DOCUMENT RESUME

ED 155 258 UD 018 264

AUTHOR Roberts, Sarah Jane
TITLE ESEA Title I Evaluation Workshops.
INSTITUTION RMC Research Corp., Mountain View, Calif.
SPONS AGENCY Office of Education (DHEW), Washington, D.C. Office of Planning, Budgeting, and Evaluation.
REPORT NO RMC-R-UR-329
PUB DATE Jan 78
NOTE 65p.

EDRS PRICE MF-$0.83 HC-$3.50 Plus Postage.
DESCRIPTORS *Administrator Education; *Compensatory Education Programs; Federal Programs; Inservice Education; *Program Administration; *Program Evaluation; Project Training Methods; Reports; State Programs; State Surveys; *Workshops
IDENTIFIERS *Elementary Secondary Education Act Title I; *Evaluator Education

ABSTRACT

During fiscal year 1977, training workshops on the new ESEA Title I evaluation and reporting system were conducted for State level program administrators and evaluators. In this report, the workshop activities are documented. The purposes and development of the evaluation system and the technical assistance efforts are summarized, as an introduction to the program. In the remainder of the report, the content, materials and procedures of the workshops are focused upon. Their role in the overall teaching strategy is discussed. Also described are the revisions of training materials, including the evaluation forms and instructions themselves, as well as technical papers on subjects such as testing of Title I students and selection of students. Some of the complexities of running an intensive training program on a nationwide scale are outlined. (Author/GC)

RMC Report
UR 329

ESEA TITLE I EVALUATION WORKSHOPS

Sarah Jane Roberts

January 1978

Prepared for

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE

Office of Education/Office of Planning, Budgeting, and Evaluation

RMC Research Corporation
Mountain View, California

The research reported herein was performed pursuant to a contract with the Office of Education, U. S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

CONTENTS

List of Figures

List of Tables

ACKNOWLEDGMENTS

The RMC workshop instructional team consisted of G. Kasten Tallmadge (Project Director), Barbara M. Fagan, A. Oscar H. Roberts, Sarah J. Roberts, and Christine T. Wood. Julia B. Marquis served as coordinator for participants' travel arrangements and materials production, and Bette C. Cameron was project secretary.

All of us are indebted to Janice K. Anderson of the U.S. Office of Education's Office of Planning, Budgeting, and Evaluation, who served as Project Officer for the contract under which the workshops were conducted. Her suggestions contributed greatly to the continuing refinement of the workshops as well as the Title I evaluation and reporting system itself.

We are also grateful to Judith Burns, Richard Fairley, Ronald Fishbein, Burma Hulten, William Lobosco, Paul Miller, Jamues Ogura, William Shoemaker, and Robert Stonehill, all of USOE, who participated in the workshops providing the federal perspective and responding to policy questions.

Valuable comments were also received from the Directors of the Technical Assistance Centers who participated in the Washington, D.C., workshop; these included Everett Barnes, Gary Echternacht, Carmen Finley, George Powell, Theodore Storlie, George Temp, Joan Troy, and William Wright.

I. INTRODUCTION

During Fiscal Year 1977, RMC Research Corporation, under contract with the U.S. Office of Education, Office of Planning, Budgeting, and Evaluation, conducted training workshops on the new ESEA Title I evaluation and reporting system for State-level program administrators and evaluators. This report documents those activities that were carried out by RMC Research Corporation under contract 300-76-0316, ESEA Title I Evaluation Workshops.

Initial development of the Title I evaluation and reporting system was begun by RMC in 1974, also under contract with USOE. In the following year further developmental work proceeded, together with an effort to present the fundamental features of the system to administrative and evaluation personnel in each of the States. This effort helped to identify needed revisions in the system, as well as to provide an estimate of the nature and scope of the technical assistance that would be required for its implementation. One of the resulting recommendations was that training assistance be provided directly to State educational agencies (SEAs) through workshops.

Following a brief summary of the purposes and development of the evaluation system and of the technical assistance efforts, the remainder of this report will focus upon the content, materials, and procedures of the workshops, and upon their role in the overall training strategy. Final revisions of the system documentation will also be described.

Background

The Elementary and Secondary Education Act of 1965 contains provisions for federal aid to education. Its largest and probably best known component, Title I, authorizes financial assistance to local educational agencies (LEAs) that have concentrations of economically disadvantaged children. Title I funds are to provide special educational programs for educationally disadvantaged children in those sites.

Federal evaluation requirements. Title I funds are allocated to SEAs for distribution to their eligible LEAs. It is the responsibility of the LEAs to design, carry out, and evaluate the effectiveness of their own Title I projects. Their evaluation results must be reported annually to their SEAs. The SEAs are required to summarize local evaluation data and submit annual State reports to USOE, which in turn reports to Congress.

In the past, a number of attempts have been made to analyze and synthesize the data in these State evaluation reports, in order to form a picture of the nationwide impact of Title I. (See Wargo, Tallmadge, Michaels, Lipe, & Morris, 1972; Gamel, Tallmadge, Wood, & Binkley, 1975.) Unfortunately, the conclusion has always been that the State reports, taken as a whole, do not provide the information needed to assess nationwide program effectiveness. There is simply too little uniformity in the type, format, and quality of information reported at each level. This predicament led to a decision to try to improve the quality of local evaluations, and to standardize reporting practices. Section 151 of ESEA Title I, as added by the Education Amendments of 1974, laid out the ground rules for the new reporting approach; in particular, paragraphs (d) through (f) specified the following:

> "Sec. 151. (d) The Commissioner shall provide to State educational agencies, models for evaluations of all programs conducted under this title, for their use in carrying out their functions under section 143(a), which shall include uniform procedures and criteria to be utilized by local educational agencies, as well as by the State agency in the evaluation of such programs.
>
> "(e) The Commissioner shall provide such technical and other assistance as may be necessary to State educational agencies to enable them to assist local educational agencies in the development and application of a systematic evaluation of programs in accordance with the models developed by the Commissioner.
>
> "(f) The models developed by the Commissioner shall specify objective criteria which shall be utilized in the evaluation of all programs and shall outline techniques

(such as longitudinal studies of children involved in such programs) and methodology (such as the use of tests which yield comparable results) for producing data which are comparable on a statewide and nationwide basis."

Development of the system. Under contract to USOE, RMC Research Corporation undertook the development of prototype evaluation models as specified in Section 151. To determine the kinds of information required by the Federal government, staff members met with educational policy makers in Congress and the Department of Health, Education and Welfare (HEW). Also, to determine the characteristics of Title I evaluations up to that date, they analyzed State reports for the years 1970 to 1974.

Based on the information gathered from these sources, preliminary objectives for the new system were developed. Reactions to the initial development were solicited in phone interviews with the State Title I Coordinators. Inputs were also received from USOE representatives, and the plans were revised on the basis of the State and USOE suggestions.

A reporting system was then developed during the winter and spring of 1975 that was designed to maintain maximum autonomy of local evaluation efforts, while still providing data that would be comparable and aggregatable on a statewide and nationwide basis. The flow of information from LEAs to SEAs would continue. Project evaluators would be able to use normed or non-normed tests of their choice, and to select from three basic evaluation models for assessing cognitive achievement gains. These three were a norm-referenced model, a control group model, and a special regression model. A new metric, the Normal Curve Equivalent (NCE),[1] was also developed that allows gains on all measures to

[1] It was later learned that an identical scale, called the Standile Scale, had been developed by Frederick B. Davis, then of the Educational Records Bureau. Unfortunately, it did not receive much publicity and disappeared from the professional literature.

be expressed in (assumedly) equal-interval units derived from national percentile ranks. The three evaluation models and the NCE score are described in greater detail in Chapter II of this report.

Once the initial version of the system had been developed, reactions and suggestions were again solicited from SEA evaluators and Title I personnel, this time in nine pilot States, which were visited and presented with the system in 1975. Their comments contributed to a revised version of the system, described in Gamel et al., 1975.

In 1976, RMC conducted a project to review the methodology of the evaluation models, visit all SEAs and a sample of LEAs in each State to discuss the system, and estimate the resources that would be required to implement the system.

The site visits to the States and their LEAs provided a great deal of information about current evaluation practices, reactions to the models, and further revisions needed in the system. All 57 SEAs were visited, and the system was presented to personnel from over 400 local districts, or an average of seven districts per State. The majority of site visits produced positive reactions to the system, but some SEA and LEA personnel were basically indifferent, and some were highly negative. A number of SEAs, particularly those that believed they already had valid evaluation practices, expressed some resentment toward a federally mandated, untested system.

Many valuable comments were received from State and local personnel during the site visits, and these contributed to yet another round of revisions in the system and the forms and instructions that had been developed for reporting data. It was this revised version of the system that was presented to SEA personnel at the workshops carried out during the project described in this report.

II. BRIEF DESCRIPTION OF THE SYSTEM

A major goal in developing the Title I evaluation and reporting system was to provide meaningful, comparable information on all Title I projects across the nation. The system was also intended to be flexible, to offer State and local evaluators considerable freedom in choosing different evaluation designs and tests. By incorporating a common metric for expressing the results from the different evaluation models and different tests, the system could at the same time allow the aggregation of impact data at the school, district, State, and Federal levels.

Definition of a Title I project. As the system was developed and refined, it became increasingly clear that the project, the molecular unit around which data collection was structured, had to be explicitly and uniformly defined. The definition of "project" that was eventually adopted reads as follows:

> A project is defined as an instructional treatment with objectives, methods, materials, personnel, and activities that are uniform for all those it serves. A project may exist in one or more schools within a district, or even in several districts. Many Title I projects are clearly defined by their application for funding. In cases where a single application results in the funding of several qualitatively different instructional treatments at different sites, the term project is reserved for the individual treatments and not for the funded composite.

Whatever definition of "project" schools or districts may currently use for administrative or accounting purposes, all must adhere to this definition when collecting and reporting data in the Title I system. Then, for each project, certain basic descriptive information is to be provided.

Types of Information Collected

The information collected in the system falls into six general categories:

- Student participation--how many students are served at each grade level, by public and non-public schools, and how many students are included in different subject-matter or service areas.

- Parent involvement--participation in parent advisory councils and types of activities carried out.

- Personnel--number and types of personnel employed by Title I projects.

- Training--areas and amounts of training received by Title I project personnel.

- Cost--overall cost of each Title I project.

- Impact--achievement gains produced by each Title I project (at present measured only in reading and mathematics areas).

This information is collected and passed up through various administrative reporting levels for use in the State report. The type of intermediate reporting levels to be used is a matter of local decision, explained more fully in the section below entitled "Forms and Instructions." The forms and instructions were developed for a prototype reporting sequence using three administrative levels: the building, the LEA, and the SEA. At each level, the information from the lower levels is aggregated; however, in the process of aggregating, project information is not lost. Instead, summary information on each project is produced at each level. At the LEA level, for instance, there is a summary of each project's gain and its characteristics. The system will always permit gains to be traced back to an individual project, thereby enabling apparently exemplary projects to be studied and, if appropriate, disseminated.

