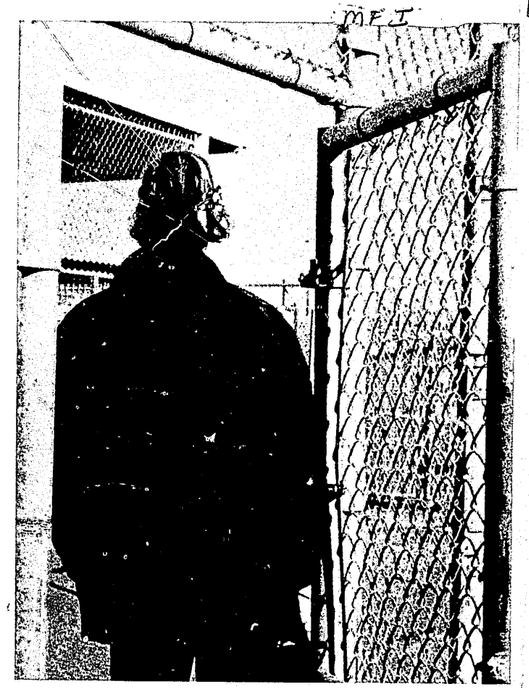

COMPENSATORY EDUCATION AND CONFINED YOUTH:

A FINAL REPORT

VOLUME 5

NATIONAL EVALUATION OF TITLE I PROGRAMS FOR NEGLECTED OR DELINQUENT YOUTH IN STATE INSTITUTIONS

COMPENSATORY EDUCATION AND CONFINED YOUTH: A FINAL REPORT

National Evaluation of Title I Programs in State Institutions for Neglected or Delinquent Youth

VOLUME 5

Judy C. Pfannenstiel
J. Ward Keesling

June 1980

This study was supported by and forms a part of the National Evaluation of Title I Programs for Neglected or Delinquent Youth in State Institutions. United States Office of Education Contract Number 300-76-0093

Education Studies and Evaluation Department
System Development Corporation
Santa Monica, California

TM-5792/015/00

Project Officer, Consultants
and Resource Persons, and
Project Personnel

Project Officer

Janice Anderson

Consultants and Resource Persons

Eric C. Brookens
Thomas J. Cook
Joyce DeMuth
LaMar T. Empey
Edmund J. Gubbins
Chester W. Harris
Tina M. Jackson
Elaine B. Jenkins
Jeremy J. Kollas
John B. Lampe
Robert L. Linn
William Lobosco
Patrick O. Mancini
Lawrence D. Penny
Melvin Seeman
George L. Smith
Frances A. Tyler
Mildred R. Walton
James R. Wickman
Alan Woods

Project Co-Directors

J. Ward Keesling
Judy C. Pfannenstiel

Former Project Staff

Ted Bartell, former Project Director
Deborah S. Brown
Sheila Cottral
Pamela DeFusco
William J. Doherty
Judith A. Hemenway
Richard J. Kwan
Kean A. Mantius
Suzy Navarrete
Alice Neal
Susan Schwartz
Brian Stipak
Ming-mei Wang
Noreen Webb

Project Secretary

M. Helen Hernandez

Field Data Collection

James Hoyt, Manager of Field data
Collection
Raydean Acevedo
Claudia Browne
Sue Brown
Cathie Feulner Farrar
Pat Fornash
Mel McCane
Elise Phillips
Doug Robertson
Mary Sewell
Cathe Talavera
Jeffrey Wallace
Brenda Ward
Renee Wolf

ACKNOWLEDGEMENTS

We would like to extend our deepest appreciation to Dr. Jan Anderson, USOE Project Monitor, for providing support when support was needed, criticism when it was needed, and consistent interest in the project over the four years of its duration. We would also like to thank the members of the Policy Advisory Group and the Research Advisory Group for their continued participation and their many valuable suggestions for the design and implementation of the project.

TABLE OF CONTENTS

TABLE OF CONTENTS (Cont'd)

TABLE OF CONTENTS (Cont'd)

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

To expand and improve the education of confined youth, the Congress has provided funding for basic-skills programs in institutions for neglected or delinquent youths, and in adult correctional institutions housing youthful offenders. This funding stems from Title I of the Elementary and Secondary Education Act of 1965, amended in 1972. Title I guidelines state that the funds should be used to provide services to residents who are under 21, have not received high school diplomas, and demonstrate educational need. These services are intended to concentrate on basic reading and math skills, and should supplement rather than supplant other educational programs in the facility.

The objectives of Congress in providing these funds are to help the institutionalized population build basic skills and to overcome the effects of prior academic failure. These Title I funds have been available to state institutions since 1967, as authorized by P.L. 89-750, passed November 3, 1966.

The National Evaluation of Title I Programs for Neglected or Delinquent Youths in State Institutions was mandated by Part (a) of Section 151 of P.L. 89-10. Under contract to the Office of Evaluation and Dissemination of the U.S. Office of Education, System Development Corporation (SDC) has conducted an evaluation of the programs since February 1976. This is the final report of that evaluation effort. This document contains a summary of the cumulative findings and a summary of each of the activities that constituted the phases of the evaluation.

The summary of findings is presented first. The findings are organized in response to the following six questions:

I) How is the Title I program operating in its environment?

II) What is the impact of the Title I program on student participants?

III) What are the important characteristics of basic-skills programs in correctional settings that have demonstrated some success?

IV) What factors impede the realization of Title I objectives?

V) What are the experiences of students after they are released from institutions?

VI) How may future Title I evaluations at the institutional and state levels be most fruitfully conducted?

The information utilized to address these questions was obtained from the various phases of the evaluation, each of which was designed with a specific study objective. The first objective, accomplished during Phase I of the evaluation, was to describe a nationally representative sample of 100 Title I programs in state-operated institutions. These descriptions provided information about the recipients of Title I services and the ways in which Title I programs had been implemented. The second objective, accomplished during Phase II of the evaluation, was to measure the impact of the Title I program on the basic-skills achievements of participants in thirty randomly selected institutions. The third objective, accomplished with the Substudy of Effective Practices, was to identify those characteristics of basic-skills programs that seem to result in greater achievement gains. The fourth objective was to conduct a follow-up study of students who had been released from institutions in order to describe their post-release experiences in terms of school entry, employment success, subsequent offense violations, and reinstitutionalization. A survey was also conducted of the availability and nature of pre- and post-release services that students receive from correctional institutions and community agencies. The fifth and final objective of the evaluation was to develop evaluation models and reporting forms to be used by the grantees themselves in conducting legislatively required evaluations of Title I programs.

In the chapters following the Summary of Evaluation Findings, the six questions addressed by the evaluation are presented. The chapter headings, it should be noted, are abstracts of the questions and correspond to them in content and number. In order to facilitate the reading of this report, much of the detail concerning the evaluation methodology has been placed in methodological summaries of each phase of the evaluation, which are presented as appendices to this report. As these are summaries, the reader is referred to individual reports for each phase of the evaluation for a complete documentation of the methodology.

SUMMARY OF EVALUATION FINDINGS

I. How is the Title I Program Operating in its Environment?

- Approximately 600 state institutions nationwide were found to be eligible for Title I funds. Based on Fall 1976 counts of participants, it was estimated that about 26,840 students nationwide are served by Title I programs as compared to about 51,310 students estimated to be eligible for services at that time.
- Apparently in response to the widespread need for services, nearly one-half of the institutions reported that they provide some Title I services to all of their eligible students. Institutions for adults, however, reported difficulties in providing Title I services because there were so few eligible residents in each institution.
- The results of achievement testing confirmed the widely held notion that students in state institutions for neglected or delinquent youth perform very poorly on tests of basic skills. The average performance of institutionalized Title I students, who are 16.5 years old on the average, was similar to that of non-institutionalized fourth- and fifth-graders, who are 9-10 years of age. Institutionalized students not in Title I (who average 18 years of age) performed at about the level of fifth- and sixth-graders.
- Most Title I programs focused on basic skills in reading and mathematics. Instruction in these programs is highly individualized.
- Students were found to have an average length of stay of six months. During their institutional confinement, basic-skills students were reported to receive about three hours of weekly reading exposure, and 2.5 hours of weekly math exposure.

II. What is the Impact of the Title I Program on Student Participants?

- The evaluation of the impact of Title I programs found no overall growth for Title I students. Achievement levels and attitudinal measures remained stable over the 12 weeks between the first and third testing.

On the average, neither the Title I students nor their peers who were not in Title I supplementary programs showed any meaningful change on these measures over this time period.

- Although no change was demonstrated on the average, some institutions were found that were effective in producing gains in reading and/or mathematics.
- Among those students who maintained their level of performance or improved over time, Title I students showed no greater average gains than did students receiving the regular education curriculum.

III. *What are the Important Characteristics of Basic-Skills Programs in Correctional Settings That Have Demonstrated Some Success?*

- Student-level analyses revealed significant, but modest, positive correlations between the amount of instructional exposure and basic-skills achievement.
- In classes where audiovisual equipment, textbooks and programmed instruction were utilized, teachers were observed to be disengaged from the instructional process, and reading and math gains were lower. When high-interest, low-skill-level materials were used, reading and math gains were larger. In other classes, where teachers or aides provided instruction directly, reading gains were larger. The fewer students per teacher and the less time teachers engaged students in non-task-related conversations, the greater were the math and reading gains.
- Positive feedback (e.g., praise) characterizing teacher-student interactions was inconsistently related to achievement gains. Positive feedback was negatively related to gains in reading, but positively related to gains in math. This may occur because positive feedback in reading is not as tied to specific skill achievements and may seem insincere or patronizing.
- The greater the proportion of students' time that was engaged in the receipt of instruction and feedback, the lesser the amount of time that was devoted to practice or drill, and the higher the ratings of

teacher supportiveness, the greater was the proportion of student task-engaged activity.

- Although positive student attitudes were not found related to gains in achievement, changes in attitudes were found to be significantly related. The greater the change towards more favorable reports of self-esteem, of school involvement, and of perceptions of the practical orientation of correctional education programs, the larger were the achievement gains.
- Institution-level analyses revealed that smaller educational programs, with low rates of staff turnover, which are directed by administrators who encourage innovations, and which operate within institutions that consider custody or security concerns to be less important than education or counseling, are more likely to demonstrate gains in basic-skills achievements.

IV. *What Factors Impede the Realization of Title I Objectives?*

- When all class time scheduled for basic-skills instruction is considered, more than one-half was consumed by non-instructional activities. In general, student absences from classes accounted for more than one-half of this non-instructional time. The predominant reasons for student absences were institutional assignments to other activities, including disciplinary procedures.
- Of scheduled class time that was non-task-related, only 25 percent was due to factors not directly under staff control, such as lack of student motivation, inattention, or absences due to illness. This finding questions the notion that failures to demonstrate effectiveness are largely a function of poor student motivation to learn.
- The remaining 20 percent of non-task-related time was found to be related to practices or occurrences within the classroom itself. Smoking breaks, clean-up activities, roll-taking and tracking missing students, the application of discipline, and the engagement of students and teachers on non-task-related conversations accounted for this non-task time.

- State and institutional Title I administrators expressed difficulties with federal guidelines (and/or state interpretations of them) that prevented programs from serving needy students or from providing needed services.

V. *What Are the Experiences of Students After They Are Released From Institutions?*

- Few students released from state institutions were found to have completed high school or to have obtained a General Equivalency Diploma (GED) at the time of their release. Approximately one-half of the 195 releasees studied were found to have entered school following their release from the institutions.
- The released students confirmed earlier reports of institutional staff concerning the expected difficulties in the transition to school. One-half of the students entering schools reported having difficulties in transferring to community schools, most often due to their perceiving themselves as performing at a level well below that expected of them by the receiving schools. Eighty percent of students returning to school dropped out prior to the completion of the school term.
- Released students indicated a widespread desire for continued schooling Eighty-five percent of releasees (including those who failed to enter school and those who entered and dropped out) intended to obtain at least a high school diploma or GED.
- Between 40 and 50 percent of released students were employed at the time of the follow-up interview. Employment was generally unskilled labor, and widespread job dissatisfaction was reported. Seventy-two percent of interviewed students had obtained employment at some point after their release, and almost all releasees had at least attempted to get a job.
- An estimated 30-40 percent of released students reported further offense violations within the three-to-six-month follow-up period. Between 10-20 percent of released students had been reinstitutionalized in either state or local institutions by the time of the follow-up reports.

- Institutions provide few pre-release or liaison services, often because they are located far from the communities to which their students will be returned, and have little contact with service agencies in those communities. Many of the institutions would have to coordinate services with a large number of agencies in several communities in order to provide all releasees with adequate services.
- Parole agencies were found to be the major providers of post-release and transitional services. The adequacy of parole services, however, is questionable, since parolees reported infrequent interactions with parole officers, and expressed desires that their interactions be even less frequent.

VI. *How may Future Title I Evaluations at the Institutional and State Levels be Most Fruitfully Conducted?*

- A review of past evaluation reports submitted to USOE revealed that more than one-half of the states do not submit the required annual Title I evaluations.
- A lack of uniformity in the content and methodologies of the reports that are submitted renders them useless as a basis for a nationwide description or assessment of the Title I program.
- Two evaluation models, the Systematic Allocation Model and the Criterion Model, were developed to remedy this lack of uniformity among state evaluations of the Title I program. A reporting mechanism for the conduct of state evaluations has also been proposed and is being evaluated.

CHAPTER I. TITLE I PROGRAM ENVIRONMENT

In 1976, slightly more than 600 state institutions nationwide were found to be eligible for the receipt of Title I funds for neglected or delinquent youths.* Eligible institutions were comprised of three types of administratively operated agencies and/or resident populations: (1) institutions for neglected children (operated by a Human Resources Agency, for instance) housing residents who had been judged to be neglected or dependent children; (2) institutions for delinquent youths (operated by a Youth Services Agency, for instance) housing residents who had been judged to be either delinquents or status offenders;** and (3) institutions for adult offenders (operated by Departments of Corrections) housing offenders who had been convicted of generally more serious offenses in the adult court system, and were committed to the more security-oriented adult institutions.

Although 600 institutions had been identified as eligible for receipt of Title I N or D funds, not all institutions were receiving them. Generally, these were institutions for adult offenders, housing small numbers of eligible residents. Indeed, in 124 of the 600 institutions, fewer than 10 residents were eligible for Title I services. In these institutions services were either not provided or were provided to so few students that an evaluation would be difficult or impossible. Thus, these 124 institutions are not represented in this national evaluation.

NUMBERS OF ELIGIBLE AND SERVED STUDENTS

Based on data obtained from 100 randomly selected institutions, it was estimated that 26,840 students residing in state institutions nationwide were receiving

---

*Title I guidelines state that funds should be used to provide services to residents who are under 21 years of age and who have not received high school diplomas. Services are to be concentrated on basic reading and math skills, and should supplement, rather than supplant, other educational programs in the institution.

**"Status Offenders" are youths convicted of crimes that are only crimes by virtue of the offender's youthful age, e.g., runaway, truancy, incorrigibility.

Title I services on a typical day in the fall of 1976. This compares to the estimated total eligible population of 51,310. About two-thirds of both the eligible and the participating populations reside in institutions for delinquent youths. Less than 10 percent of Title I participants reside in institutions for neglected youths, and the remaining 24 percent of participants reside in institutions for adult offenders.

Institutions for adults were reported to house a fairly large percentage of the eligible population (43%), yet were reported to serve only about one-third of their eligible students.* Sixty-seven percent of adult institutions have fewer than 45 eligible students and the administrative difficulties of establishing and supporting a Title I program for these few students may not seem to be worth the effort. In some cases these Title I-eligible students are served by state-supported basic-skills programs.

REGULAR EDUCATION AND TITLE I PROGRAM OFFERINGS IN STATE INSTITUTIONS

In general, state institutions were found to be operating their own regular education and Title I programs on site. A few institutions were found to have contracted with a local school district for services, such as diagnosis of special learning problems, but only in a few institutions for neglected children did some students attend public schools. Even in these situations, Title I services such as tutoring or individual counseling were also provided by institutional staff. In a few other situations, such as in community-based facilities, local school districts were reported to provide services, but all instruction was provided within the institutional setting, and the relationship with the local school district appeared to be merely an administrative arrangement where institutional teachers were paid through local school districts.

A wide variety of education programs were found to be offered in institutions operating Title I programs. State-funded educational programs as well as

*This estimate may be low because institutions for adult offenders only became eligible for Title I funds in 1973, and many projects might have been at early stages of implementation at the time of the survey. Other Title I programs might have started since the survey.

Title I programs focused on basic skills, but other academic and vocational courses were also provided within the regular education program.

Ninety percent of the institutions for neglected youths offered some form of elementary-level academic education, and 70 percent offered high school-level academic education as well. In institutions for delinquent offenders, the predominant academic program was high school-level education, while in institutions for adults the emphasis was on preparing students to obtain General Equivalency Diplomas (GEDs). Remedial programs other than Title I are offered in roughly 40 percent of the institutions for neglected or delinquent youths, and in 30 percent of the institutions for adult offenders. Reportedly, the only difference between Title I programs and other basic-skills programs in many institutions for adult offenders was the age of the participating students, since students over 21 years of age were not usually allowed into Title I classes. In addition to academic instruction, 90 percent of the institutions for neglected youths, 78 percent of the institutions for delinquent youths, and 97 percent of the institutions for adults offered vocational programs. However, vocational education courses were offered to a substantially smaller proportion of students than were academic courses.

About one-third of the classes offered in correctional education programs were Title I classes.* Without distinguishing Title I classes from regular education classes, 29 percent of classes offered in institutional education programs were in reading or language arts, 26 percent were in mathematics, 20 percent were in other academic subjects, and 25 percent were in non-academic areas.

DEFINITION OF "TITLE I"

Visits to the 100 sites nationwide revealed that Title I programs had been implemented in widely varying ways across the country. In some cases, programs were clearly structured and identifiable among the various instructional offerings. Title I recipients in these situations were a targeted population receiving

*These data were obtained from the teachers of students selected for testing in the 30 randomly sampled institutions participating in the impact evaluation.

supplementary Title I services. However, the nature of the 'program' varied considerably from place to place. In some institutions Title I students were those assigned to individualized instruction in basic-skills labs; in other institutions they were identified as those receiving instruction from teachers whose salaries were paid with Title I funds; in still other institutions Title I students were identified as those who utilized materials and equipment paid for with Title I funds. In other institutions, Title I services were identifiable, but the entire resident population received such services. All residents in these institutions were reported to be severely deficient in basic skills. Thus, the "neediest of the needy" were not exclusively targeted, and Title I services were provided to all residents. While it is difficult to argue with the practice of providing Title I services to all institutionalized students deficient in basic skills, this practice does have implications for the conduct of institution-level evaluations that will be discussed further in Chapter VI. In a few institutions, the services purchased with Title I funds, and the recipients of these services, could not be readily identified. Indeed, during the impact evaluation phase of the study, educational program administrators in two of the institutions studied were unable to make this distinction.

The presence or absence of a clear definition of what the Title I program incorporated (i.e., instructional services, instructional supplies, staff, and students) was closely related to the objective of a smooth integration of Title I and state education programs. The integration of the Title I and regular education programs was reported by institutional staff and observational teams to be well coordinated, but that degree of integration was frequently possible for the very reason that Title I services were virtually indistinguishable from other educational services. The accomplishment of the objective of program integration, while simultaneously implementing a Title I program that is distinctive and can be identified for evaluation purposes, proved to be a difficult task for several institutions.

At some sites where Title I programs were readily identifiable, the mere fact of the identification as Title I led to a poor integration of program services. Title I staff in these situations often received differential treatment that was seen as favorable and was resented by state-funded staff. For instance,

Title I teachers were allowed to participate in inservice education workshops, taught their students in smaller instructional groupings, were perceived as receiving better instructional supplies and equipment, and had no non-instructional tasks as did other teachers (e.g., student supervision during non-class time).*

Some institutions were found, however, where both objectives were successfully accomplished: Title I program services were distinct from other educational services, and yet they appeared to be well integrated with them in terms of objectives, etc. In such situations, Title I staff were perceived as resource personnel assisting in the instruction of highly deficient and/or difficult-to-handle students, providing them with intensive individualized instruction. To avoid differential staff treatment in these situations, and where state funds were not also available for inservice education for state-supported teachers, no special privileges were allowed to Title I teachers.

During the course of the three years over which data were obtained from institutional education programs, study staff noted an increasing trend by educational program administrators (EPAs) to define the Title I program and Title I participants more clearly. This trend was probably influenced by the increased attention given to Title I programs by federal program administrators, combined with the reactive effects of the evaluation itself.

ASSIGNMENT OF STUDENTS TO REGULAR EDUCATION AND TITLE I CLASSES

The EPA at each institution participating in the impact evaluation was asked to describe how residents were selected for participation in the education program (since not all institutionalized students are bound by compulsory school attendance laws). Nearly all institutions for neglected or delinquent youth regarded education as the major activity for residents, and thus assigned all students to the educational program for a varying number of courses. Most of the adult institutions had mechanisms for selecting residents, such as varying

*Program staff in the Office of Education report that the Education Amendments of 1978, which postdate much of the study, changed some of these negative aspects of distinguishing Title I from other programs.

combinations of willingness to participate, recommendations by institutional staff, and the results of diagnostic test scores demonstrating academic deficiencies.

The basis for assigning students to Title I classes proved to be more difficult to identify. Federal guidelines direct that only students eligible for Title I services (those under 21 years of age, lacking a high school diploma or General Equivalency Diploma and showing an educational need) may be enrolled in a Title I program. It is further specified that students most in need of basic skills instruction (the "neediest of the needy") should have the highest priority for receipt of Title I services. However, there is no prescribed method for determining which students are most in need of Title I instruction, and institutions vary on how students are assigned to Title I instruction or regular instruction.

The EPA and Title I reading and mathematics teachers were asked to describe how students were assigned to Title I instruction. Nearly one-half of the institutions assigned all students or all eligible students to Title I programs. The remaining institutions used a mixture of test scores, staff recommendations, and student willingness to participate in order to assign students to Title I instruction. The reports of Title I teachers describing selection procedures generally agreed with those of the EPA.

Student assignment to Title I classes was usually combined with their assignment to regular education classes. In some institutions regular education class offerings did not include basic-skills classes, specifically to avoid duplication of Title I course offerings. In other institutions regular education course-offerings specifically included basic-skills courses to avoid allegations that the Title I program was supplanting the regular education program. Thus, the group of students identified as participants in regular education only differed widely in basic-skills performance among institutions where program participants were selected. In many situations, students who received only regular instruction were not deficient in basic skills. In other situations, where a selection of the neediest students for Title I participation was reported, regular education students were also deficient, but to a lesser degree than were Title I students.

In institutions that did select students for participation in Title I (rather than serve them all), it was difficult to determine how systematically the selection process was carried out. There are several features operating in these settings that may work against systematic assignment. First, student requests for participation (usually related to the institution's system of rewards) were sometimes the basis for assignment, suggesting that assignment may be based more on motivation than on need. Secondly, institutions could not always require participation in a specific educational program because many students were beyond the age of compulsory attendance. Thirdly, students may arrive at the institution at a time when available openings in the educational program are not necessarily in the classes to which they should be assigned based on need alone.

In an attempt to assess the extent to which needier students were assigned to Title I, the average level of Title I student performances on tests of basic skills was compared to the performance of regular education students. The lower average score of the Title I students indicated that, in general, the neediest students were receiving Title I services. However, the scores of 30 percent of the students who were not served by Title I were lower than the average score of students who were served by Title I. Using criteria other than basic-skills deficiency alone could account for some of this overlap, but unsystematic assignment is also a likely cause. As systematic assignment is necessary to certain forms of evaluation, the subject will be discussed again in Chapter VI.

CHARACTERISTICS OF TITLE I AND REGULAR EDUCATION STUDENTS

Descriptions of the Title I and regular education students were obtained from questionnaires administered to students during the Phase II impact evaluation. These provide an estimate of the nationwide characteristics of Title I and regular education student populations in institutional settings.

The average age of students within correctional education programs was 17.3 years. Students in institutions for the neglected were youngest, with ages ranging from 10 to 17, averaging about 15 years. Students in institutions for delinquent youths ranged from 13 to 21 and averaged 16.5 years. Students in institutions

for adult offenders were oldest, averaging about 20 years. The range in average age across institutions was large, from about 13 to over 21. Title I students were found to be 1.6 years younger, on the average, than students who participated in regular education only ($p < .001$).

The majority of institutionalized students was male. The population of students at every sampled institution for adult offenders (as well as the population at one-half of the institutions for delinquent youths) was entirely male. One-half of institutions for neglected youths and delinquent youths housed both males and females, with males generally in the majority. Only a few entirely female institutions were found among the 100 originally sampled institutions, suggesting that only a small proportion of females are institutionalized in state facilities.

In terms of racial and ethnic differences among students, an average of 43 percent of institutionalized students were non-minority; 40 percent were black, and 17 percent were other minority members. The percentage distribution of students among racial and ethnic categories within institutions varied considerably. A significantly higher proportion of students who were assigned to regular education only in the 30 institutions were non-minority students.

Half of the offenses that students reported as the reasons for their institutionalization were property-related (e.g., burglary, larceny, auto theft, etc.). Just over one-fourth were person-related offenses (e.g., assault, battery, robbery, sexual assault, homicide, etc.). Seven percent of respondents reported drug-related offenses as the reason for institutionalization, while 14 percent reported that they were incarcerated for status offenses. The remainder were neglected or dependent children. The distribution of offenses was similar for Title I and regular education students.

On the average, students had been previously institutionalized between two and three times. Title I students had been previously institutionalized more often than regular education students ($p = .001$). The estimated average length of stay was six months overall, 10 months for students residing in institutions for

neglected youths, and 5.5 months for those residing in institutions for delinquent youths and adult offenders. These estimated averages varied greatly across institutions, with relatively large numbers of neglected and adult institutions holding residents more than 12 months.

On the average, students reported that they had last attended the ninth grade before being institutionalized. A comparison of the average age of institutionalized students (17 years) to the average age of ninth-grade students (15 years) provides a rough estimate of the length of time these students have been absent from public schools.

