DOCUMENT RESUME

ED 162 466 | EC 112 509

AUTHOR | Semmel, Melvyn I.; And Others
TITLE | The Effects of CATTS Feedback in a Preservice Special Education Teacher Training Program.
INSTITUTION | Indiana Univ., Bloomington. Center for Innovation in Teaching the Handicapped.
SPONS AGENCY | Bureau of Education for the Handicapped (DHEW/OE), Washington, D.C.
PUB DATE | Aug 76
GRANT | OEG-9-242178-4149-032
NOTE | 131p.; Several tables in Appendix I may not reproduce well due to light print
AVAILABLE FROM | Indiana University, Center for Innovation in Teaching the Handicapped, School of Education, Bloomington, Indiana 47401

EDRS PRICE | MF-$0.83 HC-$7.35 Plus Postage.
DESCRIPTORS | Behavior Change; Classroom Observation Techniques; *Computer Assisted Instruction; *Educational Technology; *Feedback; *Handicapped Children; Information Processing; Instructional Systems; Program Descriptions; Program Design; *Program Effectiveness; Skill Development; Student Behavior; Student Teacher Relationship; Teacher Behavior; *Teacher Education; *Teaching Skills
IDENTIFIERS | *Computer Assisted Teacher Training System

ABSTRACT

The effects of Computer-Assisted Teacher Training System (CATTS) feedback in a preservice special education teacher training program are discussed. It is explained that a series of studies were conducted to test the efficacy of CATTS feedback in effecting teacher trainees' acquisition and performance of specific teaching skills. Chapter 1 presents the goals and objectives of the project, an overview of the project (pilot procedures, descrimination training, laboratory classroom design, trainee evaluation), and describes both the organization of the CATTS project and an overview of CATTS itself (teaching station, observation coding station, analysis-encoding station). Chapter 2 deals with the two observation systems used for feedback of teacher and student behaviors to trainees, COG-STRAT (focused on teacher and student cognitive styles of interaction) and MAN-STRAT (focused on student on- and off-task behavior and strategies of teacher management of student behavior), and discusses such areas as observer training and evaluation of observer competencies. Such aspects of project organization and implementation as teacher education laboratory classrooms, implementation procedures (teaching and coding), scheduling observation and teachers, and the effects of feedback on trainee behaviors are studied in chapter 3. Chapter 4 summarizes the results of the study which revealed that all trainees significantly increased their rate of criterion performance as a function of CATTS feedback. The summary and conclusions of the project are presented in the final chapter. (BD)

THE COMPUTER-ASSISTED TEACHER TRAINING SYSTEM (CATTS) DEVELOPMENT AND APPLICATIONS

THE EFFECTS OF CATTS FEEDBACK IN A PRESERVICE SPECIAL EDUCATION TEACHER TRAINING PROGRAM

Project Director: Melvyn I. Semmel

Principal Investigators: Merrill C. Sitko, Dorothy S. Semmel, Ted Frick, Ted Hasselbring

Project Staff: Jerry L. Olson, Laura Meade, Judith Gehlhausen, Chad Sivasailam, Mona Ballard, Karin Meyers

Center for Innovation in Teaching the Handicapped

Indiana University, School of Education, Bloomington

August, 1976

This research was supported by grant #OEG 9-242178-4149-032 from the U.S. Office of Education, Bureau of Education for the Handicapped to the Center for Innovation in Teaching the Handicapped. 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.

Table of Contents

Tables

Figures

I. GOALS AND OBJECTIVES OF THE STUDY AND AN OVERVIEW OF THE PROJECT

The present series of studies were conducted to test the efficacy of Computer-Assisted Teacher Training System (CATTS) feedback in effecting teacher trainees' acquisition and performance of specific teaching skills. The realization of the major goal of trainee skill development through CATTS entailed the continuing development of computer hardware and software support systems for real-time processing and feedback of specific observation coding data. A computer-based system for the simultaneous collection, processing and feedback of coded observational data from several classrooms ~~at once~~ was developed. In addition, realization of the training objectives of CATTS required the development of an efficient, reliable observer training program. For this purpose, a consensus coder training package bearing the acronym "DITRMA" was designed and integrated into the total effort aimed at improving teacher-trainee skill performance.

The continuing development of CATTS as a valid performance and cost-effective program for training teachers of handicapped children rested on several major assumptions:

1. The use of observation systems for training both supervisory personnel and teachers to discriminate relevant teacher-learner behaviors provides an objective, common set of operational definitions of the objectives of a given teacher-training program. The CATTS software system has been developed to accommodate any set of observational data and has, in past applications, been programmed to process data obtained through the Flander's system (Kreider, 1969; Schmitt, 1969; Weaver, 1969), the Individual Cognitive Demand Systems, ICDS, (Lynch & Ames, 1971), and the [illegible] Behavior Management System, BMS, (Fink & Semmel, 1971). The application of CATTS to the acquisition of teaching skills rests in all instances upon [illegible] trainers'

detailed specification of the objectives of the training program through the medium of a reliable observation system.

The use of observation system coding for teacher training has in recent years gained recognition as a useful method of joining training objectives to a reliable measure of the trainee's attainment of performance objectives. Originally developed as a method for describing classroom interaction (Anderson, 1939, Smith & Meux, 1962) and for collecting such data for research purposes, the system also has great utility as a method for focusing trainees' attention on the specific interactional and classroom climate variables that are central to the given training program. In the review of research on observation system studies in teacher education which accompanies this report, Semmel (1975) held that, although there are numerous and diverse observation systems available for coding teaching-learning behaviors, the procedures used in applying these systems in training situations are essentially similar. The standard paradigm requires an observer in the classroom to record the ongoing interactions according to a prescribed code. The classroom interaction may also be recorded on audio or video tape and later transcribed and coded to conform to the system employed. The coded data is then analyzed and reported, usually as a summary of time in category or frequency of occurrence in a category of behavior, or as a matrix reflecting the sum of double-entry dyads (e.g., the number of items of behavior category x that followed category y). The data is then transmitted back to the trainee for comparison of his performance with the performance criterion set by the training program.

A limitation inherent in the typical training paradigm described above is the delay of feedback of information to the trainee due to the time lag between collection of data and feedback to trainee necessitated by the analysis processes. The CATTS system has been developed to overcome this limitation of observation systems for teacher training through the use of a coding system which utilizes the flexibility and speed of the computer to store and analyze coded data in real-time, so that information on classroom interaction may be fed back to the trainee while he/she is still engaged in the interactive teaching process (Semmel, 1975).

It may be argued that the increased use of video tape feedback in teacher training (Allen & Ryan, 1969; Borg, et al., 1970) mitigates the necessity for coding trainee's behavior. However, in a critical review of feedback in teacher-training programs, Fuller, [illegible] and Manning (1972) concluded that feedback which lacks a specific focus has not been found to change trainee behavior. They stressed the importance of cueing specific behaviors, particularly in video tape (microteaching) and self-confrontation training techniques.

On the other hand, Peck and Tucker (1971) concluded that feedback is instrumental in changing teacher behavior only when another person (e.g., supervisor or peer) participates in the feedback session. However, such performance change, if devoid of objectively measured predetermined criteria (such as provided by an observation system), may be instrumental in changing trainees' behavior as interpreted by the supervisor (e.g., subjectively), but would not offer any assurance that performance goals of the training program are met.

2. <u>The second major assumption of the CATTS training program is that, in order for trainees to generate the desired interactive teaching behaviors in the classroom, information on status of the performance must be fed back to the trainee while he/she is still teaching.</u> The importance of feedback or knowledge of results in learning is a venerable axiom in the field of human

learning. In teacher-training programs geared to observation systems and in microteaching, the principle of feedback and knowledge of results is central. However, the task of applying feedback principles, so potent in motor skill and verbal learning in the laboratory, to skill and attitude development in teacher or counselor training, is formidable. Not the least of these problems is the definition of "immediate" feedback. The typical teacher-training program is based upon a teach, supervisory feedback or video playback and reteach model (Allen, 1969; Borg, 1970). If the supervisor employs an objective assessment instrument such as an observational system, then the delay in feedback may often be several days or more. The CATTS system is designed to overcome this major limitation of the teach-feedback-reteach strategy in teacher-training programs. The feedback available though CATTS is returned to the classroom in real-time and is also stored for retrieval for post-teaching analysis and interpretation.

3. <u>The setting of teaching performance goals can be individualized so that each trainee may specify his own performance goals.</u> Individual performance goals can be determined on a logical basis if the trainee has access to a baseline of his own behaviors on previous lessons. Those categories of behavior where the individual performance on a given dimension (e.g., a single behavioral category or constellation of behavioral categories) falls above or below a desirable criterion may be selected as the individual's performance goal.

The facility of a training program to individualize performance objectives is a vital one, as the program must account for individual differences in the entry level of trainees. Also, there is little evidence that optimal teaching behaviors are unidirectional or linear (see Medley & Soar, 1975;

Semmel, Semmel & Morrissey, 1976). Thus, a training paradigm which involves simple and continuous increases in certain desirable teaching behaviors (e.g., indirect teacher questions, positive reinforcement) may lack validity. A performance objective, such as the increase in positive reinforcement of pupil responses, should be conditioned upon knowledge of the trainee's level (pre-training) of positive responses to pupils. In cases where the trainee's entry level is high, further increases may decrease the effectiveness of such behavior. Thus, a training program must be able to adjust general performance goals to the individual differences among trainees.

To summarize the major assumption underlying the application of CATTS to teacher training; the trainees acquisition of teaching skills is viewed as dependent upon several major conditions; first, the trainees development of accuracy in the discrimination of teacher-learning behaviors; second, availability of reliable feedback on teaching-learning behaviors to the trainee during the course of an ongoing lesson; and third, the availability of valid and reliable data on the trainee's own performance over time, for post-teaching analysis and review.