The Evaluation Models

The three evaluation models use a common definition of the treatment effect in measuring the cognitive impact of Title I projects. The gain attributed to the treatment is defined as the difference between the treatment group's performance on a post-treatment test and an estimate of what performance on the same test would have been had the group not received the treatment. In other words, the project's impact is the observed post-treatment performance minus the expected no-treatment performance. This relationship can be expressed as follows:

$$\text{Treatment Effect} = \begin{bmatrix}\text{Observed}\\ \text{Post-Treatment}\\ \text{Performance}\end{bmatrix} - \begin{bmatrix}\text{Expected}\\ \text{No-Treatment}\\ \text{Performance}\end{bmatrix} \qquad (1)$$

The observed post-treatment performance is always the mean or median posttest score of the treatment group. The no-treatment expectation is derived through implementation of one of the evaluation models. The norm-referenced design generates this no-treatment expectation from norms tables. The assumption is that, without the special project, the treatment group would maintain its percentile status relative to a national or local norm group from pretest to posttest. The control group model uses the posttest (or adjusted posttest) scores of a control group as the no-treatment expectation. In the special regression model, the no-treatment expectation is derived by entering the treatment group's mean pretest score in the comparison group's post-on-pretest regression equation.

The following three sections provide overviews of each model and outline the basic assumptions underlying them.

Model A: Norm-Referenced Design. Model A can be implemented with either normed tests (Model A1) or non-normed tests (Model A2). Model A1 requires that the treatment group be both pre- and posttested with a nationally or locally normed achievement test.

When tests with national norms are used, the no-treatment expectation is found by determining the percentile status of the treatment group at pretest time. It is assumed that, without the Title I treatment, the status of the group at posttest time would be the same as it was at pretest time. Thus the group's pretest percentile becomes the expected no-treatment posttest percentile. The observed post-treatment percentile rank is that which corresponds to the group's mean posttest score. If the group's posttest status is higher than the no-treatment expectation (their percentile at pretest time), then the gain is attributed to the children's participation in the Title I project.

When tests with local but no national norms are used, it is neccessary to administer a test with national norms as well. The no-treatment expectation is derived from the local norms, following the same procedure just described. Data from the nationally normed test are needed only to convert the measured gain into NCE units. (NCE conversions are discussed in a later section of this chapter.)

The norm-referenced model may also be implemented using non-normed tests (Model A2). A nationally normed test, however, must be given either at pretest or at posttest time in addition to the non-normed test. The model involves an equating of the scores on the normed and non-normed tests so that the norms may be used to derive the observed posttest performance. In essence, the procedure is as follows (assuming that the normed test was administered at pretest time). The median pretest standard score on the normed test is determined. The pretest percentile corresponding to this score is then read out of the pretest norms table. This percentile constitutes the no-treatment expectation. The normed and non-normed tests are then equated. This enables the median posttest score on the non-normed test to be converted to its normed test counterpart. This figure, in turn, is converted to a percentile using the posttest norms table. The percentile derived in this manner is the observed post-treatment performance indicator.

In the norm-referenced model, all pre- and posttesting (normed and non-normed) must be done at times that are close to empirical normative

data points for the normed test. (For Model A2 this means that the non-normed test must be given near the dates when normative data were collected for the normed test.)

Model B: Control Group Design. Model B is a control group design that requires that either a normed or non-normed test be given to treatment and control groups at pretest and posttest times. If a non-normed test is administered pre and post, a nationally normed test must also be given sometime during the school year to the treatment group. These normed test scores are used only to convert the measure of gain into NCEs.

Between pre- and posttesting, the educational experiences of children in the control group should be similar to those of the treatment group children with the single exception that the control group does not participate in the Title I project. The control group's posttest percentile (Model B1) or mean raw score (Model B2) is the no-treatment expectation. The treatment group's percentile or mean raw score on the posttest forms the observed post-treatment performance. When the treatment group's performance is superior to the control group's, it can be assumed that the project was effective.

The pretest scores of the groups are used only to verify their pre-treatment comparability, or to quantify the initial difference between the groups. Two statistical techniques are offered to adjust the post-test score difference for whatever pretest inequality may exist. Where groups were formed by random assignment of children drawn from a single population or where two pre-existing groups are enough alike to be considered random samples from a single population, the covariance method of adjustment is used. Where small systematic differences are known or assumed to exist between treatment and control groups (in other words, when the groups are best considered samples from different populations), the principal-axis or standardized-gain-score method of adjustment is used.

The accuracy of the measure of project impact depends on the suitability of the control group. Ideally, the treatment and control groups

should be equivalent on all educationally relevant dimensions such as socioeconomic status, race, sex, etc. Small between-group differences can be adequately handled through the appropriate statistical adjustment procedure. Large differences are likely to invalidate the model.

The model permits the use of different tests or different levels of the same test for pretest and posttest, and testing times need not coincide with norming dates.

Model C: Special Regression Design. Model C is a regression-based evaluation design. Any instrument, rating, or composite measure that correlates highly with the posttest may be used as a pretest. The posttest may be either a normed (Model C1) or non-normed (Model C2) test. If a non-normed test is used, a normed test must also be given to the treatment group sometime during the school year. The model requires that the selection of the treatment group be based exclusively on the pretest measure (which may, however, be a composite of test scores, teacher ratings, etc.). All pupils scoring above the cutoff score must be assigned to the comparison group while those scoring below form the treatment group. All pupils in both the treatment and comparison groups must be pre- and posttested.

Post-on-pretest regression lines are calculated separately for the treatment and comparison groups. The treatment group's regression line represents the observed mean posttest performance corresponding to various pretest scores. The comparison group's regression line, when projected across the cutoff score, provides no-treatment estimates for the same pretest scores. (Illustrations of this are presented in Chapter VI of the Title I evaluation and reporting system User's Guide.[2])

The treatment effect is defined as the distance between the regression lines and is measured separately at two points: at the treatment group's mean pretest score and at the cutoff score. For both measures

[2]Copies available from Office of Planning, Budgeting, and Evaluation, U.S. Office of Education, 400 Maryland Avenue, S.W., Washington, D.C. 20202

group's mean pretest score and at the cutoff score. For both measures it is assumed that the project had a positive impact if the observed score is higher than the expected score. However, a substantial difference between the two measures may signal a spurious apparent gain resulting from test ceiling or floor effects.

There are few constraints concerning test selection. Model C does not require the same test for pretest and posttest or that the tests be administered at norming times. It is essential, however, that the pretest and posttest scores be highly correlated and that the instruments not be so difficult or so easy that test floors or ceilings are encountered. The model assumes a linear relationship between the pretest measure and posttest scores; where floor and ceiling effects exist, the pretest/posttest relationship will no longer be linear. Finally, in Model C, scores on the pretest measure must be used as the sole basis for selecting the Title I project participants.

Normal Curve Equivalents

Normal Curve Equivalents (NCEs) are normalized standard scores with a mean of 50 and a standard deviation of 21.06. The scale is normalized because it is assumed that the characteristic being measured (e.g., reading achievement) is normally distributed in the population. To the extent that this assumption is valid, a normalized scale will be an equal-interval scale--that is, the length of the interval between any two adjacent scores on the scale is equal to the interval between every other adjacent pair of scores.

The fact that the scale is standardized means simply that the scores have been linearly transformed to give them a desired mean and standard deviation. The values of 50 and 21.06[3] were chosen to relate NCEs to the

[3]The NCE standard deviation of 21.06 was derived by dividing the distance from the mean to the 99th percentile (99 - 50 = 49 percentile points) by the same distance measured in terms of normal curve standard deviation units (2.3267 - 0 = 2.3267). The result (49 ÷ 2.3267 = 21.060) yields a scale that includes exactly 98% of the population between values of 1 and 99.

percentile scale, thereby lending intrinsic meaning to them through the association. The exact relationship is shown in Figure 1.

The two scales match at values of 1, 50, and 99 but, while the NCE scale points are equally spaced, percentiles are widely spaced at extremes of the distribution and tightly clustered near the center. It can be seen in Figure 1, for example, that the distance between NCEs of 10 and 20 is the same as the distance between NCEs of 50 and 60. In contrast, the distance between percentiles of 10 and 20 is nearly twice that between percentiles of 50 and 60.

Because percentiles are not all the same size--that is, their increments represent different amounts of change in achievement--they should not be used in arithmetic computations. The NCE metric is assumed to be an equal-interval scale, however. Because of this characteristic, it is legitimate to add, subtract, and average NCEs.

Figure 1 also illustrates the relationship between NCEs and stanines. Whereas NCEs divide the distance under the normal curve into 99 equal units, stanines divide the same distance into 9 equal-sized units. Thus, there are 11 NCEs for every one stanine. NCEs closely resemble stanines and both scales are equal-interval; however, NCEs are a more finely grained metric.

Measuring gains in NCEs with nationally normed tests. From each model, it is possible to derive a no-treatment expectation (via normative data, control group data, or comparison group data) which can be compared against the Title I children's observed post-treatment performance to yield a measure of project impact. When nationally normed tests are used pre and post, observed and expected posttest scores are typically expressed in terms of percentile ranks with respect to a national norm group or in terms of the NCE counterparts of these percentiles.

System users are provided with a percentile-to-NCE conversion table, which was derived from the table (found in most statistics books) of areas under the normal curve. Using this conversion table, observed and

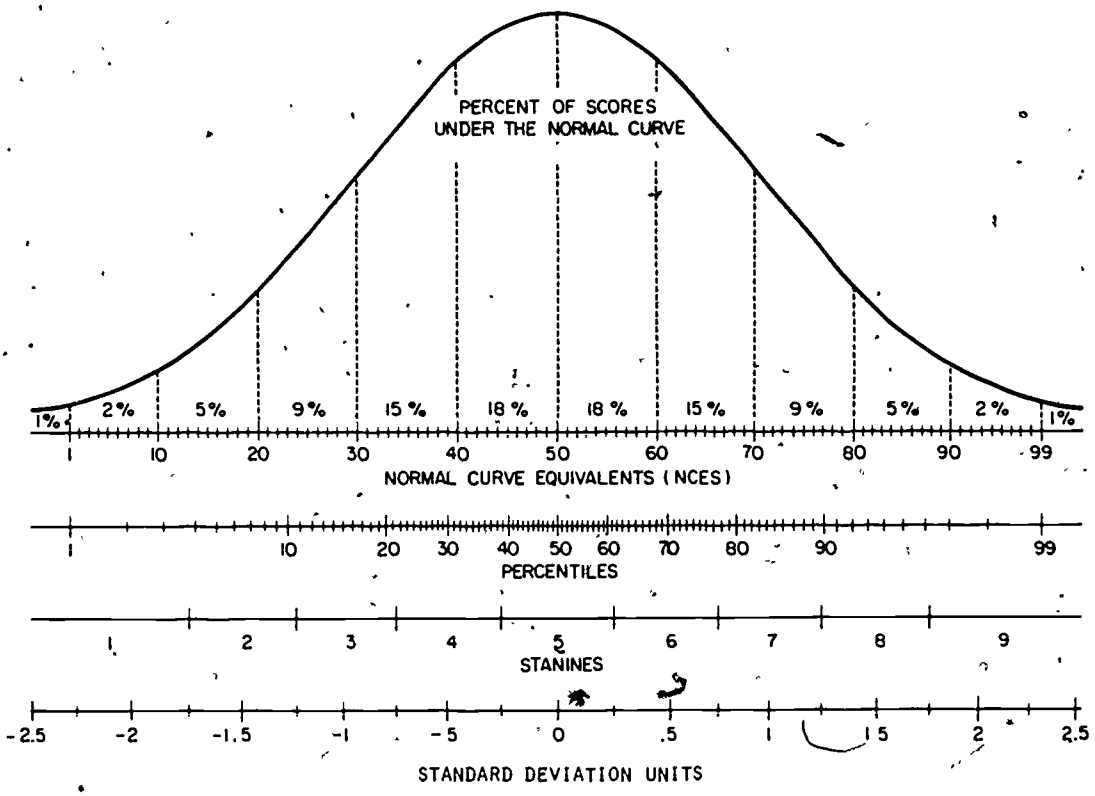

Figure 1. Area under the normal curve divided into NCEs, percentiles, staninés, and standard deviation units.

expected posttest percentiles can be converted directly to NCEs. The NCE gain is the difference between the observed NCE and the expected one.