BASIC SKILLS PERFORMANCE OF TITLE I AND REGULAR EDUCATION STUDENTS

Previous estimates of the performance level of institutionalized students, although none claimed to be nationally representative, were uniformly low (Dell'Apa, 1972; Reagen and Stoughton, 1976). The findings of this evaluation from standardized tests of basic skills obtained on representative samples of Title I and regular education students indicate that the entire population of institutionalized students performs well below their peers in public schools. Testing with the Comprehensive Tests of Basic Skills indicated that institutionalized Title I students (who are 16 to 17 years old, on the average) perform similarly to fourth-grade public school students who do not receive compensatory education, while institutionalized regular education students (who are about 18 years old, on the average) perform at the fifth-grade level. It seems clear that institutionalized students are highly deficient in basic skills compared to their non-institutionalized counterparts.

CHARACTERISTICS OF TITLE I INSTRUCTION

The content of Title I classes consisted primarily of basic-skills instruction in reading and mathematics. Other Title I program offerings included cultural enrichment courses, career education, vocational education, bilingual education, special education, and counseling services. Within the two major instructional areas, reading instruction emphasized vocabulary, literal comprehension, and the following of directions, while the emphasis in math was on fundamental

operations (e.g., addition, subtraction, multiplication and division), practical mathematics and word problems.

Classes in 100 institutions were observed by the SDC data collection team in order to obtain information on the methods employed and the physical characteristics of a representative sample of Title I and regular education classrooms. Observers provided a measure of the amount of structure existing in classroom settings. Structure was defined as the degree to which the teacher assigns tasks (structured) as opposed to the students' choosing the tasks (unstructured). Over 60 percent of the classes were rated as completely or mostly structured, although there appeared to be somewhat more emphasis on an unstructured approach in institutions for delinquent youths and adult offenders. Observations as to whether students worked individually, in small groups or as part of the entire class indicated that most instruction was individualized.

As part of the impact evaluation phase, the comparison of Title I teachers of confined youth and similar compensatory education teachers in public school settings revealed that institutional teachers place more emphasis on individualized instruction and less emphasis on small group instruction than do their public school counterparts. This is probably explained by the fact that the range in student age and achievement level in institutional settings is greater than the range in public school compensatory education settings. To respond to student needs, then, teachers of confined youth may view the assignment of tasks to individual students as the best way to match instruction to a student's achievement level.

In terms of the physical characteristics of classrooms, institutional classrooms appeared similar to classrooms in non-institutional settings. Adequate space, ventilation and lighting were observed in most classrooms. However, a significantly smaller proportion of institutions for adults was observed to have adequate working space or attractive furnishings. Special-purpose areas (e.g., media centers) were observed in approximately one-half of the observed classrooms. In terms of the availability and quality of materials and audio-

visual equipment, regular education classrooms appeared less well-equipped than were Title I classrooms.

STAFF CHARACTERISTICS

Nearly all of the institutions reportedly require state certification of their teachers. Other important criteria in teacher selection were reported to be the amount of coursework in the instruction of low-level achievers, and experience with teaching disadvantaged students. Almost all of the educational administrators were reported to have graduate degrees or at least graduate-level coursework; over 40 percent of Title I teachers had also attained this level of formal education.

Reports of teaching assignments for teachers of basic-skills classes indicated that 42 percent of teachers taught both reading and mathematics classes, 34 percent taught only reading, and 24 percent taught only math. Over one-third of the teachers taught only Title I classes, about one-half taught only regular classes, and the remainder taught some combination of Title I and regular education classes.

Ninety percent of the institutions reported problems in recruiting their educational and Title I staff. The most common problems reported were limited numbers of applicants, poor qualifications of applicants, and the geographical remoteness of institutions. Institutions for adults employed a low proportion of women as teachers and aides, reportedly because of the added security risks to women who are employed in adult correctional settings.

About one-half of the institutions reported having problems keeping their Title I staff. The percentage of Title I teachers that leave yearly varied between 10 and 26 percent. The major problems reported were poor salaries, student discipline problems, preferences for teaching in public schools, and the remote location of the institutions. In a study of public school teachers, Harnischfeger (1973) reports that over a two-year period about one-half of the schools studied had to replace between 20 and 40 percent of their teachers. This seems to indicate a rate of turnover roughly comparable to the rate reported for institutions in this study.

AMOUNT OF INSTRUCTIONAL EXPOSURE RECEIVED DURING CONFINEMENT

Among 100 nationally representative programs, the typical length of student exposure to Title I services was reported to be six months, and generally corresponded to the average length of institutional confinement reported earlier.

Only in institutions for adult offenders is instruction in basic-skills frequently discontinued for reasons other than student departures from the institution. Those institutions more frequently reported that Title I students left the basic-skills program because of disciplinary reasons, work assignments, or because of a promotion from Title I to regular education classrooms. The latter case of promotion usually occurred when Title I students had acquired their General Equivalency Diplomas and began to pursue more advanced work. Few instances were ever reported among institutions for neglected or delinquent youths where students progressed to the extent that Title I instruction was no longer required.

Educational staff estimated that during the six months they receive Title I services, students receive an average of 96 hours of total exposure to Title I reading instruction, and 89 hours of total exposure to Title I math instruction. In order to better assess the amount of instructional exposure provided to individual Title I students, teachers from the institutions participating in the impact evaluation were asked to provide instructional exposure data for each student tested. However, this data proved difficult to obtain, for many reasons. The net result of the difficulties was that the usable data may not provide accurate estimators of nationwide average exposure times, but they do provide indications that Title I programs in institutional settings may be having difficulties in delivering services to students. On the average, Title I students were reported to receive three to three and one-half hours of weekly exposure to reading instruction, and two and one-half hours exposure to mathematics instruction. Non-Title I students were reported to receive four and one-half hours of weekly exposure to reading, and two hours exposure to mathematics.

One of the striking features of the exposure data is the propensity of students receiving Title I supplementary services in both reading and mathematics to fail to attend the scheduled Title I classes. They attended only 70 percent of their Title I reading classes and 35 percent of their Title I mathematics classes. Regular education students attended nearly three-fourths of their math classes and virtually all of their reading classes, by comparison. Overall, Title I students would be exposed to an average of 82 hours of instruction in reading and 64 hours of instruction in mathematics during a six-month confinement. Chapter IV presents the reasons why more instructional time is not received by students and discusses the further loss of instructional time through non-task-related activities.

INSTITUTIONAL CONTEXT OF TITLE I PROGRAMS

The initial 100 institutions were examined to learn about the context in which Title I programs operate. Institutions receiving Title I funds ranged in size from 11 to 2500 residents. The size of institutions proved to have a highly skewed distribution, with a mean of 407 and a median of 164 residents. Institutions for neglected youths reported larger resident populations, on the average, than institutions for delinquent youths. Seventy percent of institutions for neglected youths and 47 percent of institutions for delinquent youths reported populations of more than 100 residents.

Institutions varied greatly in the ways in which they were organized to provide for the boarding needs and custody of their populations. Residents in institutions for the neglected were predominantly housed in cottages (small, self-contained living units). Institutions for delinquent youths tended to provide cottages and dorms, and institutions for adult offenders tended to house residents in dormitories and cell blocks. One-fourth of the sample institutions reported that they housed at least some of their population in individual cellblocks (a factor that generally indicates a primary concern with custody or the use of isolation). As might be expected given their larger size, institutions for adults demonstrated a greater concern with the physical restraint of residents: over 85 percent of them were characterized

by locked gates, fencing, and guards. Roughly one-third of the institutions for delinquents resembled institutions for adults in these respects.

In 40 percent of the institutions, the education unit was separated from the rest of the facility by locked doors. Such a control technique was reportedly designed to prevent the entrance of intruders as much as the exit of students from classrooms.

Overcrowding was reported to be a problem at a number of institutions, particularly in institutions for adults. In fact, one-fifth of the institutions reported a resident population more than 10 percent above the designated capacity of the institution. Interviewers observed several cases of severely overcrowded conditions where one-man cells were being occupied by two and sometimes three persons in an attempt to handle the increasing influx of residents. Another interviewer observed rooms that had previously been converted to crafts-areas again being reconverted into dormitories. This loss of classroom space was due to attempts to alleviate the overcrowding in an institution whose population had doubled over the past few years.

In addition to the reports of overcrowded conditions at several institutions, other institutions reported their populations to vary dramatically throughout the year. Large influxes of newly committed offenders were reported to be a phenomenon beyond the control of institutional staff, the influx being largely due to the sentencing patterns of judges. Classrooms were reported to double or triple in size during peak commitment months.

Although longitudinal data were not obtained on the size of institutions over the course of a year, some data were obtained that reflected the fluctuations in student populations. Over the four months of testing for the impact evaluation, the number of students newly available for selection into the study at each wave of testing increased dramatically at most institutions during the March-to-May test administrations.

OPERATING COSTS OF EDUCATIONAL PROGRAMS

Obtaining reliable data on estimated educational expenditures from correctional institutions proved to be a difficult task throughout the length of the evaluation. The major sources of difficulty were that the system of record keeping and the categories of expenditures used in maintaining records varied widely among administrative agencies and among states. Within some administrative agencies or states, record keeping was centralized, and categories of expenditures were aggregated above the institutional level. Thus, the first source of difficulty in these situations was in obtaining separate expenditures for a particular institution apart from the agency or state. Secondly, it was difficult to ascertain the extent to which variation in estimated expenditures was due to non-congruent categories of expenditures and the subsequent inability of fiscal administrators to break down these expenditures into categories of analytic interest. In addition, operational costs in all likelihood vary by type of institution, level of security, and region of the country in ways that our data are not accurate enough to ascertain. Nonetheless, the following statistics do provide a benchmark for estimating the average costs and variability of the costs incurred in operating educational and Title I programs in state institutions.

Expenditures for educational services varied widely among the 100 institutions, from $100 to $5000 per-pupil. To present some summary descriptions of funding levels, the total per-student expenditures for Title I participants were calculated by summing Title I per-student expenditures for state and other federal funds, weighted by the proportion of overlap between the populations served by these funds and the Title I population. Per-pupil figures were also adjusted to account for the rapid student turnover in some institutions. The average length of stay for students in the Title I program was calculated for each institution and used as a weighting factor for the per-student expenditures.

The results of these analyses indicate that, on the average, $456 of Title I funds, $801 of state education funds, and $171 of other federal funds are

expended per pupil. This combines to an average of $1358 per-student educational expenditure in state institutions. Clearly, Title I funds constitute the major source of federal funding for education in state institutions. The average institution's total educational budget was 65 percent state educational funds, 25 percent Title I funds, and about 9 percent other federal funds.

State education funds and Title I funds were used mainly for personnel (81 percent and 74 percent, respectively). This allocation is similar to allocations in public schools, where the single largest expenditure is for staff. Title I staff expenditures were spent almost entirely for classroom personnel (over 80 percent, on the average). Only small amounts of Title I funds were spent on administrative personnel.

The category of instructional materials was the largest single expenditure area for non-personnel funds, regardless of source. Over 40 percent of the non-personnel Title I money was spent in this area, and over 70 percent of other federal funds were expended for materials. In terms of instructional program areas, Title I funds were heavily channeled into the basic-skills area. On the average, about 70 percent of the Title I money supported instruction in reading and mathematics. State education funds were fairly evenly split among basic-skills, vocational education, and other services.

CHAPTER II. IMPACT OF THE TITLE I PROGRAM

The intent of Title I legislation was to create educational programs at institutions that would result in the improvement of the basic-skills performance of confined youth in need of remedial instruction. The first phase of the study found that the improvement of basic skills was the primary objective of Title I programs in institutions. The second phase of the study, the impact evaluation, was designed to measure the effects of the Title I program on student achievement in the basic-skills areas of reading and mathematics and on student attitudes. The following questions were addressed:

1. *Is Title I effective overall in improving student proficiency in basic skills?*
2. *Is Title I differentially effective for students of varying initial levels of basic-skills performance?*
3. *Do Title I students who improve performance over time show different amounts of growth than regular education students who improve?*

MEASURES OF TITLE I EFFECTIVENESS

Data on basic-skills achievement in reading and mathematics was obtained on probability samples of students in programs of regular-education-only and in programs with a Title I supplement. Students were tested at each of three waves separated by six weeks.

As shown in Figure 1, basic-skills performance proved to be stable across the first three waves of testing. There was no evidence of overall growth or decline over time for either Title I or regular education students. The slight increases from wave 1 to 3 shown in Figure 1 are not statistically significant.*

*To investigate whether Title I instruction was effective in improving students' proficiency in basic skills, analyses of covariance were performed, controlling for initial achievement and other student characteristics. The results of these analyses show no differences between Title I and regular education. The analysis of covariance may have questionable validity for these data, however, because of the unknown bias in assignment of students to educational programs.

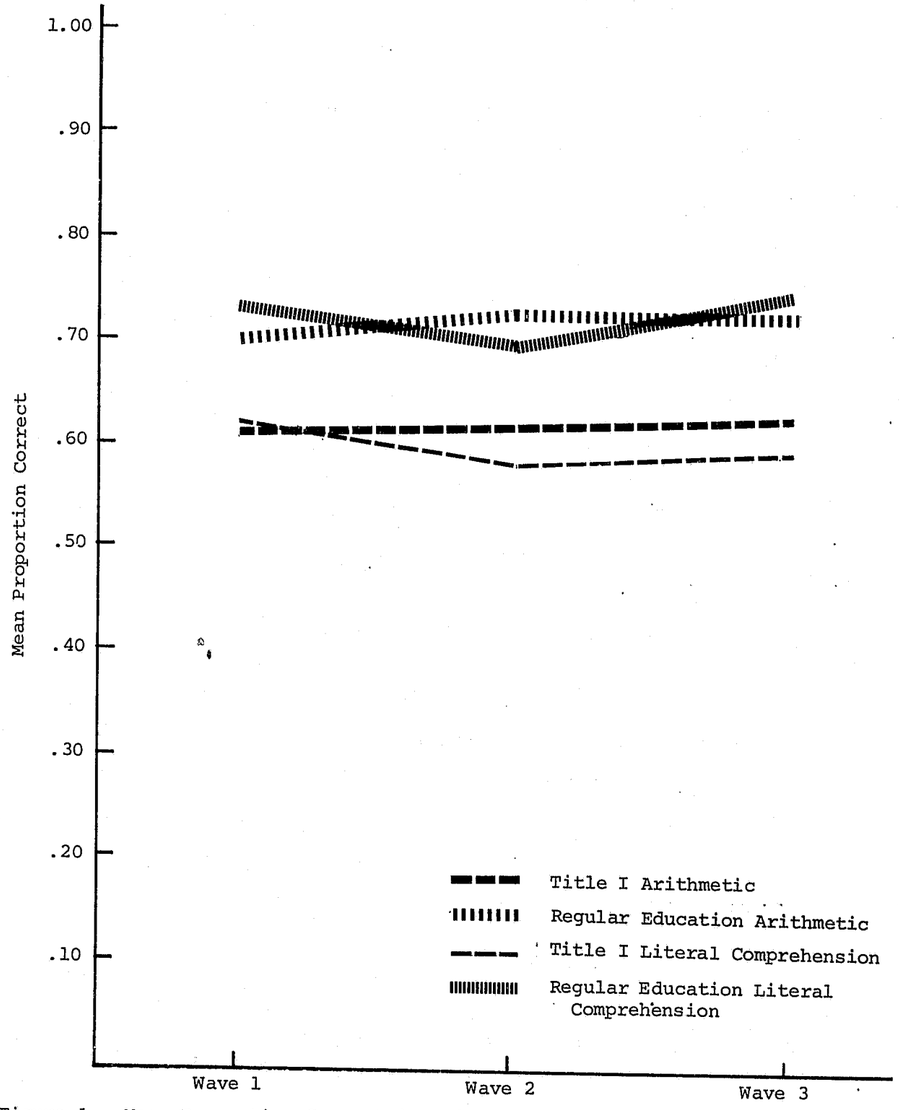

Figure 1. Mean Proportion Correct in Arithmetic and Reading (Literal Comprehension) for Title I and Regular Education Students in Randomly Selected Institutions (Weighted by Selection Probability)

Thus, Title I and regular educational programs exerted the same amount of influence; Title I students generally did not improve relative to regular education students, nor did they deteriorate relative to them. It is possible that this apparent stability masks other effects. For example, if a negative reaction to repeated test-taking is present to some degree, it may have neutralized a real increase in performance occurring simultaneously. On the other hand, it may be argued that the restrictiveness of the institutional environment is not conducive to learning and that in the absence of a strong instructional effect, performance may be expected to remain the same or even decline over time. If the latter were the case, students might be doing well to maintain their levels of performance. The data obtained in this study do not indicate which alternative is most reasonable: students may be doing well by maintaining performance levels which would be expected to decline if education programs were not offered; they may be doing just as expected--maintaining constant levels of performance in an environment which is unlikely to change their performance levels. The last hypothesis requires the least explanation, but is the most discouraging.

These overall results do not distinguish among institutions, however. Some institutions did show increases in performance among Title I or regular education students, whereas others showed decreases or mixed results. Thus, educational programs in some institutions did produce gains in basic-skills achievement over a twelve-week period. In Chapter III, some of the characteristics of effective programs are discussed, while in Chapter IV the factors impeding success are analyzed.

Detailed questions may have been masked by the overall null results. For example, do students with a low initial achievement improve over time? Or stated in another way, is Title I (or regular education) instruction effective for students with low initial achievement? Among those students who show improvement in performance over time, do Title I and regular education students show different amounts of growth?

To answer the first question, the performance of Title I and regular education students was examined separately in quartile groups defined by wave 1 scores. When ceiling effects (for high-scoring groups) and regression effects due to unreliability in the tests are discounted, there seemed to be no differential growth dependent upon initial scores.

Because scores that decrease over time may reflect effects of overtesting, and thus may not be valid indicators of student achievement, the performance of students whose scores were maintained or improved over time were examined separately. These analysis indicated that, for those students who maintained their levels of performance or improved over time, Title I students showed the same average gain as did regular education students. Similar analyses were performed for the attitude variables measured in the impact evaluation phase. Again, there were no differences found between Title I and regular education students in the patterns of change on these measures.

PERCEPTIONS OF TITLE I EFFECTIVENESS

In addition to the measured impacts of the Title I Program on achievement, staff in various roles at 100 institutions were asked their perceptions of the success of Title I and its impact on correctional education.

Most institutional administrators and teachers surveyed believed that their Title I program was effective in improving the educational performances of participating students.* At several institutions, staff reported that the length of student exposure to Title I instruction was perhaps too short for measured effects to be observed. Data presented in Chapter I, III, and IV of this report tend to reinforce this perception.

While Title I programs do not seem to cause increased growth for the students receiving their services, the institutional staff surveyed generally believed that the receipt of Title I funds and the implementation of Title I programs

*Many institutions had statistics demonstrating effectiveness.

had produced positive effects on correctional education as a whole. The expanded use of needs-assessment throughout the education program, teacher awareness of new teaching methods, improved institutional staff awareness of the need for basic-skills education, and the emphasis on individualized instruction were cited as spin-offs of the Title I program. Since the study reported here did not specifically address issues of institutional change, no data other than these impressions can be put forward to indicate whether instructional programs improved on the whole with the influx of Title I funds in these institutions.

CHAPTER III. CHARACTERISTICS OF EFFECTIVE BASIC SKILLS PROGRAMS

The Phase II impact evaluation had found, on the average, little or no change in basic-skills achievements or attitudes over a twelve-week instructional period. Some sites were found, however, that did appear to be effective in producing achievement gains and/or attitude changes. Institutional or program characteristics that were related to changes in attitudes or to gains in average student achievements were identified as part of the impact evaluation. In addition, analyses of the data obtained in the Substudy of Effective Practices (see Appendix C for a description of the methodology employed) revealed that certain instructional characteristics or practices in basic-skills classes are more likely to result in improvement in attitudes or in basic-skills performance than are others. The findings of these analyses are discussed in this chapter.

CHARACTERISTICS OF INSTITUTIONS WITH EFFECTIVE PROGRAMS

The analytic results presented in this section are based on data from the 40 institutions participating in the impact evaluation. These analyses were intended to be more exploratory than confirmatory.

Results of these analyses indicate that the number of students enrolled in the reading program is negatively related to reading achievement. Although the effect was small, it does suggest that larger reading programs seem to be more detrimental to the achievement of basic skills.

The educational program administrator's ranking of the importance of various institutional activities was found related to gains in reading achievement. The higher the ranking of educational or treatment (e.g., counseling) goals among institutional activities, the larger were the gains in reading achievement. Conversely, the higher the ranking of custody, security or institutional upkeep among institutional activities, the lower were the gains in reading achievement.

The extent to which the educational administrator encouraged teachers to be innovative within their own courses of instruction was found to be positively related to math gain scores. Also, the higher the estimated percentage of educational staff turnover, the lower were the gains in math achievements. Consistent with this latter finding, it was also found that the higher the rate of treatment staff turnover, the lower were average student gains in measures of self-esteem.

EDUCATIONAL PROGRAM CHARACTERISTICS AND ACHIEVEMENT GAINS

Educational researchers have considered many factors as potential determiners of program effectiveness. Among these are instructional materials, student background characteristics, student attitudes, teacher-student interactions, staff characteristics and student-teacher ratio. Another major factor is the amount of exposure to instruction. In this section, the relationships of the various factors mentioned above to achievement gains is assessed. In addition, the influence of various characteristics of instruction on task-engaged activity is also investigated.

Instructional Materials

The frequency of usage for different types of instructional media was investigated for its relationship to achievement gains. Types of media included textbooks, programmed workbooks, audiovisual equipment, and locally-developed and/or high interest materials. Only programmed learning materials and high interest and/or locally developed materials showed significant relationships to achievement gains. The use of programmed learning materials was consistently negatively related to achievement gains, while the use of high-interest and/or locally developed materials was positively related to achievement gains except for reading in combination basic-skills classes (where the relationship was not statistically significant).

One plausible explanation for the negative effect of programmed materials is that teachers might tend to engage themselves in less instructional activity when these materials are available. Data on teacher-student interactions in

the few classes using these materials were not numerous enough to confirm this hypothesis, however.

Student Characteristics

The individual student characteristics of age, sex, ethnicity, previous grade attended, seriousness of offense for which respondent was institutionalized, anticipated length of institutional confinement, actual length of confinement, and history of institutional confinement were analyzed for their relationships to achievement gains in math and reading. Most of the relationships were small. Those that were statistically significant are reported here. Minority students were slightly more likely to demonstrate gains in reading achievement than were non-minority.

Students were asked where the year previous to their present confinement was spent and the length of time spent in each living situation. Responses indicated that time was spent in one, or some combination, of the following situations: parental homes, foster homes, group homes, or institutional confinement. Only the amount of time spent in an institutional setting was significantly related to gains in reading achievement ($r = .42$, $p = .001$). Thus, the more time students had spent in an institutional setting in the year before present confinement, the greater were their reading gains. It should be pointed out that this correlation applies only to the 76 students who were institutionalized in the previous year.

To avoid an overinterpretation of the beneficial impacts of institutional confinement that this finding might imply, it should be pointed out that this relationship probably speaks more to the advantages of uninterrupted or continuous delivery of basic-skills services than it does to the advantages of confinement. Evidence for this interpretation is the fact that the number of times a student had been institutionalized during his/her lifetime was negatively related to gains in reading achievement.

Observers assessed individual students in terms of the existence of learning problems (e.g., visual or speech impairments, English as a second language, learning disabilities). These assessments indicated that students with learning difficulties were less likely to demonstrate gains in reading achievements than were those students with no observed problems ($r = -.14$).

Student Attitudes

Students' self-reported attitudes were examined for relationships to achievement gains. Attitudinal scales included measures of self-esteem, locus of control, attitudes toward school and learning, reading and mathematics, perceptions of the practical orientation of the educational program, and perceptions of teacher supportiveness.

These analyses revealed that, in general, more favorable student attitudes were not related to gains in achievement, with one exception. Students' ratings of perceived performance in the institutional school setting did demonstrate a significant ($r = .11$) relationship to reading gains. Students who perceived themselves as poor performers were less likely to demonstrate gains in reading than were students who perceived their institutional school performance as favorable.

Although student attitudes bore little relationship to achievement gains, changes in attitudes (over the twelve-week period) were significantly related to gains in achievement. Positive relationships were obtained for reading and math gains and changes in reports of self-esteem, the degree of school involvement, perceptions of teacher supportiveness, and perceptions of the practial orientation of correctional education programs. The magnitude of the reported relationship between perceptions of practical orientation and math gains was increased among lower-level math performers.

Teacher-Student Interactions

Descriptions of different ways in which teachers interact with students were investigated for their relationships to achievement gains. Interactions were

characterized as ones where teachers praise, commend, or provide other positive feedback; ones where teachers explain or instruct as simple statements of fact in neutral tones; and ones where teachers criticize, insult, embarrass, or provide other negative feedback. Observers rated these descriptions according to whether they described most, some, few, or none of the classroom interactions.

The results of these analyses showed consistency within class types across the second and third observational visits, but were inconsistent across class types. For both reading and combination basic-skills classes, the greater the frequency that teachers provided positive feedback in the form of praise, the less were the achievement gains. Interactions characterized by simple statements of fact, even corrections, were positively related to achievement gains.

The direction of relationships for interactions in math classes and achievement gains demonstrated different results. For math classes, the more frequently praise and positive feedback were provided, the greater were the math gains. The frequency of negative or neutral interactions was not related to achievement gains in math classes.

The relationship between observer assessment of teacher supportiveness and achievement gains had similarly mixed results for mathematics classes versus reading and combination basic-skills classes. The more supportive teachers in mathematics classes were perceived to be, the greater were the achievement gains. In reading and combination basic-skills classes, less supportive behaviors were related to achievement gains in both reading and mathematics.

While there is no obvious explanation for these conflicting findings, several possible explanations exist. Particularly among math classes, the nature of math task-related activities allows for the provision of positive feedback at many appropriate points (e.g., the understanding of an operational concept, the correct solution to a problem). Feedback in reading classes, especially where little instruction is ongoing and students engage mostly in practice, may be more difficult to provide. Students could possibly view positive feedback in these situations as patronizing, since it may not be perceived to be

related to the demonstration of any particular reading skill. This finding is consistent with other research conducted at the primary level, where it was found that some teachers tended to overuse praise, particularly insincere, perfunctory praise that was response to "praise-seeking" behavior from their students (Anderson et al., 1979; Brophy and Evertson, 1976).