Overview of the Project

Pilot procedures. Prior to the initiation of the training program described in this report, several pilot projects were carried out during the previous school year at the Teacher Education Laboratory of The Center for Innovation in Teaching the Handicapped (CITH) to test the operating efficiency and the internal integrity of CATTS, as well as to determine the effectiveness of the system in instrumentally improving the teaching performance of special education teacher-trainees. In order to evaluate the CATTS configuration in teacher training, two initial pilot studies were carried out during the previous school year. The first study was designed to ascertain the effect of two types of CATTS feedback on increasing the number of high-level questions asked by the trainee during a given lesson. The two types of CATTS feedback studied were instant video display in the classroom, compared with delayed computer printout feedback following a lesson.

The second pilot study compared the effect of the same feedback variables on trainee questioning skills in a simulated classroom setting.

The experience obtained by the CATTS development team in design and execution of the earlier pilot studies was incorporated into a full test of the effects of the two types of CATTS feedback on trainee behavior. A description of the CATTS pilot studies is presented in Appendix I of this report.

The experience gained from these initial attempts at implementing a teacher-training program under the CATT(S) system was reflected in the design of the present CATTS study in which three major areas of development and research were focal: (1) a coder-training (discrimination-training) program for trainees, (2) establishment of a laboratory classroom, (3) a formative evaluation system and assessment of the effects of two types of feedback on trainee generation of prescribed teaching behaviors.

Discrimination training. Both the pilot test phase of the CATTS project and a continuing review of the literature strongly indicated that the ability of the trainee to define and discriminate the critical classroom interactive behaviors was a necessary precondition for the trainee to subsequently generate these behaviors during teaching. The CATTS project emphasized development of discrimination skills through a series of training modules designed to reinforce the objectives of CATTS training at each stage of implementation. To this end, a coder-training program was also developed, aimed at establishing a high degree of coding reliability on the part of the participant trainees. Coding reliability was assumed to be *prima facie* evidence that the trainee could discriminate the relevant behaviors of the system.

The laboratory classroom. Several types of trainee support systems were designed and established to assist the trainees in lesson planning and pupil evaluation. Procedures for pupil selection, scheduling, and trainee consultation were also developed for the operation of the laboratory classroom.

Trainee evaluation. A system of formative evaluation of individual training goals was designed to provide trainees with an opportunity to measure and record their own teaching behaviors and to adjust their interactive patterns and goals, based on data made available through CATTS delayed (printout) feedback.

In addition to the empirical determination of the efficacy of the two types of CATTS feedback on the trainees' acquisition of specific teaching skills, several summative evaluation measures were taken, including trainee attitudes toward the program and their knowledge of teaching behaviors.

The main objectives of the study of the use of CATTS in special education preservice teacher training were as follows:

1. To develop software systems for real-time feedback of performance data to trainees while they are teaching.
2. To compare the effectiveness of two modes of feedback on trainee performance.
3. To evaluate the CATTS model of preservice teacher preparation.

Organization of the CATTS Project

Since the present project combined aspects of both an experimental study and that of a practicum course, coordination of a number of distinct operations was required. There were five major operations that characterized the project: (1) hardware and software development, (2) training of observers in the discrimination of teaching behaviors, (3) organization of a laboratory classroom, (4) generation of specific teaching behaviors with and without CATTS feedback, and (5) evaluation of trainee performance.

A detailed report on the CATTS configuration may be found in the CATTS Technical Reports (Semmel & Olson, 1977), while a brief summary of the hardware and computer software systems underlying the CATTS program is presented in the next section of this report.

Chapter II presents a detailed description of the observation systems from which trainee performance objectives were derived. It also describes the development of materials and procedures for training teachers to become skilled, reliable coders of the observation system categories.

The various phases in the development and implementation of the CATTS training program and research study are described in Chapter III.

Overview of CATTS

CATTS configuration. The system configuration of CATTS consists of three interdependent stations: Teaching Station, Observation-Coding Station, and Analysis-Encoding Station. Figures 1 and 2 illustrate this configuration with schematic diagrams of the present CATTS installation at the Teacher Education Laboratory (TEL), at CITH.

Teaching Station. The CATTS Teaching Station consists of classroom(s) which accommodate various auditory and visual feedback devices. Visual feedback devices are placed so that the teacher can have access to the feedback information during teaching, with a minimum of classroom disruption. These various feedback display devices are controlled directly by the computer.

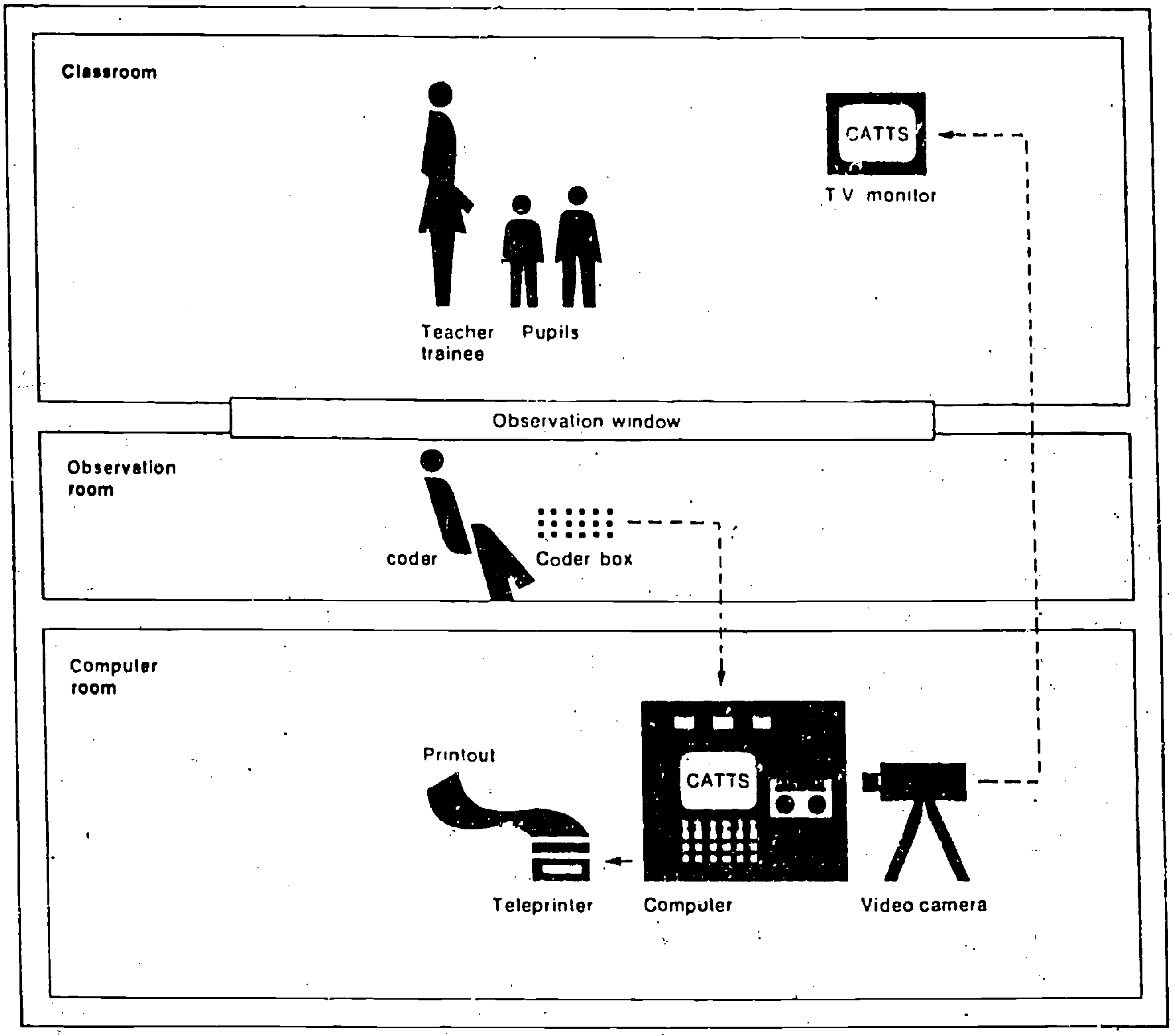


Figure 1. Arrangement of Teacher Education Laboratory for CATTS.