Measuring gains in NCEs with non-normed tests. Calculating an NCE gain when non-normed tests have been administered is a slightly more complex procedure than when normed tests are used. Without nationally representative normative data, it is generally not possible to obtain national percentile equivalents for the observed and expected posttest scores (Model A2 is an exception to this general rule). In the absence of percentiles, the measures of observed and expected posttest status cannot be translated into NCEs. It is possible, however, to express the gain in NCEs. The formula is

$$T.E. = \frac{21.06\ (\bar{Y}_t - \hat{\bar{Y}}_t)\ s_n}{\sigma_n\ s_{nn}} \qquad (2)$$

where $\bar{Y}_t$ = observed mean posttest score of treatment group pupils

$\hat{\bar{Y}}_t$ = expected mean posttest score of treatment group pupils under no-treatment conditions

s_n = standard deviation of treatment group scores on the normed test

s_{nn} = standard deviation of treatment group scores on the non-normed test

σ_n = standard deviation of norm group scores on the normed test

(The complete derivation of this formula is presented in the system User's Guide.)

Measuring gains in NCEs using tests with local norms. When tests are used that have only local and no national norms, the gain is calculated using the procedure employed with non-normed tests. A nationally normed test must be given in order to derive an estimate of a national sample's standard deviation on the locally normed test. The necessary values are inserted in Equation (2) and the NCE gain can then be calculated.

When tests are given that have both local and national norms, a somewhat different procedure is used. After the standard score difference between observed and expected posttest performance is determined, it is then divided by the standard deviation of the nationally representative sample's test scores. In contrast to the situations in which non-normed tests are used, the national sample's standard deviation does not have to be estimated; instead it can be obtained from the technical manual for the test. After the standard score gain is divided by the national standard deviation, it is multiplied by 21.06 to obtain the NCE gain.

Forms and Instructions

A system of reporting forms and instructions was designed to facilitate the aggregation of information through three administrative levels--the building, the LEA, and the SEA. Data on each project could be collected at the building level, and sent up to the LEA level, where they are aggregated to form LEA reports. The data in LEA reports could then be aggregated to produce the State reports, which are then used at the Federal level to gauge the impact of Title I nationwide.

Using the forms provided, personnel at each of the three levels can collect, analyze, and pass on exactly the information required in all six reporting areas. For the five general information areas, the forms are identical regardless of which model is used; for the impact data, different color-coded forms, appropriate to the different models, are provided. To aid users in completing the forms, detailed instructions are given at each level for both the general and the impact (model-specific) information. The instructions take the user step by step through implementation of the appropriate model, using a method that requires a minimum of technical evaluation expertise. (Further description of the forms appears in the section on "Training Materials" in Chapter III.)

Although the reporting forms are provided for system users' convenience, alternative methods of collecting and aggregating data are perfectly permissible. Thus, some States may wish to aggregate and analyze data collected directly from the buildings, bypassing the LEA altogether. Or,

a district may carry out directly the collection of district-wide data, eliminating the need for reporting from the individual buildings. Some States might have "Intermediate" or consolidated districts that would play a role in Title I evaluation and reporting. The sets of forms and instructions could be adapted to their role. The building-LEA-SEA hierarchy was chosen because it was common and yet could be easily collapsed or expanded upon to meet administrative configurations in various settings. Any data-collection procedure preferred by the schools, districts, or States is acceptable within the system. The sole requirement is that at the State level, all SEAs must report the same kinds of information.

III. DESCRIPTION OF THE WORKSHOPS

Perspective on the Training Effort

In order to help SEAs get acquainted with the requirements of the system and formulate sound evaluation plans, USOE set up a comprehensive training effort that was intended to give State personnel a basic understanding of the models and the reporting forms, to prepare them to train local evaluators, and to provide them with on-going technical assistance as they began to implement new evaluation procedures.

Technical Assistance Centers. The on-going consultation services are provided by ten Technical Assistance Centers (TACs), one in each HEW region. The TAC serves all five or six States in the region, offering training sessions for local personnel and advice in connection with specific evaluation problems identified by each State. Staff from the TACs make periodic site visits, and offer technical assistance as requested.

They also have a role in the continuing examination and refinement of the system. Each TAC was given certain methodological questions to be studied, and inputs from the actual experience of districts beginning to implement the system would contribute to the analysis of these problems. TAC directors maintain periodic contact with each other and with USOE to share information, questions, and findings.

Newsletter. In order to provide a link between USOE and all of the States as they began their implementation efforts, a newsletter was started. Issued periodically by USOE, it carries information about the progress of Federal regulations for Title I evaluation and other provisions important to State and local personnel. From time to time its contents may include letters from SEA staffs with comments or questions about the system, as well as answers from USOE personnel. An occasional parents' column appears, with comments from Title I parents. LEA personnel also contributed feature columns dealing with evaluation procedures or innovations in their settings. This newsletter is distributed to SEA and LEA staffs and other interested persons in every State.

Monograph series. USOE also intended that reference books or "monographs" be available to provide background information, more detailed explanations of methodological matters, and further references about evaluation problems. They were to be user-oriented and practical in nature and were distributed through the mail-order procedures of the Government Printing Office as well as through USOE offices. Two appeared early: A Practical Guide to Measuring Project Impact on Student Achievement and A Procedural Guide for Validating Achievement Gains in Educational Projects. Others planned for later years were to deal with assessing test bias, using criterion-referenced tests, measuring the child's affective development, and estimating project costs.

Workshops. The first intensive training of State personnel took place in the workshops described in this report. During the site visits conducted in 1975-76, most of the Title I Coordinators and evaluation personnel from the nation's 57 State Departments of Education had attended briefings at which visiting RMC personnel described the evaluation and reporting system. In addition, many became acquainted with the system through presentations, by USOE personnel and others, at Federal and regional meetings. Although a number of States and LEAs shortly decided to adopt the system, it became clear whenever implementation was tried that few were adequately prepared. When it was decided to adopt the system on a nationwide basis, it was evident that a coordinated effort would be required to give SEA personnel a thorough familiarity with the system, and to enable them to train district-level staff members in their States. The purpose of the workshops was to provide this training in the most efficient way possible within a relatively short period of time.

The workshops were planned to take place over a period of about four months in the fall and early winter of 1976. One workshop was to be held in each of the 10 regional areas of DHEW. In addition to three days of intensive training, attendees would receive materials that could subsequently be used to train other personnel within their States.

Workshop Participants

It was originally planned that each workshop should be attended by two or three persons involved in Title I evaluation in each State in the region. Various other personnel also attended, however, which enhanced the diversity of the interactions and increased the challenge of providing for different interests and areas of expertise within a single three-day workshop. On the average, there were about 33 people attending each workshop.

Personnel from SEAs. Staff members who represented State educational agencies at the workshops usually came from several different branches of their departments. The State Title I Coordinator almost always attended. Usually, one or more representatives of the State's evaluation department were also present. Occasionally, computer center staff or State assessment testing personnel also participated. The Bureau of Indian Affairs, which had Title I projects located in many States, sent representatives who were working in those States to the appropriate regional workshops. In all, a total of 243 SEA personnel attended workshops; the average number attending from each State was about four. A complete list of the SEA representatives who attended workshops is presented in the Appendix.

Personnel from Technical Assistance Centers. As a separate part of USOE's plan for providing training and consultation to State and local education agencies that were seeking to implement the new evaluation system, a Technical Assistance Center (TAC) had been set up in each of the ten HEW regions. Each workshop was attended by staff members of the Technical Assistance Center for that region. Although familiarity with the system was an important qualification for the staff of those organizations that received contracts to establish Technical Assistance Centers, the workshops provided additional training for them, as well as an opportunity to pose questions, make suggestions, and consult with RMC staff. Another important purpose of their attending the workshops was to establish personal contact with the representatives of the educational agencies with whom they would be working. Through these

meetings they could find out what evaluation plans were being made in the various States, what questions had arisen, and what kinds of training and consultation were needed.

Federal representatives. Every workshop was also attended by at least one staff member from USOE's Office of Planning, Budgeting, and Evaluation (OPBE) and by a representative from the Title I office. These Federal personnel provided workshop participants with information about the background and development of the Title I evaluation and reporting system, interpretation of the legislation, and plans for the forthcoming regulations. They were also available to respond to administrative and policy-related questions throughout the three-day sessions, and to clarify the Federal role in implementation of the system.

Staff

The workshops were presented by an instructional team consisting of five RMC staff members who traveled to all ten workshop sites. All five were thoroughly familiar with the system. Two had worked on its development from the beginning; these two and one other had been involved in the previous year's site visits to the 57 SEAs. During these visits they had presented the system to large numbers of State and local personnel, often conducting mini-workshops for small or large gatherings. The other two staff members had experience in conducting training workshops on the system.

Staff training. Since all the members of the instructional team were experienced in presenting the system to new users, staff training focused on developing and refining the specific curriculum and techniques that would be used for the ten workshops. Various activities contributed to the training of team members. Entirely new problem sets were devised for both versions of each of the three models. Creating these simulations of actual evaluations using each model helped the staff to visualize the exact steps that trainees would be going through and where they might have difficulties. Similarly, writing the technical papers provided ideas for effective ways of presenting various concepts.

Staff scheduling. Once plans for the workshop agenda were formulated, responsibility for the particular sessions was divided up among the staff members. An attempt was made to apportion the workload evenly and to provide variety in the schedule of presenters on each topic. There were several reasons for this variety. One was simply to offer some changes of pace for those attending the workshops. Another was to ensure that each team member was trained to cover several important segments of the curriculum and would be fully prepared to take over another's scheduled presentation if the need arose. Thus, for example, the overview of Model B on the first morning was presented by one person; the in-depth Model B session by a second, and the Model B overview on videotape by a third.

