DOCUMENT RESUME

ED 090 167 95 SP 007 893

AUTHOR	Boyan, Norman J.
TITLE	Improving Teacher Preparation through Training Supervising Teachers.
INSTITUTION	California Univ., Santa Barbara.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C.
PUB DATE	Apr 74
GRANT	OEG-0-70-1902; OEG-0-72-1155(715)
NOTE	54p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, Illinois, April 1974)
EDRS PRICE	MF-$0.75 HC-$3.15 PLUS POSTAGE
DESCRIPTORS	Analytical Criticism; *Cooperating Teachers; Models; Student Teachers; Supervision; *Supervisory Training; Teacher Education; *Teacher Educator Education

ABSTRACT

The University of California at Santa Barbara (UCSB) has developed a training program for supervising teachers as a strategic device for improving the preparation of teachers. Training activities and events derive from a "model" of the supervisory process which emphasizes the acquisition and use of specific supervisory skills, including the ability to observe and analyze systematically classroom instruction in order to identify behaviors needing change or maintenance. The model also emphasizes work with student teachers in developing solutions to their instructional problems. During the first two years (1970-1972), training was conducted by UCSB staff; during the third year, and "instructional package" was developed for use at distant sites. Evaluation results indicate that the targets skills are acquired by participants through both UCSB "hands on" training and through use of the "hands off" package. (Author)

IMPROVING TEACHER PREPARATION THROUGH TRAINING SUPERVISING TEACHERS

Norman J. Boyan
University of California, Santa Barbara

Prepared for Delivery at the
Annual Meeting of the
American Educational Research Association

Chicago, Illinois
April, 1974

The research and development reported herein was supported in part by the U.S. Office of Education, Department of Health, Education, and Welfare, through Grants No. OEG-0-70-1902 and OEG-0-72-1155(715). The views expressed in the paper do not necessarily reflect the position or policy of the U.S. Office of Education, and no official endorsement by the U.S. Office of Education should be inferred.

INTRODUCTION

In mid-1970 the University of California at Santa Barbara (UCSB) initiated the development of a training program for supervising teachers who work with student teachers in the schools. Training activities focused on four skill areas judged to be critical in the performance of field supervisors; namely, observing, analyzing, prescribing, and counseling. These areas were treated as sequential and systemically related rather than separate and discrete.

During 1970-71, 15 elementary school and 14 secondary school supervising teachers participated in the first round of training.* A set of 10 tests served to provide information on whether the training achieved the desired results and to generate formative feedback for appropriate program modification.**

Test scores indicated that the training program did yield acceptable results in the skill areas of observing, prescribing, and counseling.*** Only in the area of analyzing student teachers' behavior did the training fail to deliver a satisfactory pattern of desired changes in the performance of the trainees.

---

*The U.S. Office of Education supported the first year's effort through Grant No. OEG-0-70-1902.

**The following instruments were administered on a pre and post basis: (1) Observation-Specific; (2) Observation-Judgmental Language; (3) Observation-Student Teacher Problems; (4) Observation-General; (5) Prescription; (6) Analysis-Written; (7) Counseling Theory; (8) Counseling Practice; (9) Situational (10) Meaning Measurement Inventory.

***See Norman J. Boyan, et. al., A Program for Training Supervising Teachers in the Induction of New Professionals: Final Report for 1970-71, U.S. Office of Education Grant No. OEG-0-70-1902. Santa Barbara, California: University of California at Santa Barbara, November 1971.

Analysis of test data and verbal feedback from participants led the UCSB staff to modify substantially the supervision training program in 1971-72. The four-stage conceptualization of supervision used in 1970-71 (observation, analysis, prescription, counseling) gave way to an eight-step Supervision Process Model, which guided the selection of objectives, activities, and evaluation procedures in the second round of training.* The SPM included the following steps: (1) identification of an area of concern; (2) when appropriate, identification of a criterion of desired performance; (3) selection of a measuring or observing instrument; (4) observation of the selected areas of interest; (5) analysis of the obtained data; (6) identification of needed changes and alternative procedures to accomplish these changes; (7) provision of feedback to or with the student teacher; (8) selection of instructional strategies for bringing about the desired changes, with provision for the student teacher to implement the selected strategy and for recycling the entire eight-step process. Emphasis in the relationship between supervising teacher and student teacher shifted from a supervisor-dominated base to achievement of a cooperative, collegial arrangement. Transfer of supervision skills from supervising teachers to student teachers in order to enhance the capability of the latter to assess their own teaching took its place as one of the specific objectives of the training program.

Trainees in 1971-72 included 26 "new" supervising teachers (11 from the elementary school level and 15 from the secondary school level) and 10 "continuing" supervising teachers who had participated in the 1970-71 training program. The "new" and "continuing" groups were treated separately in

---

*The U.S. Office of Education supported the second year's effort through continuation of its Grant No. OEG-0-70-1902.

the assessment of their performances on seven tests, five administered on a pre and post basis, two administered on a post basis only.*

Evaluation results revealed that the training delivered nearly all of the objectives of the program at levels judged satisfactory by the training staff.** Both "new" and "continuing" supervisors demonstrated appropriate use of the revised supervision model. Like their counterparts of the previous year, the trainees increased their ability to distinguish between observational and judgmental statements, they substantially increased the number of behavioral observations made when viewing teaching behavior, and they reduced their tendency to report judgments without supporting observations. Trainees also revelaed enlarged capacity to analyze the results of observational data by specifying summary data and identifying patterns in the data. In addition they were able to specify substantially more techniques for achieving desired behavioral changes on the part of student teachers. After training, the supervising teachers demonstrated satisfactory attainment of theoretical knowledge about the process of behavioral counseling and they performed more of the counseling behaviors established as training objectives.

The 1971-72 training program included a separate experimental study of the effects of "trained" vs. "untrained" supervising teachers on the ability of student teachers to learn and demonstrate selected components of the Supervision Process Model collected into a Self-Evaluation Model. The study

---

*The instruments administered on a pre and post basis included (1) Observation-Specific Test; (2) Observation-General Test, (3) Observation-Judgmental Language Test; (4) Analysis Test; (5) Situational Test. The instruments administered on a post basis only included (1) Counseling Theory Test and (2) Field Test.

**See Norman J. Boyan, et. al., A Program for Training Supervising Teachers in the Induction of New Professionals: Final Report for 1971-72, U.S. Office of Education Grant No. OEG-0-70-1902. Santa Barbara, California: University of California at Santa Barbara, August 1972.

yielded conclusive evidence that the performance of student teachers supervised by "trained" supervising teachers was superior on target behaviors to the performance of student teachers working with "untrained" supervising teachers.*

DEVELOPMENT OF THE INSTRUCTIONAL SUPERVISION TRAINING PROGRAM

Encouraged by the results of the first two years of effort, the U.S. Office of Education contracted with UCSB to prepare a "hands-off" instructional package for improving supervision skills in Teacher Corps projects. Team leaders were identified as the primary targets for training; however, the scope of work also recognized supervisors in teacher education programs generally and supervisors in local education agencies as potential beneficiaries.

The major task of the UCSB staff was to convert the previous year's training program into a set of materials which Teacher Corps projects could use without the participation or intervention of the developers. Accordingly, the staff launched a systematic development effort to prepare, try out, and revise a self-contained package which would include all of the directions and all of the instructional units required by a user to conduct, on his own, a successful training program for supervisors.

The development work involved (1) refining the prevailing view of supervision to which training would address itself, (2) designing specifications for the instructional package, (3) preparing the preliminary units of instruction, (4) pilot testing and revising the preliminary units of instruction, (5) field testing the instructional units, and (6) revising the units

---

*See Glenn S. Pate, "Training Student Teachers to Perform Specific Self-Evaluation Behaviors", Ph.D. dissertation, University of California, Santa Barbara, March, 1973.

on the basis of field test results.

During the late Summer of 1972, the staff created a preliminary set of directions and training materials in a form judged to be specifically oriented to supervisors in Teacher Corps projects. Preparation of the preliminary form drew on (1) detailed analysis of the results achieved in the 1971-72 training program, (2) further review of pertinent literature (3) the expertise of the UCSB staff and consultants, and (4) commentary from Teacher Corps personnel about the conditions of instructional supervision especially applicable to their situations.* The preparatory activities led to the generation of a fresh set of terminal objectives, enabling objectives, training activities and materials, and associated evaluation instruments and procedures.