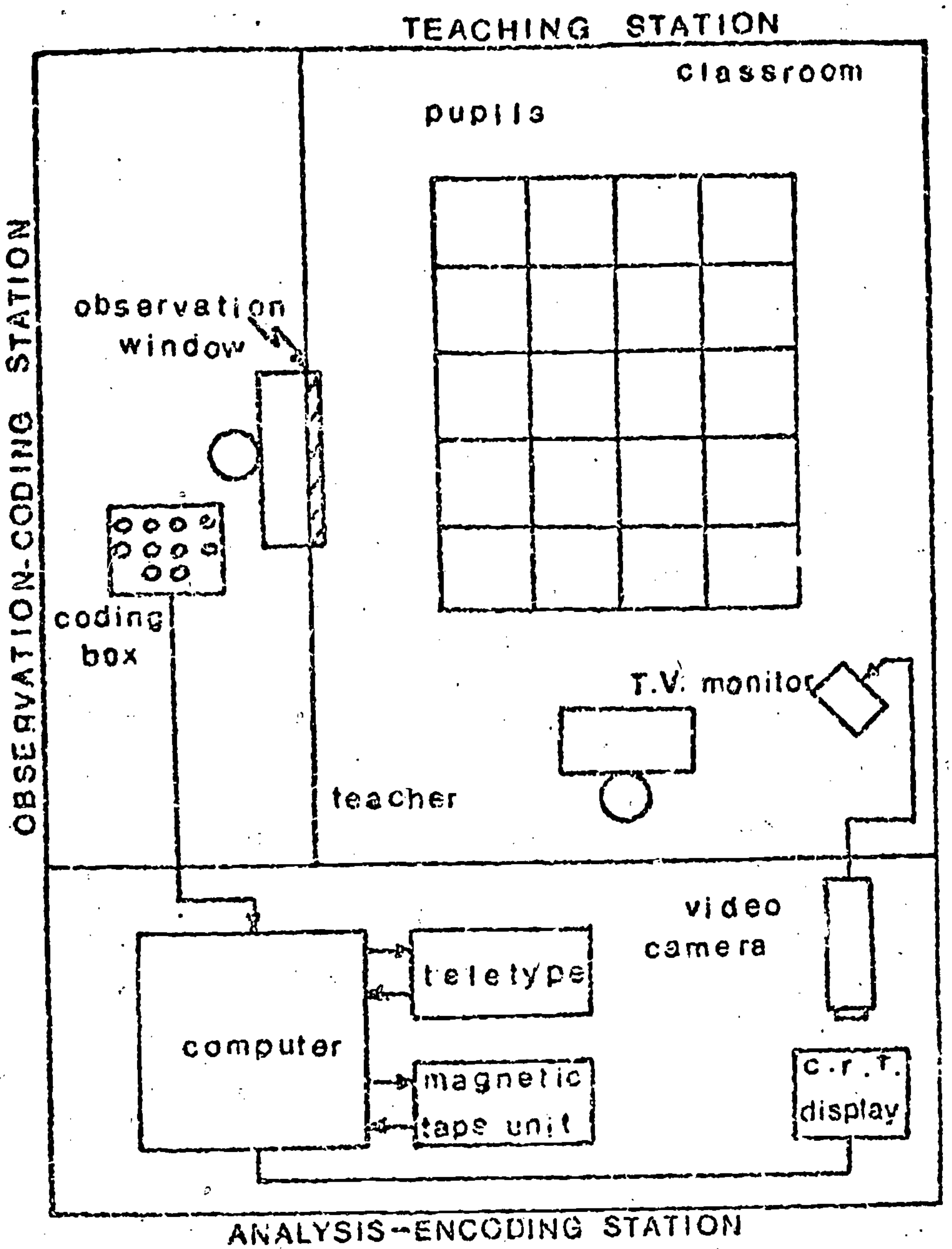


Figure 2. CATTS Configuration.

Visual displays available in the Teaching Station (classroom) vary from the closed-circuit televised images currently employed, to various external visual devices which convey feedback information to the classroom by changing light patterns.

Observation-Coding Station. The Observation-Coding Station provides the link between the events occurring in the classroom and the computer analysis of these events. Visual observation of classrooms may take place within the classroom itself, within an observation room adjoining the classroom (as in the present study), or through a closed-circuit television connection. There are three methods of data input employed by the CATTS system: a directly connected coding box device (used in the current project); a TOUCH-TONE telephone access system; and a portable recording unit for off-line data collection. The first two methods permit real-time data collection for instantaneous feedback applications. The third method extends the system's data collection ability beyond the physical limitations of a direct computer connection into the field, and permits flexible collection of base-line data and delayed feedback observations.

At present, the computer-connected coding device with which the observer enters his observations consists of a small hand-held box with 12 TOUCH-TONE pushbuttons. Currently, CATTS can accept data independently and simultaneously from 12 directly connected coding devices. This input time-sharing capability allows simultaneous data collection from various locations in order to fulfill classroom observation requirements for independent projects.

Analysis-Encoding Station. The Analysis-Encoding Station contains a PDP-12 digital computer and associated computing hardware required for the on-line processing of pushbutton-coded data which are gathered and transmitted from Observation-Coding Stations. In addition to processing the incoming data, the computer system permits the control of various display devices used in Teaching Stations and provides hard-copy printouts, storage, and transfer of data to remote computer systems for subsequent analysis. It is at this station that an operator initiates the computer program options available in CATTS. The teleprinter conso[illegible] through software program control, permits the operator to select spec[illegible] CATTS programs or options that satisfy various project data collection and feedback requirements.

The selection of the mode and conte[illegible] feedback to the teacher at the Teaching Station is initiated from the cons[illegible]e. If a CRT (Cathode Ray Tube) display is chosen as the method for feedback the operator determines the content of the display by assigning the incoming data, by code, to different computational functions for the computer to calculate and display as feedback. The nature of the display is also selected from the console and allows feedback information to be presented either in alpha-numeric or graphic form.

Upon completion of an observation session, the raw data are stored in local computer magnetic tape units and become accessible for various data analysis applications. Analysis programs are available on both the PDP-12 computer and Indiana University's central computer facility. For remote processing, the data are transferred by phone into permanent disk and tape storage files which are accessible by either batch or time-sharing analysis programs.

The associated PDP-12 teleprinter, which also serves as the communications link to the CATTS program, provides summary data printouts for inspection throughout the data-gathering stage. The printouts provide such information as the time span of coded observations and the actual feedback functions currently being displayed to the teacher. At the end of a real-time session, a selection of various data summary programs and printouts is available for delayed feedback applications.

II. OBSERVATION SYSTEMS AND OBSERVER TRAINING

The two observation systems used for feedback of teacher and pupil behavior to trainees via CATTS were the COG-STRAT and MAN-STRAT. Categories on the COG-STRAT include teacher and pupil cognitive styles of interaction, while the MAN-STRAT category is focused on pupil on- and off-task behavior and strategies of teacher management of pupil behavior (See Table 1 and also the Observer Training Manual, Appendix). It is important to note that these observation systems are neither new nor original systems. Rather, they were developed to meet teacher-trainee feedback needs for specific teaching skill development during their practicum experience.

Prior to their practicum, trainees were enrolled in a course on systematic classroom observation. During this course they were exposed to a variety of observation systems and were required to individually develop systems based on their interests. In addition, all trainees and their instructor mutually agreed on a number of teaching behaviors and corresponding pupil behaviors deemed important during classroom interaction with handicapped children. These behaviors were logically grouped into two domains, cognitive and affective management, and eventually became the COG-STRAT and MAN-STRAT systems. Moreover, specific categories or groups of categories from existing observation systems were chosen whenever they coincided with trainees' needs. Systems from which categories were specifically taken and/or modified included the Individual Cognitive Demand Schedule (Lynch & Ames, 1972), the Indiana Behavior Management System-II (Fink & Semmel, 1972) and the Flanders' Interaction Analysis System (1970). A perusal of Mirrors for Behavior (Simon & Boyer, 1970) will also evidence other systems with categories similar to those selected (e.g., Smith & Meux; Amidon).

Table 1
COG-STRAT and MAN-STRAT
Observation System Categories

COG-STRAT

Pupil Categories:

1. Initiates Talk
2. Low-Level Response
3. High-Level Response
4. No response

Teacher Categories:

5. Talk
6. Low-Level Question
7. High-Level Question
8. Positive Feedback
9. Negative Feedback
10. Incorporation/Extension
11. No Interaction
12. Interruption/Confusion/Management

MAN-STRAT

Pupil Categories:

1. On-Task
3. Off-Task

Teacher Categories:

2. Reinforcing
4. Ignoring
5. Redirecting
6. Commanding
7. Focusing
8. Norm Referencing
9. Signalling
10. Humor
11. Demeaning
12. Uncodables

Also paramount in this selection process were considerations of simplicity, clarity, and objectivity of the observation systems, since trainees would be required eventually to learn the systems and code each other when teaching. For example, one reason that entire existing systems were not chosen was that they were too complex and/or contained categories irrelevant to trainee goals of teaching skill development.

Once specific categories were selected and broadly defined, they were given to a CATTS staff member who was highly experienced in the development of observation systems and instructional programming of observer-training materials. Critical attributes of categories were identified, and the systems were then developmentally tested with a coder trainer system, "DITRMA," and subsequently revised. The operation of DITRMA will be described elsewhere in this report. Both audio and video tapes of classrooms and classroom simulations were coded on button boxes during developmental testing of the utility of the systems. End products of this activity were observer training manuals and rules for button-box coding (see Module 1).

Additional observer-training materials were then developed for use with DITRMA: (1) Two scripts were written for practice coding with feedback in order to initially familiarize trainees with the categories and numbering schema before DITRMA coding. (2) A number of video-taped, unambiguous, isolated examples were then produced via simulation for initial training with DITRMA. These examples were structured so that they became increasingly more difficult, and they included more behaviors to code as the tape progressed. By the end of the tape trainees were required to code 5 or 6 behaviors during a 15- to 20-second interval. (3) A second video tape for each system consisted of a number of continuous 5-minute segments of classroom interaction. Segments were chosen from the TEL tape library on the basis of their coding difficulty and representativeness of system categories. (4) Finally, criterion tapes were produced in order to test trainee competency in the observation systems. Each test was composed of two parts: One section consisted of randomly ordered, isolated, unambiguous simulated [illegible] classroom interaction. Approximately 10 or more examples [illegible] category were included. The second section of each video [illegible] simulation of classroom interaction which was to be coded [illegible].

Following production of these training materials, they were tested with a number of CITH personnel. In addition to training back-up coders for the practicum, the primary goals of this testing were to estimate training time for scheduling purposes and to identify specific training problems. As a result of testing, modifications of the training procedure were made; but more important, a time shortage for scheduling of observer training became evident. It was obvious that it would be unreasonable to train 27 teacher trainees on two observation systems in two weeks considering trainee class schedules and available time. In order to preserve enough time in the semester to allow each trainee to teach at least 9 times and still maintain the design of the study, it was decided that trainees could only competently learn one of the two systems in the allotted time period. COG-STRAT was chosen for the trainees since it appeared to be less difficult to learn on the basis of developmental testing. In order to utilize the other system, it was decided that CITH personnel would be trained to criterion on the MAN-STRAT, while trainees would also learn the MAN-STRAT (but not as coders) at the earliest possible time during regular class meetings.

---

[1] A rationale for this testing procedure will be discussed in a later section.

Following developmental testing of training materials, final revisions were made, and Module 1 was prepared. Module 1 essentially briefed students on what was to be expected during observer training. It included a rationale for learning observation systems, a description of DITRMA, training schedules, a training syllabus, an outside assignment, the COG-STRAT Observer's Training Manual, Rules for Button Box Coding, a practice coding script with confirmation. Module 1 also contained the MAN-STRAT Observer's Training Manual, Rules for Button Box Coding, and a practice coding script. (See Module I).

Observer Training

In the first stage of the practicum, or the discrimination stage, trainees learned the COG-STRAT system. CITH personnel also learned the MAN-STRAT system at this time. Acquisition of observation skills was facilitated by a computer-aided training device called DITRMA.

