after controlling for schooling, the literacy coefficient on literacy scores is probably capturing the effects of innate ability on labor market outcomes.

Table IV.5 displays the regression results when the test scores in all three domains were included together as control variables, and when the scores were included separately. The table displays parameter estimates, the regression R^2 values, and the ratios of implied to actual impacts on earnings and wages. The parameter estimates are very similar for program and control group members, and thus, we present findings for the pooled sample.

Our results indicate that there is some association between literacy assessment scores and labor market outcomes, but a large residual variation in earnings remains. A 10 scale point increase in quantitative scores (or one-fifth of a literacy level increase) was associated with about a $6.50 increase in earnings per week (on a base of about $180 per week for the average Job Corps participant) and with about a $0.07 increase in the hourly wage rate (on a base of about $7). The parameter estimates on the prose and document scores are not statistically significant when they are included with the quantitative scores. However, because the correlation coefficients among the test scores in the three domains are about 0.8, the parameter estimates and significance levels on the test scores are very similar when they are included separately in the models. We find very similar results if quadratic terms are included in the models (they are all statistically insignificant) or if we instead include indicator variables defined by ranges of test scores. The earnings results are also very similar if we include only those with positive earnings or the full sample. The regression R^2 values are less than 0.02.

The parameter estimates on the literacy measures are about half as large as those found by DOL (1993) for a sample of JTPA-eligible applicants. However, the DOL sample included both youths *and* adults, and thus, the range of literacy scores was greater for their sample. In our sample, the variance of quarter 10 earnings was large, whereas the variance of the test scores was relatively small

TABLE IV.5

PARAMETER ESTIMATES ON PROSE, DOCUMENT, AND QUANTITATIVE TEST SCORES USING EARNINGS PER WEEK AND HOURLY WAGES ON THE MOST RECENT JOB IN QUARTER 10 AS THE DEPENDENT VARIABLES

	Earnings Per Week in Quarter 10		Hourly Wage in Quarter 10	
Control Variables	Test Scores Entered Together	Test Scores Entered Separately	Test Scores Entered Together	Test Scores Entered Separately
Prose Scores	0.20	0.54***	0.003	0.004***
Document Scores	-0.31	0.50***	-0.007*	0.003
Quantitative Scores	0.65***	0.57***	0.007***	0.005***
R^2	0.02		0.01	
Ratio of the Implied Impact to the Actual Impact on the Labor Market Outcome[a]	0.10		0.06	
Sample Size	**2,273**	**2,273**	**1,406[b]**	**1,406[b]**

SOURCE: Literacy assessment test score data on program and control group members.

[a]The implied impact on earnings or wages was calculated as a weighted average of the impacts on the test scores, where the weights were the parameter estimates on these test scores.

[b]Figures pertain to those who were employed for at least one week during the tenth quarter after random assignment.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

because most sample members had literacy scores in Levels I and II. Thus, the association between test scores and earnings may have been smaller for our sample.

The *implied* impacts on earnings and hourly wages from the regression models are much smaller than the actual impacts on earnings and hourly wages. The 4 point impact on literacy scores implies an earnings impact of less than $3 and an hourly wage impact of less than $0.03. The actual impacts on earnings and hourly wages were about $20 and $0.30, respectively. Thus, the ratios of implied to actual impacts are about 0.10. These findings suggest that there are other important factors that Job Corps influenced that account for the earnings impacts that we observe.

We offer four possibilities. First, although there is a positive association between time spent in vocational training and test scores, the literacy test may not fully capture all the vocational skills that participants obtained. For example, the large impacts on hours spent in vocational training may have contributed to program group members being better skilled carpenters or mechanics than control group members, which led to earnings gains for the program group. However, the literacy test may not have fully captured all the skills obtained from this job-specific training. To test this explanation, we estimated a regression model of earnings on the number of hours of vocational training received and other control variables, which *included* test scores (not shown). We found that there is a strong association between earnings and vocational training, even after controlling for these scores. An extra 100 hours of vocational training led to a $3 increase in earnings per week.

Second, in addition to formal academic and vocational instruction, Job Corps offers a broad range of activities that are designed to promote life skills and workplace success. Job Corps offers social skills training, world of work classes, cultural awareness classes, and counseling services. These services may have increased participants' employability because of improvements in their attitudes about work, motivation to work, interviewing skills, and their general ability to obtain and

hold better jobs. These life skills may have increased participants' earnings, but are probably not captured in the literacy test.

Third, Job Corps offers placement assistance to all students in finding a job. Thus, program group members may have been more likely than control group members to have found jobs that matched their skills, which could have resulted in higher earnings for the program group. However, as discussed in our main impact report, few program participants reported that they received significant placement assistance. Thus, we believe that this placement assistance accounted for only a small part of the earnings impacts.

Finally, the large impacts on the receipt of GED and vocational certificates may have contributed to the earnings gains because of "signaling" effects unrelated to formal literacy levels. Employers may reward educational credentials because they may signal that the individual has characteristics (such as perseverance and motivation) that make him or her a better employee. In support of this theory, we found that those with GEDs and vocational certificates had higher earnings than those without these certificates, after controlling for prose, document, and quantitative scores.

REFERENCES

Campbell, Anne, et al. "Assessing Literacy: The Framework for the National Adult Literacy Survey." Report prepared for the U.S. Department of Education. Princeton, NJ: Educational Testing Service, October 1992.

Kirsch, I., et al. *Adult Literacy in America: A First Look at the Results of the National Adult Literacy Survey.* Washington, DC: U.S. Department of Education: National Center for Education Statistics, 1993.

Kirsch, Irwin S., Ann Jungeblut, et al. "Profiling the Literacy Proficiencies of JTPA and ES/UI Populations." Final report to the U.S. Department of Labor. Princeton, NJ: Educational Testing Service, September 1992.

Kirsch, Irwin, et al. "The ETS Tests of Applied Literacy Skills: Administration and Scoring Manual." Princeton, NJ: Educational Testing Service, 1991.