The Instructional Supervision Process

A new name, Instructional Supervision Process (ISP), was coined as a replacement for the old title of Supervision Process Model to signify the introduction of several specific alterations in the conceptualization of supervision which guided the establishment of training objectives and construction of training materials for 1972-73. A summary of changes made from 1971-72 to 1972-73 in the guiding "models" appears in Table 1.**

Assumptions invoked by the staff in building a model of the Instructional Supervision Process included (1) that instructional supervision must focus on a teacher's instructional concerns, particularly with respect to the effect

---

*For details on Teacher Corps relationships see Norman J. Boyan, et. al., The Instructional Supervision Training Program: Final Report for 1972-73, U.S. Office of Education Grant No. OEG-0-72-1155(715). Santa Barbara, California: Graduate School of Education, University of California, September 1973.

**The use of the term, "model", refers here to graphic illustration of steps incorporated into the Supervision Process Model and the Instructional Supervision Process.

TABLE 1

A SUMMARY OF CHANGES MADE IN THE CONCEPTUALIATION OF INSTRUCTIONAL SUPERVISION

Supervision Process Model (1971-72)	Instructional Supervision Process (1972-73)
1. Process initiated by supervisor identification of teaching problem perceived in the student teachcr.	1. Process initiated by supervisee or supervisor identification of perceived teaching problem.
2. Supervisor concerns emphasized in all steps of the Process.	2. Supervisee concerns emphasized in all steps of the process.
3. Observational data presented to student teacher by supervisor, with supervisor's analysis of the data.	3. Data display of observation presented to supervisee by supervisor, and supervisee guided in his own analysis of the data.
4. Supervisor suggests hypotheses and change strategies to deal with instructional problem of student teacher.	4. With supervisory guidance, the supervisee formulates hypotheses and change strategies to deal with instructional problems.
5. Supervisor serves as the authority in identification and solution of problem.	5. Supervisor guides the supervisee in the development of skills relevant to the identification and analysis of instructional problems and to proposing possible solutions.
6. Method of interaction between supervisor and supervisee dependent on abilities acquired in other settings (e.g., special training in counseling or interpersonal relations in extra-supervisory situations).	6. Supervisor employs techniques of communication and interaction specifically selected for their ability to facilitate supervisee growth in confidence and efficacy in the supervising relationship.

of the teacher's behavior on pupils; (2) that a systematic and objective approach to instructional supervision will aid supervisor and supervisee, together, to identify and resolve the latter's specific instructional problems; (3) that instructional supervision operates best when it takes the form of a nonthreatening, supportive, and helping relationship between supervisor and supervisee; and, (4) that instructional supervision operates most effectively when the strengths of a systematic problem-solving approach are combined with a nonthreatening, supportive, and helping relationship between supervisee and supervisor. Through emphasis on systematic method, ISP attempts to free the supervisor and supervisee from primary dependence upon intuitive, "one-shot" efforts for the improvement of instruction. Further, ISP calls upon the supervisor and supervisee to learn to use the process together in a joint effort to resolve instructional problems.

The process itself starts when the beginning teacher identifies with his instructional supervisor an area of teaching concern. Immediately, the supervisor helps the supervisee to define the problem (Step 1). Together, the two parties agree upon specific teaching/learning behaviors that may be related to the supervisee's problem (Step 2), and establish a method by which these target behaviors may be observed in the classroom (Step 3). The supervisor observes the supervisee in a classroom setting and sytematically records observations using personally selected, constructed, or adapted instruments (Step 4). The supervisor analyzes the data from his observation (Step 5) and determines the target behaviors which, because of their potential effect upon pupils should be changed or maintained (Step 6).

Without threat, the supervisor provides to the supervisee a clear summary of the data from the classroom observation and aids him in its analysis (Step 7). The supervisee then compares the target behaviors which are revealed by his own analysis with his previous perception of those behaviors and identifies behavior changes which may resolve the problem (Step 8).

Following the post-observation conference, the supervisee attempts to introduce the desired behavior changes in his/her classroom and may be provided with training to do so. The supervisee and supervisor can then recycle the process to verify the change attempt. When the instructional problem is resolved, the supervisor takes every effort to see that the supervisee fully understands the reasons for its resolution.

Terminal Objectives of the ISTP

Successful movement through and use of the Instructional Supervision Process require a supervisor to employ a set of skills which the UCSB staff translated into the terminal objectives of the Instructional Supervision Training Program listed below:

1. (Given a pre-observation conference setting,) together with the supervisee, to identify and define behaviorally an area of concern.
2. (Given a pre-observation conference setting,) to determine whether the establishment of a base rate or the setting of criteria is appropriate.
3. (Given a pre-observation conference setting in which the establishment of criteria is appropriate,) together with the supervisee (a) to identify criteria for the successful performance of a terminal goal;

(b) when such criteria are not realistically within the present capabilities of a supervisee, to identify criteria for the successful performance of intermediate goals.

4. (Given a statement of a supervisee's area of concern,) to select, construct, or adapt an instrument for systematically recording the occurrence of teaching/learning behaviors which may be related to the area of concern.

5. (Given an instrument for systematically recording the occurrence of specific teaching/learning behaviors,) to utilize the instrument to make non-evaluative observations in a classroom setting.

6. (Given the data resulting from the systematic observation of teaching/learning behaviors in a classroom setting,) to analyze the data by performing the following operations:

   a. Construct a data display;
   b. Make summary statements;
   c. Where criteria of successful performance have been established, determine if the criteria have been met;
   d. Identify patterns of behavior;
   e. Make comparisons between observations.

7. (Given the results of an analysis of data from a systematic observation,) to identify:

   a. Supervisee behaviors that are positive in terms of their possible pupil effects and should be maintained;
   b. Supervisee behaviors that are negative in terms of their possible pupil effects and should be changed;
   c. Alternate strategies for producing changes in supervisee behavior.

8. (Given a post-observation conference setting and given the data from the observation and analysis,) to perform the following operations:

   a. Review with the supervisee the agreed upon area of concern, its behavioral definition, the decision to obtain a base rate or the criterion that was set, and the observation instrument;
   b. Guide the supervisee in an analysis of the observed data which will approximate the complete analysis obtained by the supervisor;
   c. Assist the supervisee in identifying behaviors which are positive in terms of their possible pupil effects and should be maintained;
   d. Assist the supervisee in identifying behaviors which are negative in terms of their possible pupil effects and should be changed;
   e. Formulate with the supervisee strategies for achieving the desired changes;
   f. Determine with the supervisee if additional cycles of the ISP are necessary.

9. (Given a pre- or post-observation conference setting,) establish clear communication between supervisor and supervisee by utilizing communication skills, such as paraphrasing, perception checking, asking clarifying questions, offering information, and active, attentive listening.

10. (Given a pre-or post-observation conference setting, and an understanding of the distinction between "freeing" and "binding" statements,) to demonstrate the use of freeing and binding responses in order to encourage the growth of autonomous behavior in the supervisee.

Training Materials

After specifying terminal objectives, the staff addressed itself to the preparation of training activities and materials which would deliver the desired competencies. A "hands-off" instructional package was designed, consisting of the following pieces: (1) a Training Coordinator's Guide, (2) a set of Readings, (3) an associated set of Handouts, (4) five videotapes, and (5) four audio tapes. The preliminary package incorporated seven different instructional units involving a variety of training modes.

Preparation of the Training Coordinator's Guide was crucial to insure the likelihood that an on-site trainer who had no prior exposure to or experience with ISTP would find all of the instructions and references needed to conduct training sessions. The Guide provided the local coordinator with step-by-step directions for carrying out all activities, including instructions to participants, assignments to be made, and materials to be used in each session. The training coordinator also received a supply of Handouts for distribution to participants at designated times throughout the training period. The Handouts contained answers to exercises, role playing instructions, and other information for delivery to participants on a specified schedule.

The training program itself contained three types of instructional units: information acquisition, role playing, and practicum. Information acquisition units stressed knowledge about the various stages and steps of the ISP, with instruction presented through a multi-media array of printed materials, video tapes, audio tapes, transparencies, and written exercises. Basic Information about the ISP and its constituent skills appeared in a book of

Readings which each participant received at the beginning of the training period. Videotapes, audiotapes, and transparencies demonstrated the use of various skills, presented sample instructional situations, and illustrated concepts. A written exercise typically followed the presentation of each major concept to enhance the occurrence of comprehension. Participants usually completed the exercises on a self-paced basis, with immediate feedback available by way of an answer sheet. Role playing units emphasized development of skill in application of ISP as a whole and acquisition of analysis competencies in particular, through the use of trio and simulation exercises. Trio exercises required individual participants to interact with colleagues in a role playing design as they gained understanding of central concepts and as they practiced use of specific skills. Simulation exercises called upon participants to employ previously acquired concepts and skills in situations which closely resemble supervisory conferences, with each member of a trio rotating through the roles of supervisor, supervisee, and observer. Stimuli for both forms of role playing exercises came from video tapes, audio tapes, or written situational statements and dialogues. Practicum units focused on the development of synthesis and evaluation competencies pertinent to the use of ISP, by providing participants opportunities to employ the Process and its constituent skills in actual field settings. A practicum unit required the trainee to conduct a "real" supervisory conference with a "real" supervisee and to record the proceedings on audio tape. Critique of the performance came to the participant from his fellow trio members during a follow up training period.