DITRMA

DITRMA is based upon a simple consensus coding principle whereby individual trainees' responses from two or more coding terminals are simultaneously compared by the computer, and whether or not there is agreement across observers' entries is instantaneously fed back to trainees. Through expanded application of this simple consensus principle, the DITRMA system can be used to teach DIscrimination of relevant teacher-pupil behaviors, to TRain for reliability in recording those behaviors, and to MAintain that level of reliability (Semmel, 1975).

DITRMA operates as follows: During coder-training sessions observers code video-taped examples of the observation system on button boxes. These button boxes are similar in configuration to a TOUCH-TONE telephone and are linked by cable to a PDP-12 computer. The computer acts as an impartial judge as trainees code with a small group of their peers. If they agree with each other on a certain event, the video tape keeps playing and the computer remembers that they agreed on this example. On the other hand, if there is disagreement as to how an event should be coded, the computer automatically stops the video tape and "refuses" to start the tape until the trainees, as a group, have reached consensus on what they saw.

Physical configuration of DITRMA. For the CATTS practicum, the physical configuration of DITRMA consisted of two adjacent, interrelated stations: Training Station and Encoding/Feedback/Control Station. Included at the Training Station were a maximum of six on-line coding terminals, a closed-circuit video feedback monitor, and a video playback monitor. The Encoding/Feedback/Control Station contained the computer, a Sony EV 300 video tape recorder (VTR), the slave CRT scope and TV camera, and other associated hardware required for on-line processing of incoming data from the Training Station (see Figure 3).

The operation of DITRMA. Once an observer has initiated a code on the button box, all remaining observers must enter their codes within a given time interval. The length of this time interval can be varied, allowing the trainer to shape observer reaction time. If every entry is made before expiration of the set time interval, the computer compares the codes from each observer upon receiving the final data entry. If all observers agree, the computer enters this consensus data into its data storage buffer, and also sends an audio visual feedback signal to the trainees (see Figure 4). This signal not only serves as confirmation of agreement but also cues trainees that they can enter new codes, if necessary.

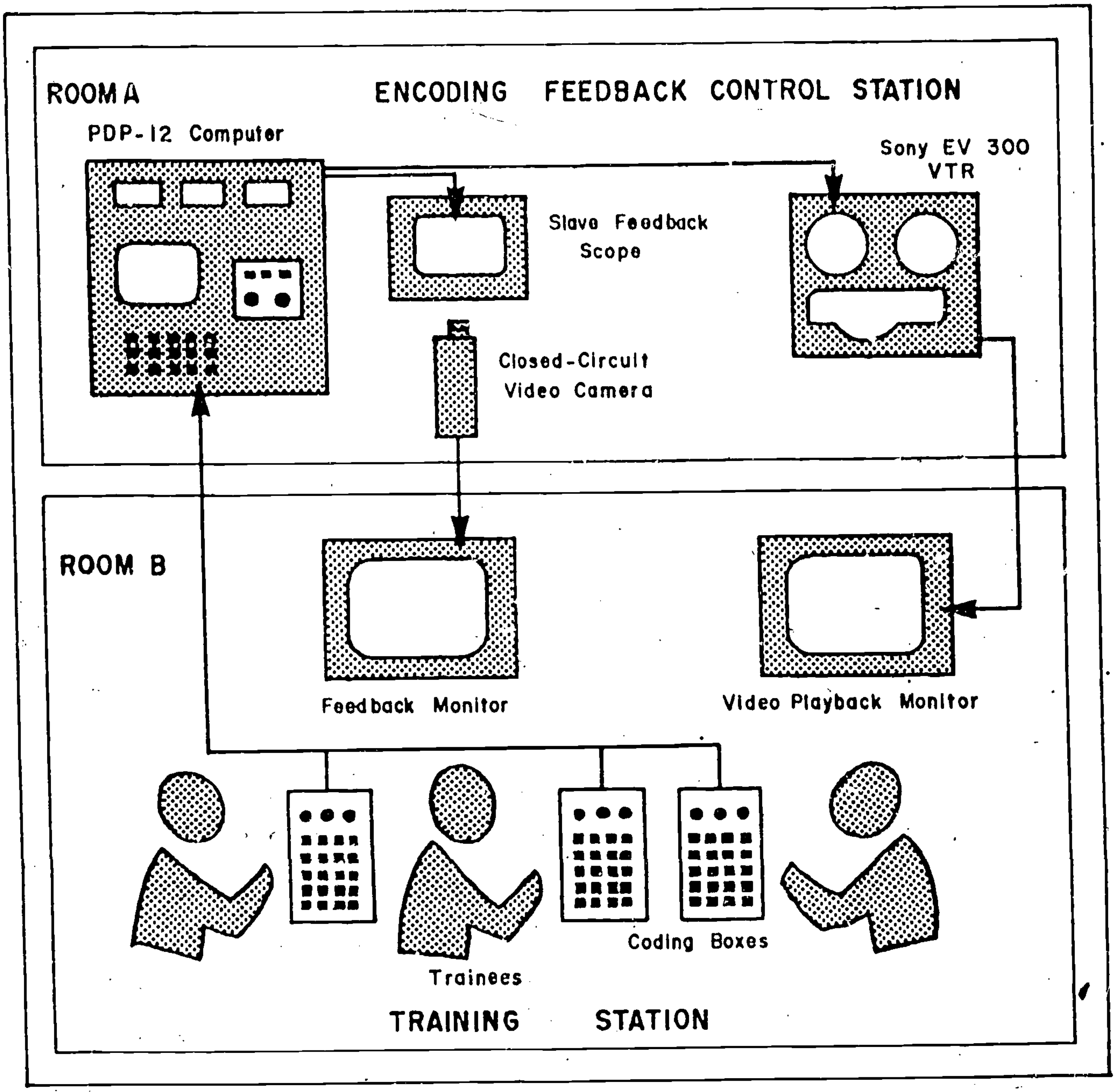
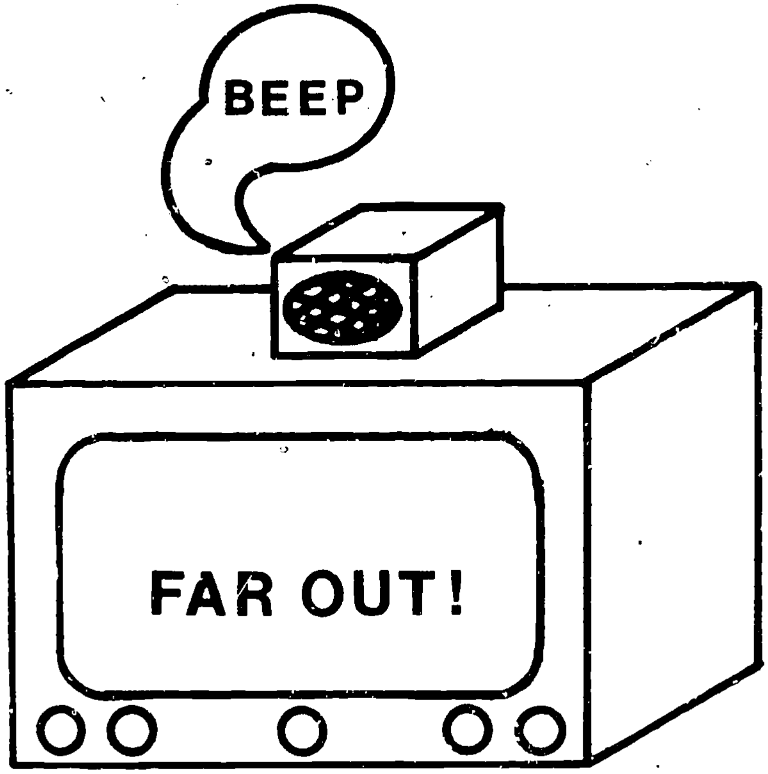


Figure 3. DITRMA configuration.

AGREEMENT



DISAGREEMENT

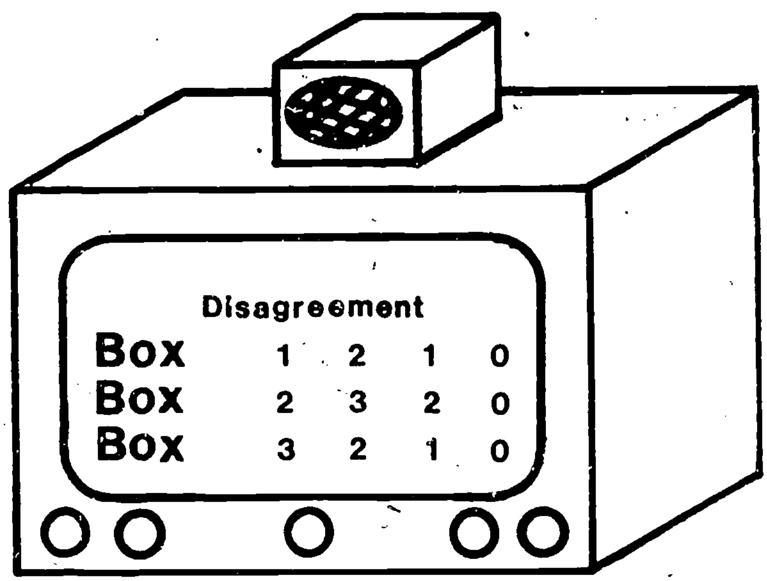


Figure 4. Examples of DITRMA Feedback Display.

If one or more observers disagree, or if the set time interval expires before all entries are completed, several events transpire in a matter of microseconds:

1. The computer instantly stops play of the video tape via a relay buffer to the video tape recorder;
2. A disagreement is entered into the data buffer, and individual entries are stored in memory; and
3. The exact nature of the disagreement is sent to trainees through a closed-circuit feedback monitor showing the data entry from each coding box (see Figure 4). At this point the computer waits for the trainer and trainees to discuss the disagreement and come to a consensus, and it will store the amount of discussion time in memory. Upon reaching an agreement, the trainer enters the consensus code on his master box in order to restart the video tape and continue the coding/feedback process.