Scheduling and Logistics

It was originally planned that a total of eleven workshops would be conducted, one in each of the ten HEW regions, and a "preview" workshop to be presented to Federal personnel in Washington, D.C. This first workshop was intended to acquaint interested persons with the system, and to serve as a rehearsal for the staff conducting the ten regional workshops. The Directors of each of the ten Technical Assistance Centers were also invited to attend the Washington workshop and to comment on the content and format of the workshops. The idea was that appropriate revisions could be made, based on reactions to the rehearsal, before the rest of the workshops began.

Workshop dates and locations. A tentative schedule for the ten regional workshops was originally designed to include all the workshops within the period from mid-September to mid-December. During the summer, letters were sent to all State Title I Coordinators, and follow-up telephone calls were made to discuss the tentative dates of the workshop and the most appropriate persons in the SEA to attend.

The preliminary schedule for the workshop staff included three two-week trips, three one-week trips, and one workshop to take place in the local area (San Francisco). It was designed to minimize the time and expense required for travel by the RMC instructional team. The two-week

trips allowed two workshops to be conducted back-to-back with a weekend between, thus requiring only one trip each to the most distant regions, the northeast, southeast, and midwest.

The cities that served as workshop sites were chosen on the basis of convenience. They were large centers within their respective HEW regions, easily accessible by major air carriers, and able to offer appropriate accommodations for participants and locations for the meetings. All sessions were headquartered in hotels that could provide sleeping rooms, eating accommodations, and conference facilities for the participants under a single roof.

Through negotiations with the various States and with Federal officials, the tentative schedule for the workshops was altered to avoid conflicts with other commitments. Owing to a meeting of Title I coordinators in Washington, workshops originally planned in Denver and Seattle had to be cancelled, and a single workshop in Portland was substituted. The number of workshops was thus reduced from ten to nine, and the designation of which States were to attend each one was revised. In revising the schedule, RMC staff tried to allow all States to attend at the most logical and convenient location. By the time the schedule was finalized, the dates had been rearranged so that the workshops actually took place from mid-October to mid-January. Except for the week of Thanksgiving and the Christmas-New Year period, one workshop was held every week beginning in November. A copy of the final schedule, including the dates and location of each workshop and the names of the States attending, appears in Figure 2.

Arrangements for participants. Travel arrangements and reservations for all persons attending the workshops were coordinated by RMC staff, who took direct responsibility for hotel reservations. Participants were responsible for scheduling their own travel arrangements to and from workshop sites. A pre-workshop packet sent to each participant contained a map showing the conference site, reservation confirmations, an airline ticket (coach), and voucher forms for recording expenses. After the workshops, these forms were mailed back to RMC. Charges for rooms and per

TITLE I EVALUATION WORKSHOPS

OCTOBER 1976

S	M	T	W	T	F	S
					1	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	29	30
31						

NOVEMBER

S	M	T	W	T	F	S
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30				

DECEMBER

S	M	T	W	T	F	S
			1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18
19	20	21	22	23	24	25
26	27	28	29	30	31	

JANUARY 1977

S	M	T	W	T	F	S
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31					

LOCATION	STATES
Washington, DC	Federal Personnel
Portland, OR	AK, ID, MT, OR, WA, WY
Milwaukee, WI	IL, IN, MI, MN, OH, WI
Kansas City, MO	CO, IA, KS, MO, NE, SD, ND
Philadelphia, PA	NY, NJ, PA, DE, MD, DC
Boston, MA	CT, ME, MA, NH, RI, VT
Albuquerque, NM	AR, BIA, LA, NM, OK, TX, UT
Charlotte, NC	VA, WV, SC, NC, KY, TN
Atlanta, GA	MS, AL, FL, GA, PR, V. Is.
San Francisco	AZ, CA, HI, NV, Guam, Samoa, Tr. Terr.

Figure 2. Schedule of workshop dates and locations.

diem costs, personal car mileage, tolls, cabs, and necessary ground transportation were then reimbursed through the RMC office.

One other item that had to be pre-arranged with the hotels was the provision of refreshments for the mid-morning and mid-afternoon breaks. This turned out to be a fairly large expenditure, but an important one. The sessions were long and intensive, and the breaks were extremely useful in letting participants stretch their legs and regroup their thoughts. Equally important, they gave participants and workshop staff a chance to chat informally about the presentations, or about participants' individual concerns.

Provision of materials. Another responsibility of the RMC staff was to arrange for the reproduction of training materials, and for shipping them to the workshop sites. The training materials included binders containing forms and instructions, problem sets, technical papers, and the User's Guide to the system. (These materials are described in detail later in this chapter.) Prior to the workshops, the User's Guide and technical papers were written, and problem sets were devised for both versions of each of the three models. Six hundred copies of each item were printed, and 600 three-ring binders were ordered. A binder for each workshop participant was assembled by RMC staff, so that everyone would have his or her own personal reference source containing all of the training materials.

The binders were to be distributed to participants on the first morning of each workshop, as they would be in constant use during the following three days. Since the traveling workshop staff could not carry all of these materials along with them, arrangements had to be made ahead of time for shipping the binders to each site. At each hotel, the manager was asked to receive the boxes of materials and to hold them until the RMC staff arrived. The binders were packed six or seven to a large box and sent by commercial parcel service to arrive a day or two before each workshop. Videotapes and overhead transparencies, also packed in boxes, were sometimes taken as airline baggage by workshop staff members.

Unlike materials to be used by participants, materials essential to the presentations were always carried by workshop staff members. These

included one set of overhead transparencies, any necessary notes and references, and an overhead projector. Since it was too fragile to be shipped or sent as baggage, the projector was taken on board all flights as carry-on baggage. Having one's own overhead projector appeared to be far preferable to trying to rely on rental equipment. Screens were supposed to be provided by the hotels, but in several cases were obtained only with some difficulty once the sessions had begun.

Workshop Curriculum

The workshop curriculum was designed with several considerations in mind. First of all, it had to provide program and evaluation personnel from each State with an understanding of the three models and the NCE metric. A basic grounding in these essential aspects of the system would be necessary for anyone who would have to take an active role in planning or providing administrative support for its implementation. Secondly, the training had to be tailored in order to accommodate the different interests and priorities of those attending; some would be interested in the specific technical and evaluation issues, while others would be concerned almost exclusively with administrative and policy-related matters pertaining to implementation of the system. More than just a basic understanding of the requirements of the models would be essential for the former group, but not necessarily justified for the latter. Finally, both groups had to be equipped not only with information that they themselves could use, but with the means to train others within their States as well.

Planning the curriculum. These goals meant that the workshops had to cover an immense amount of material in some depth over a three-day period. The diverse needs and interests also meant that not all participants would need to learn the same material at the same level of sophistication. Therefore it was decided early that the agenda should make considerable use of "streaming," with parallel sessions for the different groups, based upon their needs.

The preliminary curriculum plan called for a whole-group session on the first morning. This was to consist of some introductory remarks to acquaint participants with the legislative background and the history of the system's development, and to give them an overview of the three evaluation models and the new NCE metric. The afternoon session was to be divided into two groups. For evaluators there would be a session on the forms and instructions, both those for reporting general information (participation, parent involvement, project cost, personnel, and training) and those for recording impact data. As a result of this session, evaluators should understand the forms and be able to train others to fill them out correctly. At the same time, administrators would attend a session focusing on the overall flow of information within the system, the responsibilities of people at each level, and the interpretation of data to various audiences.

The second day was again scheduled to begin with a whole-group session, designed to acquaint all participants with Model A1, which was likely to be by far the most widely used model. Everyone would also be involved in small-group tutorial sessions on the Model A1 forms and instructions, including a problem set that would simulate the actual process of completing a Model A1 evaluation. In the afternoon, the entire group would listen to an in-depth presentation of Models C1 and C2, but then only evaluators would work through Model C problem sets, while administrators gathered in small-group tutorials organized around specific questions that might have arisen, such as how to verify proper implementation of the models, what to do about negative findings, or what to expect from audits.

The third day was originally planned to consist almost entirely of parallel sessions. This time three different groups would be run in the morning, one for evaluators interested in Model A2, one for evaluators interested in Models B1 and B2, and one for administrators on policy issues that have been raised frequently, such as test bias, use of grade-equivalent scores, and how to evaluate NCE gains. The afternoon was to begin with a whole-group session to wrap up miscellaneous technical issues

of interest to both evaluators and administrators, such as out-of-level testing, adequacy of published norms, and possible use of local norms. During the second half of the afternoon, parallel sessions would provide (a) individual consulting to those evaluators and administrators who raised special questions, and (b) a general review of the hazards associated with each model for those who had no special problems to discuss.

Results of Washington preview. The trial-run workshop held for Federal personnel in Washington served its intended purpose very well and pointed up a need for several revisions in the curriculum. One thing that became clear was that the problem sets were too long, the Model A1 problem set in particular. This problem set had been designed to include all possible variations that trainees might encounter in using Model A1. It contained simulated data from several different projects at several different grade levels, with some projects using national norms and one requiring the development and use of local norms. It simply could not be finished in the time allotted. After the Washington workshop, a shortened version was devised that included fewer projects and fewer grade levels. Also, forms were to be completed only for the Building and LEA levels; aggregation of the data at the State level was deleted.

The trial run also showed that breaking the large group into small tutorials for the problem set only served to guarantee that the groups would become increasingly farther apart from each other in completing each section of the simulation, and that eventually one group would be finished and idle while others were hopelessly behind. For this reason it was decided to walk the entire group through the simulation step by step together, rather than splitting them up into small groups. Similarly, in the administrator sessions, it became clear that there were few policy issues raised that were not of interest to all of the administrators; splitting up into small-group sessions simply produced redundancy.

On the other hand, the trial-run workshop confirmed the value of dividing the total group into two parallel sessions, one for evaluators and one for administrators. Some of those attending, such as the Technical

Assistance Center personnel and those directly involved in Title I evaluations, were eager to discuss technical aspects of the models and to explore the implications of the system for their own particular evaluation plans. Those who were not technical specialists, however, were not equipped to cope with these discussions and had little interest in them. These were generally administrators, whose first concern was with the practical and policy-related aspects of the system, problems such as how the models would affect their regular testing schedules, or the selection of Title I participants, or the manner in which results would be reported to parents. The parallel sessions met the needs of both groups while preventing the members of either one from becoming bored or frustrated.

On the whole, the Washington workshop reinforced the conviction that there was probably too much material to be absorbed, even if it could be adequately presented, in only three days. As a result, emphasis was placed on reinforcing the basic concepts of the three models and giving participants a "feel" for the technical and practical requirements of the system, rather than on mastery of the problem sets and in-depth understanding of specific methodological and policy-related issues.