Related to this explanation, and perhaps interacting with it, is the possibility that observers may have found it more difficult to distinguish among positive, neutral, or negative forms of feedback in reading classes. Observers and students may simply have differed in their perceptions of the nature and frequency of types of feedback received.

Observer assessments of student interactions were also made in terms of the positive, neutral, or conflictive nature of student interactions. Only in math classes were conflictive interactions observed, and found to be negatively related to gains in achievement.

Student-Teacher Ratio

The fewer the students assigned to each teacher in an institution, the greater the gains in reading achievement. This result is consistent with results obtained in non-correctional classrooms (Glass and Smith, 1979).

Instructional Time

For purposes of this study, a major objective was to obtain precise measurements of the amount of instructional exposure to reading and math and to assess the relationship of instructional exposure to achievement gains. The rationale for this focus was based on the results of previous research as well as on impact evaluation findings that questioned the extent to which institutionalized students were actually receiving basic-skills instruction. Past research has suggested that the amount of time needed, allowed, and spent are major factors in school learning (Bloom, 1974, 1976; Wiley and Harnischfeger, 1974). Empirical findings of a relationship between instructional time and reading achievement have been reported for non-institutionalized classroom settings (Hanson and Ross, 1975; Barr, 1974, 1975). Powell and Dishaw (1980) and Stallings (1980)

refine the concept of exposure time by including measures of student engagement in the instructional task.

In the Substudy of Effective Practices, it was discovered that absenteeism was the major source of missed time in institutional settings. When the actual attended time was related to achievement gains, significant and positive, but modest, correlations were discovered. However, a further improvement was effected by focusing on the amount of instructional time that was task-engaged time. When task-engaged time over the 12 weeks of the study was related to achievement gains, the correlations observed are about .22, on the average.

Additional analyses were undertaken to determine which task-related instructional activities seemed most beneficial. The percentages of students' class time engaged in various aspects of instruction (e.g., receipt of instruction from the teacher, practice time or drill, receiving feedback from instructors, and other activities) were investigated for their relationships to achievement gains. The ways in which students' time was spent among these categories was not found to be related to achievement gains in math classes or combination basic-skills classes. Among reading classes, however, it was found that the larger the amount of time students were receiving instruction from the teacher, the larger were the gains in reading achievement. The less time students spent in 'other' activities such as non-task-related conversations), the greater were the achievement gains.

Instructional Characteristics and Task-Engaged Activity

The amount of task-related activity within a classroom had been found to be related to achievement gains, as reported in the preceding pages. Furthermore, certain instructional processes demonstrated a relationship to achievement gains. In the interest of identifying factors related to greater task-related activity among students present in classes, instructional characteristics were investigated for their relationship to the proportion of time students engage in task-related activity.

The way that instructional time was structured within classrooms demonstrated a moderately strong correlation to the amount of time spent engaged in task-related activities. The direction of these relationships was consistent among reading classes, mathematics classes and combination basic-skills classes. The magnitude of most obtained correlations was lower for combination basic-skills classes, however. This is consistent with qualitative analyses presented by Bossert (1979).

The greater the amount of students' time that was engaged in the receipt of instruction and feedback, the greater was the proportion of student task-related activity. Conversely, the less the amount of time devoted to practice or drill, the greater the proportion of task-related activity. Quite simply, this means that when teachers work, students work. Left to their own devices for purposes of practice or drill without the provision of feedback by teachers, students are likely to engage in non-task-related activities. This finding is important in light of the fact that, on the whole, 75 percent of students' time was found to be structured for practice or drill. An experiment reported by Stallings (1980) suggests that it is possible to induce teachers to allocate class time such that activities that lead to more engaged time can be increased.

Among math classes and combination basic-skills classes, the use of praise, positive feedback or feedback characterized as simple statements of fact was positively related to task-engaged activity. Criticism was negatively related to task-engaged activity. However, as was found previously for the achievement outcomes among reading classes, the less praise or positive feedback observed, the greater was student engagement in task-related activities. Again the interpretation offered here is that praise in reading classes is perceived in a qualitatively different way from that given in math classes. Teacher supportiveness and teacher commitment to instruction were also related positively to the amount of task-engaged time in all classes (with the exception of teacher supportiveness in combination basic-skills classes). These results also suggest that the notion of praise (as distinct from support) has a different meaning for students in reading classes than in math classes.

SUMMARY OF EFFECTIVE PROGRAM CHARACTERISTICS

In summary, the combination of institutional commitment to educational objectives through the provision of an uninterrupted delivery of instructional services, the presence of teachers committed to instruction, and the engagement of students in task-related activities, is likely to result in achievement gains. In spite of formal statements of the importance of education among institutional goals, institutions for delinquent youth are often not hospitable environments for education. While they may not be actively opposed to educational activities, educational goals are often not well-integrated into the overall institutional goals. The findings of this study are that badly needed training in basic skills is sometimes offered for very small amounts of time and that student schedules are often organized on an ever-changing basis in ways that reduce class attendance. The specific nature of the obstacles to the delivery of basic-skills services to students in correctional education settings and the frequency of their occurrence are examined in the next chapter.

An equally important finding, however, is that correctional institutions do exist whereby educational goals are well-integrated and effectively pursued within the constraints imposed by the nature of institutional confinement.

CHAPTER IV. FACTORS IMPEDING BASIC-SKILLS PROGRAM SUCCESS

Throughout the evaluation, many participants identified problems that were perceived to have impeded or prevented the successful implementation of Title I programs. In addition, problems that affected the implementation of all basic-skills educational programs were also noted. Quantitative data were obtained, in the course of the evaluation, to substantiate some of these perceptions and observations of factors impeding successful program implementation.

PROBLEMS WITH THE IMPLEMENTATION OF FEDERAL TITLE I GUIDELINES

As a part of the initial phase of the evaluation, Title I administrators at the state level and at 100 institutions were asked about problems associated with Title I guidelines. They expressed considerable dissatisfaction about the lack of clarity of federal guidelines and the difficulties this provided for the implementation and operation of Title I programs. Many administrators believed that guidelines were not being closely followed in either their own state or in other states. Furthermore, they felt that widely differing interpretations of Title I objectives were made by state administrators across the country. For example, differing interpretations at the state level led to supportive services (e.g., counseling) being strictly prohibited in some states while being allowed in others. The lack of clarity in defining what constituted Title I services and Title I participants was attributed to a lack of guidelines or to misinterpretations of existing guidelines.

When state coordinators were asked about changes they would recommend in federal guidelines, many made suggestions to eliminate specific requirements about student eligibility or program content. Citing the wide variety of needs of institutionalized students, both state and institutional administrators suggested an expansion of Title I to encompass more than basic-skills objectives, the elimination of the age restriction so that more needy residents could be served in institutions for adults, the elimination of the "neediest of the needy" requirement since most institutionalized students were highly deficient in basic skills, and the removal of the requirement that Title I basic-skills programs must supplement rather than supplant state education

courses. This latter suggestion was responsive to the observation by administrators that there were duplications of course offerings (e.g., the expenditure of funds for both a Title I basic-skills lab and an Adult Basic Education lab, etc.). Others argued that education has such a low priority among some correctional agencies that the regular education courses are underdeveloped, and consequently it is difficult to use Title I funds in a supplementary way in such instances.

PROBLEMS WITH THE INSTITUTIONAL CONTEXT

Data obtained from interviews with staff respondents in 100 institutions revealed a perceived existence of constraints that are imposed on correctional education by virtue of the institutional context in which educational programs must operate.

Security measures were reported to create problems for the education staff in at least one-half of the 100 institutions, primarily among institutions for adult offenders. The most frequently cited problem was the restriction of groups of students from class attendance for disciplinary reasons. Respondents at more than half of the institutions reported that custody or cottage staff have the authority to prevent students from attending classes by assigning them to institutional work duties or by confining them to quarters. Other problems reported were the restrictions on free movement between classrooms by the preoccupation with locking and unlocking doors, the censorship of educational materials, and the continual lock-up of equipment (to prevent its disappearance or destruction). Incidents such as riots, disturbances or violent acts had caused the shutdown of education programs at some time during the preceding year in several institutions (primarily those for adults). In part due to such constraints on teaching, teachers in the more security-oriented institutions reported dissatisfaction with their present teaching positions at twice the rate (50%) that was reported among the less security-oriented institutions.

During the impact evaluation, it was discovered that nearly two-thirds of the students who had been tested on the first data-collection visit to the insti-

tutions were still present at the institutions 18 weeks later, but one-half of them were unavailable for testing. Furthermore, the exposure data collected in that phase of the evaluation suggested that students receive very small amounts of exposure to instruction. This information raised the issue of the extent to which the overall failure to demonstrate change is due to constraints placed on the delivery of basic-skills services by virtue of their operation within institutional contexts or to other factors such as the instructional practices or the motivation of the students. The Substudy of Effective Practices obtained data directly addressing these issues.

STUDENT ABSENCES

Observers maintained records of student absences for the observational days and recorded reasons for absences (reasons were generally obtained from teachers). Since continual knowledge of each resident's whereabouts is a primary task in most institutions, these data were not difficult to obtain. Estimates of student absence rates are not based on duplicated counts. For instance, a student placed in lock-up who missed several classes over the period of observation is counted once only. Thus, absence rates do not provide an accurate estimate of the amount of scheduled instruction that has been missed by students (an estimate of the total amount of instruction for which students are absent is contained in a subsequent discussion). However, reasons provided for these absences are some indication of the relative extent to which institutional factors, disciplinary procedures, and student behavior (e.g., AWOL, sick reports) account for student absences from classes.

Site-level absence rates indicated a wide range in the frequency of absences among sites, from virtually no absences in some sites to as high as an 80 percent absence rate in other sites. On the average, 40 percent of these absences were due to student participation in non-instructional, program-related activities. Institutional work assignments added to student absentee rates. The combination of these factors of institutional scheduling conflicts accounted for more than one-half of student absences for the three observational periods.

In many correctional institutions participation in educational programs is considered to be a privilege rather than an obligation. Punishment for misconduct may result in the denial of access to education. (Sometimes all the residents of a given cottage are restricted to quarters--and cannot participate in education--because of an infraction by one cottage resident.) Twenty percent of the absences were accounted for by such disciplinary measures. The rate of student lock-up at sites where such disciplinary measures are more frequently employed was much higher at the first visit than it was at the two subsequent visits. This suggests that institutional staff, in their attempts to guarantee student availability for test administration, may have utilized this method of discipline less frequently on subsequent site visits. Lock-up is judged to be a potent disrupter of educational services.

Student absences due to factors beyond the control of the institutional staff (e.g., sick reports, court appearances, AWOL) accounted for about one-fourth of total absences. Absences for these reasons are to be expected as a normal part of institutional routine.

The time scheduled for instruction in basic-skills classes was examined to determine the extent of and reasons for non-task-related activities by students. About 44 percent of the scheduled time was spent on non-task-related activities. Of this, about 57 percent was due to institutional activities that prevented students from attending classes or removed them for some portion of their scheduled class time.

Disruptions to task activities originating within classrooms themselves accounted for an average of 20 percent of non-task time. These types of interruptions included the accounting for student presence, clean-up activities, smoking breaks, interruptions caused by disciplinary activities within the classroom, and teacher or teacher aide engagements of students in non-task-related conversation. Estimates of these types of interruptions were fairly stable across observations.

Given the frequency of references to the poor motivation of institutionalized students, a surprisingly small amount of time (23 percent of non-task-related

time) was attributed to student disengagement (e.g., daydreaming, working on activities other than basic skills, or carrying on a conversation with other students or teachers).

These findings suggest that the delivery of instruction to Title I and regular education students is impeded to varying degrees by the simultaneous delivery of other services (e.g., counseling), or the pursuit of other objectives (e.g., disciplinary control or institutional upkeep). Observations of student-initiated sources of disruption indicated that student motivation at most institutions plays a relatively minor role in accounting for difficulties in delivering basic-skills instruction. This finding suggests that either institutionalized students are not as poorly motivated to learn as is often suggested, or that some institutions are quite effective in providing motivation for otherwise poorly-motivated students.

Support for this latter interpretation was provided by the finding that almost 40 percent of students had reported their school performance prior to their confinement as "poor" or "very poor", and 25 percent of the students had actually quit school prior to their institutional confinement; however, in reporting on their academic performance within the institutional education program, only 8 percent of students reported "poor" or "very poor" performances. Given the negative student assessments of their public school performances, basic-skills programs in at least some sites appear to have been effective in changing student perceptions of their performances, and perhaps their motivation to learn.

TASK-RELATED ACTIVITY AMONG TYPES OF CLASSES

One question posed by the within-site variation on task-related activity is whether different types of basic-skills classes tend to engage students in more activity by virtue of the subject matter (i.e., are students more or less task-engaged in reading or math classes?). Because of the usage of combination basic-skills classes, it was impossible to describe the extent of task activity by subject matter alone (i.e., non-task time used in making transitions to another subject matter cannot be appropriately attributed to any one of the tasks). Thus, the three types of basic-skills classes--reading only, math

only, and combination basic-skills classes--were analyzed for their relative amounts of task-related activity. The effects of these class types on student task-engagement is difficult to assess for the institutional education programs represented by the present sample, however. Combined basic-skills classes, for instance, were not an instructional feature of all nine sites. Three sites had no combination classes. Furthermore, assignment of students within institutions to combined classes or specific content classes was mixed: in some sites students were assigned to specific basic-skills content classes, and additionally to a combination basic-skills class. Time allocated for combination classes in these situations was intended to be used in whatever skill area students needed additional work. In other sites, combination basic-skills classes were characteristic of all basic-skills instruction. Thus, findings reported for types of classes are only suggestive.

Students in reading and math classes were found to be task-engaged for 65 percent of the time, on the average. Students in combination basic-skills classes were on-task only slightly more than one-half of the time.

On the average, reading classes were found to be scheduled for slightly longer periods of time than were math classes. Combination classes were approximately 40 minutes longer than were reading classes. Since combination classes lasted much longer than did reading or math classes, it is impossible to separate the influence of class length from the effects of the unspecified nature of class tasks, or other characteristics differentiating combination classes as a type. For instance, few would argue with the provision of breaks for long class periods. However, observers noted a tendency for students to disengage themselves from tasks in anticipation of these breaks, and to take long periods of time to again become task-oriented after the completion of the break.

Observers also noted a greater propensity of these largely unstructured classes to result in less task-related and more disruptive activity for <u>both</u> teachers and students. It reportedly took students and teachers a longer period of time to identify a task in which students were to be engaged at the beginning of the class. This observation is consistent with previous findings that transi-

tions disrupt time-flow and that disruptive pupil behaviors increased during unstructured transitions (Arlin, 1979). Thus, while longer periods of class-time could potentially increase instructional time by the reduced need to perform such activities as escorting students to and from classes, the frequent breaks provided within these classes are also a large source of disruption (from which both students and teachers may take longer to recover).

FACTORS INCREASING OR DIMINISHING INSTRUCTIONAL TIME

Despite successes at a few sites, the overall amount of non-task-related activity was found to be large. The importance of this finding for the successful delivery of Title I and other basic-skills programs is clear. The delivery of instruction to remediate the highly deficient basic skills of institutionalized students, already hindered by an expected six-month length of institutional exposure, is further obstructed by the high proportion of non-task-related activity. When the total amount of scheduled institutional basic-skills exposure time (which at some sites was as low as one hour per week of mathematics and one and one-half hours per week of reading instruction) is adjusted by the average proportion of non-task-related time, the actual amount of instruction delivered over a twelve-week period is about 48 hours per student. Thus, the actual amount of basic skills instruction that is expected to be provided to a student during an average length of institutional confinement is a little more than 50 hours in reading and a little more than 40 hours in mathematics.

Some suggestions of ways in which task-time could be increased and absences minimized were obtained from institutional staff and field observers. For instance, the system used to administer institutional privileges was observed in some situations to increase the frequency of poor class attendance. In these situations, institutional work was an activity for which residents were paid, and hence participation in this activity was seen as a privilege. No matter how small the salary, most residents clearly preferred the short-term objective of obtaining money to buy coveted items (such as cigarettes) over the longer-term objectives of acquiring reading and mathematics skills.

Other institutions operated, as do public schools, on a normal workday schedule. Although residents were on site at all hours of the day, all institutional activities were scheduled within the limits of a six-to-eight hour day. Thus, activities were scheduled simultaneously to meet the varied needs of institutionalized students. The net effect was that educational programs often competed for participants with counseling and work programs. Furthermore, institutional disciplinary procedures often restricted students from class attendance.

At times these alternative activities appeared to be appropriate focuses of concern, and at times they appeared to needlessly detract from basic-skills instruction. For instance, work programs that took students from classes rarely involved more than unskilled janitorial or cafeteria positions and, thus, were not expected to improve the job skills of students except as they pertained to janitorial work. It was unclear to observers that these kinds of alternatives were preferable to time spent in educational programs.

The above situations differ from past observations of education programs in some institutions for adults (Bartell et al., 1977) where educational goals were admittedly of low priority. However, the implication for the operation of the program was that classes were frequently held at times that did not compete with other institutional activities (e.g., early mornings or evenings). The result of such scheduling practices is that fewer institutional residents may receive basic-skills instruction, but the instruction that is delivered may be less subject to institutional interruption.

The therapeutic or disciplinary model that provided the philosophical basis for institutional activities also influenced the amount of time allocated for instruction and interruptions to the delivery of instruction. Many institutions utilized some variation of the positive peer-culture approach to therapy and discipline. The philosophy of such approaches--that an individual's behavior is a responsibility of the group--results in practices whereby the entire group is subsequently rewarded or punished for the actions of an individual or a few members. The implication of this approach for the operation of educa-

tional programs is that entire cottages are generally restricted from class attendance, resulting in a high frequency of class absences.

Aside from the allocation of instructional time, other institution-level factors were found that may affect program effectiveness. Although empirical evidence was not specifically obtained to support this discussion, observations over three years of data-collection activity pointed out perhaps critical factors operating in the more ineffective sites and in some of the sites demonstrating mixed results. Staff morale (and related factors such as teacher disengagement from instruction, etc.) had been a reported and observed problem among several institutions for several years. Whether poorly-motivated staff were initially hired, or whether staff developed poor motivation as a result of certain institutional factors, is unknown. Based on three years of unsystematic observations and the retrospective reporting of staff and administrators, no single, consistent factor could be isolated that was common to all the institutions reporting staff morale problems.

Common features among these sites, however, were the frequent changes on a number of dimensions, including population size, the type of residents confined, the way in which the institution was organized, administrative leadership, treatment philosophy, or some combination of these factors. Changes in the number of students confined, which is often unpredictable and quite large, may be more detrimental than a consistent, large, student population. For instance, an institution or education program designed on the basis of intensive and individualized instruction and staffed for a certain size of instructional grouping, will be overwhelmed by dramatic increases in population.* Repeated occurrences of this sort are likely to take their toll on staff morale. As teachers observed in several institutions where rapidly changing sizes in the population have occurred over several years, teachers can only "warehouse" students.

*Most institutions for delinquents of at least moderate size had reported large variations in the size of their population during the Phase II impact evaluation.

Other factors, noted in ineffective sites where staff morale was problematic, were weak or disinterested administrators, constantly changing administrations, and state-level changes in either the organization of the institution and/or the treatment philosophy pursued, the type of offender confined, etc. The implementation of such changes given the existing staffing and resources may create problems on many dimensions.

While state-level and institutional administrators may be interested in the previous findings (e.g., that Title I programs may be effectively designed and implemented, but that their services are only unsystematically provided to educational recipients) these findings are of little avail to correctional teachers who may find themselves powerless to influence the institutional factors that impede class attendance and prevent services from being delivered. The substudy also identified classroom practices that were related to a greater proportion of task-engaged class time and to gains in achievement, two criteria of instructional effectiveness.

Students in classes in which both reading and math skills were taught ("combination" classes) showed much less task-related activity than did students in classes where only one of the two subjects was taught. The combination classes seemed to have less formal structure for the use of time and the teachers seemed to be less involved with instruction.

One surprising observation was that several of the sites that strictly adhered to institutional rules and routine during non-class time were characterized by such a large degree of unstructured time within classrooms. Observers frequently noted that in these sites, educational programs were merely one more form of custody. Despite the knowledge that most of their students were highly deficient and had a limited amount of time within the institution to overcome these deficiencies, teachers in some of the more ineffective sites seemed to approach classroom time as something to "kill" rather than as something to "allocate economically" (Wiley and Harnishfeger, 1974).

In the more effective sites both students and teachers were highly task-engaged, with 80 to 90 percent of students' time observed to be devoted to reading and math instruction over the several observation points. For instance, one of the highest site-level averages for on-task time over the twelve weeks was obtained at a site where scheduled academic instructional time was most limited (e.g., students received one-half day of instruction each weekday for six weeks followed by no instruction for six weeks). The high rate of on-task time made this site effective in accumulating on-task hours.

Inceases in instructional time gained by increased time-on-task has the advantage within correctional institutions of allowing the simultaneous pursuit of other institutional objectives and of avoiding the competition of staff or programs for student time. Of course, this assumes that high levels of task-oriented behavior are combined with administrative scheduling practices that do not disrupt class attendance.

CHAPTER V. POST-RELEASE EXPERIENCES OF STUDENTS

The major objective of Title I funding for confined youth is to improve their basic skills in reading and math. The preceding chapters have shown that some institutions are successful in providing educational services that achieve this objective. However, institutional staff report that emotional and other problems associated with the transition from the institution to the community can impede or prevent students from continuing to acquire and/or employ these basic skills. While Title I legislation does not prohibit the expenditure of funds for academically related services (e.g., counseling, liaison services with community schools), the provision of these services has not been a systematic objective of Title I programs as they are presently implemented. Furthermore, there is very little factual information available about the services available to residents of correctional institutions (either before or after their release). To remedy this lack of information, the follow-up study of students released from state institutions offering Title I services was designed to answer the following questions:

1. *What are the post-release experiences of students in terms of school entry, employment, and recidivism?*
2. *What pre-release, liaison and post-release services are provided to students by correctional institutions and community agencies?*
3. *What are the relationships of measures of post-release success to student characteristics, student attitudes, the receipt of services, and problems encountered after release?*

EDUCATIONAL EXPERIENCES OF RELEASED STUDENTS

Educators and correctional personnel have long been aware of the reentry difficulties experienced by youths who try to return to public schools after their release from correctional institutions. Negative past student experiences with public schools, poor basic skills, a lack of motivation, and problematic living situations are thought to combine to prevent students from entering schools after their release from institutions.

Among the sample of students interviewed at least three months following release, one-half of the interviewed students had returned to some type of community school. For those who returned to school, there was little time delay in school entry, since the average period of transition from institutional to community school attendance was two weeks. Although the maximum length of time before school entry was five months, few students actually experienced such delays.

About one-half of all returning students were re-enrolled in schools they had attended prior to their institutional confinement; re-enrollment in the same school generally reflected the fact that the school was the only one in the geographic area. Over half of the students returning to school enrolled in regular public high schools; only 10 percent actually enrolled in vocational/technical schools, despite frequently expressed desires for vocational training. As many as one-third of released students attended specially-designed alternative schools that take into consideration the motivational differences, emotional needs, and remedial needs of students with histories of institutional confinement (e.g., work-study programs, GED preparation, adult basic education).

The perception of correctional educators that student transition from institutional schools to public schools is problematic was substantiated by respondents who returned to school. One-half of the students returning to school reported that they had problems in making the transition: the most frequently cited source of difficulty was that their basic-skills performance was at a level well below that of their classmates. One-third of the students rated their academic performance after release as "poor" or "very poor." They cited the fact that they were academically behind their peers, that they were poorly motivated and had poor attitudes toward school, as reasons for their poor performance. Further evidence of school transitional difficulties was the fact that most students who had entered schools after their release had stopped attending by the time of the interview. Eighty percent of students returning to school after their release had withdrawn from school prior to the end of the school term.

An examination of differences in individual motivations, background characteristics, attitudes, and situational constraints experienced by those who enroll in school upon release, as opposed to those who fail to enroll, is critical in terms of identifying post-release needs and the planning for means to reduce those needs. Therefore, comparisons were made of (1) those who entered school after release with those who didn't; (2) those who entered school after release but soon dropped out with those who entered and remained; and (3) those who attempted unsuccessfully to gain entry to schools--or at least desired entry--with those who had no desire and made no attempt to enter school.

Students released from correctional institutions who enrolled in schools after their release were first compared to releasees who did not enroll in school. Comparisons were made on a number of dimensions, including background characteristics, measures of self-attitudes and school-related attitudes, levels of basic-skills performance,* and the receipt of pre-release, liaison, or post-release services. It was hypothesized that students would be more likely to enroll in schools after their release if they had demonstrated higher levels of basic-skills performance while confined, if their basic-skills performance corresponds more closely to their chronological age, if they had more positive self-attitudes and attitudes toward school and learning, and if they received assistance in the transition from institutional to public schools.

Regardless of the measure utilized in assessing the level of basic-skills performance (e.g., domain-referenced tests of basic skills, standardized achievement tests, or functional literacy tests), no relationship was demonstrated between level of performance and school entry after release. Since both the basic-skills level and school entry were related to age, the influence of age was partialled out from the relationship of performance level to school entry. Still, no relationship was revealed. Better performers on basic skills are not more likely to enter school after release, regardless of their age.

*Measures of background characteristics, attitudes toward self and school, and level of basic skills performance were obtained during the impact evaluation phase.

While the receipt of Title I basic-skills instruction during confinement appeared related to school entry, an examination of this correlation (controlling for age) indicated that the relationship is explained by the younger age of Title I participants. The likelihood of enrolling in school was primarily a function of age: the younger the student, the greater the likelihood of a return to school. This finding should not be surprising given the fact that students below the age limitation of compulsory attendance attend school by virtue of legal statute as well as by virtue of any motivational aspects.

The post-release experiences and background characteristics of students who entered school and dropped out prior to the completion of the school term were then compared to those who entered school and remained in school. Dropouts differed from those who remained in school only in the respect that it had taken students who subsequently quit school longer to gain entry into school after their release: three weeks versus one week (a difference significant at the $p = .01$ level). Dropouts had not, however, entered later in the term on the average than did those who remained in school. Thus, the failure to finish the school term was not a function of a later release date for subsequent dropouts.