Lord, F. M. *Application of Item Response Theory to Practical Testing Problems.* Princeton, NJ: Educational Testing Service, 1980.

Martinson, Karin, and Daniel Friedlander. "GAIN: Basic Education in a Welfare to Work Program." New York: Manpower Development Research Corporation, January 1994.

Mislevy, R. J. "Randomization-based Inferences About Latent Variables from Complex Samples". *Physchometrika* 56, 1991.

National Assessment of Educational Progress. *Technical Report on the 1983-84 Assessment of Reading and Writing.* Princeton, NJ: Educational Testing Service (1986).

Schochet, P. "National Job Corps Study: Methodological Appendices on the Short-Term Impact Analysis, Final Report." Princeton, NJ: Mathematica Policy Research, Inc., February 2000.

Schochet, P. "National Job Corps Study: Characteristics of Youths Served by Job Corps." Princeton, NJ: Mathematica Policy Research, Inc., January 1998.

Schochet, P., J. Burghardt, and S. Glazerman. "National Job Corps Study: The Short-Term Impacts of Job Corps on Participants' Employment and Related Outcomes, Final Report." Princeton, NJ: Mathematica Policy Research, Inc., February 2000.

U.S. Department of Labor, Office of Strategic Planning and Policy Development. *Workplace Literacy and the Nation's Unemployed Workers.* Research and Evaluation Report Series 93-F. Washington, DC: DOL, 1993.

U.S. Department of Labor, The Secretary's Commission on Achieving Necessary Skills. *What Work Requires of Schools: A SCANS Report for America 2000.* Washington, DC: DOL, June 1991.

APPENDIX A

ANALYSIS OF NONRESPONSE

The universe for the analysis of Job Corps' impacts on literacy skills is the nearly 37,000 first-time applicants to Job Corps who were determined between August 1995 and February 1996 to be eligible for the program. The literacy skills sample was randomly selected from youths in the main impact study sample. A total of 3,750 youths (1,875 program group and 1,875 control group members) were selected for the literacy skills study. The sample used in the analysis contains 2,273 young adults (1,117 program group and 1,156 control group members) who completed literacy assessments. In this appendix, we provide evidence that respondents to the literacy test are representative of those in the sample universe, as well as our approach for constructing weights to adjust for the effects of nonresponse when we estimate impacts using the simple differences-in-means approach. To understand the sources of nonresponse, however, we begin with an overview of the sample design for the literacy study.

A. OVERVIEW OF THE SAMPLE DESIGN FOR THE LITERACY STUDY

The analysis sample for the literacy study was determined in three stages:

1. Youths in the main impact study sample were determined to be eligible for inclusion in the literacy assessment sample if they met the following two criteria: (1) they were eligible for a 30-month interview (see Schochet 2000); and (2) they were randomly assigned after August 1, 1995. This sampling frame of youths eligible for literacy assessment consisted of 2,902 program group members and 2,418 control group members (see Table A.1).

2. Among those eligible for literacy assessment, we selected at random 1,875 program group members and 1,875 control group members for the literacy study. These samples were selected in a way that mostly undid the unequal selection probabilities used to obtain the initial research sample. This design made the literacy skills sample nearly self-weighting.

TABLE A.1

PROCESS FOR OBTAINING THE JOB CORPS LITERACY STUDY SAMPLE

Steps	Program Group	Control Group	Total
Sample Universe[a]	--	--	36,922
Randomly Assigned to the Program Group or Control Group[a]	4,478	2,870	7,348
Eligible for a 30-Month Interview[a]	2,902	2,418	6,320
Randomly Assigned to the Literacy Assessment Sample	1,875	1,875	3,750
Completed the 30-Month Interview[a]	1,156	1,525	3,058
Completed the Literacy Assessment[b]	1,117	1,156	2,273

SOURCE: Baseline, 12-month, and 30-month follow-up interview data and Job Corps Literacy Assessment data.

[a]Pertains to applicants to the Job Corps program whose determination of eligibility was made between August 1995 and February 1996.

[b]Includes 14 control group and 20 program group members who did not complete a 30-month interview.

3. More than 81 percent of the 3,750 young adults selected for assessment completed a 30-month interview. Of those, about 74 percent completed the literacy assessment.[1] Thus, the final response rate was about 60 percent and was similar for program group and control group members.

Respondents to the literacy test may not be representative of those in the study universe for two reasons. First, the literacy sample was selected only from those eligible for 30-month follow-up interviews, and not from all those in the research sample. Thus, nonresponse bias could affect our estimates if the characteristics of those eligible for 30-month interviews differed from those in the full research sample. This issue was explored in Schochet (2000) and is not discussed here. Second, not everyone who was selected for the literacy sample completed a literacy assessment. Thus, the characteristics of test respondents may have differed from those of the full literacy sample, potentially introducing another source of nonresponse bias. The next section discusses nonresponse among those selected for the literacy study.

B. RESPONDENTS AND NONRESPONDENTS AMONG THOSE SELECTED FOR THE LITERACY STUDY

The overall (weighted) response rate to the literacy test was 60.2 percent (59.6 percent for the program group and 60.8 percent for the control group). In this section, we assess whether the analysis sample can yield reliable estimates of impacts on literacy skills. First, we examine the reasons for nonresponse. Second, we examine when the respondents completed the tests, how much time they spent completing the test, and how many items they attempted to complete. Finally, we compare the observable characteristics of respondents and nonrespondents to gauge whether the

[1]An exception is the 34 sample members who did not have a 30-month interview but did complete the literacy test. These people were not contacted until the end of the testing period, which had been extended slightly beyond the end of the 30-month interview period.

literacy skills of the respondents are likely to be representative of the literacy skills of the full sample of young adults selected for the literacy study.