In addition to receiving instantaneous feedback on coding skills, each observer can receive delayed, hard-copy computer printouts via DITRMA which summarize agreements and disagreements during the training session. Thus the trainer, as well as the trainee, can readily evaluate observer strengths and weaknesses and use this information in preparation for future training sessions. Unfortunately, delayed DITRMA feedback was not available during observer training in the CATTS Practicum because of other higher CATTS computer-programming priorities.

Observer Training in the CATTS Practicum

An intricate training schedule was necessary in order to accommodate the 36 observer-trainees' available training-time schedules, trainer availability, and free computer time. Training of different groups on both systems of DITRMA occupied as many as 12 hours per day of computer time, forcing TEL computer programmers to work night shifts during this two-week period. The 27 COG-STRAT trainees were divided into six groups of 4 or 5 members each and met for 1 to 1 1/2 hours once a day, 4 days per week. Each group was rotated through available time slots and met approximately equally often with three experienced observation system trainers in order to guard against any specific trainer biases that might develop. MAN-STRAT training was accomplished in a similar manner, although specific scheduling was worked around the less flexible COG-STRAT trainee schedules. (See Table 1.2, Appendix I).

During the first two days of COG-STRAT training, trainees familiarized themselves with the DITRMA system by coding trainer-read examples of classroom situations. The trainees had been instructed to construct these examples as a COG-STRAT familiarization exercise. Trainees also coded video-taped, isolated examples on DITRMA. The next two days were spent coding video-taped classroom simulations. Beginning with the second week of instruction, the trainees were given a criterion test. Upon analysis of this data the CATTS staff was unsatisfied with the level of observer agreement and training was continued for two additional days. At this time a second criterion measure was administered and all trainees reached a desired level of agreement. A similar training procedure was used with the nine MAN-STRAT observers. (See Table 1.3, Appendix I, and Module I).

The total time needed for trainee acquisition of the COG-STRAT system varied from 8 to 10 hours, while MAN-STRAT training took slightly longer. This includes time spent in training as well as time for taking criterion tests and retraining.

Evaluation of DITRMA

DITRMA was subjected to a thorough field test during the observer-training phase of the present CATTS study. Given the large number of trainees required to learn the observation system and the short period of time allotted for training, the DITRMA field test was conducted under less than ideal conditions. Nonetheless, several conclusions seem warranted at this time. First, the ratio of five trainees to one trainer when using DITRMA appeared to be not as efficient as anticipated. All trainers agreed post hoc that, although DITRMA can be used with six persons, a ratio of trainees to trainer of 3 to 1 is ideal for maximal exploitation of DITRMA. Second, it was concluded that the length of the training sessions on DITRMA should be limited to no more than 45 minutes to one hour at a time. Beyond this amount of time individual trainee fatigue seemed to limit the utility of DITRMA for improving group discrimination skills.

Despite the conditions under which DITRMA was used, it seemed to substantially reduce training time relative to traditional paper and pencil observer-training procedures. For example, total training time using the Indiana Behavior Management System-II observer-training package (Fink & Semmel, 1971) is typically 18-20 hours, while MAN-STRAT (a modified version of IBMS-II) training took approximately 10-12 hours with DITRMA.

In addition, the job of the trainer was obviously less burdensome when using DITRMA, since the computer monitored the trainees while coding and stored, retrieved, and analyzed observer-training data. With DITRMA the trainer's role is primarily reduced to that of a discussion leader during disagreement situations, and an ultimate authority for resolving disagreements or incorrect consensus among trainees.

Evaluation of Observer Competencies

Observer competency on the COG-STRAT and MAN-STRAT systems was important for two reasons: First, for trainees to accurately evaluate and change their teaching behavior using feedback from observers via CATTS, this feedback needed to be as precise as possible. Second, accurate coding by observers is a necessary though insufficient condition for experimental evaluation of different types of CATTS feedback.

The issue of observer agreement and reliability of observational records has been discussed extensively elsewhere (see Frick & Semmel, 1974; Medley & Mitzel, 1963; Medley & Norton, 1971; McGaw, Wardrop & Bunda, 1972). Among the major conclusions of these writers is that observers should be trained to nearly perfect agreement with a criterion or expert coder on unambiguous examples of behavioral categories before actual data collection. They should then be expected to agree on unambiguous events encountered in the field. But disagreement on ambiguous events observed in the field should also be expected, since such disagreement may help reflect a more accurate representation of what actually occurred.

Since the amount of ambiguous events occurring in the field cannot be controlled, a measure of observer agreement in that situation is difficult to interpret. Rather, the best that can be done is to document that observers can accurately code unambiguous examples. This can be accomplished by showing observers a video tape containing only unambiguous isolated examples. This type of observer agreement measure is referred to as criterion-related agreement and was used to document that observers know the observation system categories.

In addition to criterion-related agreement, Frick & Semmel (1974) have recommended that measures of intra-observer agreement be obtained by showing a video tape twice to all observers in which conditions parallel those encountered in the field. The purpose of an intra-observer agreement measure was to demonstrate the extent to which each observer can code consistently under actual observational circumstances.

Both criterion-related and intracoder measures were suggested for observer maintenance checks as well. The necessity of calculating agreement coefficients with the same type(s) of data (e.g., category frequencies, two-stage patterns, or scales) that are used in analysis of the actual data collected in the study was also emphasized (Frick & Semmel, 1974).

Thus, as previously stated, the CATTS staff produced a video tape which contained approximately 10 or more isolated examples of each behavioral category for each system, and which also contained a 15-minute simulated classroom lesson to be coded twice in order to make decisions about observer competencies.

It was decided on the basis of recommendations by Frick & Semmel (1974) that simple percentage agreement $\geq .85$ for each category would be required for the criterion-related measure, and an overall proportion of agreement (P_{o_f}) measure $\geq .75$ would also be demanded for the intracoder measure before observers would be allowed to begin classroom coding.

Since data analysis using both category raw frequencies and category percent frequencies as units of analyses was intended, initial criterion-related agreement measures were calculated on both types of data. The formula used for calculating simple percentage agreement for each category for the criterion-related measure was:

$$[1] \quad P_o = \frac{f_1}{f_2} \text{ or } \frac{f_2}{f_1}, \text{ such that } 0 \leq P_o \leq 1.00$$

where f_1 = total frequency (or percent frequency) for the observer

and f_2 = total frequency (or percent frequency) for the expert (criterion).

The formula used for determining an overall proportion of agreement (Flanders, 1967) for each observer on the intracoder agreement check was:

$$[2] \quad P_{o_f} = 1 - \sum_{i=1}^{n} \left(\left| \frac{f_{i_1}}{\sum_{i=1}^{n} f_{i_1}} - \frac{f_{i_2}}{\sum_{i=1}^{n} f_{i_2}} \right| \right)$$

where f_{i_1} = total frequency for first coding for the ith category

and f_{i_2} = total frequency for second coding for the ith category.

The reasons P_{o_f} was used rather than the simple percentage agreement measure [1] were that (a) P_{o_f} is not affected by low frequencies of categories as is P_o; and (b) since observers tend to see more and code more during the second viewing, a proportion of agreement measure such as P_{o_f} seems more appropriate than a simple percentage agreement. See Frick and Semmel (1974).

Since the video tape segment for the intracoder check contained ambiguous as well as unambiguous events, a check against a criterion was considered inappropriate and unnecessary because such a measure would be difficult to interpret (i.e., What is an acceptable level of agreement?), and documentation that observers knew the systems was already demonstrated via the criterion-related measure on solely unambiguous events.

Moreover, π_f (Flanders, 1967) was calculated for the intra-observer check. π_f gives information about observer agreement beyond that indicated by P_{o_f} in that chance agreement and categorical distributions are taken into consideration. That is:

$$[3] \quad \pi_f = \frac{P_{o_f} - P_e}{1 - P_e}$$

where P_{o_f} = proportion of agreement as in [2]

and P_e = chance agreement.

More specifically,

$$P_e = \sum_{i=1}^{n} \left(\left[\left(\frac{f_{i_1}}{\sum_{i=1}^{n} f_{i_1}} + \frac{f_{i_2}}{\sum_{i=1}^{n} f_{i_2}} \right) \Big/ 2 \right]^2 \right)$$

where f_{i_1} and f_{i_2} are the same as in [2]. Thus, π_f is affected by the number of categories in the system and the distribution of categorical frequencies whereas P_{o_f} is not, since P_e tends to increase as the number of categories decreases and/or the distribution of categories becomes more disproportional.

On the other hand, π_f tends to be positively biased when high positively correlated ratings occur (Frick & Semmel, 1974). For this reason the use of π_f for a criterion-related measure for event recording systems such as COG-STRAT and MAN-STRAT, when total category frequencies are the intended unit of analysis, can be misleading and was therefore considered inappropriate.

However, the susceptibility of π_f to correlated pairs of ratings seems to be an advantage rather than a disadvantage of using it for an intra-observer agreement measure, since observers often see more and code more during a second viewing. Thus, it was concluded that P_{o_f} and π_f were the most appropriate measures to use for making decisions about adequacy of observer consistency.

Procedure

On the basis of the preceding rationale, several types of observer agreement checks were administered before and during the study. These will be referred to as initial and maintenance observer agreement checks, respectively. (See Table 1.4, Appendix I).

For the initial checks on the two systems, both criterion-related and intra-observer tests were administered. Observers independently coded the video tapes in groups (n per group $\leq$ 6). Maintenance checks were similarly handled about one-third of the way through the study, except different video tapes were used. It should be noted that a maintenance intracoder measure was not possible with the COG-STRAT group due to severe scheduling problems and lack of trainee time to take such a test.

In addition, "live" maintenance checks were performed throughout the study on both systems in order to give the CATTS staff an indication of observer agreement with experts during actual trainee classroom lessons. This was also done for reasons beyond that of obtaining agreement estimates in situ. Since observers never knew exactly when an expert was going to double-code a given lesson, it was intended that observers would always anticipate such a possibility and come well-prepared to observe each time. Moreover, expert double-coding was almost always done with observers whose skills were found through video and/or live maintenance testing to have deteriorated beyond prespecified levels. Double-coding continued with these problem coders until they obtained satisfactory agreement with the expert. Until adequate agreement was attained, the experts' data were always used for CATTS feedback.