Training Format. A 36-hour workshop arrangement was selected as the time frame and format for conducting training. The UCSB staff chose the workshop mode to permit the scheduling of training activities under a variety of conditions, such as concentration of longer sessions in a one or two-week period or distribution of shorter sessions on a weekly basis over a

month or more, depending on the length of sessions. Sessions themselves included large group presentations for basic instruction, seminars for discussions among participants, and opportunities for extensive practice in the use of ISP and its constituent competencies through exercises in the trio mode. The trios not only enabled trainees to receive immediate feedback on their performances from fellow participants but also squared with the stress in ISTP on trainees' acquiring supervisory competencies while engaged in personal interaction with others.

Pilot Testing

During preparation of the preliminary instructional units and Training Coordinator's Guide, a group of supervising teachers in the Santa Barbara area critiqued the content and sequence of the materials. Most of the critiquers had participated in the 1971-72 UCSB training program for supervising teachers. These "reality demons" contributed enormously in making the preliminary instructional units ready for pilot testing.

Approximately 40 hours of pilot testing helped to determine the extent to which the content and procedures aided or impeded attainment of the specified objectives. The preliminary set of instructional units was administered under the directions contained in the Training Coordinator's Guide, to two groups of three participants each. The training coordinator for each group was unfamiliar with ISTP prior to conducting training sessions. Each of the coordinators was an experienced teacher and supervisor. The trainees in the pilot test were all experienced public school teachers. At the time of pilot testing, none of the participants was supervising a student or intern teacher.

The opportunity to monitor closely the administration of the instructional package by a local group of trainers to a local group of trainees led to

incorporation of numerous revisions in the training program, including the following:

(1) provision of more visual aids (e.g., transparencies) for use by trainers with trainees;

(2) provision of more detailed instructions to training coordinator and participants for practice (role playing) cycles;

(3) stipulation of sharper and more precise time estimates for completing individual training exercises; and,

(4) substantial editorial revision of written materials for improved grammar, punctuation, and syntax.

Field Testing the ISTP

Of the set of Teacher Corps project directors contacted by letter or telephone to ascertain potential interest in field testing the preliminary version of ISTP, 15 responded positively. Each received a detailed description of training and testing requirements and a sampler of the training package. After review of the requirements and sample materials, 12 directors agreed to engage in the field test. Three withdrew from participation, on the grounds of local scheduling and administrative problems, prior to actual engagement; one never completed the test battery. As a result, eight Teacher Corps sites took part in the full field test during the Spring of 1973.* The UCSB staff also arranged for the participation of a local group to insure availability of evaluative data for revising the preliminary version of ISTP in the event that logistical or diplomatic problems interfered with the collection of useful information from the eight distant sites. The training coordinator and participants at each locale received by mid-January, 1973, all of the instructional materials required for conducting the

*The sites were (1) Texas Southern University, (2) Winston-Salem State University, (3) Alabama A & M University, (4) University State College at Buffalo, (5) Idaho State University, (6) University of California at Santa Cruz, (7) Weber State College, (8) Memphis State University.

training sessions and all of the evaluation instruments required for administering the testing program.

Coordinators at each site were requested to administer the pretest prior to the initiation of training sessions and to return the test protocols to UCSB for scoring. At those sites where coordinators honored the request, pretests took place between January 15 and January 26, 1973. The proposed schedule of events called for training to proceed between the end of January and the middle of April and for posttests to be completed by the end of April. Only four of the Teacher Corps sites and the local (Santa Barbara) group adhered to the desired time table; the other four Teacher Corps sites experienced a number of problems which delayed their completion of all training and testing until the end of May. The UCSB staff scored test protocols as materials arrived from the field but pooled all test results in conducting its subsequent analyses, irrespective of the particular site in which and irrespective of the varying conditions under which training had occurred.

FIELD TEST RESULTS

An adaptation of a Solomon Four-Group research design (see Table 2) served to ascertain the extent to which the preliminary version of ISTP achieved the terminal objectives set for trainees and to provide estimates of testing effects. The expectation was that subjects who received training would score significantly higher on the exercises comprising a specially constructed Instructional Supervision Test Battery than subjects who did not receive training. The Battery consisted of six sections: a two-part Comprehension Exercise, a two-part Observation Exercise, an Analysis Exercise, and a Participant Performance Record (PPR). Parallel forms of the exercises were used for pre and posttests; the PPR was used as a posttest only, and only with the two experimental groups.

TABLE 2

Treatment and Testing Conditions
for Experimental and Control Groups

Group	Treatment and Testing Conditions		
	Pre-Tested	Received Training	Post-Tested
Experimental I	X	X	X
Experimental II		X	X
Control I	X		X
Control II			X

Subjects

Subjects included Teacher Corps team leaders (and counterparts who supervised interns) drawn from eight volunteer sites and a group of volunteer supervising teachers located in public schools close to UCSB. The manner in with subjects were recruited does not permit any claim for representativeness among all Teacher Corps personnel or all public school supervising teachers.

Moreover, the design employed for analysis of field test results assumes homogeneity of subject groups, preferably achieved by random assignment, and ideally calls for approximately equal numbers in all cells. Given the constraints which prevailed across all training sites, neither random assignment of subjects nor equal distribution of subjects according to training and testing conditions was attainable.

Homogeneity of groups. The UCSB staff conducted two separate analyses to ascertain the extent of homogeneity which did prevail across the two experimental and two control groups. The first considered some 16 personal

characteristics and previous experiences which appeared potentially salient as contributors to variability in training effects.* The second considered pretest scores on five sections of the Instructional Supervision Test Battery for subjects who did and who did not receive training.

Chi-square analyses performed on the personal characteristics revealed seven statistically significant differences in the composition of the four groups. However, when the entire experimental group was compared with the entire control group, a statistically significant difference showed itself on only two characteristics: type of credential possessed by the trainees and grade level taught by the beginning teachers supervised by the trainees. The experimental group, considered as a whole, included a larger proportion of subjects who had earned one or more credentials authorizing service in the public schools than did the control group. Also, approximately 20 percent of the experimental group was supervising, at the time of training, beginning teachers who were teaching at the secondary school level while the control group did not include any subjects who were supervising secondary novitiates.

There is little logical reason to expect lack of homogeneity on the characteristic of a beginning teacher's grade level assignment to affect materially the results of training provided his or her supervisor. In addition, evaluation results from previous years in the UCSB supervisor

---

*A Background Information Sheet was used to collect information about subjects on the following characteristics and experiences: (1) sex (2) age (3) education (4) teaching credential (5) ethnic background (6) local education agency environment (7) occupational role (8) teaching experience (9) supervision experience (10) counseling experience (11) training in interpersonal relations (12) training in systematic classroom observation (13) training in supervision (14) training in writing behavioral objectives (15) ethnic background of supervisee (16) grade level taught by supervisee. Actually the data collected for the 16 variables enabled comparisons to be made on 18 counts, because the variable entitled "occupational role" was subdivided into three roles: teacher, supervisor, administrator.

training programs had not unearthed any systematic association between subjects' abilities to demonstrate target skills after training and the grade levels at which the subjects and their student teachers worked. With respect to the difference between the proportions of credentialed and non-credentialed trainees in the experimental and control groups, some discomfort about the assumption of homogeneity does enter, on the grounds that a higher incidence of professional preparation may have provided the experimental group a potential edge in benefitting from training in supervision.

The staff also compared pretest scores on the Instructional Supervision Test Battery for the pertinent experimental and control groups (Experimental I and Control I) in order to ascertain the extent to which these subjects entered the treatment with essentially similar competencies. Statistically significant differences between the pretested experimental and control groups appeared in none of the 26 summary measures included in the Comprehension Exercise, the Observation Exercises, and the Analysis Exercise.

All told, then, direct examination of the composition of experimental and control groups and assessment of pretest scores for subsets of the two groups provide reasonable solace with respect to the assumption of homogeneity. The assumption does not, however, enjoy full support. Results reported for the field test should, therefore, be interpreted in the light of the limitation unavoidably introduced because the UCSB staff could not secure random assignment of subjects across all nine training sites.

Equal distribution of subjects in experimental and control groups. Participating sites received requests to establish control groups of approximately the same size as the groups receiving training via ISTP and to assign approximately equal numbers of trainees and controls to sit for both pretest and posttest and for the posttest only.