The final agenda. Figure 3 presents a typical agenda from a pre-workshop packet for one of the later workshops. As the first few workshops were completed, the impressions formed after the Washington trial run were strengthened. By the time of the last few workshops, staff members had incorporated some further changes not reflected in the official agenda. These changes were based on participants' responses. For example, in order to get the essential material covered during the three-day period, staff members found it more effective to avoid additional splitting of the parallel evaluator and administrator sessions into small groups. Further consolidation of the third-day parallel sessions took place, with all evaluators being given presentations on both Model A2 and Models B1 and B2, instead of having to choose between attending one or the other. They could then choose to do a problem set for whichever of these models was of greater interest and practicality for them. Several staff members were available to provide individual help while participants worked on the problem sets.

ESEA TITLE I EVALUATION WORKSHOP

AGENDA

Tuesday, January 25

9:00 a.m. - 12:15 p.m. — General Session

- Development of the Evaluation and Reporting System.

  USOE Representative

- Introduction to the System and the Evaluation Models.

  RMC Staff: Barbara Fagan, Oscar Roberts, Sarah Roberts, Kast Tallmadge, Christine Wood

12:15 p.m. - 1:30 p.m. — Lunch

1:30 p.m. - 4:30 p.m. (Parallel Sessions)

Session A for Evaluators

- 1:30-3:00 - Orientation to Training Materials and General Forms and Instructions.

  RMC Staff: Sarah Roberts

- 3:00-3:15 - Break

- 3:15-4:30 - Impact Forms and Instructions.

  RMC Staff: Christine Wood

Session B for Administrators

- 1:30-3:00 - Overview and Rationale of the Reporting System. Building, LEA, and SEA responsibilities. Definition of a project.

  RMC Staff: Barbara Fagan

- 3:00-3:15 - Break

- 3:15-4:30 - Preparing the State Report. Interpreting the data to legislators, school administrators, and parent groups.

  RMC Staff: Barbara Fagan

Figure 3. Typical workshop agenda.

Agenda - continued

Wednesday, January 26

9:00 a.m. - 12:00 p.m.	General Session 9:00-10:15 - An in-depth look at Model A1. RMC Staff: Kast Tallmadge 10:15-10:30 - Break 10:30-12:00 - Simulation of a Model A1 evaluation using both national and local norms. RMC Staff: All
12:00 p.m. - 1:15 p.m.	Lunch
1:15 p.m. - 2:45 p.m.	General Session An in-depth look at Model C. RMC Staff: Christine Wood
2:45 p.m. - 3:00 p.m.	Break
3:00 p.m. - 4:30 p.m. (Parallel Sessions)	Session A for Evaluators Simulation of a Model C evaluation. RMC Staff: All Session B for Administrators Small group tutorials on administrative issues. Audits, verification of proper model implementation, adequacy of criterion-referenced tests. How to interpret and what to do with negative findings. RMC Staff: All

Agenda - concluded

Thursday, January 27

9:00 a.m. - 12:00 p.m. (Parallel Sessions)

Session A for Evaluators

An in-depth look at Model A2 and simulation of a Model A2 evaluation.

RMC Staff: Sarah Roberts

Session B for Evaluators

An in-depth look at Model B and simulation of a Model B evaluation.

RMC Staff: Oscar Roberts

Session C for Administrators

Policy Issues: Test bias, grade-equivalent scores, evaluating NCE gains. Small group discussions on issues raised by administrators.

RMC Staff: Barbara Fagan, Kast Tallmadge

12:00 p.m. - 1:15 p.m. Lunch

1:15 p.m. - 4:30 p.m. General Session

Miscellaneous technical issues. Review of hazards. Discussion of special problems.

RMC Staff: All

The afternoon presentation on the third day became more of a wrap-up meeting for the entire group, again with parallel sessions eliminated. First the main features of the system and the three models were reviewed. Suggestions were given for using the various materials provided to each participant as a resource for further study and training. Staff members pointed out which materials would be helpful in connection with specific problems, such as sampling to get local norms for use with Model A1. Then the session was thrown open for general questions on any aspects of the system, or on policy-related matters.

The low demand for small-group tutorials showed that participants shared certain common concerns and questions much more than they felt a need to explore issues peculiar to their own situation. In some cases, however, fairly specific technical points proved to be stumbling blocks for almost everyone, and had to be given more extensive treatment than anticipated. The regression effect error, for example, became the subject of a separate lesson within the presentation on Model A.

Training Materials

On the first morning of each workshop, all participants were given a set of materials designed to serve as a permanent reference for them, and as a resource to be used in training others in their local areas. These materials were contained in a large three-ring binder, so that they could be easily removed or re-inserted. In addition to the materials provided for each individual, every State received one set of overhead transparencies identical to those used by the RMC staff in the presentations and a set of four videotapes. The following section describes each of the training materials in detail.

User's Guide. A pocket in the back of each binder contained a copy of the User's Guide to the Title I evaluation and reporting system. This Guide contains the essential information about the system. It presents, in decision-tree format, a guide to choosing an evaluation model appropriate to local circumstances and constraints. It then discusses the requirements for implementation of each of the models in turn, and provides

references to other documents within the system that present even greater detail on specific topics, for example, the statistical adjustments that may be used for nonequivalent control groups, or the selection of local norming samples. The final chapter discusses some additional considerations, such as coordinating and scheduling activities for the evaluation, analyzing the data, and the assumptions made when NCE gains are averaged.

Forms and Instructions. Each binder also contained a complete set of forms and instructions. These were set off by divider tabs and arranged by levels to correspond to the flow of information in the evaluation and reporting system: first the building level, then the LEA level, then the SEA level. Within each level, the forms were color coded. White was used for the general information forms for reporting on participation, parent involvement, project cost, personnel, and training. Forms for reporting project impact at the building and LEA levels were color coded according to which evaluation model was used. They were pink and salmon for Models A1 and A2, yellow and gold for Models B1 and B2, and blue and green for Models C1 and C2. State-level impact forms for presenting the data aggregated across all projects and evaluation models were again white. At each level, the forms were preceded by a set of instructions explaining in step-by-step detail how each page was to be completed. Terms used in a special sense were marked with asterisks on the forms, indicating that they were defined in a separate glossary section.

The forms and instructions were designed to enable each user to collect, organize, and report all the impact data required at each level for whichever evaluation model might be selected, as well as the general reporting information. Although they can be regarded only as prototypes of some future final forms, they were provided for all participants to use as they might wish, either exactly as they were, or as a starting point for revisions and additions, or in combination with other forms, or not at all (some states would plan to handle record keeping and reporting largely by computer). In any case, State representatives were informed that only at the State level would there eventually be specific requirements for the type of data reported, and that at no level were any particular forms presently a requirement.

Problem sets. Problem sets were developed for both the normed-test and non-normed-test versions of each of the three models (i.e., Models A1 and A2, B1 and B2, C1 and C2). Each problem set was designed as a simulation of the actual process of recording and analyzing the data and filling out the forms at the various reporting levels. Thus participants working through the problem sets would play the role, first of all, of a building-level evaluator, then of an LEA-level evaluator, and finally of an evaluator at the State level. (Later, when the problem sets were shortened, some of them no longer required all of these levels to be filled in by the participants.)

The problem sets presented hypothetical data for groups of students on various tests. Using appropriate norms tables (provided in the materials), participants had to make score conversions, compute means, and calculate the no-treatment expectation and post-treatment performance in the correct way for each model. They could then figure the NCE gains for the hypothetical projects, and aggregate these gains using the LEA- and SEA-level forms.

In addition to providing a comprehensive simulation of an actual evaluation using a particular model, the problem sets were designed to provide practice with each of the important variations possible with the model. For example, the Model A1 problem set required use of publisher's norms tables from a standardized test, and also required the use of local norms. The local norms had to be developed by participants, using hypothetical data from a local norming sample. The Model B problem sets involved the use of statistical adjustments for nonequivalent control groups, a principal-axis adjustment with Model B1, and a covariance adjustment with Model B2.

While participants were doing the problem sets, immediate feedback was provided by means of answer keys projected on overhead transparencies at the completion of each step. Any questions that arose about a particular step could be answered on the spot by RMC staff, who circulated throughout the room to offer tutorial assistance.

Technical papers. In order to provide further information on some of the more technical points of the system, a series of 14 brief technical papers was prepared.[4] Each one dealt with a specific topic relevant to one or more of the models, and was intended to serve as a source to which workshop participants, and those whom they would subsequently train, could refer for detailed explanations.

Varying in length from 4 to 15 pages, the technical papers were bound as a series of small, attractive booklets, with different-colored covers so that they could be easily distinguished. To facilitate reproduction, the booklets were actually made with 8-1/2 x 11 inch sheets of paper folded in half; when opened out, they could easily be reproduced by any type of copying machine.

As a result of feedback received during the course of the workshops, four additional papers were prepared for incorporation into the system. All 18 of the papers are described briefly in Table 1, which also specifies the models to which each of them is applicable.

Videotapes. To assist SEA personnel in conducting their own training workshops, four videotapes were prepared. One of these was a 40-minute overview of the system, including brief summaries of the three models and the NCE metric. The three others, each about 20 minutes in length, dealt with the models in somewhat greater depth. The first covered Model A, the norm-referenced model, including both Model A1 for tests with national or local norms, and Model A2, for non-normed tests. A second tape dealt with control group Models B1 and B2. The principal axis and covariance adjustments for nonequivalent control groups were explained, and rules were given for converting scores from non-normed tests into NCEs. The last tape was on the special regression Models C1 and C2, and included discussions of how a regression line is derived and how the gain is measured in both the regression-discontinuity and regression-projection models.

---

[4] Copies available from the Office of Planning, Budgeting, and Evaluation, U.S. Office of Education, 400 Maryland Avenue, S.W., Washington, D.C., 20202.