Results: Initial Checks

Following observer training, both criterion-related and intracoder measures were obtained respectively for COG-STRAT and MAN-STRAT observers. When criterion-related checks were first administered, most observers did fairly well, but, with few exceptions, they failed to meet the preset standard of $P_o \geq .85$ for each category. Therefore, more training ensued before the second administration of the criterion-related tests. Results of this second test are given in Table 1.5 for COG-STRAT and Table 1.7 for MAN-STRAT (see Appendix I). It should be noted that each observer code consisted of two stages: Pupil Identification and Category Identification (See Rules for Button Box Coding, Module 1, Appendix). Therefore, in each agreement check both individual pupils and individual categories were considered.

As can be seen from Tables 1.5 and 1.6 in Appendix I, there is little difference on the average between simple percentage agreement measures using raw frequencies and percent frequencies. Thus, in later maintenance criterion-related agreement measures, simple percentage agreement using percent frequencies was considered sufficient.

With few exceptions individual COG-STRAT agreement scores in Table 1.5 are greater than or equal to the preset standard of .85. If a trainee still had problems with one or two categories, he/she was quizzed by a trainer until the trainer was satisfied that the categories were completely understood. Since most trainees had some trouble with category #10 (Incorporation/ Extension), a special large group session was held in which the definition was further clarified and more examples were given for trainees to code.

As can be seen from Table 1.6 in Appendix I, the MAN-STRAT initial criterion-related agreement measures were slightly lower on the average than the corresponding COG-STRAT measures. This was in part due to low frequencies of one category (* 12). In addition, three of the coders were housewives (#7, #8, #9) who had no prior experience with observation systems and who had more difficulty learning the MAN-STRAT than did the remaining experienced CITH personnel. These three were yoked to trainers during the first two weeks of actual classroom coding until each trainer was satisfied that they were competent observers. The trainers' data were always used for CATTS feedback during this time.

In addition to initial criterion-related agreement. initial intra-observer agreement measures were obtained for both systems. As can be seen from Tables 1.7 and 1.8, P_{o_f} and π_f were calculated. All COG-STRAT coders exceeded the preset $P_{o_f} \geq .75$ requirement for both pupil and category identification. In fact, almost all exceeded $P_{o_f} \geq .85$. With one exception, MAN-STRAT observers likewise exceeded the preset standard. It can also be seen from Tables 3.7 and 3.8 that the rank-ordering of π_f's is almost identical to that of P_{o_f}'s, except that π_f's are always lower by several hundredths. P_{e_f} for pupil identification was almost always higher than P_{e_f} for category identification primarily because there were less than half as many pupils as there were categories—i.e., the probability of correctly identifying pupils by chance alone should have been higher than that for category identification. Since an attempt was made for a fairly balanced distribution of categories on the intra-observer video tape segments, the resulting P_{e_f}'s were relatively small and barely exceeded expected probabilities of chance agreement alone (assuming every category was equally likely to occur). Thus, as Frick and Semmel (1974) have concluded, the correction for chance under these circumstances should have had and did have little effect on the rank-ordering of observers, based on P_{o_f}.

Results: Video Maintenance Checks

For MAN-STRAT coders, both criterion-related and intracoder agreement measures were obtained via video tape maintenance tests. The MAN-STRAT intracoder agreement check was administered about one month into the study, and its results are given in Table 1.9 (see Appendix I). Observers #4, #5 and #7 did not meet the $P_{o_f} \geq .75$ standard. Observers #4 and #5 were used only as substitute coders at that time in the study and had been coding very infrequently. Thus, it was decided that #5 and #7 would be yoked to trainers during actual coding sessions, while #4 dropped out of the project on her own volition.

The criterion-related check was administered about 2 1/2 weeks after the intra-observer test. As can be seen from Table 1.10 in Appendix I, observers #1 and #8 did not take this test. Since observer #1 was previously involved with the development of the test, he did not take it. Observer #8 was ill during this time period and could not take the test. She was later yoked to a trainer during actual classroom coding when she recuperated.

Some of the categories in Table 1.10 occurred with low frequency. In particular, categories #2 and #4 and pupil #1 occurred infrequently. Observers #3, #6, and #9 misunderstood a special ground rule for coding category #1 on the 90 isolated unambiguous segments. This ground rule was added only because the test was broken into short segments and was not used during actual coding. Thus, after briefly quizzing these three observers, it was evident that they obviously understood category #1 but were confused about the special ground rule.

It was also evident from Table 1.10 that several observers were experiencing difficulties with a few specific categories. This prompted a general review session in order to resolve any misunderstandings about category definitions that had arisen.

Results of the COG-STRAT criterion-related observer agreement video maintenance check are given in Table 1.11 (see Appendix I). The majority of the 27 observers exceeded the $P_o' \geq .85$ standard, although it was quite evident that about one-third of the observers had developed misunderstandings on several specific categories. These observers were individually quizzed by a trainer on problem categories until he was satisfied that the categories were clearly understood. Moreover, live maintenance checks were begun with each of these observers and continued until the trainers were satisfied with their coding performances. Trainers' data were always used for CATTS feedback during these retraining sessions.

Categories #11 (No Interaction) and #12 (Interruption,Confusion/Management) were not reported in Table 1.11 because they occurred infrequently in the video tape test. This was of little concern, however, since none of the trainees had chosen these categories as teaching skill development objectives.

III. PROJECT ORGANIZATION AND IMPLEMENTATION

The Teacher Education Laboratory Classrooms

Training of teachers was conducted in laboratory classrooms which were designed for the observation, collection and feedback of coded information on trainee performance. The classrooms were linked to the Teacher Education Laboratory (TEL) computer, and each room had observation stations (one-way vision windows) used by observers to collect coded observation system data. There were three such classrooms established in which trainees conducted weekly half-hour language development lessons with small groups of pupils selected from a demonstration EMR class.

Subjects. The subjects were 27 Indiana University undergraduate students, all special education majors enrolled in K490; Curriculum and Methods for Educable Mentally Retarded. The study and practicum were administered through weekly meetings of the K490 class. The 27 trainees met in these sessions for instructions, lectures, scheduling and communication or airing of problems encountered during the practicum.

Grouping pupils for instruction. Each practicum student was assigned to teach the same group of children, once each week for the duration of the study. There were four to five children in each group and there were three such groups. These groups were drawn from a class of EMR children whose regular classroom was located in the TEL facility. The IQ range of the children in this class was 59 to 84. Age, sex and reading level data are shown in Table 2. The instructional groups were designated x, y, z, and the assignment of children to groups was permanent for the duration of the study.

The assignment of children to instructional groups was made by the classroom teacher in consultation with the Educational Director of the Developmental Training Center, Indiana University, whose responsibility included the TEL

Table 2
Assignment of Pupils for Instruction
Characteristics of Pupils in Each Group

Pupil Group X

Pupil No.	Sex	CA (Jan. 1973)*	Reading Level
01	F	7-	Pre-primer
02	M	8-	Primer
03	M	9-	Pre-primer
04	M	11-	Pre-primer

Pupil Group Y

Pupil No.	Sex	CA (Jan. 1973)	Reading Level
01	M	7-8	Readiness
02	M	7-	Readiness
03	M	7-6	Readiness
04	M	7-	Readiness

Pupil Group Z

Pupil No.	Sex	CA (Jan. 1973)	Reading Level
01	M	10-	2.[illegible]
02	F	9-	1.[illegible]
03	F	10-	2.2
04	F	7-2	1.8
05	M	11-	1.8

*In some instances the CA's are approximate, but since the basis for grouping was primarily reading level, the CA is presented for informational purposes only.

laboratory class. The grouping of pupils for instruction by CATTS trainees was based upon the similarity of instructional needs and the social maturity of the pupil.

Each group of pupils received a total of one hour of supplementary reading instruction by CATTS trainees per day. The same group of children was taught by one trainee for half an hour and by another trainee for half an hour. The three instructional groups (x, y, and z) were taught simultaneously in three different classrooms. See the weekly schedule of classes (Appendix I) for the pattern of randomization of classrooms and day and hour of teaching. From this schedule, it can be seen that each group of children received instruction from nine different trainees during any one week of the study. Each trainee always taught the same group of children, that is, a trainee always taught either x, y, or z children.

Trainee lesson planning. Self-instructional lesson-planning information was made available to the trainees in Module No. 2. The trainees were required to submit a written lesson plan for each lesson they taught. The instructional objectives for each lesson were given to the trainees two weeks in advance of the scheduled teaching time. Each trainee submitted a written lesson plan one week prior to teaching, for comment and evaluation. The lesson plans were read and evaluated by the staff member who wrote the objectives, and returned to the student several days in advance of the lesson. The module specified the items to be included in preparation of lesson plans. These items were prepared from guidelines to the course on instructional design for special education, which all trainees had taken the previous semester. Included in the checklist for writing lesson plans were suggestions on entry tests, task analysis, specification of objectives in performance terms, criterion tests for objectives and subobjectives, and teaching strategies.

In addition, several members of the staff, including the classroom teacher, were available for consultation to assist the trainees in analyzing and using the lesson objectives for preparing lesson plans, and adjusting the level of instruction to the competence of the children in their groups.

Curriculum. The curriculum selected to serve as the basis for the preparation of lesson objectives used by trainees for lesson planning was the series, Language Experiences in Reading (Van Allen and Van Allen, 1970). This series was adopted and modified specifically for the CATTS practicum to provide the basis for a supplementary reading program for the pupils in the laboratory class. The procedure for developing the objectives for each lesson included a task analysis of activities suggested in the Teachers Resource Book for Language Experiences in Reading, (Van Allen and Van Allen, 1970). Contemporary English in the Elementary School, (Tiedt and Tiedt, 1967) was also used as a resource, particularily those chapters with suggestions for supplementing and extending concepts and task. Instructional objectives that student teachers used for lesson plan preparation were stated in behavioral terms. Expected pupil terminal behaviors, which indicate the pupils attainment of a concept or ability to perform a task, were the basis for a given objective.