Unfortunately, the respective sites did not or were not able to respond faithfully to requests pertinent to the distribution of subjects by treatment and testing conditions. Two of the nine sites did not provide any control subjects; two provided information only for individuals who sat for both the pretest and posttest; and two provided only posttest protocols. Table 3 summarizes the distribution of experimental and control subjects by type of test involvement.

TABLE 3

Distribution of Experimental and Control Group Subjects by Type of Test Involvement

	Group		
	Control	Experimental	Total
Pretest and Posttest	19	61	80
Pretest Only	12	15	27
Total	31	76	107

To account for the influence of unequal cell sizes in the analyses of variance performed to assess training and testing effects, the sum of squares was estimated by using the unweighted means procedure recommended by Winer.*

Treatment

Administration of the preliminary version of ISTP by on-site training coordinators was the treatment provided to experimental subjects. The UCSB staff did not intervene directly in any training activities other than distributing full sets of ISTP materials and directions for testing. Coordinators did not have previous experience with and had not received previous instruction in the use of the materials from the developers.

The only guaranteed common element in the treatment was the substance contained in the Training Coordinator's Guide, Readings, Handouts, audio tapes, video tapes, and test instructions delivered to each training site.

*B.J. Winer, Statistical Principles in Experimental Design. New York: Graw Hill, 1962, p.222.

In other respects, the treatment varied across the nine sites according to the requirements of local conditions. Given the diverse locations and circumstances in which a revised version of the "hands off" ISTP might be used, the developers accepted wide variation in administration of the training program at the field test stage as an appropriate characteristic of the treatment.

Evaluation Instruments and Test Procedures

A specially constructed Instructional Supervision Test Battery served the purpose of generating scores for assessing the capability of the training program to meet its stated objectives. The Battery included:

(1) a two-part Comprehension Exercise to test subjects' knowledge of central concepts presented in ISTP and subjects' ability to describe classroom behavior in observable, non-evaluative terms;

(2) Observation Exercises I and II to test subjects' ability to construct an observation instrument and to use the instrument systematically in observing a classroom encounter;

(3) an Analysis Exercise to test subjects' ability to analyze the results of systematically conducted observation;

(4) a Participant Performance Record to assess the extent to which subjects appropriately utilized the approaches and skills included in ISTP when supervising a beginning teacher.

The Comprehension Exercise, the Observation Exercises, and the Analysis Exercise were administered to Experimental Group I and Control Group I as both pretests and posttests. The exercises were administered to Experimental Group II and Control Group II as posttests only. The Participant Performance Record served as a posttest only, administered at the conclusion of all training to experimental subjects only. Each training coordinator received

a full set of explicit instructions for administering the several exercises and the PPR to experimental and control subjects.

Results of the Comprehension Exercise

The first of the two parts of the exercise employed a 20-item multiple choice inventory to assess subjects' knowledge of a set of selected central concepts presented in the Readings and Handouts. The second part required subjects to classify correctly each of 35 descriptions of classroom behavior either as an observation or as a judgment, thereby providing a high inference measure of their ability to describe teaching incidents in observable, non-evaluative terms. The Comprehension Exercise yielded two scores: the number of knowledge questions answered correctly; the frequency of items correctly identified as observations or judgments.

The mean posttest scores for all four groups of subjects appear in Table 4. The experimental groups scored higher than the control groups on both measures. Training effects were statistically significant in each instance; a barely significant testing by training interaction appeared for the first measure. The data imply that training contributed materially to the ability of the experimental groups to identify correctly "knowledge" items and to distinguish between observational and judgmental statements.

A comparison of the pretest and posttest scores of Experimental Group I only (Table 5) indicates that statistically significant increases occurred in both measures assessed by the Comprehension Exercise. The extent of the changes which occurred for the subjects in Experimental Group I offers additional confirmation to the inference that the training program contributed significantly to increases in knowledge pertinent to ISP and to trainees' ability to distinguish between reports of observations and judgments.

TABLE 4

Mean Posttest Scores and Significance Test Results for the Experimental and Control Groups for the Comprehension Exercise

Variable	Exper. I	Exper. II	Control I	Control II	Testing Effect	Training Effect	Training x Testing Effect
1. No. of Correct Comprehension Items	16.93	14.87	9.26	9.83	1.36	97.96*	4.22*
2. No. of Correct Observational/Judgmental Items	31.03	29.87	22.32	21.33	0.98	63.42*	0.01

* $p \leq .05; F(1,103) = 4.22$

TABLE 5

Pre and Post Mean Scores and Significance Test Results for Experimental Group I on the Comprehension Exercise

Variable	Pretest	Posttest	F-ratio
1. No. of Correct Comprehension Items	8.71	16.93	435.79*
2. No. of Correct Observation/ Judgmental Items	21.98	31.03	134.11*

*$p \leq .05; F(1,60)=4.00$

Results of Observation Exercises I and II

In Observation Exercise I, the subjects received written background information about a teaching performance which they were soon to see. After they had completed their reading of the background information, the subjects viewed a two-minute video tape excerpt of the teaching behavior of a student teacher. They then were allowed approximately 10 minutes to plan their observations of a longer segment of the same video tape, with instructions to concentrate solely on the questioning techniques used by the student teacher. After the longer (nine-minute) video clip ran its course. subjects were permitted no more than 15 minutes to complete their observational records. All subjects viewed the same video tape.*

Observation Exercise I assessed the following training objectives: (1) developing a behavioral definition of an area of concern; (2) selecting, constructing, or adapting an instrument for recording observations; and, (3) making non-evaluative observations focused on the set of behaviors for which the observation instrument is specifically constructed. Behaviors

*The elementary grade level teaching performance presented for observation was selected from video tapes secured from the teacher preparation program, University of California at Santa Barbara.

pertinent to the first two objectives were scored for occurrence. Behavior pertinent to the third objective was scored for frequency of performance in two constituent categories: the total number of behavioral incidents (TBI) reported; the total number of judgmental statements (TJS) reported. The TBI score was the sum of (1) the number of discrete behavioral incidents recorded as specific non-judgmental statements (e.g., "The student teacher called on the boy in the red shirt.") and (2) the number of quantified behavioral statements recorded (e.g., "The student teacher called on the boy in the red shirt five times."). The TJS score was the sum of (1) the number of discrete judgmental statements included in recording classroom events (e.g., "The student teacher was warm.") and (2) the number of quantified judgmental statements included in the recording (e.g., "The student teacher was never warm.").

The purpose of Observation Exercise II was to test the ability of each subject to construct, within 20 minutes, an observation scheme which he could use if he were to observe again the same video tape viewed in Observation Exercise I. (The subjects did not actually employ the constructed scheme.) If the subject felt he had satisfactorily completed the assigned task during Observation Exercise I, he was asked to say so on his answer sheet and to submit the statement. The salient difference between the two exercises was that II addressed itself to the question of whether or not a subject could construct an appropriate observation instrument when specifically requested to do so whereas I addressed itself to the question of whether a subject actually did construct an instrument prior to conducting an observation without receiving specific instructions to do so. If a subject constructed an instrument in which not more than 25 per cent of the categories were judged by UCSB scorers to be non-behavioral in their specification, the

instrument was considered to be appropriately constructed. Similarly, if a subject designated the instrument he had actually used in Observation Exercise I as his response to Observation Exercise II, he was judged to have performed satisfactorily if not more than 25 per cent of the behaviors recorded fell outside the category system he had constructed.

The mean scores for all four groups of subjects on the Observation Exercises appear in Table 6. On all six of the constituent measures, the scores for the experimental groups were significantly different, statistically, in the desired direction from the scores of the control groups (higher on Items 1 - 5 and lower on Item 6). Table 6 also reveals statistically significant testing effects on Items 1 - 5. Thus, the influence of previous test experience on desired results cannot be discounted when comparing the posttest performances of Experimental Group I and Control Group I (the pretested groups). The extent of differences in the desired directions between Experimental Group II and Control Group II, however, attest to sizable training influence.

With respect to pretest/posttest changes, Table 7 reveals that the subjects in Experimental Group I scored significantly higher, after training, on the same six measures identified in Table 6.

The test results reported in Tables 6 and 7, like the test results for the Comprehension Exercise, support the inference that the training program yielded desired changes in those specific behaviors which represented specific terminal objectives of ISTP.