1. The Reasons for Nonresponse

The reasons for nonresponse are listed in Table A.2. Only 17 percent of those who did not complete a literacy assessment refused to take the test (16 percent of program group and 19 percent of control group members). Almost all the remaining nonrespondents could not be located. Importantly, the reasons for nonresponse were similar for program group and control group members.

2. Timing and Duration of Interviews

Respondents who completed the literacy assessment generally did so about three years (36 months) after random assignment (see Table A.3). More than half the respondents completed the assessment in months 35 through months 39. These figures are similar for program group and control group members.

Most respondents (about two-thirds of the sample) completed the literacy assessment interview in less than one hour. This interview included both a short background questionnaire, which takes about 5 to 10 minutes, and the literacy exercise, which included between 37 and 40 tasks (see Appendix B) and was not timed. The average amount of time spent on the full interview (including the background and literacy components) was about 53 minutes and was similar for program group and control group members (see Table A.4).

TABLE A.2

REASON FOR NONRESPONSE TO LITERACY ASSESSMENT INTERVIEWS, BY RESEARCH STATUS AND WHETHER COMPLETED A 30-MONTH INTERVIEW
(Percentages)

Reason	Program Group		Control Group	
	Completed Interview	Did Not Complete Interview	Completed Interview	Did Not Complete Interview
Refusal				
Sample Member Refused	14.7	10.6	16.5	16.4
Someone Else Refused or Denied Access	3.3	2.6	2.7	2.7
Total	**18.0**	**13.2**	**19.1**	**19.1**
Could Not Locate or Contact				
Unable to Locate	46.8	64.6	49.2	58.5
Moved and Unable to Locate	15.8	12.2	13.6	12.4
Located but Could Not Contact	3.5	1.3	4.3	1.2
Could Not Contact (Incarcerated)	2.1	3.5	2.9	3.9
Could Not Contact (Otherwise Institutionalized)	0.0	0.0	0.0	0.3
Unavailable (in Military)	1.9	1.3	1.1	0.9
Unavailable (Abroad)	0.2	0.6	1.6	0.3
Multiple Attempts, Case Retired	9.7	2.6	6.4	2.7
Total	**80.1**	**86.2**	**79.0**	**80.3**
Not Able to Complete				
Cannot Read or Write Adequately	0.9	0.6	0.8	0.0
Non-English Speaker	0.9	0.0	0.8	0.3
Deceased	0.0	0.0	0.3	0.3
Total	**1.9**	**0.6**	**1.9**	**0.6**
Sample Size	**423**	**311**	**376**	**330**

SOURCE: Job Corps Literacy Assessment data.

TABLE A.3

TIMING OF LITERACY ASSESSMENTS

	Program Group	Control Group	Total
Months After Random Assignment (Percentages)			
28 to 29	0.3	0.0	0.2
30 to 34	37.4	38.6	38.0
35 to 39	51.3	51.0	51.1
40 to 43	10.9	10.4	10.7
Average Months (Number)	35.6	35.7	35.6
Sample Size	**1,117**	**1,156**	**2,273**

SOURCE: Job Corps Literacy Assessment data.

TABLE A.4

DURATION OF LITERACY ASSESSMENT INTERVIEWS

	Program Group	Control Group	Total
Elapsed Time (Minutes) (Percentages)			
Less than 30	2.8	2.8	2.8
30 to 45	20.8	17.7	19.2
45 to 60	43.4	45.4	44.4
60 to 75	28.4	29.4	28.9
75 to 90	3.1	3.8	3.4
90 to 120	1.5	0.9	1.2
Average Minutes (Number)	52.3	53.2	52.8
Sample Size	**1,023**	**1,077**	**2,100**

SOURCE: Job Corps Literacy Assessment data.

NOTE: Elapsed time for some sample members could not be determined.

3. Item Nonresponse

Response rates to individual items were generally high. Table A.5 shows the average percentage of items attempted, by subgroup and research status. The item response rates for each of the key subgroups in both the program group and the control group were more than 85 percent.[2] The rate was slightly lower for those who had no high school credential when they applied to Job Corps than for those who did have a high school diploma or GED certificate. Nevertheless, the item response rates were similar for program and control groups overall and within subgroups. More than half the respondents attempted 90 percent or more of the items. Less than eight percent of the respondents attempted less than 50 percent of the items. On average, sample members gave correct answers to 48 percent of the items, including items attempted and items skipped.

4. The Comparison of Respondents and Nonrespondents

Test respondents differed somewhat from the full set of respondents and nonrespondents selected for literacy assessment. Table A.6 compares the characteristics of respondents with those of the full sample, separately by research status. This analysis was conducted using program intake data (from Form ETA-652) that are available for both respondents and nonrespondents. We use asterisks to denote whether differences were statistically significant.

We found some statistically significant differences. For example, respondents were significantly more likely to be female, to be African-American, and from nondense areas. These differences, however, were substantively very small.

[2]The effective item response rate was actually higher than 85 percent because we scored as incorrect responses the items that the respondent was presented with but did not attempt or could not answer.

TABLE A.5

AVERAGE ITEM RESPONSE RATES,
BY SUBGROUP AND RESEARCH STATUS
(Average Percentages of Items Attempted)

Baseline Characteristics	Control Group	Program Group
Full Sample	86.5	86.7
Gender		
Female	87.3	88.1
Male	86.0	85.7
Age at Application (Years)		
16 to 17	85.0	86.4
18 to 19	87.8	87.1
20 to 24	87.3	86.6
Education		
No high school credential	85.3	85.8
Had GED	92.9	95.1
Had high school diploma	90.0	89.1
Age, by Education		
16 to 17	85.0	86.4
18 to 24, no high school credential	85.7	85.3
18 to 24, with high school credential	90.8	89.8
Residential Status[a]		
Residential designee	86.3	86.7
Nonresidential designee	87.9	86.4
Sample Size	**1,117**	**1,156**

SOURCE: Baseline Interview data and Job Corps Literacy Assessment data

NOTES: All figures are calculated using sample weights to account for the sample design.