Although no significant differences in the levels of basic skills were obtained between students who completed the school term and those who dropped out, students who dropped out <u>perceived</u> themselves to be poor performers. Forty percent of those who quit school rated their performance as "poor", as compared to 16 percent of those who remained in school. Thus, while students who dropped out of community schools performed as well on tests of basic skills as the general institutional population, and as well as those who actually completed the school term after their release, perceptions of their performance were negative when they entered community schools.

Almost all students who were enrolled in vocational/technical schools, and half of those enrolled in regular schools, had quit prior to the end of the term. Almost all students who had withdrawn from school reported their intentions of returning in the fall school term. Most dropouts intended to return to regular

public schools, although a sizable 30 percent hoped to enter vocational/technical schools.

Of those students who did <u>not</u> enter school, 18 percent reported they had attempted to get back into public schools or vocational/technical schools, but that they were unsuccessful due to factors outside their control (they were barred from enrolling, trade school acceptance was highly competitive, etc.) Almost one-fourth of non-attendees reported that they had wanted to attend school upon their release but had never attempted to enroll, citing their perceived inability to gain acceptance into desired trade schools, the need to work, etc., as the reasons for their failure to seek enrollment. The remaining 60 percent of releasees who did not return to school cited their lack of motivation and dislike for school as the reason for non-attendance: for 34 percent this lack of desire was based on perceived personal inabilities (e.g., "I couldn't catch up," "I have learned all I can," "Competition is too tough"). The remaining respondents merely stated their preference for working over school attendance. Thus, only 25 percent of non-attendees (13 percent of the released student population) expressed the definitive lack of interest in school that is often thought to characterize the majority of the released student population.

A final comparison was made of those releasees who either attempted unsuccessfully to gain entry to schools--or at least desired entry--and those who had no desire and made no attempt to enter school. Because a relatively small number of students interviewed had unsuccessfully attempted entry or manifested unfulfilled desires to attend school, it was not anticipated that significant differences would be found. However, significant differences were found on the background characteristics of age and ethnicity. Older respondents more frequently expressed no desire for additional schooling. Minority members were more likely to report unsuccessful attempts or non-realized desires to enter schools than were non-minority respondents.

In order to assess student perceptions of the positive aspects of institutional education programs and their impact on the transition to the community, students

were asked to provide retrospective evaluations of their institutional school experiences. Sixty-five percent reported a helpful experience in the respect that their basic skills and survival skills improved, and that greater self-confidence and motivation resulted from the school experiences. These respondents, however, acknowledged that in spite of such improvements they did not have enough instructional exposure to gain competency in basic skills or to qualify for a GED. The 35 percent who found institutional schooling unhelpful claimed they didn't learn anything and that the instruction offered by institutions was below their own level of ability.

Although the average age of the post-release follow-up students was 17--which made them at least chronologically near equals to public school graduates--80 percent of the interviewed releasees had no diploma or General Equivalency Diploma (GED). In fact, institutional schooling had provided this credential for only 13 percent of those released from correctional institutions. Interviewed students expressed desires for such a degree: 40 percent intended to complete high school, 25 percent intended to acquire a GED, and 20 percent hoped to acquire a trade-school certificate.

Pre-release assistance in gaining entry into school was at a low level. Only one-fourth of respondents reported assistance from either institutional personnel, community agents, or parole officers in entering school. For one-half of those receiving pre-release services, the services consisted simply of encouragement; others reported receiving actual assistance in handling administrative details in making the transition. This finding supports recent conclusions (GAO, 1977) that institutions assume few responsibilities in assisting youths in their transition from correctional institutions to schools in the community. Post-release services were similarly reported at a low level of delivery (30%). No students from the two adult institutions received post-release services. The receipt of either pre- or post-release services was reported by 43 percent of respondents; 7 percent of respondents received both pre- and post-release services in making transitions to non-institutional schools. As would be expected, those who received pre- or post-release assis-

tance were more likely to enter school. Controlling for age did not change the magnitude of this relationship. In general, parole agencies continue to be the major providers of post-release school transition service to students (Glaser, 1969).

The findings on the post-release educational experiences of students are suggestive of targets of concern. Those who might be of immediate concern, for example, are the 75 percent of the non-attendees who cited personal inabilities or situational constraints as reasons for their failure to attend school, and those students who entered school after release but subsequently dropped out. For these releasees, services could be offered that would build on existing motivations (e.g., stated intentions of entering school) and provide opportunities to overcome fatalistic expectations of never being able to catch up academically with one's peers.

For the one-fourth of respondents who did not return to school and who unambiguously cited personal preferences for working rather than school attendance, service delivery could possibly be more difficult. However, even the statement of negative attitudes regarding school entry should not be taken at face value; to do so assumes that nothing could change these preferences. If negative attitudes toward school are based on the fact that these respondents have had no positive educational experiences in institutional settings or elsewhere, then the "choice" to discontinue schooling may not be a real choice, since it is made without alternatives.

EMPLOYMENT EXPERIENCES

Interviewed students reported a fairly high degree of success in obtaining employment after their release. Seventy-two percent of interviewed students reported having obtained employment at some point following their release. While this 72-percent rate may be indicative of the motivation and capability of released students to obtain some type of employment, it is not a good indicator of the level of employment at a given point in time. Seventy percent of the interviewees who obtained jobs reported that they were no longer working

at the job they initially obtained at the time of the interview. Only slightly over one-half of these initially employed respondents were re-employed at the time of the interview. Thus, only 47 percent of interview respondents were employed at the time of the interview.

The types of employment obtained by those released from institutions consisted of unskilled or semi-skilled types of labor, such as waiters, busboys, maintenance, construction assistants, service station attendants, etc. In similar findings of types of jobs obtained by releasees, Glaser (1969) found a practice of releasees initially taking clearly unsatisfactory jobs to fulfill parole obligations, with the intention of changing jobs for more favorable employment at the first opportunity. Job turnover, therefore, would not necessarily be an indicator of poor adjustment. Considering the age of the respondents, jobs obtained by this group of ex-offenders are probably not unlike the types of employment obtained by their non-offender counterparts. The important issue in the long run, however, is whether this group is more likely to remain indefinitely in these categories of employment by virtue of their inability to enter trade schools or to complete high school diploma requirements.

RECIDIVISM AND REINSTITUTIONALIZATION

Within the three-to-seven month period of release, 40 percent of interviewed releasees reported further offense violations. About 10 percent of the recidivists reported that their violations were officially dismissed by the subsequent actions of police or prosecutors. Calculating recidivism rates using the alternative method of number of offenses (rather than number of offenders) revealed a 72-percent rate of further offense commission for releasees.

Approximately 30 percent of the recidivists' most serious offenses consisted of (1) violations that are legally noncriminal in nature if such offenses were to have been committed by an adult (e.g., runaway); (2) violations of conditions of parole (e.g., failure to participate in a treatment program); or (3) violations for which a great deal of variation exists in whether or not arrests are made given the pervasiveness of such behavior among society members as a whole (e.g., disorderly conduct, drunkeness, and possession of marijuana).

An examination of background characteristics, attitudes, the receipt of transitional services, and problematic situations encountered after release revealed that only the age of the respondent, the respondent's feelings of control, and a history of placement in non-parental homes (e.g., foster or group homes) significantly discriminated recidivists from non-recidivists. Thus, younger releasees, who had histories of foster home placements, and who felt that factors beyond their own control determined events in their lives, were more likely to report further offense violations after their release from institutional confinement.

Thirteen percent of releasees who reported offenses after their release had been reinstitutionalized at the time of the follow-up interview. Reinstitutionalized respondents were found to differ from other recidivists only with respect to age: older recidivists were more likely to be reinstitutionalized. The official response to further violations of younger releasees was more likely to be continuation on probation, or participation in some form of community-based treatment.

CURRENT SERVICES FOR RELEASED STUDENTS

The influence of pre- and post-release services upon post-release adjustment--a great concern of policy makers--has been the subject of several investigations. The rationale for the provision of pre-release, liaison, or post-release services to released offenders is that these services are supposed to facilitate a transition to school or employment and should reduce the likelihood of poor post-release adjustments.

The present study sought to investigate whether the personal needs of releasees, in terms of school transitional and entry assistance, employment assistance, and counseling support for emotional adjustments, were being met by state correctional institutions, by parole agencies, or by community support agencies.

Institutional Pre-Release Services

Regardless of the perceptions of the desirability of providing program services from a humanistic perspective alone, correctional administrators disclaimed the implied association between institutional program services or treatment and a releasee's subsequent success in a community. Many factors have been demonstrated to interact with institutional treatment while students are confined, or after their release, to affect post-release success (Kwaitler, 1977). Nonetheless, the need for pre-release services has been recognized in a small proportion of institutions and has manifested itself in the development of specially designed pre-release programs to facilitate the transition to non-institutional living. About one-fifth of the institutions identified pre-release planning as a special program objective.

These pre-release programs pursue objectives such as the development of job-finding skills, the setting of long-term and short-term career objectives, the identification of methods for controlling drug and alcohol-related problems, the development of social skills, and the setting of expectations for post-release difficulties in terms of psychological adjustment (e.g., the impact of moving from a highly structured to a relatively unstructured setting; the need for self-discipline, etc.).

Despite the development of such integrated approaches to releasee assistance among a few institutions (where the entire institutional program is designed to be an integrated approach to release rather than to confinement), correctional institutions as a whole do not offer much in the way of systematically administered pre-release services. One reason for relatively few pre-release activities is the often distant location of institutions from communities, and a lack of communication with community service agencies. The combination of physical distance, the unknown identity of community agencies, and the large number of communities to which students return has--at least in the past--prevented the development of a systematic and integrated post-release service delivery system.

Community Agency and Student Reports of Post-Release Service Delivery

Less than one half of respondents reported to have had any contact with a community counselor after their release and, using an estimate based on the average site-level percentage of students receiving services, a figure of less than 33 percent is obtained. This latter estimate removes the extreme influences of particular states that have extensive systems of helping-agencies, and other states where no such agencies exist: neither of these situations seems to be characteristic of states or communities on the whole.

For the one-half of releasees who received counseling services, contact with a community counselor was reported to have been made within one week after release. Sixty percent of those who were in contact with counselors had been sent to them by their parole officers; less than 20 percent of those receiving services were contacted by the agency providing services. In the remaining situations, students themselves initiated the contact (but generally only for job counseling services) or parents of the student initiated contact (generally for the student's emotional, alcohol, or drug problems). Counselors were seen an average of once a week.

Most of the students receiving services cited the counselor's emotional supportiveness, and the fact that they had someone to talk to, as the most positive aspects of counseling. Fewer than ten respondents who received counseling felt that the counseling was unhelpful. Fifteen percent of those who did not receive counseling services (mostly releasees not on parole) actually wanted and perceived a need for such services. The type of services requested was related to employment and vocational training for 50 percent of respondents, and psychological services for another 50 percent. Those who did not receive post-release counseling services and did not desire services felt that they had no need for services (71%), that counselors in general are of no help (15%), and that the releasees were reluctant to discuss personal problems with others. Despite a general reluctance to admit a need for assistance for oneself, 30 percent of respondents reported that most students released would need emotional support and someone to talk to in order to relieve anxiety upon their release.

Post-release services offered by community agencies ranged from narrowly focused programs working on only one aspect of adjustment problems (e.g., drug addiction) to comprehensive therapeutic community programs designed to service every need of released offenders (e.g., housing, employment, schooling, psychological counseling, financial assistance).

Helping agencies most frequently serving released offenders were classified by three types: (1) agencies serving the job assistance needs or vocational rehabilitation of offenders; (2) neighborhood clinics generally operating as offshoots of mental health centers; and (3) programs serving the special needs of clients, such as drug, alcohol, or substance abuse centers. Other helping agencies identified were primarily regional developments. For example, one state has a well-developed network of group homes; in two states alternative schools (with the same goals as public schools but recognizing the skill deficiencies and motivational difficulties of released students) were operative; another state utilized volunteer programs (e.g., "Big Brothers"); and in still other states the recreational program approach was used by several agencies to reach those ex-offenders with adjustment problems who were not likely to seek the assistance of a community agency.

In three of the six states where students were tracked, parole services appeared to be the major source of available assistance. These parole agencies expanded their activities beyond a purely supervisory role. They were operating job-assistance programs, work-release programs, vocational rehabilitation departments, group homes, and crisis-intervention centers. In still another state, a network of community-based treatment centers had more or less taken over the supervisory and treatment roles of parole officers. In these situations, the parole agency served primarily as an administrative and legal agency, assigning releasees to various treatment centers.

PERSISTENT PROBLEMS

In spite of current failures to provide pre-release and transitional services to meet student needs, most institutional staff were aware of the problems

students face in making the transition to non-institutional life, particularly with respect to school-related failures. Institutional staff at all sites were extremely consistent in their identification of the following problems:

1. The emotional needs and basic-skills deficiencies of formerly institutionalized students require a highly individualized form of attention that is not provided by public schools.
2. The interruption of schooling when students enter and are released from institutions at mid-term makes it difficult, if not impossible for them to successfully re-enter regular community schools.
3. The fact that correctional educational standards for course requirements are frequently not accepted by public schools presents difficulties for students when they transfer to public schools.
4. The self-discipline problems and motivational problems of students often result in poor school attendance after their release.
5. The fact that public schools have not identified school deficiencies and have not developed realistic expectations for students with special needs in terms of factors such as required course load is thought to prevent released students from entering and succeeding in schools.
6. It was thought that schools would be reluctant to admit students who have a history of school-related problems or that students with institutional records might be stigmatized; however, these problems were not substantiated by released students themselves.
7. The competitive nature of acceptance into the highly desired vocational or trade schools, and the inability of institutionalized students to compete for these positions, leaves students with few alternatives after their release.

Many of these problems are not easily remedied. Others could be remedied, but by sources other than federal programs. For example, the problems posed by the unacceptability of institutional coursework is a matter for which state

educational agencies must provide a solution. The extent of this problem was documented in the Phase II impact evaluation. More than ten states have attempted to remedy this problem by the creation of correctional education as a separate school district under the jurisdiction of the state educational agency rather than the state correctional agency.

Solutions to the problems of transferring to schools or other transitional problems depend to a large extent on the type of institution in which residents are housed. Students residing in urban facilities are more easily served, since the community to which they are returned may be at hand, and the community helping-agencies might be familiar to institutional personnel. Among institutions physically isolated from communities, and from which students return to many communities, the solution is more difficult.

Community agency respondents also reported problems that limited their effectiveness in meeting released offenders' needs. Service-oriented bureaucracies were described as bogged down in red tape; inter-agency competition for funds was reported, and piecemeal service delivery was the admitted outcome. Administrators in one state, for example, pointed out that releasees may be certified to receive counseling services, while their parents are not. Thus, family therapy cannot be provided. Competition for funds and the need to provide positive program evaluations has been said to result in the practice where "...most service agencies disallow the provision of services to the more difficult clients--the ones who fill correctional institutions and require extensive, integrated services."

In many situations, the extent to which community agencies were serving released offenders' needs was impossible to determine from the community agency perspective because releasees from correctional institutions were not always readily distinguishable among clients, and because little inter-agency communication existed. Nonetheless, almost all service agency respondents called for major reorganizations of the service delivery system to replace the present piecemeal intervention.

CHAPTER VI. THE FUTURE OF TITLE I EVALUATION

As a part of its contract with the Office of Education, SDC evaluated the present system for reporting evaluations of Title I and proposed an alternative system. A review of Title I evaluations submitted to the U.S. Office of Education found that, apart from the sheer failures of one-half of the states to report data at all, there was no uniformity in content or format to the reports (Hoyt, 1977). An over-all assessment of the program was thus virtually impossible. Furthermore, less than half of the evaluation reports provided information about project expenditures, project duration, instructional content, or the student participants. More than one-fourth did not describe the evaluation instrument or methodology. To remedy this lack of information, SDC proposed a reporting system (Webb and Keesling, 1976) that would provide uniform information about several aspects of the Title I program at each institution and would permit aggregations of this information up to the state and federal levels. In the SDC proposed system, data are to be collected on student participation, on educational personnel (and their training), on program characteristics, and on student test performance.

Before this system is presented in any detail, however, it should be noted that the previous chapters of this report have documented that Title I programs in institutions are sometimes difficult to distinguish from other educational programs. If the services are not distinctive and are not offered to a distinct group of students, then evaluation becomes virtually impossible. In addition, many students receive small amounts of the services that are offered. Detecting the small effects such low-level treatments are likely to produce would require grouping together data from many similar institutions and may not be worth the effort.

In the following discussion, two different models for collecting achievement data at the institution level are presented, followed by an outline of the entire reporting system (from institutions through to the U.S. Office of Education).

The emphasis of the data collection in the proposed reporting system is on the assessment of the effects of Title I programs on the basic skills of the students served by the program. Two models were developed for collecting and analyzing the achievement data: the Systematic Allocation Model and the Criterion Model. The Systematic Allocation Model was designed for use in institutions that provide Title I instruction to a reasonably large group of eligible students, and provide only regular education instruction to another reasonably large group of eligible students. The Criterion Model was designed for use in institutions that provide Title I instruction to most students, and would not have, therefore, an adequate comparison group to permit the use of the Systematic Allocation Model.

SYSTEMATIC ALLOCATION MODEL

The Systematic Allocation Model compares the performance of students receiving Title I instruction (in addition to regular education instruction) to the performance of entirely dissimilar students receiving only regular education instruction. In this evaluation model, the neediest students are systematically assigned to the Title I program. This model requires that students be assigned to the Title I program or regular education program according to a single cutoff score on an 'index of need.' This index may be a test score or a composite index of test scores and other ratings, such as teacher recommendations. All students exhibiting initial performances below the cutoff score must be assigned to a program of Title I supplementary instruction. All students exhibiting performances above the cutoff must be assigned to regular education instruction only. This model can be used only when the cutoff score is strictly enforced.

The Systematic Allocation Model was designed to assess whether Title I instruction, combined with regular education instruction, is more effective than regular education alone. It can also be used to determine how close Title I students come to a pre-established criterion of performance, and whether (and how much) the performances of Title I students improve. The impact of Title I supplementary instruction compared to regular education alone is assessed by comparing the observed performance of Title I students to their expected performance in the absence of Title I instruction (i.e., with regular education

instruction only). The expected performance in the absence of Title I supplementary instruction is determined from the performance of students receiving only regular education instruction by adjusting their performance to account for differences in the index of need. If the observed performance of Title I students is better than their expected performance in the absence of Title I instruction, then the conclusion is that Title I instruction (supplementing regular education instruction) is more effective than regular education alone.

CRITERION MODEL

In some cases students cannot be systematically allocated to Title I or regular education instruction. Programs that plan to give most students Title I instruction fall in this category. The Criterion Model was designed to assess the performance of Title I students when a sufficiently large comparison group is not available.* This model allows the observed performance of Title I students to be compared to a pre-established criterion of performance; it also assesses the magnitude of improvement in performance over time. Only the performance of students receiving Title I instruction is examined in this model.

The major drawback to this model is the arbitrary nature of the criterion to be met. An appropriate standard should evolve over time, after experience with this model has been obtained.

It should be noted that the Criterion Model can be used to supplement the Systematic Allocation Model when students are systematically assigned to educational programs. The Criterion Model would provide information on the performance of Title I students compared to a pre-determined criterion, and on improvements in performance over time.

*To provide a valid comparison, the group not receiving Title I instruction must be formed as described under the Systematic Allocation Model. Any other group of students at the institution, e.g., vocational education students, would not provide a valid comparison for the purpose of evaluation.

OTHER STUDENT INFORMATION

In addition to the test scores, the age at the time of assignment to the program (or pretest) is given to determine the range in student ages and for use in determining percentile scores, where appropriate. Test dates are given to show the variation in length of duration of instruction across students and to show the variation in entry and exit dates.

DATA ON THE CONTENT AND OPERATION OF TITLE I PROGRAMS

In order to provide information to all levels of program operation about the content and operation of the Title I programs, additional data are to be collected about several characteristics of each project. The information on program characteristics is obtained for several purposes. First, a program evaluation could relate program characteristics to student achievement in order to determine whether certain kinds of programs facilitate academic progress more than others. Second, teachers are given a vehicle for relaying problems and suggestions to administrators, as well as providing a further insight into what happens in the classroom. Third, administrators at the state level are provided with information on how programs are operating.

Student participation is recorded by program (Title I or regular education) and by subject matter (reading or mathematics). To determine how many students are enrolled in academic education programs, the number of eligible students who are enrolled in Title I and regular education programs during the reporting period is recorded. Enrollment is determined separately by subject matter and by program to show the programs' concentrations in different basic-skills areas. Because enrollment fluctuates during the year, the number of eligible students enrolled in each program is determined at two points during the year, approximately six months apart. A comparison of enrollments at the two points in time to the total enrollment during the year shows the amount of turnover in the programs. The number of students eligible for Title I services during the year or reporting period is given to show the proportion of eligible students enrolled in academic programs.

Information on educational personnel and training is determined separately by program. Employment is separated into two categories: aides, and all other personnel. The (relative) number of aides indicates generally how much support teachers have in classrooms. Participation in inservice training is also given separately for the two personnel categories.

Data on program characteristics is provided by teachers and principals, separately for reading and mathematics. Principals provide information on the hours of exposure to reading and mathematics instruction for a typical student in a typical week, and the instructor-to-student ratio, separately for Title I students and regular education students. The hours of exposure indicate how much educational instruction is given at an institution, and the instructor-to-student ratio shows the size of the average instructional unit.

The principal also describes several features of the structure of the Title I reading and mathematics programs. Information on objectives provided by the principal shows the orientation of the program. The kinds of materials, staff available, special services, and classroom facilities show the richness and variety of resources available at an institution. Teachers provide information on what happens in the classroom or educational unit. The emphasis is on assessing the degree to which individualized programs are developed for students. Descriptions of instructional and institutional problems and solutions attempted indicate problems specific to this setting that hinder academic progress. Information on new teaching methods tried or planned and suggestions for improvement show the variety of teaching approaches.

REPORTING SYSTEM

In this section the system for gathering, analyzing, and reporting the data described above is presented.

The State Applicant Agency (SAA) gathers the information described above for all institutions in its region. The SAA conducts a statistical analysis for each institution to determine the impact of Title I programs. The results of this analysis are summarized and sent back to the institution with instructions

for interpreting the results. The results of all analyses for an SAA are summarized site-by-site for transmittal to the State Education Agency (SEA). The SAA summarizes the teachers' descriptions for each institution and sends the summary to the principal.* The SAA sends an overall summary of institution descriptive information and Title I program characteristics to the SEA.

The SEA collects the SAA summaries of institution descriptive information, Title I program characteristics, and statistical analyses and sends the package to the Office of Education.

CONDITIONS AND CONSTRAINTS ON THE USE OF THE SDC PROPOSED REPORTING SYSTEM

In developing the reporting system, SDC considered several factors that would influence the choice of models and the formats to be used in reporting results (Webb and Keesling, 1978a). The advisory panels suggested that a discussion of these conditions and constraints be removed from the handbook itself in order to make the presentation of the models clearer. However, these issues are important, and their effects will have to be examined in the future. They are summarized below.

Student transiency affects the reporting system profoundly. The brief average length of stay (median of about seven months) means that there is a considerable turnover in the student population at most institutions. Because students are constantly entering and exiting the institutions' educational programs, it is difficult to standardize both the testing dates and the length of exposure to the educational programs. Fixing the testing dates produces great variation among students in the amount of exposure, while testing each student at entry and at exit creates a complex testing schedule (without necessarily entirely standardizing the exposure time). The Handbook recommends the strategy of testing upon entry and again upon exit because some entry-level testing is almost always done at the institutions and the panels felt that an exit test would not be too difficult to arrange. They did point out, however, that testing too close to the day of admission to the institution, or too close to the day of exit, may produce aberrant scores because of the student's state of emotional stress. In addition, the teachers at some institutions indicated that they were often unaware of the exact date at which a student was to leave--making an exit test impossible to obtain.

*Our Policy-advisory panel strongly recommended that principals not see these descriptions so that teachers could respond in confidence.

Information collected in the impact evaluation and the subsequent substudy of effective practices indicated that there was considerable variation in the receipt of instructional services among students within an institution. In order to build a properly specified statistical model for analyzing data from either the Systematic Allocation Model or the Criterion Model, it may be necessary to include a measure of exposure-to-instruction in the statistical model. The measures of duration of instruction is a start in this direction. However, more sophisticated measures, adjusting for all missed instruction (and, perhaps, for non-task-engaged instructional time) may have to be developed.

Domain-referenced tests (DRTs) must be used in the Criterion Referenced Model and may be used in the Systematic Allocation Model. Generally speaking, the results of the impact evaluation phase (Keesling, Webb, and Pfannenstiel, 1979) suggest that DRTs are generally much more revealing about the levels of basic skills that students have attained than are standardized, norm-referenced tests. For this reason, DRTs are preferred for use in both models.

Rather than require that the specific DRTs used in the SDC data collections be used in the evaluation reporting system, it is suggested that DRTs be developed locally, making them more suitable to local program goals. Appendix A of the Phase II report (Kessling, Webb, and Pfannenstiel, 1979) gives the details of the construction of the DRTs used by SDC and could serve as a model for a DRT development activity. Pages 33 through 38 of the Handblook list resources for ready-made DRTs. The extent to which institutions will require technical assistance in this activity will have to be determined through field trials of the evaluation and reporting system.

Analysis issues were also discussed by Webb and Keesling (1978a). In general, SDC's advisory panels felt that analyses should be performed at higher levels in the educational system (e.g., USOE or the SEAs) rather than lower levels in (e.g., SAAs or institutions) because the resources for performing the analyses would not be available at the lower levels. Two other considerations reinforce this point of view:

1. The background characteristics of the students and the goals of the programs vary over a very broad range. To attempt to aggregate these disparate programs into a general assessment of the effect of Title I could only be performed at a higher level.

2. Because many SAAs and even SEAs have very few institutions to deal with, they would be limited in their attempts to assess the effectiveness of different program characteristics on student achievement or attitude outcomes.

Compromises must be made between too much aggregation (which obscures effects by blending cases and variables that act differently) and too little aggregation (which obscures effects by representing outcomes as if they were the idiosyncratic results of each institution with respect to each variable). Some years of data collection are required before a clear idea of what such compromises should emerge.

Additional issues were brought out in the advisory panel's reviews of the Phase II and Phase III reports. These reviews led to suggestions that the evaluation-reporting system should be modified to provide for collecting additional data in three areas: affective outcomes, the provision of educational services, and the provision of services intended to smooth the transitions from the institutions to the communities.