Reliability of ratings on Observation Exercises. Two expert raters, after reaching an agreed upon set of standards for each category, jointly rated the tests blind in two groupings: (1) pretests intermixed for experimental and control subjects, and (2) posttests intermixed for experimental

TABLE 6

Mean Post Test Scores and Significance Test Results for the Experimental and Control Groups on Observation Exercises I and II

	Measures from Observation Exercises I & II	Exper. I	Exper. II	Control I	Control II	Testing Effect	Training Effect	Training x Testing
1.	% Behaviorally Specifying Area of Concern	96.55	92.3	50.00	0.00	17.54*	114.97*	12.49*
2.	% Selecting Observation Instrument	96.55	92.3	41.07	8.33	7.37*	100.58*	4.42*
3.	% Focusing on Specified Behavior	93.10	74.9	50.00	8.33	12.56*	42.20*	0.00
4.	% Constructing Observation Instrument	100.00	76.91	75.00	8.33	47.21*	101.46*	29.64*
5.	Total No. of Behavioral Incidents (TBI)	59.78	53.69	41.25	10.92	6.35*	17.99*	2.62
6.	Total No. of Judgmental Statements (TJS)	3.16	4.85	11.42	11.58	0.36	23.75*	0.25

*$p \leq .05; F(1,91) = 4.00$

TABLE 7

Mean Pretest and Posttest Scores and Significance Test Results for Experimental Group I on Observation Exercises I and II

	Measures for Observation Exercise I and II	Pretest	Posttest	F-ratio
1.	% Behaviorally Specifying Area of Concern	8.6	96.55	415.29*
2.	% Selecting Observation Instrument	10.34	96.55	356.25*
3.	% Focusing on Specified Behavior	3.45	93.10	494.00*
4.	% Constructing Observation Instrument	22.41	100.00	197.30*
5.	Total No. of Behavioral Incidents (TBI)	13.26	59.78	116.56*
6.	Total No. of Judgmental Statements (TJS)	9.69	3.16	50.34*

*$p \leq .05; F(1,57) = 4.00$

and control subjects. A test-retest reliability on a random sample selected from both groupings of the tests yielded a mean Pearson correlation coefficient (r) of .99. Approximately three weeks elapsed between the test (original scoring) and retest (rescoring of random sample) for the posttest. The second scoring of the random sample of the pretest took place several months after the first scoring.

Results of the Analysis Exercise

The exercise provided scores for assessing the capability of ISTP to deliver the following analysis skills: (1) identifying a criterion for the successful performance of an intermediate goal; (2) constructing a data display; (3) making summary statements; (4) determining if a criterion for successful performance was met; (5) identifying patterns of behavior; (6) identifying antecedents and consequents of specific behavior; (7) specifying and justifying, in terms of anticipated pupil effects, which teaching behaviors are positive and which are negative; (8) identifying positive teaching behaviors which should be maintained; (9) identifying negative teaching behaviors which should be changed; (10) identifying alternative strategies for producing the desired changes.

All subjects received an identical written protocol which contained a set of raw observation data and related background information. The observation records deliberately included examples of both patterns and individual acts of teaching behavior which would elicit application of the analysis skills to which training addressed itself. Subjects were requested to analyze the observation records according to instructions directed to each of the relevant skills.

Target abilities were scored for occurrence, for increase in frequency of performance, or for both, as follows:

(1) for occurrence of correct identification or determination -- identifying a criterion for the successful performance of an intermediate goal; constructing a data display; determining whether a criterion for successful performance was met;

(2) for frequency in number of instances identified -- making summary statements; identifying patterns of behavior; identifying antecedents and consequents of specific behavior; specifying and justifying, in terms of anticipated pupil effects, which teaching behaviors are positive and which are negative;

(3) for both occurrence and frequency -- identifying positive teaching behaviors which should be maintained; identifying negative teaching behaviors which should be changed; identifying alternative strategies for producing the desired changes.*

Table 8 presents the mean posttest scores for each group of subjects on the Analysis Exercise. A statistically significant training effect in the desired direction showed itself for 11 of the 17 constituent measures. On three measures (Items 11, 15, 16) the differences between experimental and control groups did not reach statistical significance but were in the desired direction. On two measures (Items 6 and 13), differences were inconsistent across the four subject groups. Item 2 showed no difference

*In addition to scoring the subjects' responses for the number of instances where they identified positive and negative teaching behaviors on the grounds of specified student effects, scores were also tallied for the identification of positive and negative instances where no mention of possible student effect was made. The sum of the two scores provided a total number of identified teaching behaviors.

TABLE 8

Mean Post Test Scores and Significance Test Results for Experimental and Control Groups on the Analysis Exercise

	Analysis Test Items	Exper. Group I	Exper. Group II	Control Group I	Control Group II	Testing Effect	Training Effect	Testing Training Effect
1.	No. of Summaries	13.98	11.43	4.84	3.50	1.77	34.11*	0.17
2.	% of Subjects Analyzing Criterion	100.00	100.00	100.00	100.00	0.00	0.00	0.00
3.	No. of Patterns	2.98	2.43	1.83	1.08	4.41*	15.94*	0.11
4.	No. of Antecedents/Consequents	1.43	1.00	0.28	0.50	0.13	7.37*	3.10
5.	Total No. of Statements of Classroom Behavior	4.42	4.36	2.67	2.92	0.03	9.19*	0.09
6.	Total No. of Statements: With Reason	3.62	1.86	1.94	2.33	1.28	0.74	3.68
7.	Total No. of Statements: Without Reason	0.77	2.50	0.72	0.58	5.13*	7.89*	7.08*
8.	% of Subjects Specifying Changes	90.57	85.71	77.78	50.00	3.50	7.73*	1.73
9.	No. of Changes	2.13	1.98	1.17	1.00	.52	6.32*	0.11
10.	% of Subjects Specifying Maintenance	66.04	42.86	27.78	33.34	0.58	4.29*	1.55
11.	No. of Maintenance Behaviors	1.17	0.93	0.39	0.83	0.14	2.55	1.56
12.	% of Subjects Specifying Strategies	79.25	78.57	70.59	50.00	1.02	3.12*	0.89
13.	No. of Strategies	3.32	1.86	1.78	2.42	0.43	0.61	2.78
14.	% of Subjects Specifying Intermediate Goal	88.68	71.43	55.55	41.67	2.53	10.33*	0.03
15.	% of Subjects Specifying at Least One Strategy for Each Change	81.13	57.14	66.67	33.33	7.22*	3.22	0.18
16.	% of Subjects Specifying a Change for Each Negative Teaching Behavior	56.60	50.00	55.56	16.67	3.71	2.12	0.17
17.	% of Subjects Specifying Maintenance for Each Positive Teaching Behavior	52.83	28.57	5.56	25.00	0.05	5.46*	4.02

*$p \leq .05$; $F(1,93) = 4.00$

across groups because all subjects satisfactorily met the requirement of actually counting the particular behaviors which they had incorporated into their chosen criteria.

A statistically significant testing effect was evident for Items 3, 7, and 15; only for Item 7 was there also a statistically significant testing by training interaction. The unusual behavior of Experimental Group II on Item 7 deserves special comment. The training program emphasized identification of both positive and negative incidents of teaching performance and presentation of justification for assigning positive or negative weight on the basis of possible effects on students. Inspection of Items 6 and 7 together reveals that both experimental groups made more identification statements than both control groups, but for some unexplained reason Experimental Group II offered many more statements without specifying reasons than any other group.*

Table 9 shows the pretest and posttest scores for Experimental Group I. On 15 of the 17 measures, statistically significant changes followed exposure to the training program. In the two instances where statistically significant differences did not appear, changes occurred in the desired direction.

The data which appear in Tables 8 and 9 provide evidence that the training program was successful in achieving desired results on the majority of the measures contained in the Analysis Exercise. The significant testing effects associated with Items 3, 7, and 15 could have occurred because of the very explicit nature of the instructions used in the exercise. If experimental and control groups are used again for further revisions of ISTP, it would appear appropriate to modify the test instructions.

---

*Note that lower scores are desirable for Item 7.

TABLE 9

Mean Pretest and Posttest Scores and Significance Test Results on Analysis Exercise for Experimental Group I

	Variable	Pretest	Posttest	F-ratio
1.	No. of Summaries	6.06	13.98	60.31*
2.	% of Subjects Analyzing Criterion	88.68	100.00	6.64*
3.	No. of Patterns	1.51	2.98	64.04*
4.	No. of Antecedents/Consequents	0.36	1.43	28.04*
5.	Total No. of Statements of Classroom Behavior	3.00	4.42	18.47*
6.	Total No. of Statements: With Reason	1.85	3.62	30.18*
7.	Total No. of Statements: Without Reason	1.00	0.77	0.67
8.	% of Subjects Specifying Changes	62.26	90.57	20.53*
9.	No. of Changes	1.15	2.13	24.80*
10.	% of Subjects Specifying Maintenance	26.42	66.04	23.16*
11.	No. of Maintenance Behaviors	0.32	1.17	24.59*
12.	% of Subjects Specifying Strategies	49.06	79.25	16.56*
13.	No. of Strategies	1.13	3.32	40.24*
14.	% of Subjects Specifying Intermediate Goal	49.06	88.68	23.16*
15.	% of Subjects Specifying at Least One Strategy for Each Change	32.08	81.13	38.46*
16.	% of Subjects Specifying Change for Each Negative Teaching Behavior	37.74	56.60	3.76
17.	% of Subjects Specifying Maintenance for Each Positive Teaching Behavior	11.30	52.83	28.15*

*$p \leq .05$, $F(1,53) = 4.03$

Reliability of ratings. Two expert raters jointly rated the analysis test blindly in two different groupings: (1) pretest intermixed between experimental and control subjects, and (2) posttest intermixed between experimental and control subjects. A test-retest reliability on a random sample of combination of these two groups yielded a mean Pearson correlation coefficient (r) of .99. A delay of approximately one month occurred between the test (original scoring) and retest (rescoring of a random sample) of the posttest scoring; several months elapsed between the two scorings of the pretest.