[a]Figures were obtained using data on OA counselor projections about the centers that youths were likely to attend.

TABLE A.6

CHARACTERISTICS OF RESPONDENTS TO THE LITERACY ASSESSMENT AND THE FULL SAMPLE OF RESPONDENTS AND NONRESPONDENTS, BY RESEARCH STATUS
(Percentages)

	Control Group		Program Group	
	Respondents	Respondents and Nonrespondents	Respondents	Respondents and Nonrespondents
Demographic Characteristics				
Male	56.9	60.7***	56.2	59.3***
Age at Application (Years)				
16 to 17	40.9	39.7	39.9	39.3*
18 to 19	33.8	33.4	33.5	33.3
20 to 21	16.9	17.2	15.4	16.4
22 to 24	8.3	9.7	11.1	11.1
Race/Ethnicity				
White, non-Hispanic	28.1	28.5**	27.6	28.6***
Black, non-Hispanic	54.2	51.1	53.4	50.8
Hispanic	13.0	15.0	15.0	16.0
Other	4.8	5.5	4.0	4.6
Region				
1	4.0	4.6***	4.2	4.7***
2	6.2	7.5	4.4	6.8
3	12.6	12.9	12.2	12.6
4	23.6	22.9	23.3	22.1
5	11.0	10.2	11.4	10.5
6	15.0	14.4	15.2	15.3
7/8	11.7	11.1	13.8	12.2
9	9.7	11.4	10.4	11.2
10	6.2	5.0	5.1	4.5
Size of City of Residence				
Less than 2,500	11.1	9.8	10.2	9.6
2,500 to 10,000	10.7	10.4	11.7	10.9
10,000 to 50,000	17.7	17.0	18.7	18.0
50,000 to 250,000	18.1	19.0	17.0	18.1
250,000 or more	42.5	43.7	42.4	43.4
PMSA or MSA Residence Status				
In PMSA	30.4	32.6***	30.4	32.7**
In MSA	47.2	46.6	44.6	45.8
In neither	22.4	20.9	25.0	21.5

TABLE A.6 *(continued)*

	Control Group		Program Group	
	Respondents	Respondents and Nonrespondents	Respondents	Respondents and Nonrespondents
Density of Area of Residence				
Superdense	35.4	36.8***	33.1	35.9**
Dense	26.3	27.2	27.3	27.8
Nondense	38.3	36.0	39.6	36.4
Lived in 1 of 57 Areas with a Large Concentration of Nonresidential Females	34.1	33.6	31.7	32.1
Legal U.S. Resident	98.8	98.8	98.4	98.3
Job Corps Application Date				
11/94 to 2/95	0.3	0.5	0.5	0.4*
3/95 to 6/95	2.9	3.7	2.7	3.3
7/95 to 9/95	50.7	51.7	53.5	52.2
10/95 to 12/95	46.0	44.2	43.2	44.1
Fertility and Family Status				
Had Dependents	17.4	16.1*	15.6	14.8
Family Status				
Family head	12.8	12.5	13.5	13.4
Family member	60.6	60.7	62.2	61.3
Unrelated person	26.6	26.8	24.3	25.2
Average Family Size	3.2	3.1**	3.3	3.3
Education				
Completed the 12th Grade	23.3	22.2	23.5	22.6*
Welfare Dependence				
Public Assistance Receipt				
Received AFDC	29.1	27.2	28.0	27.5*
Received other assistance	15.4	15.1	16.5	16.1
Did not receive assistance	55.5	57.8	55.5	56.4

TABLE A.6 *(continued)*

	Control Group		Program Group	
	Respondents	Respondents and Nonrespondents	Respondents	Respondents and Nonrespondents
Health				
Had Any Health Conditions that Were Being Treated	4.8	4.3	4.0	3.4***
Crime				
Arrested in Past Three Years	12.2	12.3	12.2	12.5
Ever Convicted or Adjudged Delinquent	6.9	6.9	6.3	6.3
Completion Status to Previous Interviews				
Completed the 12-Month Interview	96.1	90.8***	96.0	92.3*
Anticipated Program Enrollment Information				
Designated for a Nonresidential Slot	15.4	14.4	15.0	14.2
Designated for a CCC Center[a]	15.2	14.4	15.1	14.0
Designated for a High- or Medium-High-Performing Center[a]	57.9	55.8**	57.8	56.0
Designated for a Large or Medium-Large Center[a]	66.1	64.3*	64.8	64.7
Sample Size	**1,156**	**1,875**	**1,117**	**1,875**

SOURCE: ETA-652 and ETA-652 Supplement data.

[a]Figures are obtained using data on OA counselor projections about the centers that youths were likely to attend.

*Difference between respondents and the full sample is significant at the .10 level, two-tailed test.
**Difference between respondents and the full sample is significant at the .05 level, two-tailed test.
***Difference between respondents and the full sample is significant at the .01 level, two-tailed test.

We also examined differences in the initial literacy skills of respondents and nonrespondents in the program group who enrolled in Job Corps, using data from the Test of Adult Basic Education (TABE). At entry, Job Corps students take the TABE, including the total reading portion and the mathematics computation subtest, and these scores are available in the Job Corps Student Pay, Allotment, and Management Information System (SPAMIS) enrollment data. Thus, we compared average TABE scores for program group participants in the literacy sample who completed the literacy tests and for those who did not complete the tests.

Table A.7 compares the TABE scores of respondents with the scores of the full sample of respondents and nonrespondents. The TABE scores of the respondents were very similar, about 1/10 of a grade equivalent higher, compared with the set of respondents and nonrespondents for whom TABE scores were available. This finding suggests that the respondents were representative of the full sample in baseline achievement.

Overall, our findings suggest that the test respondents appear to be representative of those in the population universe. However, as discussed next, we constructed weights to adjust for potential nonresponse biases to account for the small differences in the characteristics of respondents and nonrespondents. We used these weights to estimate impacts with the simple differences-in-means approach.