Every attempt was made to assure the continuity of lesson objectives for each instructional group. To this end, modifications in the long-range curriculum plan were made as the need arose. This process was monitored by a staff member responsible for curriculum development and by the classroom teacher. Lesson objectives were designed for appropriateness to the instructional level of each group, and a specific objective was written for each lesson. A complete set of daily lesson objectives for each group is found in the trainee's Instructional Module.

Implementation Procedures: Teaching and Coding

Table 3 indicates the main operations of the project, which were constrained by the multiple nature of the training and research objectives and also by temporal considerations. In place concurrently were the necessary computer software, training manuals for observer training, instructional modules for lesson-planning skills, instructional materials on the use of CATTS feedback and performance evaluation, and teaching and observation-coding schedules.

The preservice trainees in the program were required to plan lessons, to teach one half hour per week, to evaluate their teaching performance based upon CATTS feedback, to develop individual goals for modification of teaching performance as indicated by CATTS data, and to reliably code other trainees engaged in teaching. As described in Chapter II, a block of time was allocated for the training of the teachers to become reliable coders of MAN-STRAT/ COG-STRAT Observation Systems.

Baseline period; daily procedures. Following the coder-training phase of the program, each trainee received an individual schedule for teaching and coding, and also instructional objectives for preparing the next lesson. On the assigned day of teaching, the trainee arrived at the TEL and signed in. Following information posted in the laboratory, the trainee met the children in his/her group, and went to an assigned room. Trainees scheduled for coding followed posted information on specific observation station and button box assignments. These schedules were designed to counter-balance classroom and coder effects.

Prior to teaching, most trainees began by setting rules, establishing routines or motivation for the lesson. As soon as teaching started, coders began recording categories of behavior on the coding boxes. The coding

Table 3

CATTS Project Operations and Time Line for Development and Implementation

OPERATION	DECEMBER	JANUARY	FEBRUARY	MARCH	APRIL	MAY	JUNE
Hardware/Software Development		DITRMA Ready	CATTS Feedback Ready	CATTS Feedback to Trainees	CATTS Feedback to Trainees	CATTS Data Transmission for Analysis	
Observation System Development & Discrimination Training	Selection of COG-STRAT MAN-STRAT Categories. Development of Training Manuals & Materials →	Observer Training to Criterion. (Module 1) ←	Begin ← Observation Coding →	Conduct Reliability Checks →	→		
Curriculum & Lesson Planning	Group Pupils. Select Curriculum.	Develop Lesson Objectives & Sequence (Module 2: Lesson Planning).	Staff Support System → Consultation on Lesson Planning, Selection of Performance Objectives.		→		
Teaching 1. Baseline Observation Period 2. Teaching with Feedback (Generation Phase).		Schedule Teachers and Observers	Start Baseline → -Selection of Teaching Objectives -(Module 3 & 4) Begin Graphing & Interpreting Behaviors.	← CATTS F.B. (Treatment) →	→		
Evaluation			Formative Evaluation of Baseline Teaching, CATTS Post-Teaching F.B. Questionnaire → Trainee Evaluation of Teaching Performance (Graphing) →	Trainee Evaluation of Practicum (Group Meetings)	→	Final Trainee Evaluation of Practicum	Data Analysis

continued for thirty minutes, after which the computer automatically shut off the coding terminals. Even in those instances where teachers continued the lesson, only the first thirty minutes of coded data was taken. In those cases where the teacher trainee terminated the lesson before 30 minutes, the coders ceased recording when it was clear that the lesson was over. One half hour after teaching, the trainee took his place at the observation station and coded another trainee's lesson on the COG-STRAT System. In this manner, trainees reversed the coder-teacher roles.

Depending on a preset schedule, which was in turn based upon the research design, each trainee taught at least three and up to six times, without receiving any feedback on teaching. This period of teaching without feedback resulted in the establishment of a baseline of individual teaching behaviors. All baseline data transmitted by the coders on the COG-STRAT and MAN-STRAT systems were retrieved, printed out, and copies were distributed to the trainees before feedback treatment began.

Feedback phase. Prior to the first teaching session with CATTS feedback, each trainee received a copy of the COG-STRAT and MAN-STRAT printout for each of the baseline lessons taught. Immediately after the trainees' final baseline teaching session, CATTS staff members instructed the trainee in the interpretation of the printout information. This was supplemented with Module No. 3, which provided the same information in a self-instructional format.

Graphing baseline behavior. The trainees were instructed in procedures for using the printout information to prepare graphs showing the percent frequency of occurrence of behaviors for each of the observation system categories. They were also assigned the task of graphing all categories in the baseline period and required to evaluate their teaching performance on the basis of their graphs and to also select three behavioral categories that they judged as most important to their teaching. They were then instructed to consider their baseline performance on the three preferred categories and choose one category from among the three that had shown a relatively stable pattern from lesson to lesson during the baseline period. This choice became, in most cases, the category that was targeted for improvement in subsequent lessons. The trainee also had to indicate whether s/he wanted to accelerate or decelerate the percent frequency of occurrence of the target behavioral category in subsequent lessons. The graphing assignment and choice of teaching behavior category was completed prior to the first lesson taught with feedback.

All trainees received the graphing-printout instructional module (No. 3) after their final baseline teaching sessions. This was used to supplement the information given to the trainees in small group instructional sessions held immediately after the last baseline lesson. These sessions were held so that each trainee would know how to interpret the baseline data and make an informed decision on which category of behavior to choose for improvement.

VTR feedback. Half of the trainees received instruction on interpreting the video scope feedback. Module 4 (Appendix) was developed for self-instruction in video scope interpretation. The trainees assigned to receive video scope feedback were also given video scope instruction by a CATTS staff member in the TEL, which included viewing a simulated data display.

Instruction in interpreting and using CATTS data, either from printout or from printout and video display, always occurred between the trainees' last baseline lesson and first lesson with feedback.

During the feedback stage, the weekly lessons continued as during baseline and, in addition, all trainees received printouts on their latest lesson within two hours after teaching. The group receiving video feedback always had in their classroom a video monitor which showed the cumulative percent frequency of occurrence of the targeted behavioral category that the trainee had chosen for teaching behavior change. In addition, a video monitor was placed in the rooms of teachers who received printout feedback only. In these rooms, the VTR showed the same graph as the VTR feedback group had but with the baseline arbitrarily set at about 20% and a continuous point of light flashing on the x axis at short time intervals. The cumulative points constituted a straight line on the ordinate over the 30-minute period. Figure 5 shows VTR display for scope feedback and printout-only feedback groups.

Focusing the trainees' attention on the targeted behavioral category. Several steps were taken to assure that the trainees would attend to the category chosen for feedback and behavioral change. That is, measures were taken to focus the trainees' attention to the skill development task that was imposed as the primary goal of the practicum experience. The printout-only group was required to graph the percent frequency for the targeted category after the printout was received. This data was recorded in a cumulative, individual graph for the category and included the trainees' baseline percentages In addition, all trainees kept a running record of the percent frequency in category for each successive lesson by entering the information for the previous lesson on a graph in the laboratory. The trainee did this immediately before joining the pupils and beginning to teach. This, in effect, was a redundant activity since the trainee had already entered the percent frequency information in his own record book after teaching, and then repeated the same entry in another book immediately before entering the classroom to teach another lesson.

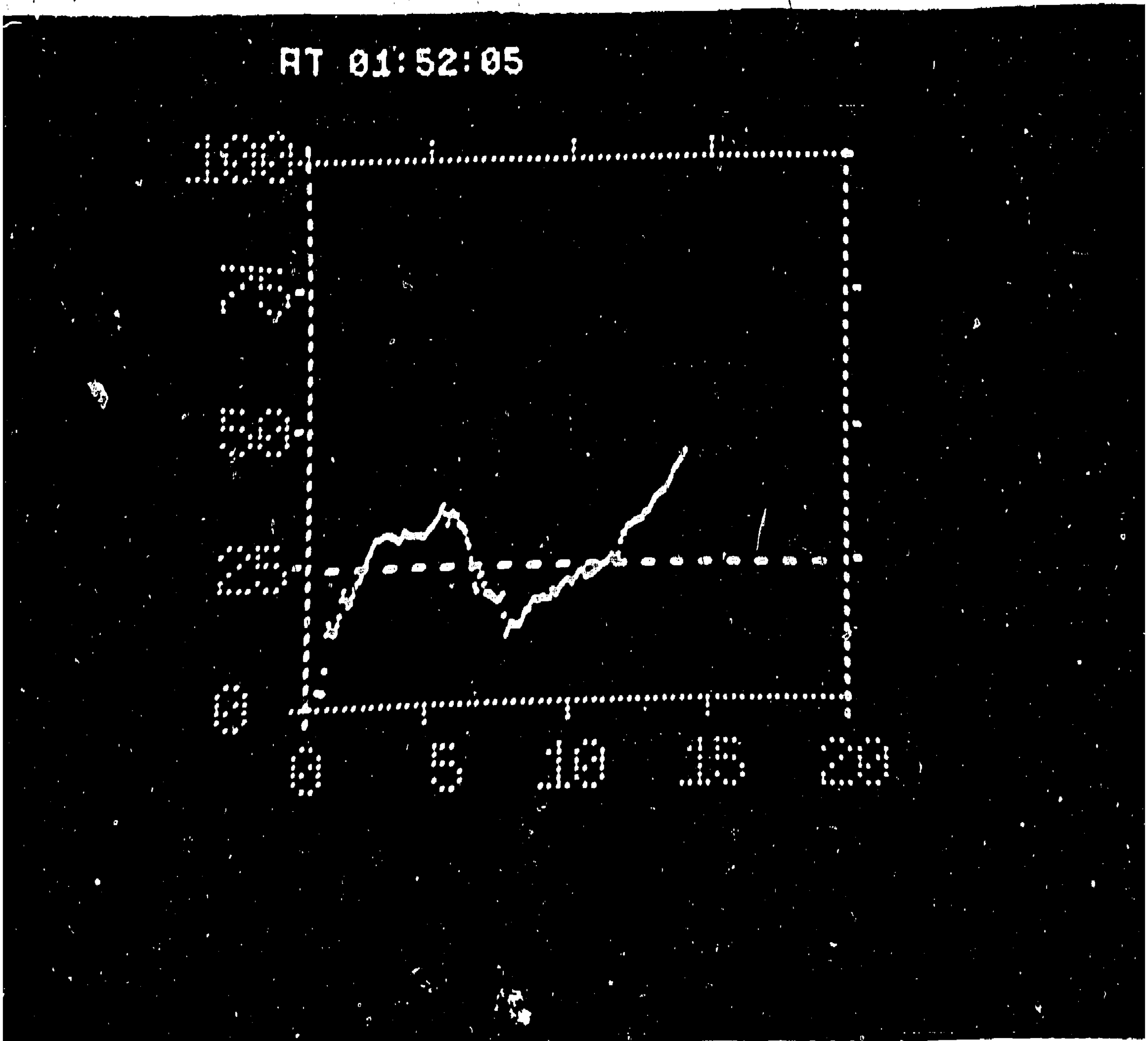


Figure 5. VTR display for scope feedback group and printout group.