Results of the Participant Performance Record

The Participant Performance Record (PPR) assessed the results of the overall training program on the subjects' performance in actual supervisory settings. Each supervisor received instructions (1) to record a Pre- and a Post-Observation with one of his supervisees in the school where the supervisee was teaching; (2) to observe the supervisee's class using an observation instrument constructed for that purpose; (3) to analyze fully the recorded data with particular attention to identification of positive teaching behaviors to be maintained, to identification of teaching behaviors to be changed, and to specification of strategies for achieving the desired changes. The supervisor's trio members evaluated the tapes and written materials for feedback purposes as the final activity in the training program prior to delivery of the protocols to UCSB for scoring. Each subject was extended the option of re-doing the entire PPR if he was not satisfied with his first effort. Fully completed protocols were available for only 54 of the 76 experimental subjects.

Each terminal objective of the training program was scored according to the percentage of program participants in the combined experimental groups who successfully performed the desired supervisory behaviors in actual conferences with their supervisees. The scoring also noted whether or not the trainees demonstrably employed the eight steps of ISP.

Table 10 shows the percentage of 54 subjects who used the steps and who exhibited each of twenty-five constituent behaviors. Actually 27 separate scores appear in the table. However, items 3a and 8a represent summary statements rather than indicators of constituent behaviors per se. Specifically, item 3a says that 18.5 per cent of the trainees decided (with their supervisees) to set a criterion of performance. The proportions reported for behaviors 3ai and 3aii refer only to the 18.5 per cent and not to the entire trainee group. Item 8a says that all trainees did review the intern's performance with the intern. The constituent behaviors cited indicate how the trainees performed in those particular respects. Of the 25 behaviors scored, the proportion of subjects who performed as desired was higher than the proportion who performed otherwise in 22 cases. In 17 of the 22 instances, the differences between the proportion who did and the proportion who did not exhibit the expected behaviors were statistically significant. In the three instances where the proportion of trainees who did not perform as desired was greater than the proportion who did so (6e, 8aii, and 8c), the differences were significant for the first (6e) and the third (8c). The scores for steps 3 and 4 were derived directly from the sole behavior subsumed under each step and were, therefore, identical to the step itself.

The scoring of the terminal objectives under Step 6 (Analyze the Observation Data) deserves special note. The two constituent behaviors which all trainees were expected to exhibit were 6a (Construct data display)

TABLE 10

Summary of Percentage of Participants Attaining
Terminal Objectives as Measured by
Participant Performance Record (PPR)

THE INSTRUCTIONAL SUPERVISION PROCESS

Terminal Objectives	Yes	No	Not Applicable	Chi-Square**
STEP 1 - Identify the area of concern in behavioral terms	92.6	7.4	---	37.40*
1a. To identify area of concern	100.0	0.0	---	52.02*
1b. To behaviorally define area of concern	92.6	7.4	---	37.40*
STEP 2 - Decide to obtain a base rate or criterion	87.0	13.0	---	28.17*
2. To determine whether the establishment of a base rate or the setting of a criterion is appropriate	87.0	13.0	---	28.17*
3a. If criterion is appropriate,	(18.5)	---	(81.5)	---
(i) to behaviorally define criterion	100.0	0.0	---	8.10*
(ii) to set terminal goal	80.0	20.0	---	2.50
3b. To identify intermediate goal	5.5	1.9	92.5	0.25
STEP 3 - Select, modify or construct an observation instrument	100.0	0.0	---	52.02*
4. To select, construct or adapt an instrument for the area of concern	100.0	0.0	---	52.02*
STEP 4 - Observe the behavior representing the area of concern	100.0	0.0	---	52.05*

* $p \leq .05$

** $X_1^2 = 3.841$, with 1 d.f. at $p = .05$

(Table 10, continued)

Terminal Objectives	Yes	No	Not Applicable	Chi-Square
5. To utilize the instrument to make non-evaluative observation in a classroom setting	100.0	0.0	---	52.05*
STEP 5 - Analyze the observation data	81.5	18.5	---	20.17*
6. To analyze the data by performing the following operations:				
a. Construct data display	98.1	1.9	---	48.17*
b. Make summary statements	98.1	1.9	---	48.17*
c. Determine if criterion met if appropriate	24.1	3.7	72.2	3.33
d. Identify patterns of behavior	55.6	44.4	---	0.47
e. Identify antecedents/ consequents of specific behaviors	25.9	74.1	---	11.58*
f. Specify positive and negative behaviors, justify in terms of anticipated pupil effects	81.5	18.5	---	20.17*
STEP 6 - Identify behaviors needing maintenance or change	57.4	42.6	---	0.91
7. To identify:				
a. positive behaviors to be maintained	61.1	31.5	7.4	4.50*
b. negative behaviors to be changed	64.8	33.3	1.9	4.83*
c. alternate strategies for producing changes in intern behaviors	61.1	37.0	1.9	2.71

(Table 10, continued)

Terminal Objectives	Yes	No	Not Applicable	Chi-Square
STEP 7 - Provide feedback to Intern	94.4	5.6	---	40.90*
8. To perform the following operations:				
a. Review with intern:	(100.0)	(0)	---	(52.02*)
(i) agreed upon area of concern and its behavioral definition	64.8	35.2	---	4.17*
(ii) criterion or base rate predetermined	42.6	55.5	1.9	0.93
(iii) the observation instrument	92.6	7.4	---	37.40*
b. Guide intern in analyzing the data	98.1	1.9	---	48.17*
c. Assist intern to identify behaviors that are positive in terms of their possible pupil effects and should be maintained	24.1	75.9	---	13.50*
d. Assist intern to identify behaviors that are negative in terms of their possible pupil effect and should be changed	94.4	3.7	1.9	44.31*
STEP 8 - Develop strategies for behavior change	88.8	9.3	1.9	33.23*
e. Formulate with intern strategies for achieving the desired changes	88.8	9.3	1.9	33.23*
f. Determine with the intern if additional cycles of the Instructional Supervision Process are necessary	75.9	24.1	---	13.50*

and 6f (Specify positive and negative behaviors with justification based on anticipated pupil effects). Otherwise, the UCSB staff did not expect each trainee would be able to, or necessarily should, perform all of the other four constituent analysis skills during the course of his Pre- and Post-Observation Conferences, on the grounds that the situation might not justifiably call for the employment of these particular skills. In view of the situational factors at work, PPR scores of the trainees on items 6c, 6d, and 6e did not reach the same level as scores on other items in the area of Analysis.

Reversal of expected behavior on 8c, as compared to performance on 8d, squares with previous experience in the UCSB training of supervisors. Apparently it is easier for supervisors to cite instances of negative behavior by student teachers and interns than to identify positive teaching behaviors. Also of interest is the difference between the written and verbal behavior of the trainees on items 8e and 8f. The scorers noted that while the subjects often did not put in writing their views on desired changes in supervisees' teaching performance, and associated strategies, a very large proportion were obviously prepared to discuss both the changes and strategies when they met the supervisee for verbal feedback sessions.

The audio tapes of the Pre- and Post Observation Conferences also permitted the UCSB staff to score supervisor responses to supervisee as "freeing" or "binding". Table 11 summarizes the mean scores, standard deviations, and ranges for the two categories of responses. The scores reveal the participants used a far greater number of "freeing" than "binding" responses in conducting their conferences.