Affective outcomes are emphasized in many institutions and could be given greater recognition in the reporting system. Of course, this would necessitate a good deal of developmental work to generate appropriate measures and might require a modification of the evaluation models. Expanding from the exclusive

focus on basic skills would probably reflect the concerns of program staff at the institutions more closely.

The Phase II and Phase III reports documented the rather low levels of provision of educational services (when measured by the hours of instruction received by students). It was felt that by collecting additional data on this aspect of the educational programs, a better understanding of the service delivery system would be obtained and, perhaps, some modification to the institutional settings that lead to these low service levels might be effected.

The post-release study (see Chapter V and Pfannenstiel, 1979) documented the lack of services provided to students (prior to their release or after their release) to ease their resumption of 'normal' activities in the community. If Title I money becomes available to provide for these services, which seem to be badly needed, then the reporting system will have to be modified to provide data on these services and their effects on students.

THE FUTURE OF THE EVALUATION REPORTING SYSTEM

The document, known as The Handbook, which describes the SDC proposed evaluation reporting system, has been circulated for review and comment for about one year. A very few comments have been received, mostly of a technical nature. The basic structure has not been questioned.

Before the reporting system can be recommended for wide-scale usage, however, field trials are needed to assess whether or not the data can be captured, reported and analyzed in an accurate and meaningful fashion. Field trials will also identify areas for which technical assistance should be provided.

After a cycle of field trials and revisions, it is expected that the Office of Education would recommend the adoption of the revised reporting system by the Title I program. Eventually, all institutions receiving Title I funds could be required to use a version of the reporting system.

Based upon the findings of the national evaluation study, the proposed reporting system should be a feasible activity for all levels of program operation and review. Additional review and comment as well as field trials of the system will provide guidance in creating an improved version of this system.

REFERENCES

Anderson, L. M., Evertson, C. E., and Brophy, J. E. An experimental study of effective teaching in first-grade reading groups. Elementary School Journal, 1979, 79, 193-223.

Arlin, M. Teacher transitions can disrupt time flow in classrooms. American Educational Research Journal, 1979, 16, 42-56.

Barr, R. C. Instructional pace differences and their effect on reading acquisition. Reading Research Quarterly, 1974, 9, 526-554.

Barr, R. C. How children are taught to read: grouping and pacing. School Review, 1975, 83, 479-498.

Bartell, T. et al., Compensatory Education and Confined Youth: A National Evaluation of Title I Programs in State Institutions for Neglected or Delinquent Youth, Volume 1. TM-5792/005/00. Santa Monica, CA: System Development Corporation, 1977.

Bloom, B. S. Time and learning. American Psychologist, 1974, 29, 682-686.

Bloom, B. S. Human Characteristics and School Learning. New York: McGraw-Hill, 1976.

Bossert, S. T. Tasks and Social Relationships in Classrooms: A Study of Instructional Organization and Its Consequences. ASA Rose Monograph Series, Cambridge University Press, 1979.

Brophy, J. E. and Evertson, C. M. Learning From Teaching: A Developmental Perspective. Boston: Allyn & Bacon, 1976.

Cressey, D. P. and McDermott, R. A. Diversion from the Juvenile Justice System. Washington, D.C.: National Institute of Law Enforcement and Criminal Justice, 1974.

Dell'Apa, F. Education for the Youthful Offender in Correctional Institutions. Western Interstate Commission for Higher Education, Boulder, Colorado, 1972.

Glaser, D. The Effectiveness of a Prison and Parole System. Indianapolis: Bobbs-Merrill, 1969.

Glaser, R. and Nitko, A. J. Measurement in Learning and Instruction. In Thorndike, R. L., ed. Educational Measurement. Washington, D.C.: American Council on Education, 1971.

Glass, G. V. and Smith, M. L. Meta-Analysis of Research on the Relationship of Class Size and Achievement. San Francisco, CA: Far West Laboratory for Educational Research and Development, 1979.

Hanson, R. A. and Ross, S. M. The adequacy of instructional time allocation. Studies in Educational Evaluation, 1975, 1, 23-29.

Harnischfeger, A. Personal and Institutional Characteristics Affecting Teacher Mobility, IV: Schools Do Make A Difference. Mehr Licht! Studies of Educative Processes, Chicago, 1973, 6.

Heckmann, J. J. Sample selection bias as a specification error. Econometrica, 1979, 47, 153-161.

Hoyt, J. Review of ESEA Title I N or D Evaluation Reports. TM-5792/006/00. Santa Monica, CA: System Development Corporation, 1977.

Keesling, J. W., Webb, N., and Pfannenstiel, J. Compensatory Education and Confined Youth: Volume 2. TM-5792/011/00. Santa Monica, CA: System Development Corporation, 1979.

Kwaitler, L. Behind Bars: Prisons in America. New York: Vintage Books, 1977.

Lipton, D., Martinson, R., and Wilks, J. The Effectiveness of Correctional Treatment. New York: Praeger Publishers, 1975.

Millman, J. Criterion-Referenced Measurement. In Popham, W. J., ed. Evaluation in Education: Current Applications. Berkeley, California: McCutchen Publishing Corporation, 1974.

Pfannenstiel, J. Post-Release Experiences of Students From State Correctional Institutions, Volume 3. TM-5792/012/00. Santa Monica, CA: System Development Corporation, 1979.

Pfannenstiel, J., Sewell, M. M., and Keesling, J. W. Compensatory Education and Confined Youth: The Substudy of Effective Practices, Volume 4. TM-5792/014/00. Santa Monica, CA: System Development Corporation, 1980.

Powell, M. and Dishaw, M. A realistic picture of reading instructional time. Reading Research Quarterly, 1980 (in press).

Reagen, M. and Stoughten, D. School Behind Bars: A Descriptive Overview of Correctional Education in the American Prison System. Metuchen, NJ: The Scarecrow Press, Inc., 1976.

Stallings, J. A. Allocated Academic Learning Time Revisited, or Beyond Time on Task. Paper Presented at the Annual Meeting of the American Educational Research Association, Boston, April 1980.

Vitner, R. D., Downs, G., and Hall, J. Juvenile Corrections in the States: Residential Programs and Deinstitutionalization. Ann Arbor: National Assessment of Juvenile Corrections, University of Michigan, 1975.

Webb, N. and Keesling, J. W. Some Issues in the Design of a Handbook for the Evaluation of Title I Programs in State Institutions for Neglected or Delinquent Youth. TM-5792/009/00. Santa Monica, CA: System Development Corporation, 1978a.

Webb, N. and Keesling, J. W. Handbook for Evaluation of Title I Programs in State Institutions for Neglected or Delinquent Youth. TM-5792/013/00. Santa Monica, CA: System Development Corporation, 1978b.

Wiley, D. and Harnischfeger, A. Explosion of a myth: quantity of schooling and exposure to instruction, major education vehicles. Educational Researcher, 1974, 3, 7-12

APPENDIX A. METHODOLOGICAL SUMMARY OF PHASE I

Phase I of the evaluation was designed to collect descriptive information about representative Title I programs in state institutions, as well as management information from all state coordinators of the Title I N or D program. The decision to select a representative set of programs in state institutions for data collection, rather than visiting all Title I programs, was motivated by a desire to keep research costs within reasonable limits, while still providing unbaised estimates of national program characteristics.

This discussion focuses on four major methodological areas of the research undertaken in Phase I: sampling (including project and respondent selection), instrument development, field operations (including communication with participants, selection and training of data collectors, data collection procedures and debriefing), and coding. Each section of the chapter describes the principal features of the study design as implemented, the rationale for the particular strategies chosen, and any problems encountered that might affect the interpretation of the resulting data.

SAMPLING

The basic Phase I study design called for two kinds of activities in the field: interviewing all State Education Agency (SEA) and State Applicant Agency (SAA) Coordinators of the Title I N or D program in the United States and the territory of Puerto Rico, and visiting a representative sample of 100 Title I projects so as to be able to describe the delivery of program services at the institutional level across the nation. Within the selected institutions, the overall design called for a probability selection of individual respondents.

Since all state coordinators were interviewed, the data obtained at this level involve no problems of statistical inference. The percentages contained in this report are precise indicators of the activities and opinions of state coordinators, within the limits of measurement error. The sampling of projects and of individuals within projects, on the other hand, required somewhat more complex procedures, which are described in some detail below.

Project Selection

Discussions with federal Title I program personnel during the spring of 1976 indicated that the only institutional and project information available for stratification purposes was the type of institution (i.e. adult, delinquent, or neglected), the geographic region, the size of the resident population, and the number of residents eligible for Title I services. Data were unavailable for such project characteristics as the actual number of students served by the Title I program, the funding allocation to individual projects, per-capita student expenditures, and characteristics of instruction. Since size of resident population was found to be highly correlated with number of students eligible for Title I, the use of only one of these two variables for stratification was necessary.

Before selecting the 100 project sites, Title I-eligible institutions were stratified on three dimensions:

(a) type of institution
(b) geographic region
(c) number of Title-I-eligible students

The type of institution variable consisted of three categories: adult institutions, delinquent institutions, and institutions for neglected children. For the regional stratifying variable, the following four geographic areas were used:

Northeast:	Maine, New Hampshire, Vermont, Massachusetts, Connecticut, Rhode Island, New York, Puerto Rico, New Jersey, Pennsylvania, Delaware, Maryland, District of Columbia
Mid-Atlantic/ Southeast:	Virginia, West Virginia, Kentucky, Tennessee, North Carolina, South Carolina, Florida, Georgia, Alabama, Mississippi, Arkansas, Louisiana
Midwest:	Michigan, Ohio, Indiana, Illinois, Wisconsin, Minnesota, Iowa, Missouri, Kansas, Nebraska, South Dakota, North Dakota

Mountain/West:	Montana, Idaho, Wyoming, Utah, Colorado, Texas Oklahoma, New Mexico, Arizona, Nevada, California, Oregon, Washington, Alaska, Hawaii

Two sets of categories for number of eligible students were used for stratification: (a) less than 70 students, and 70 students or more (for delinquent institutions); and (b) less than 45 students, and 45 students or more (for adult institutions). The groupings indicated for the latter two stratifying variables were chosen so as to distribute cases equally by cell, to the extent possible.

It was thought that institutions having extremely small numbers of students eligible for Title I would be likely to either: (a) have no Title I program actually operating at the facility; or (b) have a Title I program consisting mainly of equipment purchases, with little or no investment in instructional personnel. Conversations with state administrators confirmed this expectation. Due to the relatively low or non-existent Title I funding at such institutions, it seemed advisable to eliminate them from the sampling frame before selecting the sample. Setting a criterion level of eleven eligible residents or more had the effect of removing 124 facilities from the study population. It should be noted that the use of this procedure disproportionately eliminated adult institutions with relatively small numbers of residents. On the other hand, regional variations in the proportion of institutions removed from the frame were relatively minor.

Once the 460 remaining institutions had been stratified, projects were selected systematically with known probability so as to produce a nationally representative sample of all Title I projects in facilities having eleven or more eligible students. Table A-1 shows the number of facilities selected from each stratum.

In order to test for possible bias introduced by eliminating institutions having small numbers of eligible students, a comparison was run between the characteristics of institutions that participated in the Title I program in

Table A-1. Number of Sample Selections by Stratum

	Neglected Institutions	Delinquent Institutions		Adult Institutions	
		Eligible Students < 70	Eligible Students ≥70	Eligible Students <45	Eligible Students ≥45
Northeast	2	15	6	3	3
Mid-Atlantic/ Southeast	2	6	9	7	6
Midwest	2	4	7	4	4
Mountain/ West	4	5	6	3	2

1975-76 (as reported to SDC during Phase I data collection), and institutions that met USOE's requirements for eligibility in 1975-76 but did not participate in the program. Table A-2 indicates that non-participating delinquent institutions were more often small institutions, with correspondingly small numbers of Title I-eligible students, confirming our a priori expectations. For example, whereas 13.3 percent of all Title I-participating institutions (N=437) were delinquent institutions in the northeast with fewer than 100 residents and less than 70 eligible students, over one-fourth (26.2 percent) of all eligible but non-participating institutions in the country (N=184) had this set of characteristics. Most of the non-participating adult institutions had small numbers of eligible students, regardless of the size of the resident population. The procedure of eliminating institutions with small numbers of

CONTINUED

1 OF 2

Table A-1. Number of Sample Selections by Stratum

	Neglected Institutions	Delinquent Institutions		Adult Institutions	
		Eligible Students < 70	Eligible Students ≥ 70	Eligible Students <45	Eligible Students ≥45
Northeast	2	15	6	3	3
Mid-Atlantic/ Southeast	2	6	9	7	6
Midwest	2	4	7	4	4
Mountain/ West	4	5	6	3	2

1975-76 (as reported to SDC during Phase I data collection), and institutions that met USOE's requirements for eligibility in 1975-76 but did not participate in the program. Table A-2 indicates that non-participating delinquent institutions were more often small institutions, with correspondingly small numbers of Title I-eligible students, confirming our a priori expectations. For example, whereas 13.3 percent of all Title I-participating institutions (N=437) were delinquent institutions in the northeast with fewer than 100 residents and less than 70 eligible students, over one-fourth (26.2 percent) of all eligible but non-participating institutions in the country (N=184) had this set of characteristics. Most of the non-participating adult institutions had small numbers of eligible students, regardless of the size of the resident population. The procedure of eliminating institutions with small numbers of

Table A-2. Comparison of Title I Participating and Non-Participating Facilities[a]

	Neglected Institutions		Delinquent Institutions								Adult Institutions								Total Row Percentages	
			<100 Residents				≥100 Residents				<100 Residents				≥100 Residents					
			Eligible Students <70		Eligible Students ≥70		Eligible Students <70		Eligible Students ≥70		Eligible Students <45		Eligible Students ≥45		Eligible Students <45		Eligible Students ≥45			
Northeast	1.1[b]	0.0[c]	13.3	26.2	0.9	0.5	0.9	0.5	6.2	2.2	0.7	3.8	0.0	0.5	4.8	1.1	4.1	5.5	32.0[b]	40.3[c]
Mid-Atlantic/ Southeast	1.1	0.0	5.3	3.3	1.8	0.5	0.2	1.1	8.3	0.5	1.1	2.7	0.2	0.0	4.6	24.5	5.0	2.7	27.6	35.3
Midwest	1.1	1.1	2.8	0.5	1.6	0.5	0.5	0.0	6.4	1.1	1.4	4.3	0.2	0.0	3.9	2.7	3.0	1.1	20.9	11.3
Mountain/ West	2.3	0.0	3.2	3.3	0.7	0.0	2.7	0.0	5.7	0.5	0.0	0.5	0.0	0.0	2.5	8.7	2.3	0.0	19.4	13.0
Total Column Percentages	5.6[b]	1.1[c]	24.6	33.3	5.0	1.5	4.3	1.6	26.6	4.3	3.2	11.3	0.4	0.5	15.8	37.0	14.4	9.3		

[a] Entries in this table are expressed as percentages of the total number of facilities reported.

[b] Participating Facilities FY76 (N=437)

[c] Eligible But Non-Participating Facilities FY76 (N=184)

eligible students prior to sample selection is thus unlikely to have produced any bias in the representativeness of the sample for Title I programs nationwide.

Since separate estimates were desired for projects in institutions for neglected children, these institutions were oversampled so that a minimum of 10 state institutions for neglected children (of the 27 such institutions nationwide) would be selected for the Phase I site visits. The oversampling of institutions for neglected children required the use of weighting procedures in the analysis of overall project characteristics (due to unequal selection probabilities across strata), and reduced somewhat the number of cases available for estimation in the remaining types of institutions (delinquent and adult institutions). The actual selection probability for neglected institutions was .37, while the selection probability for delinquent and adult institutions was .21.

It should be emphasized that the sampling frame for the Phase I study was constructed from eligible institutions, not institutions actually participating in the program. At the time of the study, no information existed at the federal level on which institutions were currently operating Title I programs. It was therefore necessary to draw replacements for institutions that were selected for the sample, but turned out to have no Title I program. Replacements were drawn at random from the same stratum as the original selection. Seventeen of the original selections were replaced in this manner.

Evidence for the representativeness of the resulting sample of institutions (with respect to the universe of institutions actually participating in the Title I program) is shown in Table A-3. Taking into account the intentional oversampling of neglected facilities (by a factor of 1.7) discussed earlier, it can be seen that the distribution for sample institutions closely approximates the characteristics of institutions that were implementing Title I programs in 1975-76.

Table A-3. Comparison of Title I Participating Facilities and Sample Facilities[a]

	Neglected Institutions		Delinquent Institutions				Adult Institutions				Total Row Percentages	
			Eligible Students <70		Eligible Students ≥70		Eligible Students <45		Eligible Students ≥45			
Northeast	1.1[b]	2.0[c]	14.2	15.0	7.1	6.0	5.5	3.0	4.1	3.0	32.0[b]	29.0[c]
Mid-Atlantic Southeast	1.1	2.0	5.5	6.0	10.1	9.0	5.7	7.0	5.3	6.0	27.7	30.0
Midwest	1.1	2.0	3.2	4.0	8.0	7.0	5.3	4.0	3.2	4.0	20.8	21.0
Mountain/ West	2.3	4.0	6.0	5.0	6.4	6.0	2.5	3.0	2.3	2.0	19.5	20.0
Total Column Percentages	5.6[b]	10.0[c]	28.9	30.0	31.6	28.0	19.0	17.0	14.9	15.0		

[a]Entries in this table are expressed as percentages of the total number of facilities reported.

[b]Participating Facilities FY76 (N=437)

[c]Sample Facilities (N=100)

Respondent Selection

The SEA Coordinators and all SAA Coordinators of the Title I N or D program were interviewed in each state, the District of Columbia, and Puerto Rico. The SEA Fiscal Officer in each location was also given a questionnaire on program cost information. At the institutional level, all facility directors, fiscal officers, educational program administrators, and Title I supervisors were interviewed. Where possible, the following persons were interviewed in each institution: one academic education teacher, one vocational education teacher, one Title I math teacher, one Title I language arts teacher, two Title I adult or student teacher aides, one custody staff member, and one treatment staff member. In each instance, these persons were selected at random from the roster of all persons in these roles. In addition, ten students from the list of all students participating in the Title I program in each sample institution were selected at random and interviewed in groups of four to five. Observations were also conducted in randomly selected Title I and non-Title I classrooms.

INSTRUMENT DEVELOPMENT

Since the basic objective of Phase I was to collect information about state agencies and representative Title I programs in state institutions, areas of interest to this description were delineated in the instrument development process. Identified areas of interest included: the characteristics of the institutional, educational, and Title I populations; the pattern of money flow; lines of federal/state institutional communication; institutional organization and climate (including custody vs. treatment orientation, institutional communication and conflict, decision-making mechanisms, and relative support for education and other institutional activities); the implementation of eight designated program components in Title I programs (needs assessment, performance objectives, program operations, program management, inservice education, testing and evaluation, program integration, and program spin-offs); and the comparison of Title I and regular correctional education programs. Items were constructed to yield information relevant to policy issues such as the nationwide role of Title I programs in institutions for neglected or delinquent

youth, the various ways in which Title I programs are managed, and the appropriateness of the basic design of the Title I program for institutionalized youths.

Instruments were pretested in New York, Michigan, Georgia, and Washington. Several findings of the pretest influenced the configuration of the instruments. First, pretest results indicated large variations across institutions on most items. For many items, an exhaustive set of responses could not be specified that would adequately describe the complexity of actual Title I service delivery. Therefore, instruments were revised to a partially open-ended format; that is, questions were posed without a predefined set of possible response categories. For these items the respondent's answer was simply recorded verbatim by the interviewer.

A second pretest finding concerned the manner in which data would need to be obtained from various respondents. Respondents in the pretest indicated that, due to the sensitive nature of some of the questioning that related to the operations of correctional programs and staff interactions, a personal interview where rapport and trust could be established was essential. Personal interviews were seen as having the added advantage of reducing respondent burden by avoiding additional paperwork.

A final pretest finding was that Title I students were more likely to respond to questioning if they were interviewed in groups and their individual responses were made less salient.

Instruments took one of three forms: self-report questionnaires (SEA Fiscal and Facility Fiscal); interview instruments (State Coordinator, Facility Director, Education Program Administrator, Teacher, Adult/Student Aide, Custody/Treatment Staff, and Student); and observational forms (Institutional Environment, Classroom Observation, and Title I Materials Observation).

In the final state of development, the instruments were reviewed by the Policy Advi ory Group; the Research Advisory Group; the Office of Planning, Budgeting,

and Evaluation (OPBE) in USOE; the Division of Education for the Disadvantaged in USOE: the forms Clearance Office in USOE; the Education Data Acquisition Council (EDAC) in the Education Division of DHEW; the Office of the Secretary of DHEW; and the Office of Management and Budget (OMB).

FIELD OPERATIONS

Two major tasks were performed by the field operations staff during Phase I. First, staff members served as the focal point for all communication between the study staff and the institutions and agencies selected to participate in the study; second, they organized and managed the data collection teams and their activities.

Communications with the Participants

Shortly after the award of the contract, the Honorable Terrell H. Bell, then United States Commissioner of Education, wrote to all Chief State School Officers to notify them of the study. His letter discussed the reasons for the study, the major purposes of the study, and the study design. Richard L. Fairley, Director of the Division of Education for the Disadvantaged, then sent an informational notice to all state Title I ESEA Coordinators. After the 100 institutions had been selected, Dr. John W. Evans, Assistant Commissioner of USOE's Office of Planning, Budgeting, and Evaluation (OPBE), wrote to all of the participating agencies and institutions to notify them of their selection as participants, to explain the purpose of the study, and to provide preliminary information on how the study would be conducted.

Shortly after these letters were sent, and during the time SDC was initially contacting the participants, the project staff provided additional information about the study to the correctional education community by publishing information articles in professional corrections journals and attending the annual meetings of professional associations. This attempt to make the correctional education community aware of the study also enabled SDC to recruit data collectors with correctional and/or educational experience.

A member of the field operations staff then contacted each State Education Agency (SEA) and discussed the study with the individual responsible for Title I N or D programs within that state. The purpose of these calls was to introduce SDC, schedule the visit, and determine what other state agencies also had some responsibility for Title I programs within the N or D institutions in that state.

After these initial calls to the SEA, each State Applicant Agency (Department of Corrections, Youth Authority, Agency for Neglected Children) within the state was contacted. During the discussions with the SAAs, an attempt was made to determine whether that agency had responsibility for the Title I programs in the facilities that reported to the agency. In addition, schedules for interviews were established and permission to contact the participating facilities in the state was obtained. The next step in this series of communications was to contact each institution in the sample. These calls provided SDC with the name of the person in charge of Title I programs, the number of students being served by Title I, the number of teachers in the Title I program and in the regular education program, and the focus of the Title I program. It is worth noting that all of the sampled institutions agreed to participate in the study.

Follow-up mailings were sent to each agency and to the institutions selected to participate in the study. These mailings confirmed the schedule for the visit and provided suggested time allotments for the activities required during the visit, a list of types of staff who would be interviewed and/or observed, the method of selecting participants, and other information needed to enable the agencies or institutions to prepare for our visit. The fiscal questionnaire, which would be collected by the SDC data collection team during the visit, was also distributed.

A few days prior to each visit, one of the data collectors called the on-site contact person from the field to confirm the schedule and answer any remaining questions about the visit.

In order to encourage communications between the field operations staff and the participating agencies and institutions, a toll-free phone was established in the field operations office, and all participants were encouraged to call at any time. Due to certain delays in obtaining instrument approval, the schedule for data collection changed several times. Each such change required communications with the SEAs, SAAs, and institutions to reestablish a convenient time for the visits. Despite these delays, excellent cooperation was received from all states and institutions.

Selection and Training of Data Collectors

The collection of data in Phase I required 13 two-person teams to visit all SEAs and SAAs as well as the 100 institutions in the sample. This was accomplished in 10 weeks of actual data collection. The 26 team members were recruited, hired, and trained in three separate groups.

Data collectors were selected from among applicants having experience in corrections, correctional education, or compensatory education, and/or experience in collecting data on other large national education evaluations. Before the data collectors were selected, notices about the study and the need for data collectors were sent to professional organizations, universities and colleges, state correctional agencies, and minority placement agencies. The response was very gratifying; among the respondents were practitioners from the field of correctional education, individuals engaged in graduate degree programs in both corrections and education, and people who had taught in compensatory education programs in the public schools. Many of these individuals had also been previously associated with other large educational research projects.

Three one-week training sessions were conducted during the fall of 1976. Each training session had two major objectives: (1) to introduce the data collectors to the objectives and design of the study; and (2) to prepare them to collect data using the study instruments. Various training methods were used including role playing, one-on-one encounters, and reviewing the instruments

item by item. Trainees were expected not only to be able to administer the study instruments, but also to understand why each item was important to the analysis plan.

Data Collection

All Phase I data were collected during the fall of 1976, and represent the characteristics of the Title I N or D program at that time. During the first week of data collection, a member of the project staff spent the full week in the field with each of the data collection teams. The teams were thus monitored to ensure that they were well prepared for their tasks, and that an experienced staff member was available for questions during the team's first visit. After the first group of teams had been in the field for three weeks, they were split up and paired with members of the second group of trainees, who were going into the field for the first time. Similarly, the third group of trainees, starting three weeks later, were paired with experienced persons from both the first and second groups. The design of the data collection effort thus ensured that there would always be someone on each team with previous experience in collecting data for this study. Four separate data collection trips were made, with new teams being formed for each trip.

In addition to accompanying the teams to the field during the first week of data collection, the project staff monitored the teams during other field visits. These visits were scheduled so that each new data collector was observed toward the beginning of his/her first data collection trip, and periodically thereafter.

The data collection teams made two types of visits, one to state agencies and the other to institutions. Each visit to a state agency lasted about four hours and included a short briefing on the study. State-level visits took about two days per state, and typically included visits to the SEA and two SAAs, and preparation of a case study integrating the information obtained at the various state-level offices. Each visit to a institution took about three days, and included selecting and notifying the participants, conducting

the necessary interviews, observing various classes and activities, writing a detailed case study, and editing the data to be returned.

Participants were assured of the confidentiality of their responses and of anonymity for themselves and their agency or institution. The procedure of using numbers, rather than names, to identify respondents, and requesting the institution to retain the selection list, was designed to guarantee that once the team left the institution, no outside record of who participated in the interview or observations would exist.