C. CONSTRUCTION OF SAMPLE WEIGHTS

To generalize from our literacy assessment sample to the broader population of eligible Job Corps applicants, we calculated sample weights to account for the sample design and for nonresponse. The weights were constructed by computing the probability that sample members were eligible for the 30-month interview (P_1); the probability that a sample member was selected for

TABLE A.7

TABE SCORES OF RESPONDENTS AND NONRESPONDENTS
TO THE LITERACY ASSESSMENT[a]
(Grade Equivalents)

	TABE Form 5/6[b]			TABE Form 7/8[b]		
Subject	Respondents	Non-Respondents	Respondents and Non-Respondents	Respondents	Non-Respondents	Respondents and Non-Respondents
Reading	8.0	7.8	7.9	6.9	6.2	6.7***
Math	7.5	7.3	7.4	6.6	6.3	6.5
Sample Size	**460**	**326**	**786**	**339**	**190**	**529**

SOURCE: SPAMIS data on program participants.

[a] Data pertain to Job Corps participants who were administered TABE shortly after enrollment and who were later selected for inclusion in the Job Corps adult literacy assessment sample.

[b] TABE is the Test of Adult Basic Education. Form 5 and Form 6 of the test were used by Job Corps until 7/1/96, after which the program switched to Forms 7 and 8. Scale scores on Forms 5 and 6 are comparable to each other, but not to those from Forms 7 and 8.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

APPENDIX B

ADMINISTRATION AND SCORING OF THE JOB CORPS ADULT LITERACY ASSESSMENT

This appendix explains how we estimated the literacy skills of different groups for descriptive analyses and impact analyses, using responses to the Job Corps adult literacy assessment instrument. The item development, item sampling, and scaling were performed by the Educational Testing Service (ETS). Kirsch et al. (1992) and Mislevy (1991) provide detailed descriptions of the underlying methods, which we summarize here, along with an explanation of how we adapted the methods for the Job Corps study.

A. BIB SPIRALING FOR ITEM ADMINISTRATION

The Job Corps literacy assessment was meant to cover a wide range of literacy proficiencies without placing too much burden on any one respondent. In order to do this, we used a powerful item sampling procedure called "balanced incomplete block" (BIB) spiraling. BIB spiraling takes a large pool of exercises (items), groups them into blocks, and administers to each respondent a block of core items and a randomly selected subset of additional item blocks. The BIB spiraling design ensures that each item is administered to a representative sample of respondents, even if any given individual faces only a subset of items.

Table B.1 shows the block design used for the Job Corps literacy assessment. We administered to each respondent one of seven randomly chosen booklets, containing a block of six core items common to each booklet and three blocks of additional items. The blocks are arranged so that each block is paired with every other block and the relationships among every pair of items can be estimated. In addition, the order of the blocks is varied to control for positional effects or test fatigue effects. In fact, every block appears once in each position. Because the blocks can have 10, 11, or 12 questions each, the total number of items administered to each respondent varies from 37 to 40 across booklets.

TABLE B.1

BALANCED INCOMPLETE BLOCK DESIGN FOR
JOB CORPS LITERACY ASSESSMENT

Booklet	1st Block	2nd Block	3rd Block	4th Block	Number of Items
1	Core	1	2	3	38
2	Core	2	3	4	37
3	Core	3	4	5	37
4	Core	4	5	6	38
5	Core	5	6	7	40
6	Core	6	7	1	40
7	Core	7	1	2	40

SOURCE: Job Corps Literacy Assessment data

NOTE: The core block has 6 items. Blocks 1, 2, 5, and 6 have 11 items. Blocks 3 and 4 have 10 items. Block 7 has 12 items.

This item sampling method makes estimation of any individual's proficiency difficult. However, this type of design is highly efficient for estimating the literacy levels of most groups and major subgroups, including those defined by Job Corps research status (program or control group).

B. IRT SCALING

Another important feature of the literacy skills analysis is the method for summarizing examinee responses. Conventional methods, which typically use the number or proportion of correct responses (successfully completed tasks), would be inappropriate for test takers who receive different sets of items. Moreover, reporting the item-by-item proportion correct does not take full advantage of the information possibly conveyed by our matrix sampling design. Having multiple examinees take overlapping tests provides important information about the items, the relationships among items, and the performance of the same individual or groups of individuals on common items.

The limitations of conventional methods can be overcome by using item response theory (IRT) scaling. When several items require similar skills, the response patterns should have some regularity. This regularity can be used to characterize both examinees and items in terms of a common standard scale, even when examinees do not attempt identical sets of items. In this way, it is possible to place all respondents *and* tasks on a common scale based on their ability to successfully complete increasingly difficult tasks.

The scaling model used for the Job Corps assessment is a two-parameter logistic model from IRT (Lord 1980).[1] It is a mathematical model expressing the probability that an examinee with given proficiency will give a correct response to a particular item from a common domain of items. This

[1]Lord's formulation is more general, as it includes a third parameter that describes a "guessing floor," or the probability that a person with even extremely low proficiency might give a correct response. This parameter is necessary with multiple choice tests but can be ignored in the current context, because the Job Corps literacy assessment used free-response (open-ended) items.

probability depends on the test taker's proficiency, 2, and the two parameters characterizing the properties of the test item. The two parameters used in the Job Corps analysis are the item's discriminating power (sensitivity to proficiency), *a,* and the item's difficulty, *b*. Indexing individuals by *i* and items by *j,* we have the following expression:

$$(1) \quad P(Y_{ij} = 1 | 2_i, a_j, b_j) = \frac{1}{1 + \exp(-1.7 a_j (2_i - b_j))},$$

where Y_{ij} is an indicator for a correct response of person *i* to item *j*. By supposing that 2_i is common to all items on the test for person *i* (in other words, that the items represent a unidimensional skill domain), different items can be combined to estimate the proficiency of individuals or groups.