TABLE 11

Summary of Participant Scores in Demonstrating the Use of Freeing and Binding Responses Measured by the Participant Performance Record (PPR)

Pre- & Post-Observation Conferences	Mean	Standard Deviations	Ranges
Total Freeing Response	39.70	30.21	4-150
Total Binding Response	1.30	1.34	0-6

Use of the steps of the Instructional Supervision Process. Table 12 summarizes, in convenient fashion, the performance of the trainees on the eight steps of ISP. The pertinent behaviors were scored for occurrence on one or more occasions during the taped Pre- and Post-Observation Conferences. On seven of the eight steps of ISP, over 80 per cent of the participants performed the appropriate behaviors at high levels of statistical significance. The inability of the ISTP to deliver on Step 6, as compared to the other steps, is conspicuous. This datum, together with the data from two earlier supervisor training efforts at UCSB confirms the common sense observation that analysis of the teaching act is a complex and confounding activity and pointed up the need for further conceptualization and implementation of the Analysis stage in revising ISTP.

Table 13 indicates that approximately 78 percent of the trainees employed at least seven of the eight steps of ISP in their live supervisory sessions. The mean number of steps completed by all participants for whom scores were available was 7.02.

Taken together, the data contained in Tables 10, 11, 12, and 13 offer confirming evidence to analyses of the written exercises contained in the Instructional Supervision Test Battery that ISTP contributed substantially

TABLE 12

Percentage of Participants Completing the Individual Steps of the Instructional Supervision Process and Significance Test Results

Instructional Supervision Process	Yes	No	Not Applicable	Chi-Square
Step 1	92.6	7.4	---	37.40*
Step 2	87.0	13.0	---	28.17*
Step 3	100.0	0.0	---	52.02*
Step 4	100.0	0.0	---	52.02*
Step 5	81.5	18.5	---	20.17*
Step 6	57.4	42.6	---	0.91
Step 7	94.4	5.6	---	40.90*
Step 8	88.8	9.3	1.9	33.23*

*X^2 significant at $p \leq .05$

TABLE 13

Total Number of Steps of the Instructional Supervision Process and the Number and Percentage of Participants Completing Them

No. of Steps	Number of Participants	Percentage
8	24	44.4
7	18	33.3
6	6	11.1
5	3	5.6
4	2	3.7
3	0	0.0
2	1	1.9
1	0	0.0
0	0	0.0
Total	54*	100.0%

Mean No. of Steps 7.02

*PPR data were available for 54 of the 76 training participants

to the development of specified supervisory skills among subjects who received training.

Reliability of ratings for the Participant Performance Record. The individual audio tapes were divided and scored by two expert raters. After the initial completion of the scoring, a random sample of sixteen tapes was selected and each rater scored that half of the tapes which he had not previously rated. A Pearson correlation coefficient (r) of .99 was obtained between the ratings of the two scorers.

Summary of Results of the Field Test

The treatment consisted of administration of a "hands-off" Instructional Supervision Training Program in nine sites (eight Teacher Corps projects across the nation and one Santa Barbara group), involving 76 experimental and 31 control subjects, all volunteers. On-site coordinators who had no previous exposure to ISTP administered the training. The UCSB staff did not intervene directly in the training or in administering the Instructional Supervision Test Battery and associated evaluation instruments.

An adaptation of the Solomon Four Group research design established conditions of pretest and posttest to two experimental groups and two control groups. The selected design assumes homogeneity of group composition. Because the UCSB staff was unable to insure random assignment of subjects at the several training sites, the composition of the groups was examined directly to ascertain the extent of similarities and differences. Chi-square analyses on 18 personal and experience characteristics revealed that on seven of the 18 characteristics statistically significant differences appeared in the composition of the four groups.

However, when the entire experimental group was compared with the entire control group, a statistically significant difference showed itself on only two characteristics: types of credential(s) possessed by the trainees and grade levels taught by the supervisee(s) supervised by the trainees.

Pretest scores also permitted a degree of comparison between one experimental and one control group. No statistically significant differences appeared in the 26 summary measures incorporated in the Comprehension Exercise, in Observation Exercises I and II, and in the Analysis Exercise.

While direct examination of selected personal characteristics of all subjects and comparison of pretest scores for selected experimental and control subgroups unearthed only a few salient differences, the assumption of homogeneity pertinent to the research design does not enjoy full support. The results of the field test should be interpreted in the light of the limitation unavoidably introduced because the UCSB Project Staff could not secure random assignment of subjects. Further caution in the interpretation of results is pertinent because of the large differences in the sizes of the four subject groups.

The Comprehension, Observation, and Analysis Exercises provided one primary batch of posttest data for assessing training effects. On the Comprehension Exercise, both experimental groups scored at statistically significant higher levels than both control groups on the two measures included. On all six of the measures incorporated into the Observation Exercises mean scores for the experimental groups were significantly different, statistically, in the desired direction from the scores of the control groups. The influence of testing and interaction effects cannot be discounted in five of the statistically significant differences between Experimental Group I and Control Group I (the pretested groups). However, differ-

ences in the desired direction between Experimental Group II and Control Group II (not pretested) on each of the same five items attest to a sizable and consistent training influence. On the Analysis Exercise, a statistically significant training effect appeared for 11 of the 17 constituent measures. On three other measures the differences ran in the desired direction although they did not reach statistical significance. On three measures a statistically significant testing effect appeared, but there was only one instance of a statistically significant testing by training interaction.

The second primary batch of data for assessing training effects came from scores of the protocols included in the Participant Performance Record (PPR). Audio tapes and written materials from 54 of the 76 subjects in the two experimental groups were returned to UCSB for scoring the extent to which the trainees were able to perform the terminal objectives of ISTP within a live supervisory setting outside of the training environment. Of 25 behaviors scored, the proportion of subjects who supplied evidence of desired performances was higher in 22 cases than the proportion who did not perform as desired. In 17 of the 22 instances, the differences between the proportion who did and the proportion who did not exhibit the desired behaviors were statistically significant. In only three instances were the differences in the undesired direction; two of these three differences were statistically significant. The PPR also provided evidence that the trainees used a far greater number of "freeing" than "binding" responses in conducting conferences with interns, a behavior pattern which represented one of the 10 terminal objectives of ISTP. The PPR scores, finally, revealed that over 44 percent of the 54 trainees for whom full protocols were available used all eight steps in conjuction with their Pre-Observation and Post-Observation Conferences and that 78 percent of the trainees employed at least seven of the eight steps.

Taken together, the data provided by the Comprehension Exercise, the Observation Exercises, the Analysis Exercise, and the Participant Performance Record offer substantial evidence that the preliminary "hands-off" training materials developed at UCSB contributed measurably to improving the supervisory performance of subjects who received training. All told, 60 different scores were available for analysis of training effects; 53 of the scores were in the desired direction, and 45 were statistically significant. Only two of the few scores which were in the undesired direction were statistically significant.

FEEDBACK FROM USERS

Reaction from training sites to the employment of ISTP as an intervention strategy was primarily positive, and identified a number of specific points where revisions of the materials were pertinent. Feedback came in the form of questionaire responses from seven of the sites and of verbal input from training coordinators who assembled for a Field Test Evaluation Conference in Santa Barbara in April, 1973. The conferees agreed both that there was room for improvement in ISTP and that the materials did deliver training which was appropriate and helpful for the supervision of beginning teachers. According to verbal reports received at the conference, trainees themselves had positively endorsed the training program. As a group, the assembled coordinators urged further revision of the ISTP package and distribution of the revised materials to all interested Teacher Corps projects. They agreed also that ISTP was pertinent to use by supervisors in both teacher preparation programs and local educational agencies. As individuals, several coordinators submitted letters to the Teacher Corps program staff in the U.S. Office of Education in support of the UCSB project.

Specific positive comments from users included the observations that (1) the training program was useful and relevant; (2) the Instructional Supervision Process itself serves as a model of good teaching technique; (3) the Instructional Supervision Process does "free" beginning teachers to make self-evaluations; (4) ISTP does address the improvement of classroom instruction through the acquisition and application of pertinent supervisory skills; (5) ISTP should retain its focus on specific skills needed by instructional supervisors in their work with teachers on classroom instructional behavior, even though these skills represent only a limited portion of supervisory behavior broadly conceived.

Users proposed a total of 28 specific revisions, 15 in the Training Coordinator's Guide and 13 in the other components of ISTP.* Recommendations for the Guide stressed (1) provision of more examples of role playing and observation instruments, (2) improvement of instructions for and substance in the area of Analysis, (3) clarification of the meaning of the concepts of autonomy and collegiality, (4) introduction of a "live" practicum experience in lieu of a second rule playing practice exercise, (5) improvement of instruction for choice of strategies by beginning teachers to improve their performance in selected areas of concern, (6) elimination of excessive directions in the Guide when the directions could be placed in the Readings, (7) addition of bibliography on observational instruments, (8) addition of a glossary of key concepts and terms, and (9) provision of more flexibility for the coordinator to incorporate optional activities into the overall training program.

Other recommendations for revision included (1) presenting the "model" to depict more exactly the several stages incorporated in ISP, (2) tightening

---

*See Boyan et. al., 1973, pp. 37-39, for a full listing.