During Phase I, the teams administered the state Title I Coordinator instrument to a total of 52 SEA Coordinators and 73 SAA Coordinators. The Facility Director instrument was administered to all 100 directors or wardens. The Education Program Administrator instrument was used with individuals who were in charge of the entire education program, or who supervised the Title I program, or both--a total of 120 persons.

The total number of teachers interviewed was 313, of whom 143 were Title I teachers (including 82 Language Arts teachers and 61 Mathematics teachers); 96 were non-Title I academic teachers; and 74 were non-Title I vocational teachers. Nine remedial specialists or diagnosticians were also interviewed using the Teacher questionnaire.

The Adult/Student Aide instrument was administered to 91 Title I aides. The Custody/Treatment Staff questionnaire was administered 186 times, including 79 interviews with custody personnel and 107 with treatment staff. The Student questionnaire was administered in group interviews to 857 Title I students.

The Class Observation form was used in 339 classrooms, of which 167 were Title I classrooms and 172 were non-Title I classrooms. The Materials Observation form was used 220 times and the Institutional Climate Observation form 95 times. A State Fiscal questionnaire was returned by every SEA, and every facility in the sample completed a Facility Fiscal questionnaire. A summary of the number of interviews and observations is provided in Table A-4.

Table A-4. Numbers of Interviews and Observations Conducted

Role or Instrument	N
SEA Coordinators	52
SAA Coordinators	73
Facility Directors	100
Education Program Administrators	120
Title I Language Arts Teachers	82
Title I Mathematics Teachers	61
Academic Teachers	96
Vocational Teachers	74
Remedial Specialists and Diagnosticians	9
Title I Aides	91
Custody Staff	79
Treatment Staff	107
Title I Students	857
Title I Classroom Observations	167
Non-Title I Classroom Observations	172
Materials Observations	220
Institutional Climate Observations	95
State Fiscal Questionnaires	52
Facility Fiscal Questionnaires	100

Debriefing Sessions

Upon the return of each team to the home office, a member of the field staff and a member of the editorial staff held a debriefing session with the team and reviewed the case studies for the visits they had just completed. These debriefing sessions served to alert the staff to any data problems, and helped to ensure that all important data points were included in the case studies.

In addition to these specific debriefings, general debriefings were held with all data collectors after their last trip. During these group sessions, the data collectors were asked to provide general feedback on the study, the instruments, and their impressions of particular aspects of the overall Title I N or D program.

At both state agencies and institutions, the data collection teams were well received. Most respondents viewed the evaluation as a way to express themselves on Title I and felt that a national study of Title I N or D was long overdue. The data collectors reported that most people appeared to be very open and willing to discuss both the positive and the negative aspects of their programs.

CODING

Ten coders were selected, all of whom had prior experience in the coding of interview data. A week was spent in familiarizing the coders with the logic and intent of each questionnaire item, and in training them on the forced-choice items (items requiring a choice among predesignated response categories). Categories for the open-ended responses were developed somewhat later, after a sufficient number of interviews had been completed, and a second week of training was spent on open-ended responses at that time. As the coding got underway, quality-control procedures were established. A minimum of 20 percent of the open-ended coding on each questionnaire was reviewed for individual accuracy and consistency by the two supervisors.

All data were also reviewed by a second coder to ensure that the correct numerical values were being entered in the appropriate column field and that the forced-choice coding was correct. Edits for logical consistency were also done at that time.

To establish the level of coding accuracy, a reliability check was conducted about three weeks after open-ended coding began. Coding reliability was defined as the percent agreement among coders on a given item. Completed items from all instruments except the state and fiscal questionnaires were included in the coder reliability check. Because of the large number of open-ended items on each questionnaire, a subset of completed items was selected from each instrument. Items were chosen on the basis of uniqueness (whether the item was in that questionnaire only) and analytical importance. Items that were replicated across instruments were selected from the most appropriate instrument. For example, a question about the strongest aspects of the Title I program that appeared in both the Adult/Student Aide questionnaire and the Education Program Administrator questionnaire was pulled from the latter instrument. The rest of the items were selected on a random basis. The total proportion reviewed was approximately 50 percent of all the open-ended items.

The ten coders all coded the same material. Eight examples of each type of questionnaire were coded, to ensure that at least five answers to a given item would be available from each coder. Answers that were either "no answer," "don't know," or "no problem/no change" were also not counted, as they could be expected to raise the percent agreement figure spuriously.

Items that consisted of "multiple mentions" (i.e., several different answers to a given question) were considered to be one item: the percent agreement for all codeable mentions was totalled, and an average score for the item was computed. Some collapsing of discrete coding categories was done within logical groups of codes to approximate the types of aggregated classification groupings used for the Phase I analysis. After one month of open-ended coding, the overall coder reliability score for the ten coders was a very respectable 79 percent.

Attempts were made to correct data that were found to be inconsistent, illogical, or missing during editing. An example of illogical or inconsistent data would be a sample sheet indicating no Title I aides at a facility and a completed interview schedule for an aide. Generally, a phone call to one of the data collectors who visited the site would clear up any problems with incomsistent data. Missing information for analytically important items (except open-ended items) was completed by first probing the memory of the data collectors who had visited the particular agency or institution. In cases where it was necessary, and for all open-ended items, the respondent was recontacted. For items involving actual figures (e.g., number of participants in the institution's Title I program or amount of Title I funds), the data were obtained directly from the appropriate respondent.

APPENDIX B. METHODOLOGICAL SUMMARY OF PHASE II

This summary presents an overview of the Phase II impact evaluation of the National Evaluation of Title I Programs in State Institutions for Neglected or Delinquent Youth. More detailed information is contained in the full report, which is Volume 2 of the series, entitled Compensatory Education and Confined Youth.

The first phase of the study (reported in Volume 1) provided a description of Title I programs at 100 randomly chosen institutions. The second phase of the study was designed to address the question of whether Title I programs are effective in improving students' basic skills and attitudes, and to determine, if so, which components of educational programs and which characteristics of institutions are related to such improvements. The study was non-experimental in nature, surveying current practices in state institutions across the nation.

In the following sections, the sampling plan, instrumentation, field operations, development of domain-referenced tests and methodological problems encountered are described.

SAMPLING PLAN

From among the 100 Phase I sites, a nationally representative sample of 30 institutions was selected for participation in the impact evaluation. A purposively selected sample of 10 additional sites that were judged to be of particular interest based on Phase I data was also included in the impact evaluation.

The sampling plan for this study was based upon knowledge of project characteristics discovered during Phase I (Bartell et al., 1977). The plan was reviewed and revised according to suggestions made by the Policy and Research Advisory Groups.

Two variables were used for stratification: degree of emphasis on basic skills, and average per-student expenditure. The Phase I study indicated that Title I programs in institutions for neglected or delinquent youth could be characterized as providing instruction of one of the following types: 1) instruction in reading and math only; 2) instruction emphasizing reading and math; 3) instruction that placed a minor emphasis on reading and math. These program types constituted the index of "degree of emphasis on basic skills." Title I per-student expenditures were used as the other stratification factor to assure that the full range of such expenditures was included in the study.

Some of the Phase I facilities were excluded from this new sampling frame due to special problems that would make student testing impractical. Sample facilities in Puerto Rico (N-3), where only Spanish is spoken, were dropped from the sampling frame in order to eliminate the problems of ensuring comparability of test items across languages. Three sites with programs that lasted fewer than four weeks were eliminated because such short-term programs are likely to evidence effects too small to be detected, and few students would be available for follow-up testing. Three more sites were dropped from the frame because they were no longer operating Title I programs.

Since the data from Phase I indicated only slight program differences between Title I programs for the neglected and programs for the delinquent, the former were undersampled by a factor of 0.56 for this study in order to equalize their overall sampling fraction with that of adult and delinquent institutions--thus producing a self-weighting sample for our analyses.* This procedure resulted in approximately 6 of the 10 Phase I institutions for the neglected being eligible for selection in this study.

*Neglected facilities were oversampled by a factor of 1.79 in the earlier study. Undersampling selected facilities by a factor of 0.56 produces an overall sampling fraction for these sites in the impact study equivalent to that of delinquent and adult facilities (1.79 X 0.56 = 1.00).

The remaining 87 sites visited in Phase I were distributed as indicated in Table B-1. Thirty of these sites were selected with equal probability. This approach yielded the numbers of sites by cell, as indicated in Table B-2.

In addition to the 30 sites sampled with strict probability methods, 10 additional sites were purposively selected for their likelihood of being effective. Criteria established in the Phase I study (such as strong staff commitment to the program, staff cohesiveness, high quality administrative leadership in the educational context, positive student and staff rapport, and the utilization of innovative approaches) were used to select these sites.

Individual students within sample facilities were selected with probability methods. The goal was to test 35 Title I students in each site where that number of students was available, and to test all Title I students in smaller programs. Comparison students from other remedial and regular academic education programs were sampled and tested in those sites where such programs existed.* Students participating only in post-secondary education programs, or only in vocational education programs, were eliminated from the sampling frame. The intent was to test a total of approximately 25 comparison students in sample institutions where that number of students was available. In the smaller facilities and programs, all residents were tested at each point in time to minimize disruptions to these facilities caused by the selection of all but a very few students.

One important implication of not taking a constant proportion of students from each sample facility is that, while institutions were selected with equal probability, the total set of individual students did not constitute a self-weighting sample. When students were pooled across facilities to do individual-level analyses, weighting factors had to be introduced to correct for unequal selection probabilities.

*It should be pointed out that the designation of students as Title I students or as recipients of Title I services was based on differing criteria among the institutions. In some institutions Title I students were those assigned to individualized labs; in other institutions they were identified on the basis of their receipt of instruction from teachers whose salaries were paid by Title I funds; in still other institutions Title I students were those students utilizing materials and equipment paid for with Title I funds.

Table B-1. Number of Facilities Eligible for Selection into Phase II Samples, by Strata

Title I Program Emphasis	Title I Per-Student Expenditure $501+	$201-500	$1-200
Reading, Math Only	10	7	10
Strong Emphasis on Reading, Math	11	13	8
Weak Emphasis on Reading, Math	12	8	8

Table B-2. Number of Facilities in Phase II Sample, by Strata

Title I Program Emphasis	Title I Per-Student Expenditure $501+	$201-500	$1-200
Reading, Math Only	3	2	3
Strong Emphasis on Reading, Math	4	5	3
Weak Emphasis on Reading, Math	4	3	3

The weighting factors ranged from 1.0 to 5.2 across the 30 random sites, for Title I students, and from 1.0 to 8.3 across sites for non-Title I students. At 20 of these sites all Title I students were tested, while at 21 of these sites all regular education students not receiving Title I instruction were tested. In addition, one of the randomly sampled sites provided Title I services to all students. Thus, only eight of the 30 randomly sampled sites had weights other than one (1.0) for the non-Title I students.

One of the major problems in collecting data from these institutions was the variable availability of students for testing. It was anticipated that the constant release of students from the facilities would reduce the number of students available for the planned four waves of testing. In order to compensate for this attrition, lists of alternate students from the first rosters were provided during the first wave of data collection. In institutions where sampling was performed, substitutions were made from those lists in order to retain, as nearly as possible, 35 Title I and 25 non-Title I students. These students were then maintained in our sample throughout the study. No more selections were made from students who were entered on the first rosters.

To maintain a representative sample across testings, students who entered an educational program between testings were added to the roster and selected at the same rate as previously tested students (per program, per institution) and retained for subsequent test administrations. A subsequent section of this chapter examines the retention rate of students tested in the first visit.

INSTRUMENTATION

Table B-1 provides a guide to the use of the instruments in this study. Teachers, educational program administrators and facility directors provided information about the goals and operational characteristics of the educational program and the Title I components in particular. Information on the costs of the educational program was obtained from each institution's fiscal officer. Institutional staff maintained records of dates of student entry to and exit from educational programs, as well as the time students actually spent in instructinoal settings.

Table B-3. Content of Instruments Used in the Phase II Study

Instrument	Content
Facility Director Questionnaire (FDQ)	Characteristics and goals of the institution and the educational program
Facility Fiscal Questionnaire (FFQ)	Costs of operating the regular education and Title I instructional programs.
Educational Program Administrator Questionnaires (EPA)	Characteristics and goals of the educational programs
Teacher Questionnaires	
Part A (TQA)	Characteristics and goals of the educational programs
Part B (TQB)	Implementation and process information about the Title I reading program
Part C (TQC)	Implementation and process information about the Title I mathematics program
Student Questionnaire (SQ)	Student background information
Student Data Sheet (SDS)	Record of student entry into and exit from educational programs
Student Participation and Attendance Record (SPAR)	Record of student time in instructional settings
Classroom Observations (CO)	Data collectors' assessments of program characteristics and goals
Test Session Observations (TO)	Data collectors' record of testing activities
Attitudes Toward School and Learning (SAA)	Paper and pencil survey of student attitudes
Practical Achievement Scale (PAS)	Paper and pencil test of student achievement related to practical settings
Domain-Referenced Tests (DRT)	Paper and pencil tests of student achievement in basic skills
Comprehensive Test of Basic Skills (CTBS)	Standardized, norm-referenced, paper and pencil test of achievement

FIELD OPERATIONS

The field operations staff performed three major tasks during Phase II of this study. First, they were responsible for communication between SDC and the 40 institutions participating in the study; second, they selected and trained the data collectors who conducted the actual institution visits; and third, they performed all the tasks associated with coordinating and scheduling the arrival of personnel and materials at each institution. In addition to these major tasks, the field operations staff monitored the activities of the data collectors during each of the four data collection trips.

During the fall of 1977, Dr. John W. Evans, Assistant Commissioner of USOE's Office of Evaluation and Dissemination (OED) wrote to all State Administrators of the ESEA Title I Programs in institutions for neglected or delinquent children. His letter explained the second phase of the study and enclosed the list of selected institutions. A second letter was then sent to the 40 institutions from the project director explaining in more detail the intent of the Phase II study and the basic involvement of each institution.

The field operations staff then called the Phase I contact person at each institution to provide more information about the Phase II study, specifically regarding the test and data collection activities scheduled for that institution. During this preliminary discussion the staff members accomplished the following objectives:

- The institutional contact person for the Phase II evaluation was identified.
- A broad overview of the institution's task involvement in Phase II was presented.
- Dates and schedules for the first visits were established.

After each institutional visit was scheduled, a Procedures Manual for institutions participating in the second phase of the impact study was mailed to each contact person. This manual contained information on 1) the background and

objectives of the study; 2) the purpose of the Phase II site visits; 3) procedures for assuring the confidentiality of student data; 4) student selection procedures; 5) site preparations prior to visits; 6) site activities during visits; and 7) site activities between visits.

Enclosed with the Procedures Manual was a set of blank student rosters used to identify the population of students from which the participants in the study would be selected. Prior to each subsequent visit the institutions were provided new rosters upon which to list students who had entered the education program since the completion of the previous roster. These rosters were designed so that the contact person could link student names to identification numbers. Only the student identification number and program data were sent to SDC. In this way student anonymity was ensured. These rosters were then used to aggregate students by Title I and regular education program types and to select those to be tested.

Institutions were visited four times at six-week intervals and students were tested each time. The six-week cycle for visits was formulated on the basis of two factors that are important to the evaluation of Title I programs in neglected or delinquent institutions: (1) test administrations should be scheduled at intervals of sufficient length so that gains in achievement can be detected; and (2) testing intervals should be relatively short to enhance the availability of sample students for subsequent test administrations. Each visit required two days and was conducted by a two-person team. Eight data collectors, working in two-person teams, were able to visit the 40 sites within a four-week schedule. Institutional visits were scheduled at six-week intervals, as follows:

Site Visit	From	To
1	November 28, 1977	December 23, 1977
2	January 9, 1978	February 3, 1978
3	February 20, 1978	March 17, 1978
4	April 3, 1978	April 28, 1978

During the two-day visit, each data collection team was responsible for three activities:

1) Questionnaires

   The team was involved in administering, distributing, and assisting institutional staff in the completion of six questionnaires and two institutional reporting forms.

2) Observations

   During each visit the team recorded their observations about test administrations. During the second visit they recorded observations about the institution and the educational program.

3) Test Administration

   The team administered achievement tests, attitudinal measures, and questionnaires to students during the four visits to the sites.

Table B-4 shows the schedule of activities during each site visit.

Prior to the arrival of the data collection team at the institution, the field operations staff mailed all testing materials to the institution. Also prior to the team's arrival, the institution scheduled students for testing. The following guidelines for test scheduling were used: test sessions should not have more than fifteen students each, and each student should be scheduled for two different sessions during the visit. If possible, the two test sessions for a student were to be scheduled on different days. The first activity for data collectors upon arrival at an institution was to confirm that all of the pre-visit activities had in fact been accomplished.

During the first test session the data collector explained to students the purpose of the testing, how they were selected, confidentiality procedures, and the voluntary nature of the testing. The general procedure followed by the data collectors was to read all items and responses--except when reading was part of the tested outcome. Only the Comprehensive Tests of Basic Skills required a timed administration. The reading aloud of items on other tests did tend to set a pace, however. The first session usually required about one and

Table B-4. Schedule of Data Collection Activities for Phase II Study

ACTIVITY	VISIT			
	1	2	3	4
Questionnaire Administration or Distribution				
Facility Director (FDQ)	X*			
Facility Fiscal (FFQ)	X			
Educational Program Administrator (EPA)	X			
Teacher Questionnaire				
Part A (TQA)	X			
Part B (TQB)	X			
Part C (TQC)	X			
Student Questionnaire (SQ)	X			
Student Data Sheet (SDS)	X	X	X	X
Student Participation and Attendance Record (SPAR)	X	X	X	X
Field Observations				
Classroom Observations (CO)		X		
Test Session Observations (TO)	X	X	X	X
Test Administrations				
Attitudes Toward School and Learning (SSA)	X		X	
Practical Achievement Scale (PAS)		X		
Domain-Referenced Tests (DRT)	X	X	X	X
Comprehensive Tests of Basic Skills (CTBS)				X

*An "X" indicates that this instrument was administered during this visit.

a half hours from the time the students entered until they left; the second session was usually shorter. Between 1500 and 1700 students were tested during each visit, using SDC-developed tests of performance in basic skills. These tests will be described in more detail in the following section.

After the test administration, the team collected, packaged and mailed all materials to SDC. Figure B-1 displays the order in which tests were administered during the four sites visits.

DEVELOPMENT OF DOMAIN-REFERENCED TESTS

Domain-referenced testing is an outgrowth of recent developments in the field of educational evaluation. With the growing interest in individualized instruction and the attendant emphasis on the diagnosis of deficiencies and the prescription of remediations, it has become clear that standardized, norm-referenced tests are not adequate in assessing individual patterns of proficiency. At the same time, as interest has grown in comparing curricula, and in assessing the impact of educational programs, it has been suggested that the standardized, norm-referenced test might not fairly represent the content of the curriculum or program being implemented.

The purpose of domain-referenced tests is to provide a clear, rigorous definition of an item in a specific content area. The performance of each student in terms of "measurements that are directly interpretable in terms of specified performance standards" (Glaser and Nitko, 1971) can thus be assessed. Furthermore, this enables the curriculum designer and the teacher to assess the correspondence between the curriculum and the test. Millman (1974) has provided an explicit definition of a domain-referenced test:

> Any test consisting of a random or stratified random sample of items selected from a well-defined set or class of tasks (a domain) shall be referred to here as a DRT [domain-referenced test]... "Well defined" refers to an explicitly stated domain or items or tasks or to an item-generating procedure.
>
> The major advantage of the DRT is that it permits the user to make a special kind of CR [criterion-referenced] interpretation, namely an estimation of an examinee's domain score or level of functioning, defined as the percent of the population of items the examinee could answer correctly. . . . (p. 315).

SESSION	INSTITUTION VISIT: FIRST	SECOND	THIRD	FOURTH
I	Student Questionnaire (All Students)	Student Questionnaire (New Students)	Student Questionnaire (New Students)	Student Questionnaire (New Students)
	Attitudes Toward School and Learning	Practical Achievement Scale	Attitudes Toward School and Learning	Domain-Referenced Test Mathematics
	Domain-Referenced Test Mathematics	Domain-Referenced Test Mathematics	Domain-Referenced Test Mathematics	Domain-Referenced Test Reading
II	Domain-Referenced Test Reading	Domain-Referenced Test Reading	Domain-Referenced Test Reading	CTBS

Figure B-1. Sequence of Test Administrations

The domain-referenced tests for use in Phase II were developed in three steps: domain identification and definition, domain-specification, and item/test generation.

The identification and definition of the domains to be tested was a crucial step because the validity and usefulness of the tests is a function of the importance and meaningfulness of the domains to be tested. An extensive search of the available curriculum and test materials in the basic skills area was undertaken, coupled with inputs from members of the Policy Advisory Group and the Research Advisory Group. This resulted in the selection of the basic skills domains for use in the Phase II study. The second step involved the specification of the content to be tested, the item form to be used, and the rules for generating individual items. The third step, test generation, involved the sampling of items from domains and assembling the items into test forms. This step included the overall testing strategy, the design of test instruments, and the characteristics (number of items, length of time, etc.) of individual tests.

The reader is referred to Appendix A of the Phase II report for further details on these procedures. The content of the resulting DRTs is summarized below.

The Domain-Referenced Test of Arithmetic Basic Skills

The DRT of arithmetic basic skills was constructed to include the four basic operations of addition, subtraction, multiplication, and division. Each of these operations was presented in a separate test booklet. Within each booklet there were four discrete parts. The first part consisted of eight arithmetic 'fact' items--operations on single digit integers--typically found on 'numbers tables' posted on elementary school classroom walls. The second part consisted of items involving integer numbers, at least one of which had more than one digit. The third part consisted of operations on decimals, and the fourth part consisted of operations on fractions.

A fifth test booklet was produced in which basic-skills operations were applied in descriptions of practical settings. This test had 12 word problems. The

content ranged from making change to computing simple interest. The score on this test was considered separately from the other arithmetic DRTs in assessing outcomes. Unlike the other arithmetic tests, these problems required reading skills.

The Domain-Referenced Tests of Language Arts

The domains included in the language arts DRTs were word recognition, vocabulary, and literal comprehension. The word recognition and vocabulary tests each yielded one score. The literal comprehension test consisted of eight passages with certain words deleted. The student had to find the most appropriate word to insert from a list of words provided at the point of deletion. The eight passages were graded on a scale of reading difficulty. Passages from a variety of content areas were included, selected from materials found in grades 2 through 10.

Form-to-Form Difficulty and Reliability

The domain-referenced tests were administered four times at six-week intervals. To minimize the effect of practice, four parallel forms of each test were constructed. At the first test administration, forms were randomly assigned to students so that each form was assigned to one-fourth of the students in the sample. To assess the magnitude of differences across forms, analyses-of-variance were conducted comparing scores on the four forms during the first test. Table B-5 gives the mean proportion correct for each form of a test and the significance of the F ratio for the analysis-of-variance of each test. Different forms of a test showed highly comparable levels of difficulty. The mean proportions correct were nearly identical across all forms of the arithmetic tests. Among language arts tests, paragraph comprehension and vocabulary showed differences across forms. An examination of the means for those tests revealed that Form 2 was significantly easier than the other forms in paragraph comprehension. In vocabulary, Form 2 was significantly harder than the other forms. Thus, scores in paragraph comprehension and vocabulary were adjusted for the form differences in subsequent analyses. Because analyses of the other tests showed no statistically significant differences between forms, those scores were not adjusted.

Table B-5. Differences Across Parallel Forms of the Domain-Referenced Tests at Wave 1*

Test	Mean Proportion Correct for Form				Significance of F Ratio
	1	2	3	4	
Addition	.79	.78	.79	.78	NS
Subtraction	.73	.72	.72	.72	NS
Multiplication	.59	.56	.58	.57	NS
Division	.56	.54	.53	.53	NS
Practical Mathematics	.29	.30	.30	.29	NS
Word Recognition	.94	.95	.94	.95	NS
Vocabulary	.66	.63	.67	.69	.001
Paragraph Comprehension	.67	.76	.68	.69	.001

*All students who had taken a test were included.

Because the domain-referenced tests were administered more than once, information on the reliability of the test scores was assessed in several ways. First, a measure of internal consistency, Cronbach's coefficient alpha (α), was computed. This coefficient is an index of the homogeneity of items within a test. Second, a measure of consistency over time, the correlation between scores at different test administrations, was computed. This correlation shows how much test scores change relative to the change in other scores over time. It should be noted that this correlation is not a simple test-retest correlation, which is usually interpreted to be influenced only by random fluctuations in the condition of the student or the testing environment. This correlation includes the influence of instruction on the knowledge of the content of the domain-referenced tests. Because the exposure to instruction differed across students and institutions, one could not expect a perfect correlation over the six-week

intervals even if the test correlated perfectly on two occasions close-in-time. Both measures of reliability are given in Table B-6. The measures of internal consistency are fairly high for all tests. Each test seems to be homogenous in content. Thus, although the arithmetic tests each consisted of items on facts, integers, decimals, and fractions, the central arithmetic operations, i.e., addition, produced consistency among these different types of items.

Table B-6. Reliability of Domain-Referenced Tests

Test	No. Items	Reliability Coefficient						
		Internal Consistency (α)*	Correlations Between Tests at Waves**					
			1,2	1,3	1,4	2,3	2,4	3,4
Addition	28	.86	.66	.75	.70	.76	.73	.81
Subtraction	25	.90	.74	.73	.73	.77	.76	.78
Multiplication	24	.89	.78	.75	.73	.80	.80	.76
Division	24	.93	.84	.78	.75	.80	.80	.82
Practical Mathematics	12	.78	.73	.69	.68	.74	.75	.78
Word Recognition	16	.90	.56	.64	.49	.75	.63	.65
Vocabulary***	24	.81	.60	.67	.56	.65	.64	.57
Reading Comprehension***	75	.97	.65	.74	.69	.71	.74	.75

*Coefficient alpha (α) was calculated for test scores at the first wave of testing.

**The coefficients in this table were computed on a minimum of 280 cases and a maximum of 840 cases.

***Because form-to-form differences were found for reading comprehension and vocabulary, coefficient alpha (α) was computed for a single form and across forms. All analyses produced nearly identical results.