TABLE 1. SUMMARY OF TECHNICAL PAPERS

Title of Paper and Brief Description	Models to which Paper is Applicable: Norm-Referenced A1	Norm-Referenced A2	Control-Group B1	Control-Group B2	Special Regression C1	Special Regression C2
Psychometric and Interpretational Problems with Grade-Equivalent Scores - Summarizes anomalies inherent in G-E scores and problems with their use to measure project gains.	X		X		X	
Interpreting NCEs - Explains rationale of Normal Curve Equivalent scores and how to use them for project evaluation and reporting gains.	X	X	X	X	X	X
The Regression Effect - Gives the theoretical basis of the regression effect error, and details how and why it requires use of separate selection and pretest measures in Model A, special care in selecting treatment and control groups in Model B.	X	X	X	X		
Selecting a Norm-Referenced Test - Presents guidelines for choosing an appropriate test to use in norm-referenced evaluations.	X		X		X	
Characteristics of Eight Commonly Used, Nationally Normed Tests - Summarizes what forms and levels are available, when testing must be done to coincide with normative data points, what score conversions are needed.	X		X		X	
Out-of-Level Testing - Discusses when out-of-level testing is appropriate, how to convert out-of-level raw scores to in-level percentile scores.	X		X		X	
Local Norms - Explains how local norming samples can be obtained and tested.	X					
Types of Test Scores - Summarizes the different types of scores and their uses.	X		X		X	
Score Conversions - Gives instructions for using publishers' norms tables to convert raw scores to standard scores, to percentiles, etc.	X		X		X	

TABLE 1 (continued)

Title of Paper and Brief Description	Models to which Paper is Applicable: Norm-Referenced A1	Norm-Referenced A2	Control-Group B1	Control-Group B2	Special Regression C1	Special Regression C2
Composite Scores - Tells how to combine several measures to get selection/pretest scores in Model C.					X	
Criterion-Referenced Tests - Explains properties of criterion-referenced and norm-referenced tests, how non-normed tests can be used with each of the models.		X		X		X
Statistical Adjustments for Nonequivalent Control Groups - Presents the rationale and instructions for use of the covariance analysis and principal-axis adjustments.			X	X		
Collecting Achievement Test Data - Outlines essential procedures for proper administration of tests.	X	X	X	X	X	X
Common Evaluation Hazards - Gives capsule summaries of twelve methodological mistakes that frequently invalidate results.	X	X	X	X	X	X
Test Floor and Ceiling Effects[5] - Provides guidelines for identifying presence of ceiling and floor effects, and choosing an appropriate test level.	X	X	X	X	X	X
Selecting Students for Title I Projects[5] - Describes effects of student selection procedures on each of the evaluation models.	X	X	X	X	X	X
Assessing the Adequacy of Normative Data[5] - Tells how to judge the adequacy of test norms in terms of representativeness and freedom from bias.	X		X		X	
Factors that Influence Test Results[5] - Discusses sources of spurious results in testing, and ways of avoiding them.	X	X	X	X	X	X

[5]Developed after the workshops to meet expressed needs.

These presentations, given by members of the workshop staff, were essentially similar to the detailed sessions on each model given at the workshops. As in the workshops, diagrams were used to illustrate the conceptual basis of the models. Special video techniques allowed implementation rules to be superimposed on the screen while staff members were explaining them.

Each State received one set of videotapes, in cassette or reel-to-reel format, whichever SEA personnel specified as appropriate for their playback equipment.

Overhead transparencies. Each State was also provided with one set of overhead transparencies identical to those used by RMC staff throughout the workshops. The transparencies included:

- Copies of forms to use in showing how they were to be filled in.
- Rules for implementing each of the evaluation models.
- Diagrams illustrating such things as how the no-treatment expectation is derived from percentile growth curves in Model A, how the covariance and principal axis adjustments affect Model B, or how the no-treatment expectation is obtained from the comparison group regression line in Model C.
- Answer keys for the problem sets.

In addition to the set of transparencies given to each State, every person attending received, in his materials binder, a set of paper copies of all the transparencies. Participants could use these copies to follow along and take notes during the workshop presentations, or they could use them to generate additional transparencies.

IV. REVISION OF MATERIALS

Once the workshops were completed, the second major task under the contract was the revision of the training materials, including the forms and instructions themselves, as well as the technical papers. During the sessions, participants had provided RMC staff with considerable feedback on the usefulness of the materials, and these comments and suggestions were most helpful during the revision process.

Comments and Suggestions from Workshop Participants

The reactions of participants to the elements of the system itself were basically similar to those encountered during the visits to SEAs in the previous year. (For a full account of these visits, the reader is referred to Bessey, Rosen, Chiang, & Tallmadge, 1976.) Many SEA representatives were pleased with the possibilities that the system appeared to offer for improving their evaluation practices. Of these, some were eager to begin working with the models as soon as they could; others were somewhat apprehensive about the extent of the changes they would have to make in their present evaluation procedures and whether they had the expertise and resources to accomplish these changes successfully. Both groups were anxious to get as much training as possible. Finally, there were a few States that viewed the system as having little to offer beyond what their present evaluation practices could already do for them. These people were interested in learning how they could adapt or modify the system to produce minimum interference with their own, and to demand the least additional effort in order to meet eventual Federal reporting requirements.

Response to presentations. Regardless of their feelings about the evaluation system, participants were generally positive in their reactions to the manner in which material was presented during the three-day workshops. They were particularly impressed with the materials that had been designed to aid them in their own training efforts. While most felt that the workshops had been extremely valuable, however, it was also clear that few were equipped to absorb all that was covered during the three days.

It appeared that a very important function of the sessions, in fact, was to make participants aware of the areas where they would need to seek further help and to work closely with their Technical Assistance Centers.

Suggestions for the system. The SEA personnel had a number of suggestions to offer on how certain aspects of the system (mainly the reporting of general, rather than impact, information) might be modified to match more closely their current practices or to make their data collection job more manageable. One comment, for example, was that using a full-time equivalent count of students served would place too great a record-keeping burden on projects and would not produce reliable data. The need for consistency in the definitions of terms used in the forms for the Title I evaluation and reporting system, and other forms SEAs and LEAs are required to submit, was also pointed out.

There were also a few suggestions for improving some of the impact reporting forms. For instance, during the problem-set sessions several participants pointed out that whenever a project summary form required worksheets to be completed before a certain number could be filled in, a brief instruction to that effect should appear on the summary form itself, not just in the instructions. In general, participants seemed to feel that the more self-explanatory the forms could be made, the better. This had been a guiding principle of forms development originally, so such suggestions could be easily incorporated.

Revision of Forms and Instructions

The revisions that were made in the forms and instructions were intended to reflect the most up-to-date information and suggestions from SEAs about what kind of reporting forms and documentation would be most workable for them. It was recognized that any "final" forms would be changed still further in the process of government clearance, etc., but the revised versions developed after the workshops would represent the completion of RMC's developmental effort. There were two main activities involved in the revision of the forms and instructions:

(1) a terminology review, which led to changes in the terms and definitions used in the system; and (2) discussions of other issues with impact on how data were to be collected, broken down, and reported. These activities are described in the following sections.

Terminology review. As a number of SEA representatives had pointed out in the workshops, the usage of some terms in the Title I forms and instructions varied from the usage in other currently required forms, which was derived from the official handbooks published by the National Center for Educational Statistics (NCES), Washington, D.C. It was decided that an attempt should be made to achieve as much consistency as possible between the forms and instructions and the NCES standards. Copies of the relevant NCES handbooks were obtained, and were searched for definitions of all terms that appeared in the glossary of the Title I evaluation and reporting system, and of all major terms in the forms. The volumes included in this review were:

Harris, Y. Y., & Seibert, I. N. The State Education Agency (VII). 1975.

Property Accounting (III) (Draft). 1974.

Putnam, J. F. Student/pupil accounting (V). 1974.

Putnam, J. F., & Chismore, W. D. Standard terminology for curriculum and instruction in local and State school systems (IV). 1970.

Roberts, C. T. Staff accounting (IV). 1974.

Roberts, C. T., & Lichtenberger, A. R. Financial accounting (II). 1973.

Seibert, I. N. Educational technology (X). 1975.

For each term from the glossary or the forms, a check was made to see (1) whether there was an official NCES definition of the term, and (2) whether that definition was inconsistent with the definition proposed in the system. In general, the recommendations for standardizing the Title I terminology to that of NCES were based upon the following guidelines:

- If a term in the forms and instructions was not an NCES term, no revisions (other than editing) were recommended.

- If an NCES term appeared in the forms and instructions and had a definition inconsistent with NCES' usage, a new term was recommended.

- If a term unique to the forms and instructions could be replaced with an already existing NCES term and the definitions were consistent, it was recommended that the NCES term be adopted.

- If a similar term appeared in both the NCES handbook and the forms and instructions, an alternate definition might be recommended to preserve the unique purpose of the term.

- If a term had become obsolete because of a proposed revision, it was recommended that the term be deleted.

Other issues. Several other issues relevant to revision of the forms and instructions had arisen during the course of the workshop project. These issues, and their resolution, are discussed in the following sections.

1. How to count student participants--Several different breakdowns of student participation were requested in the proposed reporting system. The counts of students included:

a. The number of students served by Title I at each grade and in each service area (form B-1).

b. The number of students in each project (form B-3).

c. The number of students at each grade level in each project (form B-5).

These counts were requested at the building level and at other reporting levels. There was a variety of ways to define how these counts could be computed. The choice of which definition to use depended on both

practical considerations and the purposes for which the data were to be collected. In addition to providing information about the number of students served at various levels, the participation counts would be used to calculate project cost per pupil and to judge the adequacy of the size of the evaluation sample. Thus, it was important that the method used to compute the participation counts produce data compatible with these purposes.

The advantages and disadvantages of using each of the following definitions were reviewed

- full-time equivalents
- average Title I membership
- average Title I attendance
- a continuous body count

It was decided that a body count was not suitable to use in determining average project cost per pupil or in assessing the adequacy of the data sample. Furthermore, aggregating body-count data at each reporting level would produce duplicated counts.

Full-time equivalent counts were an unattractive choice because, as workshop participants had pointed out, they required the largest amount of record keeping. Also, it was doubtful that any satisfactory definition of a full-time-equivalent Title I student could be found. In addition, counting the number of students with test data in FTE units to judge data representativeness would be conceptually unsatisfactory to many system users and would probably result in more errors.

Computing average attendance data would provide very similar participation figures to those in FTE units, but would not require as much record keeping. However, the use of either average attendance or FTE units to compute costs would produce overestimates of the amount of money spent on pupils in a project and the adequacy of a project's data sample was likely to be overestimated when judgments were based on attendance.

The most promising method for computing participation counts appeared to be the use of average Title I membership. Membership counts

would be less difficult and time consuming to calculate than attendance or FTE counts. Membership figures would be more appropriate to use in computing project cost per student and in deciding the adequacy of the size of the data sample. Unlike body-count data, average membership counts when aggregated would not become duplicated. For these reasons, it was concluded that average Title I enrollment would be the most useful type of participation data to collect in the reporting system.

2. What, if any, cost data should be collected--At the time when the workshops were conducted, the system called for project cost information on the LEA-level impact forms. Only a gross breakdown was required (Title I funds, other supplementary funds, and total), and LEAs were not required to follow any specific accounting practices to arrive at the numbers they provided. Nothing as sophisticated as a resource cost approach (which would be the only way the figures could be made comparable across different parts of the country) was even hinted at.

It seemed impossible, within the scope of the system, to collect sophisticated cost statistics. There appeared to be two reasons why the collecting of cost information should either be abandoned or postponed until more adequate cost-accounting procedures could be prescribed. The first reason was that the data provided might be so crude and error-laden as to bear little relationship to the truth. The second reason was that the data might be over-interpreted, leading to unfair cost-effectiveness comparisons among projects.

While neither of these possibilities could be denied, it was felt that there would be a substantial correlation between the cost data actually reported and those that would be reported under an ideal system. Thus the data would have some utility. Some overinterpretation of the data would be likely, regardless of all the cautions that might be offered. It was concluded, however, that the benefits to be gained from comparisons that did not go beyond the precision of the data could outweigh the negative impacts of comparisons that did exceed those limits.