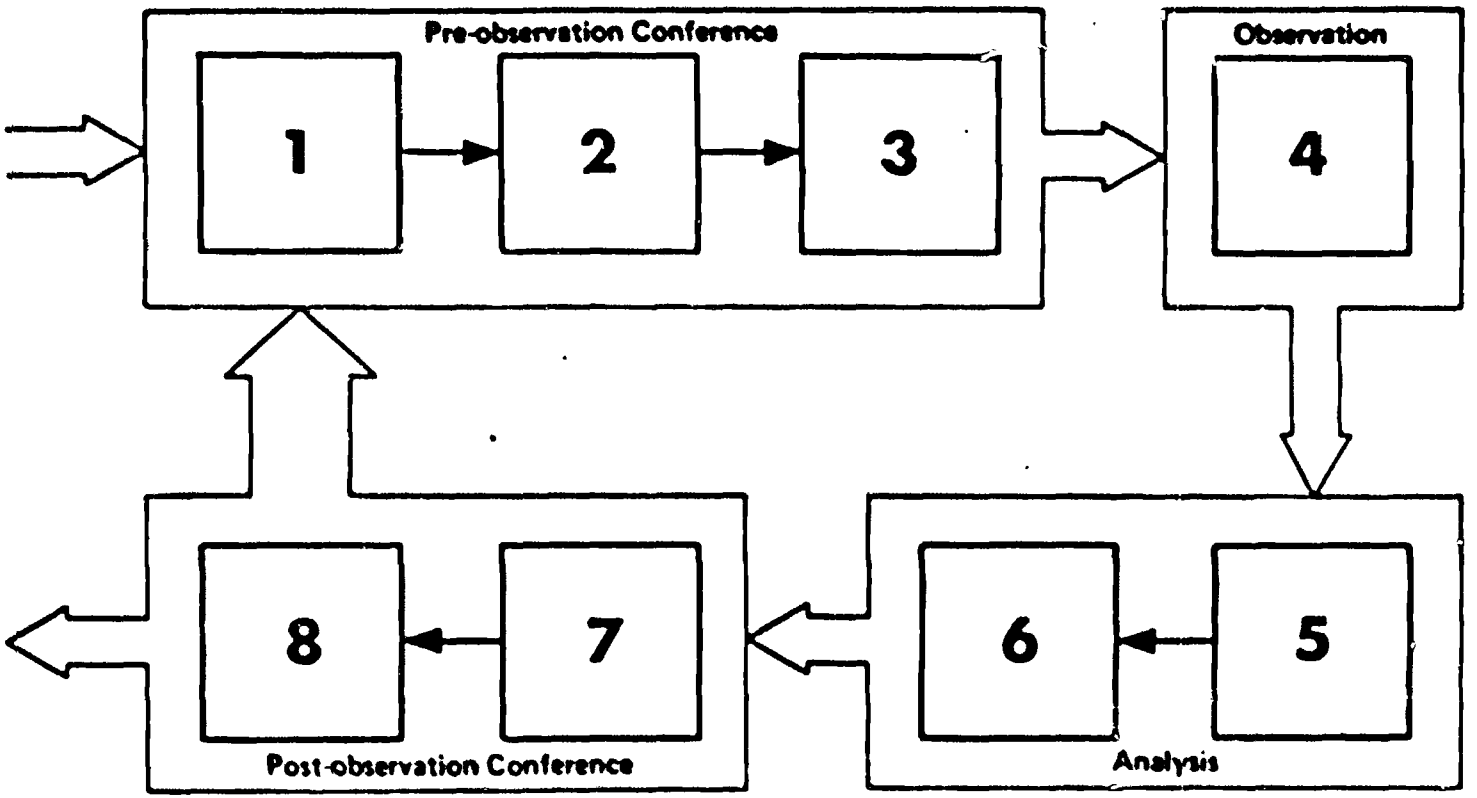
the materials to communicate more content in less space, (3) paginating the Readings and Handouts in an integrated form, (4) adding some large group activities to stimulate broader discussion opportunities for participants, (5) increasing the variety of illustrative concerns of beginning teachers, (6) providing for switching trio members over time, (7) improving role playing exercises, (8) improving the substantive content and technical quality of the audio tape and video tape exemplars and stimuli, and (9) adding a number of descriptive transparencies for large group presentations.

CONCLUDING NOTE

After analysis of field test results, review of questionnaire responses from training sites, and consideration of verbal assessments from Teacher Corps trainers, each segment of the ISTP materials (Training Coordinator's Guide, Readings, Handouts, audio tapes, and video tapes) experienced extensive revision.

The revised package groups the eight sequential steps of instructional supervision into four stages: Pre-Observation Conference, Observation, Analysis, and Post-Observation Conference. The change in the representation of the "model" of ISP will hopefully provide users with clear understanding of the relationship of the "steps" to the "stages". More specifically, the grouping of the "steps" into "stages" should assist trainees to apply the constituent elements of ISP in systemically related fashion in real supervisory situations. (See Table 10 for identification of the steps and Figure 1 for representation of the stages.)

The original seven instructional units have become 16, through reconstitution and rearrangement. Each of the revised units employs one of the following training modes: information transmission and acquisition,



THE INSTRUCTIONAL SUPERVISION PROCESS

Figure 1

role playing, or practicum (see Table 14). Estimated advantages of the rearrangement are (1) that each instructional unit addresses itself to the development of a specific skill and that the amount of training time for each unit is approximately the same; (2) that trainees receive more opportunities throughout the training for practicum experiences; (3) that, prior to each practicum experience, trainees demonstrate their competency with respect to knowledge about key concepts in instructional supervision and to skills in applying the concepts.

The field test results pointed to the need for major revisions in the Analysis Stage of ISP, and associated training, despite salutary evidence secured from the written Analysis Exercise. Data from the PPR, especially scores for Steps 5 and 6, indicated that in a live supervisory situation, trainees required further assistance in mastering the Analysis domain. Training for the Analysis Stage of ISP (1) has dropped use of the notions of antecedent and consequent events as strategic, (2) places heavier stress on the citation and maintenance of positive teacher behaviors while retaining emphasis on eliminating negative behaviors, and (3) introduces the employment of an entirely new Analysis protocol form to insure inclusion by trainees of appropriate justifications related to probable effects on pupils in classifying a teaching behavior as positive or negative.

Instruction for the Post-Observation Conference highlights (1) the essential need for supervisor review with supervisee of their initial agreements in the area of concern, (2) the importance of the supervisor's guiding the supervisee through analysis of the classroom observation data in a way which approximates the analysis conducted separately by the supervisor, (3) the necessity for identifying instructional strategies which attend to the maintenance of positive behaviors as well as the elimination of

TABLE 14

Instructional Units of ISTP According to Instructional Mode

TITLES	FOCUS: Information Acquisition	Role Playing	Practicum
I. Introduction to Instructional Supervision	X		
II. Interpersonal Communication	X		
III. Behavioral Language	X		
IV. Pre-Observation Conference: Step 1 - Behaviorally Define the Area of Concern	X		
V. Pre-Observation Conference: Step 2 - Decide to Obtain a Base Rate or Set a Performance Criterion	X		
VI. Pre-Observation Conference: Role Playing Exercise		X	
VII. Pre-Observation Conference and Observation: Steps 3 and 4 - Select and Use an Observation Instrument	X		
VIII. Pre-Observation Conference and Observation: Role Playing Exercise		X	
IX. Pre-Observation Conference and Observation: Practicum			X
X. Analysis: Step 5 - Analyze the Observation Results	X		
XI. Analysis: Step 6 - Identify Behaviors Needing Maintenance or Change	X		
XII. Pre-Observation Conference, Observation and Analysis: Practicum			X
XIII. Post-Observation Conference: Steps 7 and 8 - Feedback the Data Results and Determine Strategies	X		
XIV. Analysis and Post-Observation Conference Role Playing Exercise		X	
XV. Instructional Supervision Process: Role Playing Exercise		X	
XVI. Instructional Supervision Process: Practicum			X

negative behaviors, and (4) the value of insisting on following a decision to recycle the Process with observation of additionally specified supervisee behaviors. The reconstructed Post-Observation Conference Stage omits reference to discrimination training because training coordinators in the field test reported such training to contain little utility.

The revised set of materials expand the capability for participant self-evaluation. At the start of each training unit, the participant receives a statement of objectives to be mastered. During the course of the unit, presentation of lead questions enables the participant to check his progress toward mastering the stated objectives. At the end of each unit, the objectives appear again, for review purposes.

Revision of ISTP occupied the UCSB staff during the summer months of 1973. A fully revised training package, entitled the "Instructional Supervision Training Program", was delivered to the U.S. Office of Education in September, 1973.