The intercorrelations among test administrations shown in Table B-6 are lower than the measures of internal consistency, which is to be expected. These figures show that scores on most of the tests were fairly stable over time. The lower correlations for practical mathematics and word recognition may partially be explained by the restriction in the range of scores for any one test administration. Many students performed at or near the maximum score on word recognition, showing evidence of a ceiling effect. In practical mathematics, many students performed at or near the minimum score, showing evidence of a floor effect. Because the range in scores on these tests was less than that of the other tests, somewhat lower correlations might be expected.

In summary, the results indicated that, generally speaking, different forms of the DRTs were equivalent. Furthermore, analyses of internal consistency and test-retest reliability showed that the test scores were reliable indicators of student performance.

It should be emphasized that these tests were generated by the application of specific item construction rules. They were not generated by the usual item analysis procedures which eliminate items based on their difficulty levels or correlations with total test scores. The technical quality of these tests indicates that the exclusive use of item construction rules can result in useful, reliable instruments.

Methodological Problems

The most serious methodological problem in the Phase II study was the rapid rate of attrition of students from the sample. Table B-7 presents the percentage of students selected for the first testing who were still at the institution for subsequent testings.

These percentages represent the students who had not been released from the institutions. However, physical presence at the institution was not the only factor determining a student's availability for subsequent testing.

Table B-7. Percentage of Students Selected for First Testing Who Were Present at Institutions for Later Testings

Visit	Percent of Students Remaining
Second	83
Third	74
Fourth	67

Table B-8 shows the breakdown of initially-selected students tested during the four visits.

Table B-8. Students Selected for Testing During Wave 1 Who Were Tested at Subsequent Site Visits

	FIRST	SECOND	THIRD	FOURTH
Selected for testing*	1265	__**	__**	__**
Tested	1222	825	604	413
Not tested				
Not at institution	0	218	324	414
At institution	43	222	337	438

*Data is for the 30 randomly selected institutions.

**Only those students initially selected on the first roster for participation are considered here.

By the fourth visit, 33 percent of the students originally selected to participate had left the institution. Thirty-five percent of the original group were still at the institution at the fourth testing, but for a variety of reasons did not attend the test sessions, as noted above. In fact, over half of the students expected at test sessions--but who did not appear--were still at the institution.

An analysis of the Test Session Observations for the second visit indicated that, of the students who were at the institutions but were not available for testing:

- 28 percent were in "lock-up";
- 21 percent were reported to be sick;
- 23 percent had a conflict appointment which could not be broken (parole board, work release, furlough, etc.);
- 28 percent refused to take the test.

Of the students recorded as no longer at the institution:

- 74 percent had been released or transferred from the institution;
- 13 percent had left the educational program;
- 13 percent were AWOL from the institution.

These data are presented to show the types of problems faced in collecting complete data on these students. These data should not be treated as true estimates of the magnitude of these problems, because reasons for absences were recorded somewhat sketchily. Indeed, one large institution did not explain the unavailability of some 39 previously tested pupils. The data from subsequent testings is not as reliable, in our estimation, because students and institutions could have learned how to provide an excuse other than refusal when, in fact, a student refused to take the test.

These data provide a benchmark for assessing the amount of interference that educational programs may experience in institutional contexts due to competing institutional activities that may be given priority over educational activities. They also indicate that subsequent studies of such institutions must be aware of these conflicts and try to work around them as effectively as possible in order to obtain complete data.

The second most important methodological problem concerned the choice of analysis method to respond to the questions about the impact of Title I programs. At the time this report was prepared there was no widely accepted methodology for treating data in which the assignment to treatment conditions is biased and based upon errorful measures. Had there been a clearly and uniformly established procedure for identifying Title I recipients, an appropriate analysis would have been possible. However, most data analysts believe that data such as that collected in this study could only lead to biased comparisons of Title I and regular education--where the direction and extent of the bias would be unknown. Recent work suggests for future evaluations that an appropriate methodology may be available under certain conditions (Heckman, 1979).

APPENDIX C. METHODOLOGICAL SUMMARY OF THE SUBSTUDY OF EFFECTIVE PRACTICES

The Substudy of Effective Practices was designed to provide for the dissemination of information to correctional teachers, site administrators, and state administrators about projects or practices that produce or hinder achievement gains in basic skills among institutionalized students.* While the Phase II impact evaluation had found that overall basic-skills achievement levels remained stable over a twelve-week period, site-level averages for several institutions revealed some degree of demonstrated effectiveness in producing gains in reading achievement and/or math achievement. Programs at other institutions were not effective.

Questions addressed by this study included the following:

1) What is the interaction of educational program characteristics, student characteristics, and institutional characteristics that best explains student growth in basic skills?

2) How much exposure to basic-skills instruction do students actually receive in institutional settings?

3) What types of instructional approaches seem to work best with institutionalized students?

4) How do some educational programs successfully operate within the context of institutional constraints to produce achievement gains?

On the basis of the empirical findings of the impact evaluation, nine sites that demonstrated success or failure in producing gains in reading and/or mathematics over a twelve-week period of instruction were selected for the indepth Substudy of Effective Practices; student outcome measures utilized in the impact evaluation were readministered during the substudy to validate

*The complete report for Substudy of Effective Practices is Volume 4 of this series of reports.

the previous findings of effectiveness and ineffectiveness for the nine sites, and to investigate the relationship between instructional processes and achievement outcomes.

Findings of the impact evaluation largely influenced and formed the basis of the design of the Substudy of Effective Practices, and were suggestive of factors that may be important to the operation of effective programs. However, the incompleteness of Phase II data regarding the receipt of instructional services presented the need for more accurate measurement and observations of variables and processes that appeared to be critical to the effective operation of basic-skills programs.

An important issue posed by these findings was whether students who spend more time in a given subject area learn more than students who do not. A reasonable hypothesis generated from Phase II data was that the overall lack of growth in basic skills among institutions receiving Title I funds was due to the fact that instructional time competes with other institutional activities. To test this hypothesis, the Substudy of Effective Practices was designed to obtain several measures of instructional exposure, including student class absences and amount of exposure to reading and math-related tasks over a twelve-week period. Furthermore, observations were conducted at three points in time to verify reports of instructional exposure, including observations of the amount of class time actually devoted to task-related and non-task-related activities.

In the following sections, site and student selection procedures, instrumentation, and data collection procedures employed in the Substudy of Effective Practices are discussed.

SITE AND RESPONDENT SELECTION

Site Selection

Selection of the nine sites for the Substudy of Effective Practices was based on reading and math achievement results from the forty sites participating in

the Phase II impact evaluation.* Average site-level achievement gains obtained during the Phase II study were not different for Title I students or regular education students. Thus, all students for whom pre- and post-tests were obtained over a twelve-week period were the basis for the calculation of site-level averages. The determination of whether the site was characterized as gaining, losing, or remaining the same in its reading and math achievement was based on these averages.

The initial sample of 40 sites available for selection for the in-depth study was constrained for several reasons. Constraints included the following situations:

1) Sites were eliminated where too few cases of changes in basic-skills achievement were obtained to justify with any confidence their losses or gains in achievement (i.e., less than 5 students). The types of sites eliminated on this basis were generally those sites having a small student population because other institutional actvities, such as work assignments, gained the attention of students and resulted in a short length of educational program participation (e.g., institutions for adults); those sites experiencing a rapid turnover of student population (e.g., short-term institutions); and those institutions where cooperation from the administration and student population during Phase II was lacking, and where few repeated tests were thus obtained.

2) Sites housing residents that were non-comparable to the general institutionalized population on important background characteristics (e.g., students who were 10-12 years of age, or who had demonstrated severe emotional disturbances) were eliminated.

3) Sites were eliminated where the educational program characteristics were atypical for the institutional setting (e.g., institutionalized

*Because no claims of representativeness are made in this study, both the thirty randomly selected and the ten purposively selected Phase II sites were eligible for selection.

students who attended public schools), and thus would not provide information useful to most correctional educational programs.

4) Sites that had lost their Title I funds during the course of the evaluation were eliminated, as were sites where plans were operant for an institutional reorganization having a dramatic impact on the operations of the educational program. Since the objective of this study was identification of practices related to effectiveness, the effects a major reorganization might have on instruction were unknown. It was anticipated that reorganization could well result in sites that were previously defined as ineffective demonstrating subsequent effectiveness.

5) Sites were eliminated from consideration where empirical losses in achievement were quite large, and where the cooperation of institutional staff with Phase II evaluation objectives had been of sufficiently poor quality to seriously question the validity of the obtained data.

The combination of these factors removed all adult institutions and all neglected institutions from consideration for the Substudy of Effective Practices. Ten Sites remained: (1) three were identified as effective with respect to both reading and math gains; (2) three were effective in reading but ineffective in math;* and (3) four were not effective in terms of either reading or math. Because of constraints that determined the selection of only nine sites, one of the sites that was identified as not effective in either reading or math and that had program characteristics similar to two other ineffective sites (e.g., size) was eliminated. The distribution of achievement gains and losses for the nine sites selected are provided in Table C-1.

Although sites known to be undergoing major organizational changes were not eligible for selection, the problem of reorganization persisted nontheless.

*"0" gain was considered an effective program.

Table C-1. Proportional Site Gains and Losses in Reading and Mathematics for Phase II and Substudy Findings *(Numbers in Italics are Substudy Results)*

Site Number	Reading Gains	Math Gains	Reading Gains	Math Losses	Reading Losses	Math Gains	Reading Losses	Math Losses
1	.010 (16)* *.124* *(33)*	.03 (15) *.068* *(27)*						
2	.29 (8)	.06 (6)					*-.179* *(39)*	*-.005* *(23)*
3	.010 (30) *.023* *(46)*	.020 (27) *.011* *(44)*						
4			.110 (14)	-.050 (11)	*-.020* *(22)*	*.029* *(20)*		
5			.020 (14)	-.050 (14)			*-.001* *(25)*	*-.025* *(20)*
6			.030 (14)	0 (13)			*-.033* *(51)*	*.024* *(47)*
7							-.070 (20) *-.048* *(38)*	-.050 (17) *-.009* *(34)*
8			*.067* *(62)*	*-.020* *(50)*			-.020 (24)	0 (24)
9			*.019* *(31)*	*-.042* *(23)*			-.180 (10)	-.020 (10)

*Numbers in parentheses are the numbers of students on which proportions are based.

During the course of the study, two sites underwent organizational changes so overwhelming that it would be impossible to conclude that the educational program observed was similar to the one found during the Phase II impact evaluation.

Notification of Sites and Meeting of Site Personnel

Because of the need to provide solutions for problems encountered during the data collection effort for the Phase II impact evaluation, it was felt that a meeting with institutional administrators would prove to be beneficial in terms of assistance they could provide in planning data collection strategies and in soliciting support from institutional staff and residents. The institutional administrator and educational program administrator at each of the nine selected sites were contacted to request their participation in the Substudy of Effective Practices. Both administrators at all sites were asked to attend an orientation meeting. The purpose of this meeting was threefold:

1) Although these sites had been accomodating in past phases of the study, continued goodwill and cooperation were essential. The small number of sites and students involved in this study prescribed a strong need for cooperation from institutional staff and students in accomplishing the difficult task of obtaining repeated and valid test measures on institutionalized students. Since Phase II evaluation experiences had revealed that a combination of institutional factors and the lack of individual student motivation had resulted in student unavailability for testing, the orientation for the institutional administrator and the educational administrator was intended to mobilize their support and cooperation in accomplishing study objectives.

2) Both Phase I and Phase II data collection efforts had revealed a large amount of variation on operational characteristics of educational programs that predetermine the types of data that can be obtained and the data collection methodology that can be reasonably employed. For instance, the types of classes offered, the instructional groups formed, length of class time, length of school day,

and days of the week in which the educational program operates all had implications for the study design. Thus, study staff met individually with administrators from each site to obtain information specific to their program and class schedule.

3) Phase II testing had revealed considerable variation among sites in students' willingness to attend the testing sessions and in providing valid test scores. Specific recommendations were obtained from sites on ways in which data collectors could invoke the institutional reward system to motivate students to participate in the testing situation.

During the course of these meetings, site administrators were briefed on the results of the evaluation efforts to date. They were apprised of the difficulties encountered during Phase II and the proposed solutions to these problems. Institutional staff expressed enthusiasm for the study once the purposes of the study and the rationale for collection of certain data were made clear. The net result was that institutional administrators gave their assurance that they would exercise their authority to the extent possible to prevent competing institutional activities from interfering with test administration (though not with the daily operation of the classes to be observed).

Class and Student Selection

In order to replicate the findings of effectiveness or ineffectiveness demonstrated at sites in the previous year, selection of students on a basis similar to Phase II student selection was desirable. Respondents during Phase II represented either all students in small institutions (N=70), or a random sample of students in larger institutions.

While it was possible in this study to observe all the basic skills instruction for all students in smaller institutions (N=3), it was impossible to do so at the larger sites. Furthermore, limitations imposed by the amount of field time and number of personnel available, as well as constraints imposed by the widely varying class schedules of the nine sites, made it impossible

to randomly select students, classes or teachers for this study. In the six institutions where classes and students were selected, SDC staff arrived on site in advance of the data collection team in order to select classes and to arrange the observation and testing schedule.

Since the identification of factors accounting for basic skills achievement gains was the major objective of the study, it was necessary to observe all the basic skills instruction of students to be tested. The combined constraints of limited observational time, a limited number of observational staff, conflicts in class schedules or restricted class schedules (e.g., many basic-skills classes taught at the same time, or classes scheduled for half-day sessions only), and the need for repeated observations for validation purposes limited the basis on which classes and students could be selected.

Wide variation on the ways in which instruction was organized among the nine sites prescribed that class and student selection would differ somewhat at each of the six sites where selections occurred. Selections were made with the following objectives in mind:

1) To reduce expected problems with attrition in the post-test sample size (as was experienced in Phase II), classes and students were selected to optimize the likelihood of obtaining repeated measures, with the added constraint that <u>all</u> basic skills instruction for selected students had to be observed.

2) Somewhat conflicting with the first objective of a sufficiently large sample size was the desirability of selecting small instructional groups so that valid observational measures could be obtained. Basic-skills instruction was scheduled for large groups of students at some sites. It was expected that observers could identify and record the behavior of five students. Generally, no more than five students were ever observed by any one observer. In some cases, however, classes were large and observers recorded the behavior of more than five students.

In these situations it was felt that the added difficulty of observing more than five students was not as problematic as would be the disruption caused by the presence of many observers in the classroom.

3) Attempts were made to select classes and students representing the site's full range of basic-skills instruction. At sites where classes were formed on the basis of measured level of performance, classes were selected from the lower, middle, and upper levels of basic-skills performance.

4) Attempts were made to select classes that additionally assured representation of the basic-skills teaching staff, and the different methods by which instruction was provided (e.g., one-to-one instruction, utilization of instructional media, etc.).

While the meeting of all these objectives in selecting classes and students was desired, they were not always met. The basis on which students were assigned to classes, for instance, at times determined the representativeness of students participating in the study. Students frequently attend classes by "cottage" residential groupings. Such groupings may be formed on the basis of measures of maturity, aggressiveness, or other behavioral indices. In these cases, basic-skills classes at times contained students of widely varying skill levels--some of whom did not appear to be low-level performers.

The population of classes from which selections were made was compiled by site personnel. They were instructed to identify all basic-skills instruction, regardless of whether or not it was designated as a Title I class. The reason for this decision was two-fold: (1) Since it was the objective of this study to identify effective practices that Title I programs might employ rather than to investigate the effectiveness of Title I itself, it was expected that effective practices would not be confined to one funding source. Rather, factors such as the types of basic-skills instruction provided, or the amount of instructional exposure provided, were hypothesized as accounting for the accomplishment of achievement gains; and (2) Based on information received during previous phases of the overall evaluation, it was unclear

whether "Title I" consistently implied much more than a source of funding, since the Title I designation seemed to be based on either student performance, source of teacher funding, materials and equipment bought with Title I funds, or some combination of these factors.

For the first site visit, students selected for participation in the study and for observation generally reflected entire instructional groupings. The composition of classes or groupings at several sites, however, proved to be highly unstable at subsequent visits. Schedule changes, institutional reorganization, and the constant entry and exit of residents often resulted in continually changing instructional groups.

The generalization of substudy findings to educational programs in state institutions as a whole is partly a function of the extent to which the present sample is comparable to the population of students and the population of educational programs in correctional institutions nationwide. Since sites participating in the Substudy of Effective Practices were purposively selected, there is no basis for assuming that site characteristics, individual student characteristics and basic-skills performance of students participating in the present study are comparable to the characteristics and performance of the randomly sampled students and sites that participated in the impact evaluation.

REPRESENTATIVENESS OF THE SAMPLE

Institutional Characteristics

Site selection procedures employed in the substudy clarified the types of state institutions that are not represented by this study. Institutions for adults housing some residents under twenty-one years of age and institutions for neglected children, for instance, are not represented by this study. Furthermore, shorter-term institutions and those institutions concentrating on Title I objectives other than basic-skills performance are not represented in this study. The institutions selected for study come from the population of state institutions for delinquent youths. These institutions comprise

about 60 percent of all institutions receiving Title I Neglected or Delinquent funds, receive two-thirds of Title I Neglected or Delinquent funds, and house approximately two-thirds of the recipients of such funds.

To provide some assessment of the extent to which the nine sites participating in the substudy represent state institutions for delinquent youths nationwide, comparisons were made on institutional descriptors obtained during Phase II. In terms of size of institution, randomly selected institutions for delinquents participating in the Phase II impact evaluation averaged 180 residents. This compares to an average population of approximately 190 residents for sites participating in the Substudy of Effective Practices. Educational administrators at some sites again reported overcrowded conditions. A site-by-site comparison of institutional populations indicated that almost one-half of the institutions participating in the substudy have been operating in the past, and continue to operate at or beyond capacity. At least one-half of the institutions reported housing an additional 30 percent or more students this year than were institutionalized during the Phase II evaluation.

Student Characteristics

The characteristics of individual students participating in the Substudy of Effective Practices were compared to average values on characteristics of students from randomly selected sites participating in the Phase II evaluation. Substudy participants averaged almost 16 years of age, with age ranges of 10 to 19 years. This compares to the average age of 16.6 years for Title I students participating in Phase II, 18.2 years for non-Title I students, and 17.3 years for students in institutions for delinquent youths. For both samples, students had last attended the ninth grade prior to their institutionalization.

The racial/ethnic breakdown of Phase II sites found approximately 50 percent of the students to be of non-minority status. This compares well to the average of 49 percent non-minority students in the substudy sample.

Percentage comparisons of the reasons for institutionalization (status offense, drug offense, property offense and person-related offense), again, reflect very little difference between the students in the nine sites in this sample and the Phase II population. Property offenses were given most often as the reasons for institutionalization.

Students participating in the Substudy of Effective Practices were compared to the Phase II sample of randomly selected Title I and non-Title I students in terms of basic-skills performance. This comparison group of both non-Title I and Title I students was considered the appropriate comparison since the focus of the substudy was basic-skills students rather than students designated as Title I recipients. Although not all non-Title I students were receiving basic-skills instruction during the Phase II impact evaluation, findings had revealed that approximately 30 percent of these students were at least as needy as were Title I students.

The comparisons on mathematics and reading performance level of these samples indicate a surprising degree of similarity. The mathematics performance level of students participating in the substudy was identical to the performance level of students participating in the Phase II study. This similarity of performance was found on each mathematics subdomain of addition, subtraction, multiplication, and division. Substudy students in general also performed similarly on operations on the subdomains of math facts (operations on single digit numbers), integers, decimals, and fractions. Only on the division of decimals did the substudy sample differ from the representative population (performing more poorly on this subdomain).

The performance of students on literal comprehension was similar to the performance of Phase II students. On the average, substudy students provided correct responses for 65 percent of reading passages, while Phase II students responded correctly to 62 percent.

Students scored somewhat comparably to the Phase II sample on the Practical Achievement Scale, a measure of functional literacy. On the average, substudy

students responded correctly to 19 of the 30 items testing math- and reading-related skills with practical problems (e.g., the computation of costs and savings). This compares to correct responses on an average of 17 items for Title I and 21 items for non-Title I students in Phase II.

Characteristics of the sites and students participating in the Substudy of Effective Practices were compared to the characteristics obtained on these same nine sites for the previous year. With respect to size of population and length of stay, data obtained in the substudy were almost identical to data obtained for these sites during the Phase II evaluation. Similarly, average age of the student population and proportions of minority and non-minority members were almost identical for both years.

The average performance level for Title I students for both reading and mathematics obtained in Phase II at the nine sites was compared to the average performance level for site-level substudy averages. For most sites, averages obtained in the substudy were consistent with averages obtained in the previous year for both reading and mathematics. At three sites, however, widely discrepant averages in reading for the two samples were found. For two of these sites, the Phase II averages had been based on the entire student population. However, a sample of students had been selected for the substudy. It appears that poorer performers had been selected from these sites for participation in the substudy. At the third site, the "neediest of the needy" students had been identified as Title I students. Since substudy selections focused on basic-skills students rather than Title I students, the site-level reading performance obtained in the substudy was the average of the Title I and regular education student performance obtained during Phase II.

Average scores on attitudinal measures obtained on substudy students at each site were similar to previous findings at these sites. Only on the attitudinal measures of locus of control, liking of reading tasks, perceptions of teacher supportiveness, and perceptions of the practical orientation of education programs did students in the substudy consistently report more negative attitudes than did the Phase II sample at these same sites.

In general, then, sites selected for their demonstration of effectiveness or ineffectiveness in producing gains in achievement in reading and/or mathematics appeared to have similar samples of students represented in the Substudy of Effective Practices. Except for the effects that the selection of poorer reading performers at two sites, better reading performers at one site and students with more negative attitudes on some dimensions at several sites might have on gains, there is no a priori reason to expect that site and student characteristics of substudy participants are different for the two years. The demonstration of a reasonable amount of accuracy in differentiating effective from ineffective sites is thus expected.

Furthermore, the similarity of site characteristics of the substudy sample to the population of institutions for delinquent youths indicates that study findings of the receipt and delivery of basic-skills services might be applicable to these institutions in general. The comparability of performance levels of substudy students and the representative sample of students in Title I Neglected or Delinquent institutions indicates that the identification of instructional characteristics effective for the substudy sample might also be appropriate for basic-skills students generally.

INSTRUMENTATION

Data collection instruments used in the Substudy of Effective Practices consisted of student outcome measures of achievement and attitudinal measures, staff and administrator conversation guides, and observational protocols utilized by data collectors in assessing the amount of task-related activities occurring in basic-skills classes for reading and mathematics. Contents of instruments used in the study are outlined in Table C-2.

Student Measures

Student outcome measures employed in the Phase II impact evaluation were also utilized in the Substudy of Effective Practices. These measures included most of the domain-referenced tests of basic skills in reading and mathematics,* and the Practical Achievement Scale, a test of functional literacy.

*See the Phase II Report for a complete description of the construction of these tests.

Some sections of the domain-referenced test administered during the impact evaluation were not used in the substudy because of their failure to distinguish among site-level performances. For example, the "word problems" subdomain of the mathematics domain-referenced test was simply too difficult for almost all students. On the other hand, the "word recognition" subdomain of the language arts domain-referenced test was too simple for most students. Hence, neither was used in the substudy.

Mathematics subdomains tested for the substudy were the operations of addition, subtraction, multiplication, and division as performed on "math facts,"* integers, decimals and fractions. Reading comprehension was tested using the eight reading passages utilized in Phase II testing.

Four alternate forms of the domain-referenced tests had been developed for the Phase II impact evaluation. Because one alternate form had empirically been proven more difficult than the others during the Phase II evaluation, it was not utilized in this substudy. Furthermore, because of the fear of a high rate of student attrition at some sites prior to the post-test, one alternate form was reserved for the possibility that some post-tests might be administered by site personnel. The two remaining forms were used at each administration, and one-half of the students received each form. Students were then post-tested on the alternate form.

The impact evaluation had utilized the Practical Achievement Scale, a test of functional literacy. Phase II test administrators had reported favorable student response to this pictorial test focusing on practical experience (e.g., computing the amount of savings on items advertised on sale). Because of these reports of student interest, the Practical Achievement Scale was administered for both the substudy's pre- and post-test, although it had not been used as a change measure during the impact evaluation.

*The performance of math functions on two single digit integers.

Table C-2. Measures Obtained in the Substudy of Effective Practices

Student Outcome Measures	Source/Content
Domain-Referenced Test	Measure of student achievement in basic skills reading and math.
Practical Achievement Scale	Measure of functional literacy, student knowledge related to practical settings.
Attitudinal Measures	Measure of student self attitudes and perceptions of school and institution related activities and student characteristics (e.g., age, ethnicity, type of offense, frequency and length of prior institutional confinement and prior educational experience).
Measures of Instructional Time	
Student Participation and Attendance Roster (SPAR)	Measure maintained by teachers on the amount of individual student exposure to basic-skills instruction over a twelve week time span between pre- and post-test.
Classroom Observation Protocol	Data collector observation measures of the amount of instructional time individual students devote to task and non-task related activities within basic-skills classrooms.
Student Enrollment Data	Record provided by educational administrators on individual student entry and exit date information, hours of academic, vocational and Title I instruction.
Measures of Instructional Process and Program Characteristics	
General Classroom Observation	Data collector assessment of class content, group size, class time allocations for instruction versus other institutional activities, teacher attitudes within classrooms, student teacher interaction, and use of materials and resources.

Table C-2. (Cont'd)

Measures of Instructional Process and Program Characteristics (Cont'd)	Source/Content
Staff Questionnaires	Background characteristics on various educational staff members (e.g., educational administrators, teachers, aides) provides descriptions of individual background characteristics, time allocations, availability and use of resources. Individual assessments of staff support and program strengths and weaknesses.
Site Observation Summary	Observer assessments of staff support and interaction, perceptions of program strengths and weaknesses.
Measures of Student Behavior	
Assessment Summaries	Teacher and observer assessments of student motivation and individual characteristics related to learning and behavior problems.

Measures of student self-attitudes and perceptions of school and institution-related environment that were utilized during the impact evaluation were also administered in the substudy. Students additionally responded to questions on age, ethnicity, seriousness of offense for which confined, frequency and length of prior institutional confinement, and prior educational experiences (as had students in the Phase II study).