The method for IRT scaling proceeds in two steps. The first step is item calibration. In this step, background characteristics of respondents are used to estimate the item parameters in equation (1). See Kirsch et al. (1992) for estimation details. Normally the conditioning variables used to estimate item parameters are taken from a background characteristics questionnaire that ETS includes with most of its literacy assessments (for example, in the National Adult Literacy Survey). Our assessment included a similar background questionnaire, but program research status was also used as a conditioning variable. This change enabled us to control for any real differences between the program and control groups. We also analyzed the results using proficiency estimates from a procedure that excluded program group status from the set of conditioning variables. This made no difference in the proficiency estimation. The second step is proficiency estimation, discussed in the next section.

C. PROFICIENCY ESTIMATION BY PLAUSIBLE VALUES

After item parameters have been estimated, the pattern of correct responses can be used to estimate the proficiency of any examinee or group of examinees. Specifically, we can express the pattern of item responses as a joint probability of each individual's set of responses to each item, modeled using equation (1). Assuming independence of these events (conditional on the individual's θ_i), the likelihood function becomes:

$$(2) \quad p(Y|\theta_i,a,b) = \prod_{j=1}^{n} [p_j(\theta_i)]^{y_j}[1-p_j(\theta_i)]^{1-y_j},$$

where Y is the vector of item responses $\{y_1, y_2, ...\}$.

We used a method known as "plausible values methodology" to report and analyze individual test scores. As noted, the BIB spiraling design used in the Job Corps literacy assessment was an efficient way to estimate the proficiency of groups but generally yields imprecise estimates of the proficiency of any one individual. Thus, instead of using point estimates of θ_i to calculate the statistics of interest (such as the average proficiency of program group members), we use plausible values methodology. Plausible values methodology begins with a "posterior distribution," an expression that relates every possible proficiency level to the probability that it accurately describes the individual. The posterior distribution is estimated using the model in equation (2). If the posterior distribution, which typically can be thought of as a bell-shaped curve, is more precisely estimated, it would have a higher and narrower center and smaller tails, indicating a higher probability of the person's proficiency being in a narrower range.

In practical terms, it is difficult to compute statistics describing groups of individuals using information from their full posterior distribution. Instead, Mislevy (1991) shows how one can use a small sample of randomly drawn values from this distribution for each person to summarize both

the estimate of the respondent's proficiency and the uncertainty about that estimate and to calculate statistics (functions of those random variables). These randomly drawn scores from the posterior proficiency distribution are called "plausible values." Plausible values are useful because they are sufficient statistics to reasonably approximate any function of proficiency, such as the average proficiency level of a group or the proportion of group members scoring above some cutoff score (for example, to place them in level II or higher.) Normally, to estimate a function of proficiency using the posterior distribution would require evaluating the following integral:

$$(3) \quad E[h(2,X)|Y,X] \;'\; \int_{m} h(2,X)p(2|Y,X)d2.$$

Mislevy shows how this can be approximated using a limited set of "plausible values" drawn from the posterior distribution. For the analysis of Job Corps data, ETS produced five plausible values for each domain for each examinee. We analyzed the data by arbitrarily choosing the first plausible value from the vector. This represents an unbiased estimate of the person's proficiency, while preserving the variability due to individual-level imprecision.

Thus, the estimate of average proficiency for group *j* is:

$$(4) \quad \hat{\mu}_j \;'\; \sum_{i'1}^{n_j} \frac{T_{ij1}}{n_j},$$

where T_{ijk} is the test score for individual *i* in group *j,* using plausible value *k,* and n_j is the number of people in group *j*.

The variance of this mean has two parts, the within-person variance and the between-person variance. Note that the between-person variance uses only the first plausible value:

$$
\begin{aligned}
\sigma_j^2 &= Var_k(\bar{T}_{.jk}) + Var_i(T_{ij1}) \\
&= \frac{1}{4}\sum_{k=1}^{5}(\bar{T}_{.jk}-\bar{T}_{.j.})^2 + \frac{1}{n_j-1}\sum_{i=1}^{n_j}(T_{ij1}-\bar{T}_{.j1})^2. \qquad (5)
\end{aligned}
$$

Alternatively, one can compute the mean using all five plausible values and compute the variance of the mean (or group means) using all five variances for the between-person variance component. In other words, the group mean and variance would be estimated using:

$$
\begin{aligned}
\hat{\mu}_j &= \bar{T}_{.j.} \\
e\sigma_j^2 &= Var(\bar{T}_{.j.}) + \frac{1}{5}\sum_{k=1}^{5} Var(\bar{T}_{.jk}). \qquad (6)
\end{aligned}
$$

These approaches are equivalent. Therefore, we conducted all analyses in this report using the first plausible value but repeated it for each plausible value and for the average of the five. The results were robust to the choice of plausible value, so we only report those results.

The main difference between the plausible value approach and conventional estimation using point estimates of observed test scores is the first variance term in equations (5) and (6). Ignoring this term still provides unbiased estimates but could fail to account for variation at the individual level when conducting hypothesis tests and other forms of inference that rely on the standard error of the estimate. In practice, the adjustment of standard errors for within-person variance in our sample made almost no difference.