3. What subject breakdown to use for impact data--Originally the system requested information on two types of cognitive skills--verbal and numerical. This breakdown was used both for reporting pupil participation by project area and for reporting impact. While the categories verbal and numerical were considered adequate for administrative purposes such as reporting pupil participation, there were several alternatives that some felt offered a more useful breakdown of information for reporting impact. It was pointed out, for example, that the terminology used for impact reporting purposes should be meaningful and appropriate in terms of curriculum and of measurement. For this reason, a change from "numerical" to the more familiar "mathematics" was recommended. In the verbal area there was some question as to what sort of breakdown should be used. "Verbal" covered several more or less clearly defined subject areas, which could be specified individually. The disadvantage of "verbal" was that it did not isolate certain subcategories that were clearly identifiable and were of interest to decision makers, the most important of these being reading.

There appeared to be no reason why reading could not be successfully broken out as a separate category. It was unambiguously defined both as a curriculum area and as a measurement area. Since so much effort within Title I focused upon the improvement of reading skills, it would be useful to know how many pupils were receiving services in this area. Furthermore, when achievement test data were reported, a separate category for reading would mean that the results aggregated there were all derived from reading tests and represented the outcomes of instruction in reading. More meaningful interpretations could be made of the information in the "Achievement Gains by Project Characteristics" matrix on form S-5, because comparisons could be clearly specified as involving reading projects only. Therefore it was decided that impact data should be categorized into "Reading," "Language Arts Other than Reading," and "Mathematics."

4. Whether to aggregate project impact data by grade level or across grade levels--Impact data were to be aggregated at the State level on two forms, S-4 and S-5. One would provide achievement data by grade level across the State; the other would give statewide achievement gains by

project characteristics. The question of how impact data should be aggregated involved mainly the gains-by-project-characteristics matrix on S-5. There were two possibilities:

Option 1--Aggregating gains by project characteristics across grade levels.

Option 2--Aggregating gains by project characteristics separately for each grade level.

Requiring both aggregations was not considered because of the burden it would place upon State personnel, who would simply have to break down the same data in two different ways.

Aggregating gains by project characteristics across grade levels was the procedure originally used in the system. On form S-5, NCE gains, broken down into six ranges, would be reported for four types of project characteristics: hours of instruction per week, total hours of instruction, instructor-to-student ratio, and cost per student. Requiring that gains be aggregated separately for each grade level would have meant that States would fill out form S-5 once for each grade level at which achievement data were reported. This would entail as many as 12 or 13 S-5 forms for each State, depending on how many grade levels included cognitively oriented Title I activities. Although this would obviously mean more work for State personnel, it would also provide more information. In deciding which procedure to require within the system, it was necessary to consider whether the more detailed breakdown would have enough usefulness to decision makers to be worth the extra effort.

The most critical requirement was that the gains reported be as reliable as possible. Small sample sizes could threaten the interpretability of gains from many Title I projects as it was; breaking results down by grade level would simply exacerbate this problem. In many cases, the gains reported would represent fewer than 30 children, making the likelihood much greater that random errors would distort the true picture of a project's grade-level impact. On the other hand, gains reported for

an entire project would represent the maximum number of children, allowing greater confidence that the figure for each project was not distorted by random errors, but was an accurate reflection of the instructional treatment's true impact. It was therefore recommended that aggregation of gains by project across grade levels be continued.

5. Whether to add a summary of total-project data--The possibility of using "project vectors" represented an altogether new addition to the system. A project vector might look something like the following:

#4328 GR 4-6 LA 70 S/G H/W 2.5 TH 75 I/4S $285 3NCE

This would translate to: "Project Number 4328 provides language arts other than reading instruction to students in grades 4 through 6 with an average enrollment of 70 students per grade. The treatment is 2.5 hours per week for a total of 75 hours per year. There is one instructor for every 4 students and per-pupil cost is $285. The project produced a gain of 3 NCEs."

Such vectors would be quite simple for a computer to produce. The task would place a considerable clerical burden on SEAs that did not generate their report data by computer. However, with minimum redesign of the LEA reporting form L-13, all the vector information could be arranged in an easily transcribable format.

Since the State reporting forms S-4 and S-5 provided no total-project data, it was recommended that project vectors be incorporated into the system. This resulted in the addition of a new reporting form at the State level, form S-6.

Technical Papers

The 14 technical papers that were produced as part of the training materials for the workshops were described in Chapter III of this report. After the workshops had been completed, these technical papers were revised. In addition, four new technical papers were produced, making a total of 18.

Revised technical papers. New, revised versions were produced for the 14 original technical papers. Revisions were based on comments and suggestions from persons within USOE, from Technical Assistance Center staff, and from those who attended the workshops. Some of the technical papers had only minor editorial changes; in other cases major revisions or additions of new material were made.

New technical papers. Four new technical papers were also added to the system. These dealt with subjects that, from observations of those attending the workshops, appeared to need more extensive treatment in the system documentation.

A new paper entitled Test Floor and Ceiling Effects deals with the problems caused by these effects and how to avoid them. Several methods are discussed for detecting the interference of the floor and ceiling when analyzing a group's test scores, and tables appended to the paper show percentile levels at which floor and ceiling effects occur on eight commonly used, nationally normed tests. The principles discussed apply to all tests, however. This paper is intended to help evaluators in choosing a test level that will be appropriate for their Title I group, and should be useful in connection with all of the models.

A second paper, Selecting Students for Title I Projects, explains how the implementation rules for each of the three models affect the method of selecting Title I participants. It discusses the practical implications of model-dictated restrictions--for example, the fact that the pretest cannot be used to select treatment group participants in Model A, but must be so used in Model C.

Assessing the Adequacy of Normative Data discusses aspects of the norming procedures that should be examined when selecting a norm-referenced test, in order to judge the adequacy of the norms. Among the topics included are freedom from bias in the norming sample, size of the sample, and some of the sources of problems faced by test producers trying to obtain unbiased norms. The considerations outlined in this paper apply in judging the adequacy of either nationally or locally developed tests.

The fourth paper, Factors that Influence Test Results, examines some of the things that contribute to a student's test score. It considers the role of actual knowledge and of test wiseness. Guessing and other test-taking strategies are discussed, as well as situational factors and administrative procedures that can exert spurious influences on scores. Recommendations are made on how to avoid these problems or how to minimize their biasing influences.

Revision of the User's Guide

The User's Guide to the system was also revised. In addition to minor changes in the various chapters, a new chapter entitled "Troubleshooting and Refinements" was produced. This chapter discusses some additional analyses that local evaluators may wish to undertake in addition to following the requirements of implementing the particular model chosen. It also discusses steps that may be taken to "salvage" evaluations that have been incorrectly implemented or flawed in some other way (e.g., test ceilings or floors are encountered).

Some new appendices were also added, and the tables for converting percentiles to NCEs and vice versa were revised to show all numbers to two significant figures. This meant that percentiles and NCEs from 1 to 9 were carried out to one decimal place, while all others were shown as whole numbers.

V. CONCLUSIONS AND RECOMMENDATIONS

Training Problems and Issues

During the course of the workshops, much was learned about the problems and the complexities of running an intensive training program on a nationwide scale. Some of this experience is summarized in the following sections.

Diversity of participants' expertise. The major problem facing the workshop staff was how to deal with a range of different levels of expertise and different interests. Part of the problem here was that the amount of knowledge each participant would bring to the workshops was not known beforehand. All that was known was that there would be a great variety. In the face of this situation, it was decided that the best approach would be to plan the opening presentations for persons with little knowledge of the models and little background in the technical aspects of evaluation. For such persons, the sessions should be understandable and helpful, while for those with more sophisticated training they should at least provide a useful summary of the system.

Once the overviews were finished, the plan was to individualize the curriculum by dividing the groups into parallel sessions for administrators and evaluators, and by further subdividing these sessions into small-group tutorials. The first division was definitely essential; interestingly enough, the second division did not appear to be. As discussed in Chapter III, these small-group tutorials were often not actually used; rather, the administrator and evaluator sessions seemed to be sufficiently individualized as they were. This was perhaps due partly to the fact that two or three RMC staff members were always present at each parallel session. Thus, for example, if a problem set was being completed in an evaluator session, several staff members could circulate among the trainees to answer individual questions, even as one staff member was taking the whole group through the problems using overhead transparencies to display the answers.

It seems likely that the reason why the small-group tutorials were not as useful as anticipated was that, despite the variety in their training and interests, the participants were practically all approaching the system as a new development. Many of the SEA personnel had been briefly introduced to the system during the site visits a year before. However, few actually had experience using the models, or even in formulating definite plans for implementation. Undoubtedly if those attending had been further along in their own use of the system, there would have been more demand for small-group sessions focusing on particular local problems and issues. As it was, the evaluators and, to an even greater extent, the administrators, seemed to share a number of common concerns, which could be most satisfactorily dealt with in the single sessions designed for each group.

Allocation of time. Another question about the training strategy involved the allocation of time. How much time was to be spent on each of the technical points was determined basically by their relative importance and anticipated usefulness to participants. Thus, for example, more time was spent on Model A1, which was likely to be the most widely used model. A second and more difficult aspect of this question involved how to allocate time between presentations by workshop staff and questions or comments from the audience. Because of the large amount of material to be covered, most of the time was originally assigned to presentations. Allowance was always made, however, for questions to be asked at the end of each presentation. Since the technical material was generally new to most participants, staff members felt that they had been successful if there were a few technical questions, but not too many. While some questions could indicate that trainees had acquired a good basic grasp of the subject, a large number of questions, particularly elementary questions, signaled that the presentation had failed to give the listeners an adequate understanding of the material.

As the workshops progressed, it became apparent that participants favored having a large general session open to questions at the end of the last day. The original plan for that afternoon was to have small-group tutorials on issues of interest to various groups of participants.

By the last day, however, the groups seemed to want to meet together and to air their various questions and comments. The tutorial sessions were therefore dropped in order to provide a whole-group question-and-answer session.

Miscellaneous training problems. One fact that became clear during the workshops was that the amount of preparation required should not be underestimated. Staff members spent many days writing technical papers; devising problem sets; designing graphs, drawings, and diagrams for the overhead transparencies; and preparing their presentations for the videotapes and for the workshops themselves. Needless to say, the presentations improved with practice at each successive session. Another aspect of the preparation effort was the arrangement of all necessary support services. In addition to the efforts of the traveling workshop staff, one full-time coordinator was required to handle the hotel reservations, airline tickets, reimbursements, and the printing, packing, and shipping of materials.

Another factor that should not be underestimated is the importance of good facilities. The best efforts to produce top-notch materials and polished presentations can be negated if workshop participants cannot see or hear them easily. In some cases, difficulties were encountered in getting screens, chalk or erasers, etc. There were some rooms where the heating or air conditioning made so much noise that a microphone had to be used--an inconvenience when one is also moving around, using a pointer, or trying to change overhead transparencies. Despite careful attempts to ensure adequate facilities in advance, meeting rooms were not always ideal. At one workshop, the agreed-upon conference room was apparently preempted by a larger group, and the hotel offered a substitute room that was too small. Fortunately, the session was able to be moved to conference facilities kindly offered by the regional Technical Assistance Center. In another case, the meeting rooms were so cold that people sat in their coats, as the heating system tried vainly to cope with unusually harsh winter weather.