Measures of Instructional Time

Phase II impact evaluation findings had suggested that the amount of time devoted to basic-skills instructional activities varied by subject area and by site. In order to address the question of whether students who spend more time in a given subject area learn more than students who do not, measures of instructional time of exposure were obtained using several definitions of "time," and using several methods to obtain the measures. "Instructional time" was measured as (1) the amount of time allocated or scheduled for basic-skills instruction in reading and mathematics; (2) the amount of time individual students receive instructional services (i.e., attend class) in reading and mathematics over a twelve-week period; (3) the amount of engaged time, or time devoted to basic-skills tasks; and (4) the amount of time spent on specific reading- and math-related tasks (e.g., addition, subtraction, multiplication, division).

The measure of time allocated to basic-skills instruction was obtained from the educational program administrator and cross-validated by data collector observations. Scheduled class time was used as a rough approximation of allocated time. Measures of allocated time are problematic, however, in situations where instruction is highly individualized and services are provided to meet the immediate needs of students. In such cases, prior estimates of allocated time are difficult to calculate.

Measures of the amount of continuous exposure to basic-skills instruction for individual students over the twelve-week period separating pretests and post-tests were obtained from all teachers who provided basic-skills instruction to any of the sampled students. Teachers maintained the Student Participation

and Attendance Record (SPAR) for each of two six-week periods separating the pre- and post-tests. These SPARs recorded the number of minutes students attended classes in reading, mathematics, or a combination of basic-skills instruction. SPARs were updated for schedule changes after the first six-week period, and new SPARs were distributed for the last six weeks of instruction.

Measures of the amount of instructional time devoted to task-related activities were obtained from observers using the class observation protocols. In five-minute increments, observational protocols specified the amount of observed time devoted to non-task-related versus task-related activities. Non-task-related time was accounted for according to the source of the interruption. Sources of non-task-related behavior were institutional interruptions (e.g., students were disciplined by the confinement of an entire group to their cottage; students were assigned to work activities, etc.); class-related interruptions (e.g., clean-up activities; personal conversations of teachers and students); and student-related interruptions (e.g., students reported in sick, daydreamed, or conversed with fellow students, etc.).

Task-related time was recorded separately for math- and reading-related activities for each type of class (reading, mathematics or combination basic-skills classes). Additionally, observations of the amount of time spent in subdomains (e.g., addition, subtraction, multiplication, and division) were recorded on protocols. In order to obtain a measure of the quality as well as quantity of engaged time, rating scales for level of student motivation were provided on observational forms.

Measures of Instructional Processes and Program Characteristics

The demonstration of effectiveness in producing gains in basic skills may be a function of the nature as well as the amount of instruction. Since the relative importance of "kind" and "amount" of instruction is unknown, measures of "kind" of instruction were developed. Various kinds of instructions were measured on the basis of class content, group size, teacher behaviors and

instructional methodologies, materials and audio-visual resources used, perceptions of school and institutional environment, and class scheduling arrangements. Descriptions were obtained through staff questionnaires and field observations. Observations on the characteristics of instruction for each student observed were obtained on each observational visit. Repeated observations of instructional characteristics were originally designed for validational purposes. It became clear throughout the course of the study, however, that these characteristics changed frequently for students at several sites. Repeated measures were thus measures of changes in instructional characteristics.

Educational program administrators, teachers, and observers provided estimates of teacher behaviors and attitudes, such as time allocations for instruction versus other institutional activities, and supportiveness of students. Staff background characteristics and teaching experiences were obtained with questionnaires, as were teacher and administrator opinions about institutional factors impeding or enhancing the delivery of basic-skills services.

Program descriptions such as the amount of educational expenditures, staff support and cooperation, and the perceived importance of education were obtained from educational and institutional administrators.

Measures of Student Behavior

The demonstration of effectiveness in producing gains in basic-skills achievement could be a joint function of individual student characteristics and program characteristics. Teacher and observer assessment forms were developed to obtain the additional individual measures of student motivation, learning problems, and behavioral problems.

SELECTION AND TRAINING OF FIELD PERSONNEL

During the three prior data collection phases of the evaluation, SDC had employed a group of practitioners and evaluators with previous experience in correctional and educational settings. Five of the six data collectors who conducted the observations and testing for the Substudy of Effective Practices had been engaged in the previous phases of this study. These five data collectors had acquired almost a full year's exposure to residents and programs in all types of correctional institutions that operate nationwide. The sixth data collector had previous data collection experience and was experienced in institutional settings, but had not worked on previous phases of this evaluation.

Since the data collection requirements of this primarily observational substudy were different from previous data collection requirements, a one-week training session was held prior to the site visits. Data collectors were briefed on the design and objectives of the study and were familiarized with data collection instruments and testing procedures.

Considerable attention was paid to a discussion of and training in observational techniques of field data collection. The use of class observational protocols was practiced in compensatory education classes in a local public school. The classrooms observed were selected because of their similarity to correctional education classes (e.g., highly individualized instruction, lab settings). These practice sessions allowed observers to develop techniques for the rapid and easy identification of students they were to observe, and to reach agreement among themselves and with the study's staff as to the appropriate use of instruments.

DATA COLLECTION PROCEDURES

Data was collected from each of the nine institutions participating in the study on three site visits conducted at six-week intervals over the January-May, 1979 data collection period. Teams of data collectors were assigned to institutions to conduct repeated observations of instructional time and instructional processes, and to administer questionnaires and achievement tests for a five-day period for each of the three site visits.

During the first site visit, data collectors scheduled a meeting with educational staff participants in the study. The meeting was scheduled prior to the conduct of any class observations so that participants were provided an

opportunity to become acquainted with the objectives of the study, to request staff cooperation in providing instructional exposure data over the twelve-week instructional period, and to discuss the manner in which students should be briefed on the purposes of the study.

Since it was necessary for observers to learn rapidly the identity of the five students they were to observe, teachers or their aides were asked to provide some means to identify students as they were assembling for class (e.g., blue shirt, etc.). Data collectors stressed the fact that classroom observers were not to become involved in the instructional process, nor were they to interact with students or teachers once the rationale for their presence was explained.

Students and classes selected for observation were observed on two days for each of the three visits. Thus, for validation purposes, measures of task-related activity were obtained on six occasions. Using observational protocols, data collectors recorded student behavior in five-minute increments for the duration of the class period. The length of class time observed varied from 20 minutes to 2.5 hours, depending on the scheduled class time at each site.

At least for the first site visits, all students in selected classes were observed. Thus, the aggregated data provided descriptions of basic-skills classes as well as individual student measures. On subsequent site visits, all students in a class were not always observed, since pretests had not been obtained on newly admitted students.

Student Test Administration

Experience gained during the Phase II impact evaluation revealed the difficulty of obtaining valid pre- and post-tests on students confined in correctional institutions. Confined students, although they may be motivated to learn, are often not well motivated to perform on achievement tests, and quickly tire of testing situations.

During the final two days of the first and third visits and immediately following class observations, achievement tests, functional literacy tests, and attitudinal measures utilized in the previous impact evaluation were administered to all students observed. To minimize problems with student absences from testing sessions, arrangements were made with institutional staff to avoid the assignment of students to other institutional programs or work activities on the day of the testing. Data collectors obtained permission to test students locked up or restricted from class and confined to residential cottages (at times a rather large percentage of students to be tested).

Prior to the testing days, lists of student names and testing schedules were distributed among institutional staff members so their assistance could be solicited in monitoring attendance. This staff assistance was helpful since groups as large as 60 students at times were assembled in a location selected for testing.* In other instances students were residentially grouped according to certain behavioral problems and could be tested only in those special groupings. In these situations many more testing sessions were required than had been originally anticipated.

Tests were administered in a timely sequence that provided a mid-point break to alleviate problems of student restlessness. Students were offered some previously determined motivational reward (e.g., smoking breaks, soft drinks, etc.) as an incentive for completing the test. Sessions required two-and-one-half hours on the average.

The need for observers to identify students in order to observe and record their behavior, and the frequency of contact between observers and students over repeated visits was generally reported to result in the establishment of favorable rapport. Students and observers reportedly would interact between class periods or during meals on days of classroom observations. Thus, data collectors reported their ability to successfully coax reluctant participants to attend test sessions and to complete subtests.

*In some situations, serious space problems made it difficult to find a suitable testing location.

The combination of these factors produced post-test results on 61 percent of the students pre-tested. Failure to obtain post-tests on students was mainly due to their release from the institutions rather than to their refusal. Students who were present in the institution but unavailable for testing accounted for only 3 percent of the cases where post-tests were not obtained.

At only one of the nine sites were fewer than 50 percent of the students post-tested. At this site the initial sample contained a large number of students for whom it was subsequently decided by the court that their commitment to the institution would terminate at the end of a sixty-day diagnostic period.

APPENDIX D. METHODOLOGICAL SUMMARY OF THE POST-RELEASE STUDY

The follow-up study of post-release experiences of students released from Title I N or D institutions that is the subject of this report was designed (1) to describe the adjustment of students upon release and their degree of success in re-entering society; (2) to describe the types of pre-release, liaison, and post-release services that are provided to institutionalized students by correctional institutions and community agencies; and (3) to investigate the relationship of institutional activities, receipt of pre- and post-release services, and student attitudes and charactertistics to post-release experiences in terms of school entry, employment, living arrangements, the commission of subsequent offenses and reinstitutionalization.

In the following sections, procedures for site and respondent selection, the data collection methodology, and interviewer training for the Study of Post-Release Experiences of Students from State Correctional Institutions* are discussed.

SITE AND RESPONDENT SELECTION

Site Selection

Since only eight sites were to be included in the follow-up study, the population of institutions from which sites were selected was constrained to represent a narrower range of institutions than was represented in the Phase II impact evaluation. Institutions for the neglected, for example, tended to demonstrate release patterns and transitional services that are applicable only to that relatively small population of neglected or dependent youth. These institutions demonstrate a pattern of post-release service delivery that includes public welfare agencies and social service agencies rather than parole agencies, which are generally thought to serve as the focal point for post-release service delivery for delinquents. Consequently, it was expected

*The complete report of this study is contained in Volume 3 of this series of reports.

that releasees from facilities for the neglected would describe the receipt of services on a frequency basis that could not be generalized to the delinquent institutional population.

In addition, institutions for the neglected are generally located within communities and utilize public schools for at least some of the educational needs of their students. The rate of success in school re-entry upon release from neglected institutions is thus likely to be higher for neglected than for delinquent students, since the rate may merely reflect a continuation in the schools that students had been attending while they were institutionalized. Since only two neglected sites were represented in the thirty random sites, these institutions were purged from the sampling frame.

Cost-benefit considerations for a site-level sample size of eight argued for the further elimination of sites that released few students over the four-month period that defined the study's respondent selection interval. Sites releasing fewer than ten residents over the four-month period were thus eliminated from the sampling frame.

One issue to be addressed by this study was the relative merit of institutional (or other service agency) attempts to facilitate successful re-entry into the community. In order to enhance the likelihood that some sites providing pre-release services and some sites not providing many pre-release services would be selected, the 30 randomly selected sites participating in the Phase II impact evaluation were stratified on the basis of the presence or absence of pre-release or liaison activities. Information on which to base a site stratification was obtained from the data collection team, which, as part of the on-site activities during their third visit of Phase II, prepared a short description of the procedures that the institution utilized in preparing students to return to the community. This information was obtained from informal discussions with administrators, educational and treatment staff, and residents.

It was additionally anticipated that post-release success in terms of school entry might well be related to the level of student achievement, since stu-

dents having the ability to perform well on basic skills might demonstrate, for example, more success in entering schools. The average level of student achievement was utilized as the second variable for site stratification. This stratification procedure yielded eight sites providing pre-release services and housing higher performers, five sites providing pre-release services and housing lower performers, three sites providing few pre-release services and housing higher performers, and three sites providing few pre-release services and housing lower performers. The eight sites were then randomly selected from this stratified list of nineteen sites.

Student Respondent Selection

Another objective of the follow-up study was to estimate the rate at which residents released from correctional institutions experience success or failure on a number of outcome measures. To expand the sample size beyond the approximately 200 students who were expected to be released from the eight sites, and in order to obtain a broader representation of students released from correctional institutions than were represented by the eight sites alone, parole agents (or aftercare workers as they are sometimes called) from sites where students were not to be interviewed were asked to provide information on the school enrollment, employment success, living arrangements, offense recidivism and the reinstitutionalization of students released under their supervision.

All students tested at the sites during the Phase II impact evaluation and released from the institution over the four-month period of December, 1977 to March 31, 1978 were identified as the sample of students to either participate in follow-up interviewing or to be reported on by their parole officers. These site and student selection procedures yielded exhaustive samples of releasees numbering 200 students from the eight institutions to be interviewed, and 584 students from 28 sites about whom data were to be collected from parole officers.

The definition of a four-month period during which respondents were released and thus eligible for selection was based on two considerations: (1) releasees selected will have been present for at least one test administration during the Phase II evaluation so that some background information on releasees was available; and (2) students will have had to be released at least three months prior to the conduct of the interview in order to allow for a reasonable amount of time for adjustment to the community. Because of the continual entrance and exit of offenders from correctional institutions--which defined a different release date for each respondent--it was impossible to standardize the length of time since release for interview scheduling. Respondents interviewed will have thus been released from institutions a minimum of three months and a maximum of seven months at the time of the interview.

Selection of Service Agencies

Directors of community service-oriented agencies providing services for the students being interviewed were respondents for the survey of community helping agencies. Because no known population of service agencies exists (and indeed, one of the objectives of this study was to identify those agencies providing services to releasees from correctional institutions), it is impossible to clearly define the population of agencies that our respondents represent. The problem of defining the population of agencies, and of obtaining a description of post-release service delivery systems, is that the task is as difficult and as imprecise as the agencies themselves are fragmented and lack a systematic pattern of service delivery. In recent observations on the existence of community helping agencies that provide services to delinquents it has been found that criminal justice personnel do not know what services are available even within their own community (Silberman, 1978; Cressey and McDermott, 1974; Vintner et al., 1975).

Since the objective of interviews with agency administrators was to describe the post-release services received by the follow-up sample of students, the identification of helping agency respondents was obtained from several sources. Representatives of parole agencies were interviewed in each city to which a

student had been released, since parole agencies have previously been documented as the major providers of post-release services (Glaser, 1969). These representatives provided a list of agencies in each community that parolees commonly used in obtaining services. All releasees interviewed were additionally asked to provide the names of service agencies with whom they had interacted since their release from the institution.

Another problem in the identification and selection of helping agencies was that the geographic dispersion of releasees was wide. In many cases, no more than one or two students returned to the same community. The implication of this dispersion for the identification of service agencies and the interviewing of agency respondents was that the list of all agencies in every community to which students were returned might have been endless. Consequently, where the number of communities to which residents had been released from an institution was large, communities were randomly sampled for the conduct of agency interviews.*

DATA COLLECTION PROCEDURES

Procedures used to collect data for the follow-up study differed depending on whether the institution was selected as one of the eight institutions from which the students were to be personally interviewed, or whether information about students was to be obtained from parole officers supervising the released student. From those institutions where parole officers were to provide post-release outcome measures, data collection procedures included the following:

- Students were identified who had been tested at least once during the impact evaluation but who had left the institution prior to March 31, 1978. This date allowed a minimum three-month period for community adjustment before reports of adjustment were obtained. Data collectors

*Communities were weighted for the number of returning residents. Thus, communities with a larger number of returning releasees had a higher probability of selection than did communities with fewer releasees.

completed and returned a form to SDC that contained all student identification numbers assigned to each parole officer, as well as the name and address of the parole officer who would be requested to supply information on those students. This form was to be used for follow-up on nonresponses from parole officers.

- The form to be sent to parole officers was partially completed by the data collector. This form listed the student's name, study identification number, and student birth date. It was designed so that each parole officer would receive a single form that would list all of the follow-up students assigned to his/her supervision. After they had provided needed information on the form, parole officers were instructed to remove all student names and to return only the student identification numbers with the data in stamped self-addressed envelopes.
- Forms were to be held at the institution awaiting our instructions to mail them on a date to be specified. This procedure was established to ensure that all institutions would mail the forms during the same time period over which interviews were to be conducted.

For those institutions where interviews were to be conducted with the students the following procedures were used:

- Each data collection team identified students who had been tested at least once, but who had left the institution before March 31, 1978. From institutional records the staff obtained the home addresses of students and, where extant, the area parole office or aftercare agency responsible for their supervision.
- One member of the data collection team was designated as the field coordinator for the follow-up studies. Forms completed on site, which contained names and addresses, were mailed directly from the institution to the residence of the field coordinator. Thus, at no time was SDC able to form linkages among student names, identification numbers, and data.

- The field coordinator was provided with a letter to send to each student to be interviewed, including a notice that the student would soon be contacted by phone to arrange for a personal interview. A postcard was included whereby the student could indicate a prospective change of address or phone number. Students were informed that they would be paid ten dollars for the interview; they were additionally asked to call us collect* if they had questions.
- The field coordinator was responsible for providing the names, addresses and identification numbers of students to the follow-up interviewers on site.
- Address corrections were requested from the post office for those letters that were undeliverable as initially addressed. Interviewers were updated daily on new addresses or phone numbers.

Legality and Confidentiality Issues

The anonymity of juveniles committed to correctional institutions is in many states safeguarded by laws prohibiting the release of names of juveniles who have been arrested or who have been committed to a correctional institution. The nature of these laws varies from state to state, as does the legal definition of "juveniles" (ranging from 15 to 19 years of age). Since a sizable proportion of students who were to be interviewed were juveniles, and since it was essential for follow-up purposes that these names be obtained, legal permission to allow the release of names was sought from each state in which the study was conducted. The procedure for obtaining this release varied almost on a state-by-state basis and generally required the consent of a legal representative of the state correctional agency, a representative of the state educational agency, or a Children's Court judge. Similar follow-up studies had been previously conducted in several states; thus the request was not an unusual one and few difficulties were encountered.

*Instructions were included that described how to make such a call.

In order to obtain study approval for the institutions where parole officers were to provide post-release data, a letter from the study's project director was sent to each parole officer from whom information was to be requested. This letter included a copy of the study brochure describing overall evaluation objectives and a description of the post-release study objectives. The letter stressed that student names were to be removed from the form after it had been completed and before it was returned to SDC. In some states area parole officers requested that we obtain approval from state parole agencies which were responsible for applying the release of data.

To further safeguard the confidentiality rights of respondents, SDC limited the acquisition of post-release reports on students to reports from criminal justice-related personnel (e.g., parole officers), where the identity of respondents as parolees was already known. Neither school representatives nor employers were contacted in the evaluation of post-release adjustment in order to guard against the disclosure of student institutionalization records.

Interview Content

The focus of the semi-structured student interview schedule was a description of the experiences and problems students encountered after their release in terms of school, employment, living situations, parole experiences, recidivism and reinstitutionalization. Because of releasees' participation in Phase II testing, data on respondent background characteristics, length of institutionalization, measures of self attitudes, attitudes toward school and expectations for post-release success and measures of basic-skills performance were available for the relational analyses with regard to post-release outcomes. Additionally, attitudinal measures were readministered to identify any changes in these attitudes that post-release experiences may have produced, and to assess the stability of attitudes over different situations and experiences.

INTERVIEWER TRAINING

A three-day training session was held to familiarize interviewers with the study objectives and the instruments to be used. Four of the five interviewers

employed had collected data and administered tests to students during Phase II of the evaluation. The fifth interviewer was a teacher in a correctional institution who had extensive experience in interacting with students similar to the respondents in this study.

Items on the semi-structured interview schedules were explained on an item-by-item basis. Interviewers were instructed to record as much information as possible so that the post-release adjustment of students could be adequately assessed. They were advised to establish an informal, conversational approach to data collection.

Procedures to be utilized in establishing the first crucial contact with students were explained. Because of the difficulties that might be encountered in obtaining respondent cooperation when the initial contact was to be by telephone, interviewers were provided with several approaches for conducting positive, nonthreatening conversations. As part of this training session, the initial phone calls to students to be interviewed were made.

The Conduct of Student Interviews

Past follow-up studies reporting less-than-favorable response rates often question the methodological procedures employed since they may influence respondent cooperation. For that reason, the methodology employed in this study is described in some detail.

Students were provided with an opportunity to select the location for the interview so that a comfortable, non-threatening environment would be self-selected. Instructions to the interviewers directed them to be flexible about locations and to attempt to ensure student privacy and comfort with the choice of the location, with the additional caution that interviewer safety could be monitored. These procedures, of course, could not be realistically implemented where released students had been reinstitutionalized in a state or local detention facility at the time of the interview.

The most frequent problem encountered in the initial conduct of the interviews was in ensuring privacy for the interview. Many of the interviews were scheduled to take place in the home, and when data collectors arrived they often found other family members present. In these situations interviews were conducted outside the home. Since the weather in June and July allowed for the use of the outdoors, this solution was frequently utilized.

Interviewers emphasized the importance of each respondent's contribution to the study, which reportedly fostered a feeling of self-importance among respondents. Interviewers frequently reported that respondents expressed disbelief that anyone would travel from another state to talk to them.

Conduct of Service Agency Interviews

In each area where students were released, an attempt was made to obtain information from all agencies that students indicated had provided post-release services, and from agencies most often used by parole agencies in referring parolees for services. Additionally, all directors of parole agencies in each area were interviewed. Community service agency directors were identified and contacted. Because of the easy access to the helping-agency directors by phone, and because these interviews could be scheduled during times that students were unavailable for interviews, service agency data was gathered at little additional cost to the study.

Response Rates Obtained

In previous post-release follow-up studies, newly released residents of correctional institutions have demonstrated a high rate of mobility during the first year after release. This mobility presents logistical problems for post-release studies because of the manifold difficulties in tracking releasees. Gaining the intended respondent's cooperation is a problem above and beyond the problem of locating respondents. Of the exhaustive sample of 195 students from the eight institutions to be interviewed, 175 were contacted and 174 were interviewed for a 90 percent response rate. Thus, we have been able to overcome many of the usual problems accompanying less than desirable response rates. The

high response rate enhances the likelihood that respondents interviewed are representative of the population released from institutions within a four-month time period. Aside from the one nonresponse due to respondent refusal, all incompleted interviews resulted from staff inability to locate respondents. Thus, completed interviews do not merely reflect the more easily located and cooperative releasees that are thought to be reflected in studies where lower response rates and non-cooperation are reported.

The favorable response rate was probably due to several factors. All interviewers made repeated attempts to contact the students by phone; if that method failed, they either sent a mailgram or visited the address itself to attempt to make contact or locate the student. Attempted contacts ranged from one attempt to 44 attempts, and averaged five attempts per completed interview. An adequate amount of time was allowed in each state so that interviewers could exhaust all possible tracking procedures. Additionally, the payment for interview generally appeared to be a motivation for respondents to participate once they have been located.

Response rates from parole officers and aftercare workers who reported on non-interviewed releasees were somewhat lower than interview response rates, but were still impressive considering the geographic dispersion of released inmates and the large number of parole agencies contacted. Of 590 students released from institutions over the four-month period, data were returned by parole officers on 437 students. This 74 percent response rate was accumulated on the basis of reports from 178 parole officers and aftercare workers; thus each parole agency and aftercare respondent reported on only two or three released residents.* It should be noted that although parole officers were generally reported to be knowledgeable of further law violations for parolees, a lack of knowledge was often reported for outcomes such as school entry, employment success or living situations obtained.

*No data were received on any student from four of the sites in the original Phase II samples due to (1) institutional policies of straight releasees (where inmates are not released on parole and thereby lack respondents to provide follow-up data); and (2) the reluctance of one site to participate due to ongoing litigation.

Representativeness of Follow-up Studies

The similarity of background chacteristics of students interviewed and those reported on by others is of special interest because of different group composition based on whether students were released on parole or whether students were straight releasees from the institutions. Unlike the students (who were all released on parole) reported on by parole officers, students interviewed represented the population of releasees that included both parolees and those released directly from institutions. While it may be argued that the inclusion of directly released students in the interviewed sample rendered the two samples noncomparable on a critical dimension, the decision to include these students was in response to criticism that previous follow-up studies have only assessed releasees' adjustment to parole rather than adjustment to communities (Lipton et al., 1975). The expectation is that the rate of failure on important outcome measures may be underestimated for those who do not receive parole supervision.

In order to ascertain the representativeness of both samples--those interviewed and those reported upon--average values on background characteristics and attitudinal measures for both samples were compared to average values for the entire Phase II sample that had been selected to represent the institutional student population nationwide. These comparisons assess the comparability of a sample of parolees to a sample composed of parolees and direct releasees. In a later section, parolees and direct releasees among the interviewed sample will be compared to determine whether differential patterns of post-release adjustment occur.

Students interviewed after their release from institutions, and those reported on by parole officers, were similar to the general institutional population in terms of age and grade placement. Since the sample of students interviewed and those students reported on by parole officers represent the population of students released from correctional institutions over a four-month span of time, they would be expected to differ from the general institutional population of students for whom the length of institutional confinement is likely to be shorter than the length of stay for the more general population. This expectation was borne out by the data. Both samples of releasees had anticipated

significantly shorter lengths of stay than did the general population at the time of Phase II testing, and interviewed students anticipated even shorter lengths of stay than did releasees on whom parole officers reported.

Those released from correctional institutions over a relatively short length of time (constrained as the four-month period of December, 1977 to March, 1978) differed from the general institutional population in other respects. Minority residents, for example, are under-representated, as are offenders institutionalized for the more serious person-related offenses, which is usually correlated with length of confinement. Females are not represented in the sample of students interviewed. Since institutions tend to house either all male or all female populations (and since relatively few all-female institutions were represented in the Phase II sample), it is not surprising that the random selection of only eight sites would not include an institution housing females.

Releasees tended to be better performers in terms of reading and math skills than were their counterparts in the general institutional student population. Furthermore, institutions selected for the follow-up study demonstrated significantly higher average levels of basic skills than did institutions that were not selected.

Both samples of releasees were very similar to the general population in terms of self-attitudes (e.g., self-esteem, locus of control, and sociability) and school-related attitudes (e.g., liking of school, school involvement). The conclusion drawn from these sample comparisons is that both samples appear to be similar to the institutional student population on a number of important dimensions. However, where generalizations to the population are made on the basis of post-release experiences reported by follow-up students or their parole officers, caution must be exercised in interpreting the influence of variables on which samples differed significantly (e.g., the impact of length of confinement and ethnicity on post-release experiences).

END