APPENDIX C

SUPPLEMENTARY TABLES

TABLE C.1

IMPACTS ON LITERACY SKILLS USING THE SIMPLE DIFFERENCES-IN-MEANS APPROACH

Outcome Measure	Program Group	Control Group	Estimated Impact Per Eligible Applicant[a]	Program Group Job Corps Participants	Estimated Impact Per Participant[b]
Prose Literacy (Percentages)					
Level I	23.7	27.7	-4.0**	23.8	-5.5
Level II	50.3	46.2	4.1	50.0	5.5
Level III	22.3	23.8	-1.4	22.4	-1.9
Level IV	3.7	2.3	1.4	3.9	1.9
Level V	0.0	0.1	-0.1	0.0	-0.1
Average Prose Proficiency	250.3	248.1	2.3	250.8	3.1
Document Literacy (Percentages)					
Level I	19.1	20.3	-1.3	18.7	-1.8
Level II	47.4	47.4	0.0	47.3	0.0
Level III	28.7	28.7	0.0	29.2	0.0
Level IV	4.7	3.4	1.3	4.7	1.8
Level V	0.1	0.2	0.0	0.0	0.0
Average Prose Proficiency	257.5	256.2	1.3	257.9	1.8
Quantitative Literacy (Percentages)					
Level I	40.3	44.0	-3.6	38.8	-4.9
Level II	39.7	37.3	2.4	42.0	3.3
Level III	18.1	16.8	1.4	17.1	1.9
Level IV	1.7	1.7	0.0	1.9	0.0
Level V	0.2	0.2	0.0	0.3	0.0
Average Quantitative Proficiency	234.6	231.3	3.3*	235.0	4.5
Sample Size					

SOURCE: Job Corps Literacy Assessment data.

NOTES: 1. All estimates were calculated using sample weights to account for sample and survey designs and interview nonresponse.

2. Level I scores are between 0 and 225; Level II scores; Level III scores are between 275 and 325; Level IV scores are between 325 and 375; Level V scores are between 375 and 500.

[a]Estimated impacts per eligible applicant are measured as the difference between the weighted means for program and control group members.

[b]Estimated impacts per Job Corps participant are measured as the estimated impacts per eligible applicants divided by the weighed proportion of program group members who enrolled in Job Corps. Standard errors for these estimates were inflated to account for the estimation error in the Job Corps participation rate.

[c]The significance levels pertain to statistical tests for differences in the distribution of the outcome measure for the program and control group members.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

TABLE C.2

PARAMETER ESTIMATES FROM MODELS ESTIMATING PROGRAM IMPACTS, BY LITERACY DOMAIN
(T-Statistics in Parentheses)

	Dependent Variable		
Control Variables[a]	Prose	Document	Quantitative
Intercept	227.4	233.3	208.6
	(18.8)	(9.5)	(14.9)
Assigned to Job Corps Program Group	2.7	1.1	3.6
	(1.7)	(0.7)	(1.9)
Months Since Random Assignment	-0.3	-0.1	0.1
	(0.9)	(0.2)	(0.3)
Female	9.6	9.1	6.6
	(5.7)	(5.5)	(3.4)
Age at Application to Job Corps (16 to 17 is the omitted category)			
18 to 19	2.3	0.2	1.1
	(1.1)	(0.1)	(0.4)
20 to 24	1.6	0.9	1.6
	(0.6)	(0.3)	(0.5)
Race/Ethnicity (White is the omitted category)			
Black	-15.5	-19.0	-23.7
	(7.7)	(9.5)	(10.1)
Hispanic	-5.1	-6.3	-10.6
	(1.5)	(1.9)	(2.7)
American Indians, Alaskan Natives, Asians, and Pacific Islanders	-4.8	-7.6	-11.4
	(1.3)	(2.1)	(2.6)
Native Language (English is the omitted category)			
Spanish	-8.9	-6.1	1.0
	(2.2)	(1.5)	(0.2)
Other	-13.1	-15.5	-5.2
	(3.0)	(3.7)	(1.0)
Mother Has a High School Diploma	1.6	1.6	1.2
	(4.3)	(4.1)	(2.8)

TABLE C.2 *(continued)*

Control Variables[a]	Dependent Variable		
	Prose	Document	Quantitative
Educational Attainment (completed less than grade 9 and has not GED is the omitted category)			
Completed grades 9-11 and has no GED	3.8	1.8	4.8
	(1.5)	(0.7)	(1.6)
Had a GED	14.0	13.8	21.7
	(4.3)	(4.3)	(5.7)
Had a High School Diploma	34.0	27.5	37.4
	(7.1)	(5.8)	(6.7)
Had a Vocational Certificate	-2.0	-6.4	-10.8
	(0.4)	(1.2)	(1.8)
Number of Education Programs Attended in the Past Year	1.7	2.1	1.9
	(1.5)	(1.8)	(1.4)
Most Recent Program Attended in the Past Year was a Vocational Program	7.1	5.7	6.2
	(2.3)	(1.9)	(1.7)
Number of Jobs in the Past Year	3.5	3.4	3.4
	(3.4)	(3.4)	(2.8)
Usual Weekly Hours Worked on the Most Recent Job in the Past Year (<20 is the omitted category)			
20 to 29	-3.0	-1.5	-5.1
	(0.9)	(0.5)	(1.4)
30 or more	-6.2	-3.5	-7.8
	(2.3)	(1.3)	(2.5)
Hourly Wage on the Most Recent Job in the Past Year (<$4.25 is the omitted category)			
$4.25	4.0	2.6	3.8
	(1.2)	(0.8)	(1.0)
$4.25 to $5.00	3.6	1.8	4.9
	(1.1)	(0.5)	(1.2)
$5 to $6.50	7.9	6.2	10.9
	(2.5)	(2.0)	(3.0)
$6.50 or more	12.3	8.9	16.3
	(2.9)	(2.2)	(3.4)

TABLE C.2 *(continued)*

Control Variables[a]	Dependent Variable		
	Prose	Document	Quantitative
Missing	-3.3	-4.0	-1.0
	(0.8)	(1.0)	(0.2)
Amount of Time Received Welfare While Growing Up (Never is the omitted category)			
Most of the Time	-0.5	-0.8	-3.6
	(0.2)	(0.3)	(1.3)
Sometimes	4.0	2.9	2.0
	(2.1)	(1.6)	(0.9)
Welfare Receipt in the Past Year			
Received Welfare	-2.7	-0.5	-1.3
	(1.5)	(0.3)	(0.6)
Missing	-7.6	-5.9	-7.8
	(2.3)	(1.8)	(2.0)
Health Status (in poor or fair health is the omitted category)			
Good	6.2	4.9	3.1
	(2.4)	(1.9)	(1.0)
Excellent	5.0	4.7	3.4
	(1.9)	(1.8)	(1.1)
Had Health Problem for >5 Years	-9.3	-9.7	-9.8
	(1.9)	(2.1)	(1.8)
Months From Random Assignment Until Skills Test Was Completed			
R^2 Value	.13	.13	.14
Sample Size	**2,273**	**2,273**	**2,273**

SOURCE: Job Corps Literacy Assessment data on 1,117 program and 1,156 control group members.