The best way to avoid unpleasant surprises with the conference facilities is to inspect them in person or to choose hotels that are already

known to staff members. In three of the ten workshop locations, this was possible, and the facilities were entirely satisfactory. If this is not possible, it is a good idea to get as much detailed information as possible about the meeting rooms. For example, actual room dimensions should be requested, rather than just the hotel's estimate of how many people the room will seat.

Materials and equipment are as important as facilities. Arrangements for microphone, podium, screen, blackboard, etc., should be clearly confirmed. Among the materials used in presenting these workshops, the overhead projector was one of the most valuable. The transparencies were large, bright, and easy to see. They appeared to be extremely useful in helping participants to visualize and remember the concepts underlying the models, and to understand practical examples. Another advantage was that the illustrations of technical material could be carefully made up ahead of time, rather than hastily drawn on a chalkboard while a presentation was in progress. When impromptu illustrations were needed in the midst of a presentation or in response to a question, special pens could be used to draw or write on blank transparencies. Chalkboards were sometimes used as a supplement to the overhead transparencies, but they were found to be less effective--too small, poorly lighted, and in general more difficult to see. Another useful aid was the set of hard copies of the transparencies that was included in each person's binder.

Cost of Training

The cost of providing three days of workshop training for representatives of all SEAs is summarized in Table 2. The figures for materials provided to each State are kept separate from those for items provided to each participant. The former materials were, in fact, intended for use by State personnel in conducting their own training efforts, rather than for use as a part of the workshops.

As can be seen from Table 2, the combined cost of travel and per diem expenses was the largest part of the expense for individual participants, averaging nearly $212 per person. Materials costs averaged

TABLE 2. SUMMARY OF WORKSHOP COSTS

	Total Costs	Unit Costs
MATERIALS (600 sets)		
Binders	$ 1,480	
Index dividers	587	
Sheet protectors	259	
Buttons	116	
Forms and Instructions	3,225	
Problem sets	1,167	
Technical Papers	5,725	
Transparencies (hard copies)	936	
User's Guide	1,565	
Total	$15,060	$ 25.10/set
TRAVEL (for 167 participants reimbursed by RMC)		
Air	17,465	
Local	3,221	
Total	$20,686	$123.86/participant
PER DIEM (for 167 participants reimbursed by RMC)		
Meals/Lodging	$14,705	$ 88.05/participant
MISCELLANEOUS (for 300 workshops attendees)		
Telephone	1,016	
Postage	980	
Staff Mileage	212	
Xerox	557	
Supplies	391	
Screen rental	24	
Hotel facilities rental	500	
Refreshments (coffee break)	1,890	
Total	$ 5,570	$ 18.57/attendee
STATE COSTS (57 sets)		
Transparencies	1,343	
Videotapes	10,492	
Total	$11,835	$207.63/state
GRAND TOTAL	$67,856	

$25 per person, and miscellaneous costs another $19. Extra materials were produced, allowing each State to be given additional sets for use in their own training.

Costs not shown in Table 2 were basically the salaries and travel expenses of the five persons on the workshop instructional team, one full-time scheduling and materials coordinator, and office support staff.

APPENDIX

ESEA TITLE I EVALUATION WORKSHOPS

1976 - 1977

SEA Participant List

Alabama

Rod Hildreth
David Nettles
Edward Spears

Alaska

Ruth Harris
Steven Hole
Lawrence Schutt

Arizona

Donald Kearns
Gary Mossman

Arkansas

Dean Andrew
Robert Kerr
Clarence Morris

California

Edward Bispo
Manual Ceja
James Fulten
Claude Hansen
Jane Vinson

Colorado

Robert Cheuvront
Terry Lawson
Barry Shaffer
Charles Shaffer

Connecticut

Alice Bordonaro
Ernestine Brown
James Burke
Shirley Foster
Patrick Proctor
Edward Ricciuti

Delaware

Atwood Badman
William Corkle
Edwin Skinner
Janet Wall

District of Columbia

June Bland
Jan Dell Conley
Earl Hunter

Florida

Robert Friedman
Halley B. Lewis
Rolland Mielke
Juanita Parks
Clyde Stevens
James Temple

Georgia

R. C. Beemon
Sarah Moore
William Tidwell
Susan Underwood
Ann White

Hawaii

Pao-Ming Tchou
Rose Yamada
Harold Wilfong

Idaho

Michael Brunner
Donald Carpenter
Ruth Sydell
John Taggart

Illinois

Norman Stenzel
Paul Taubr
Connie Wise

Indiana

John Heseman
Barbara Pashos
William Strange
Fausto Vergara

Iowa

Gilbert Hewett
Oliver Himley
Ronald Huff

Kansas

Jayanne Angell
Kenneth Gentry
Donald Hardesty
Ann Harrison
William Lange

Kentucky

Jacqueline Cantrell
William Field
Donald Hart
Donald Van Fleet

Louisiana

Virginia Gerace
Elizabeth Hensley
Charles Jarreau
Richard Owen

Maine

Donald Christie
Charles Cosgrove
Donald Graham
Donald Zambri

SEA Participant List (continued)

Maryland

Eugene Adcock
Charles Burns
George Lisby
Lisa Spurrier
Joan Stevenson
Percy Williams

Massachusetts

Jack Baptista
Robert Consalvo
John Howell
Robert Reid
Richard Zusman

Michigan

David Donovan
Eugene Paslov
Daniel Schooley

Minnesota

Clyde Bezanson

Mississippi

A. C. Bilbo
Albert Comfort
Frank Drummonds
R. B. DuBoise
W. L. Herd
Samuel Parker

Missouri

Charles Blackman
Gary Brummitt
John Jones
Amos Morris
Ruth Pair
Donald Snyder

Montana

Daniel Ferriter
Dean Lindahl
Jay McCallum
Robert Ruthemeyer
Gerald Shanley

Nebraska

Elizabeth Alfred
Jack Baillie
Douglas Colberg
Ervin DeBoer
Jess Medina
Mert Smith

Nevada

Richard Gunkel
R. H. Mathers
Harold Sayler

New Hampshire

Michael Brophy
MacKnight Black
Richard Hodges

New Jersey

Thomas Corcoran
Sherwood Gordon
Eugene Mason
Joseph Moore

New Mexico

Livie Duran
Charles Epler
Gilbert Martinez
Paul Rost
Carol Ross
Kerry Wenger

New York

George Cronk
William Flannigan
Jack House
Paul Hughes

North Carolina

John Bolton
William Brown
H. T. Conner
Gerald Donnely
William Hennis
Weaver Rogers
Harold Webb

North Dakota

Warren Borchert
Alton Koppang
David Lee
Minard McCrea
Ronald Torgeson

Ohio

Arlie Cox
Carl Evans
Ken Taylor

Oklahoma

James Casey
Frank Hobbs
Edward Huey
Robert Maxwell

Oregon

Gordon Ascher
Fred Buehling
Robert Clemmer
Jerry Fuller
Mark Greene
Marshall Herron
Barbara Hunt

Pennsylvania

Frank Reardon
Thomas Schurtz
Greg Shannon

Rhode Island

Gini Bilotti
Mary Lynne Bourque
Henry D'Aloisio
Pat DeVito
Gladys Thomas
Phillip Zarlengo

South Carolina

Garlan Hicks
Charles Statler
Lane Tranthen

SEA Participant List (continued)

South Dakota

Clint Berndt
Dennis Gibbs
Robert Travis

Tennessee

Joleta Reynolds
Chris Satterfield
William Tomlinson

Texas

Frank Contreras
Richard Hardebeck
Joseph Lopez
Oren Poage
Marvin Veselka

Utah

Jay Donaldson
David Nelson

Vermont

Gerry Asselin
Marshall Knight

Virginia

L. Bruce Johnson
W. H. McCann
W. E. Newell

Washington

Marion Cupp
William Hulten
Daniel Organ
John Schlotsfeldt

West Virginia

Charles Duffy
Edward Moran
David Purdy
Robert Taylor
James Thompson
Philip Thornton

Wisconsin

Clem Baime
Frank Burkholder
Frank Evans
Gail Krc

Wyoming

George Bohl
Jerry Lewis
Kathleen Verville

American Samoa

Sili Atuatasi
Carol Golanbeski
L. Tagoilelago

Guam

Joseph Cruz
Anthony Kallingal
Lillian Lujan

Puerto Rico

Marta Barros
Blanca Cacho
Fanny Freytes
Domingo Ortiz
Aida Rodriguez
Vidal Velez Serras

Trust Territory of the Pacific

Edward Klingberg
Birch Robison
Damian Sohl

Virgin Islands

Austin Donovan
Janet Griffeth
Agatha Jarvis
Kurt Komives

Bureau of Indian Affairs

Verlin Belgarde
Jan Dreibilbis
Sandra Fox
Robert Hall
Gene Knight
William Long
Noel Malone
Paul Melchoir
Heinz Meyer
Stewart Munz
Tom Patterson
Gilbert Rogers
Neil Reece
Annabell Rosenbluth
Earl Yeahquo

NOTE: Each SEA was invited to send up to three participants whose expenses would be covered by RMC under the terms of the contract. Additional participants attended at their own expense.

REFERENCES

Bessey, B. L., Rosen, L. D., Chiang, A, & Tallmadge, G. K. Further documentation of State ESEA Title I reporting models and their technical assistance requirements, Pahse I (Part One). Mountain View, CA: RMC Research Corporation, August 1976. (RMC Report UR-308)

Gamel, N.N., Tallmadge, G. K., Wood, C. T., & Binkley, J. L. State ESEA Title I reports: Review and analysis of past reports and development of a model reporting system and format. Mountain View, CA: RMC Research Corporation, October 1975. (UR-294, out of print)

Horst, D. P., Tallmadge, G. K., and Wood, C. T. A practical guide to measuring project impact on student achievement. Washington, DC: U.S. Government Printing Office, 1975. (Stock No. 017-080-01460, $1.90)

Tallmadge, G. K., & Horst, D. P. A procedural guide for validating achievement gains in educational projects. Washington, DC: U.S. Government Printing Office, 1976. (Stock No. 017-080-01516, $2.10)

User's Guide: ESEA Title I Evaluation and Reporting System. Washington, DC: Office of Planning, Budgeting, and Evaluation, U.S. Office of Education, 1977.

Wargo, M. J., Tallmadge, G. K., Michaels, D. D., Lipe, D., & Morris, S. J. ESEA Title I: A reanalysis and synthesis of evaluation data from fiscal year 1965 through 1970. Palo Alto, CA: American Institutes for Research, March 1972. (AIR-27400-3/72-FR)