NOTE:

[a]Explanatory variables refer to baseline characteristics.

TABLE C.3

PARAMETER ESTIMATES FROM MODELS ESTIMATING THE ASSOCIATION BETWEEN HUMAN CAPITAL MEASURES AND SKILL SCORES, BY LITERACY DOMAIN
(T-Statistics in Parentheses)

	Literacy Doman		
Control Variables	Prose	Domain	Quantitative
Intercept	225.1 (18.4)	232.1 (19.2)	203.0 (14.3)
Human Capital Variables Pertaining to the 30 Months After Random Assignment			
Total Hours Worked (in 1,000s)	1.8 (3.8)	1.6 (3.3)	2.6 (4.7)
Total Hours in Academic Classes (in 1,000s)	-1.3 (1.0)	-1.4 (1.1)	-0.3 (0.2)
Total Hours in Vocational Training (in 1,000s)	5.3 (3.4)	4.9 (3.2)	6.4 (3.5)
Baseline Variables			
In the Program Group	0.6 (0.4)	-0.7 (0.4)	1.0 (0.5)
Female	10.1 (6.0)	9.5 (5.7)	7.3 (3.7)
Age at Application to Job Corps (16 to 17 is the omitted category)			
18 to 19	1.3 (0.6)	-0.8 (0.4)	0.1 (0.1)
20 to 24	-0.0 (0.0)	-0.6 (0.2)	-0.1 (0.0)
Race/Ethnicity (White is the omitted category)			
Black	-14.1 (6.9)	-17.7 (8.7)	-22.1 (9.3)
Hispanic	-4.1 (1.2)	-5.4 (1.6)	-9.4 (2.4)
American Indians, Alaskan Natives, Asians, and Pacific Islanders	-3.2 (0.8)	-6.1 (1.6)	-9.4 (2.1)

TABLE C.3 *(continued)*

Control Variables	Literacy Domain		
	Prose	Domain	Quantitative
Native Language (English is the omitted category)			
Spanish	-9.0	-6.1	0.6
	(2.3)	(1.6)	(0.1)
Other	-13.0	-15.4	-5.3
	(3.0)	(3.6)	(1.1)
Mother Has a High School Diploma	1.6	1.5	1.1
	(4.1)	(4.0)	(2.6)
Educational Attainment (completed less than grade 9 and has not GED is the omitted category)			
Completed grades 9-11 and has no GED	3.5	1.6	4.3
	(1.4)	(0.6)	(1.5)
Had a GED	32.6	26.3	35.5
	(6.8)	(5.5)	(6.4)
Had a High School Diploma	12.3	12.3	19.4
	(3.8)	(3.8)	(5.2)
Had a Vocational Certificate	-2.9	-7.2	-12.1
	(0.5)	(1.4)	(2.0)
Number of Education Programs Attended in the Past Year	1.4	1.8	1.4
	(1.2)	(1.6)	(1.1)
Most Recent Program Attended in the Past Year was a Vocational Program	6.9	5.5	6.0
	(2.3)	(1.8)	(1.7)
Number of Jobs in the Past Year	3.1	3.1	2.9
	(3.0)	(3.0)	(2.4)
Usual Weekly Hours Worked on the Most Recent Job in the Past Year (<20 is the omitted category)			
20 to 29	-3.3	-1.7	-5.4
	(1.0)	(0.5)	(1.4)
30 or more	-6.6	-3.9	-8.3
	(2.5)	(1.5)	(2.7)

TABLE C.3 *(continued)*

Control Variables	Literacy Domain		
	Prose	Domain	Quantitative
Hourly Wage on the Most Recent Job in the Past Year (<$4.25 is the omitted category)			
$4.25	3.7	2.4	3.3
	(1.1)	(0.7)	(0.9)
$4.25 to $5.00	3.0	1.2	4.0
	(0.9)	(0.4)	(1.0)
$5 to $6.50	7.0	5.5	9.6
	(2.2)	(1.8)	(2.6)
$6.50 or more	11.1	7.9	14.7
	(2.7)	(1.9)	(3.1)
Missing	-4.0	-4.5	-2.0
	(1.0)	(1.1)	(0.4)
Amount of Time Received Welfare While Growing Up (Never is the omitted category)			
Most of the Time	-0.5	-0.7	-3.5
	(0.2)	(0.3)	(1.3)
Sometimes	4.1	3.0	2.2
	(2.2)	(1.6)	(1.0)
Welfare Receipt in the Past Year			
Received Welfare	-2.2	-0.1	-0.7
	(1.2)	(0.1)	(0.3)
Missing	-7.6	-5.8	-7.8
	(2.3)	(1.8)	(2.0)
Health Status (in poor or fair health is the omitted category)			
Good	5.0	4.7	3.2
	(1.9)	(1.9)	(1.1)
Excellent	6.1	4.9	2.9
	(2.4)	(1.9)	(1.0)
Had Health Problem for >5 Years	-9.3	-9.7	-9.7
	(1.9)	(2.1)	(1.8)
Months From Random Assignment Until Skills Test Was Completed	-0.3	-0.1	0.1
	(0.9)	(0.3)	(0.4)

TABLE C.3 *(continued)*

	Literacy Doman		
Control Variables	Prose	Domain	Quantitative
R^2 Value	.16	.15	.16
Sample Size	**2,273**	**2,273**	**2,273**

SOURCE: NALS test data on 1,117 program and 1,156 control group members.