The video feedback group was shown a post-teaching questionnaire during the scope training period and informed that it would be administered after each lesson they taught. This information was reiterated in Module 4, CATTS Video Feedback. It was expected that the questionnaire would serve to reinforce the trainees attention to the scope during teaching. Among items in the questionnaire where attention to scope feedback was requisite, was question No. 2, in which the trainee was required to estimate the number of times he looked at the monitor. Question No. 3 required the trainee to reconstruct the video feedback display for the entire 30-minute lesson (See questionnaire in Module 4).

The number of times a trainee looked at the feedback monitor was also independently tracked by a third observer in the coding booth, as a cross-validation measure.

Scheduling of Teachers and Observers

Conditions for the preparation of the schedule for the CATTS teaching sessions were determined by the constraints of the statistical design of the study and required appropriation of the following variables:

1. Assignment of trainees to pupil groups.
2. Assignment of trainees to feedback conditions.
3. Rotation of room assignments.
4. Assignments of trainees to day of teaching.

Assignment of trainees to pupil groups. The trainees were assigned to pupil groups (x, y or z) on a random basis. Table 4 shows the three pupil groups and the identification number of the trainees assigned to teach each group.

Assignment of trainees to feedback conditions. There were three controlled conditions under which the trainees taught. These were: (1) the baseline, (2) the printout feedback only, and (3) the scope and printout

Table 4
Assignment of Trainees to Pupil Groups

Pupil Group	X	Y	Z
Teacher Trainee ID Number	14	05	20
	01	26	02
	06	09	12
	24	03	21
	10	18	22
	04	16	23
	19	07	17
	15	11	28
	1[illegible]	08	25

feedback condition. In the baseline condition, all trainees taught the lessons they had prepared without receiving feedback of any kind. The number of lessons taught without feedback varied from trainee to trainee. The baseline teaching always included at least the first three lessons and varied up to the first six lessons taught.

The second controlled condition was teaching after the trainee received computer printout information about his/her baseline teaching performance and a set of instructions (Module No. 3) on how to read the computer print-out. At this stage the trainees were also required to evaluate their base-line teaching and to select by listing in order of personal preference three of the observation system categories they wished to work on for improving their teaching performance. The instructions for choosing a behavioral category for teaching performance improvement was included in Module No. 3. Within a week of receiving baseline information, the trainees were required to select a single behavioral category for subsequent improvement. The final selection was arrived at in consultation with a staff member. All of these activities were carried out after the baseline period and before the trainee taught with feedback.

In addition to these procedures, the trainees were randomly assigned to feedback conditions. Half the trainees were assigned to a printout-only feedback group, and half to a scope and printout feedback group. The scope feedback group was presented with a self-instructional module (Module No. 4) and also received instructions from a staff member on how to use CATTS video scope feedback. This step was performed prior to the trainees teaching with feed-back. Table 5 illustrates the variable baseline and the two types of feed-back conditions.

Table 5
Schedule of Baseline,
CATTS Instruction and Feedback Conditions

Week	1	2	3	4	5	6	7	8	9	10
Design										
1	A	A	AX	B	B	B	B	B	B	B
2	A	A	AXY	C	C	C	C	C	C	C
3	A	A	A	AX	B	B	B	B	B	B
4	A	A	A	AXY	C	C	C	C	C	C
5	A	A	A	A	AX	B	B	B	B	B
6	A	A	A	A	AXY	C	C	C	C	C
7	A	A	A	A	A	AX	B	B	B	B
8	A	A	A	A	A	AXY	C	C	C	C

Conditions:

A = Baseline, no feedback

B = Printout feedback only

C = Video scope and printout feedback

X = Instruction on CATTS printout and graphing

Y = Instruction on video scope use

Assignment of trainees to baseline and feedback conditions was random. Table 6 shows assignment of trainees to feedback conditions and the intervention of instruction on the use of CATTS printout and graphing, as well as instruction on the use of the video scope feedback.

Coder scheduling. The scheduling of COG-STRAT coders, all of whom were trainees, depended upon the trainees' teaching assignment. For example, if a trainee taught during the first half hour, she/he coded during the second half hour. The trainee who coded the first half hour, taught during the second half hour. The MAN-STRAT coders were either staff members or trained MAN-STRAT coders employed for the project. Assignment of trainees to day of teaching was random. Coders were also randomly assigned to teachers such that each teacher was observed by a different coder in her/his group each time. Teaching room assignment was likewise rotated between three rooms in the TEL facility. An assignment schedule was prepared in advance of actual teaching and distributed to each trainee and coder as well as to the staff. The master schedule for the entire project was posted in the TEL so that the day, hour and conditions (F.B., no F.B.) were known in advance to coders, trainees and staff. From time to time, adjustments in schedules were necessary due to absences. These were made up under conditions which matched the original assignment as closely as possible.

Control procedures. The teaching and coding schedule was prepared in advance of the first teaching session, and a few adjustments were made to accommodate some conflicts in assignment with other personal committments. Once the schedule was finalized, a master schedule was posted in the TEL and copies were given to each trainee. All persons participating in a teaching session signed in before beginning to teach or code.

Table 6
Assignment of Trainees to Feedback Conditions

Week	Feedback Condition 1	2	3	4	5	6	7	8	9	10
Trainee ID Number										
01	A	A	A	A	AX	B	B	B	B	B
02	A	A	A	AX	B	B	B	B	B	B
03	A	A	A	A	AXY	C	C	C	C	C
04	A	A	A	A	AX	B	B	B	B	B
05	A	A	A	A	AX	B	B	B	B	B
06	A	A	A	A	A	AXY	C	C	C	C
07	A	A	A	AXY	C	C	C	C	C	C
08	A	A	A	A	A	AX	B	B	B	B
09	A	A	AXY	C	C	C	C	C	C	C
10	A	A	AXY	C	C	C	C	C	C	C
11	A	A	A	A	A	AXY	C	C	C	C
12	A	A	A	A	A	AXY	C	C	C	C
13	A	A	AX	B	B	B	B	B	B	B
14	A	A	A	A	A	AX	B	B	B	B
15	A	A	A	A	AXY	C	C	C	C	C
16	A	A	A	A	AXY	C	C	C	C	C
17	A	A	A	A	AXY	C	C	C	C	C
18	A	A	A	AX	B	B	B	B	B	B
19	A	A	A	A	A	AX	B	B	B	B
20	A	A	AXY	C	C	C	C	C	C	C
21	A	A	A	A	AXY	C	C	C	C	C
22	A	A	AX	B	B	B	B	B	B	B
23	A	A	A	A	A	AX	B	B	B	B
24	A	A	A	AXY	C	C	C	C	C	C
25	A	A	A	A	A	AXY	C	C	C	C
26	A	A	AX	B	B	B	B	B	B	B
27	A	A	A	A	-	-	-	-	-	-
28	A	A	A	A	A	AX	B	B	B	B

Conditions:

A = Baseline, no feedback
B = Printout feedback only
C = Video scope and printout feedback
X = Instruction on CATTS printout and graphing
Y = Instruction on video scope use

The Study of the Effects of CATTS Feedback on Trainee Behaviors

The empirical questions addressed in the present study were as follows: 1) What is the effect of providing CATTS instant (video scope) feedback displaying trainees' performance in a target behavioral category, and also providing post-teaching printout summaries of coded teaching behaviors, upon the rate of generation of targeted behavior as compared to the base rate of the behavior? 2) What is the effect of receiving only post-teaching feedback (printout summaries) upon the rate of generation of the targeted behavior, compared to the base rate of behavior? 3) What is the relative effectiveness of CATTS instant and post-teaching printout feedback, compared to post-teaching printout only? 4) What are the effects of targeting and feedback of behaviors on the rate of generation, compared to rates generated in categories of behavior not targeted for feedback?

IV. RESULTS AND DISCUSSION

The statistical design used to analyze the CATTS project data was a repeated measures ANOVA design with two between-block factors--Group (G) and Treatments (T)--and one within-block factor, Periods (P). Groups refer to the three reading groups (X, Y, Z) to which the classroom children were assigned. Eight teacher trainees were then randomly assigned to each reading group, totaling 24 subjects overall. Treatments refer to the two types of feedback treatments randomly assigned to each half of the subjects within each group; T_1 refers to the group of subjects who received the hard-copy printout only, and T_2 refers to the group who received the instantaneous scope display in addition to the printout. Periods refer to the baseline and treatment periods of the study. P_1 refers to the average of the baseline teaching trials and P_2 refers to the average of the treatment teaching trials for each subject. This brings the total number of observations for the analysis to 48. The original design contained a fourth factor, different variations of baseline and treatment trial combinations, but due to inconsistent teaching schedules and some missing observations, this baseline/treatment factor was not fully completed. A preliminary analysis, including an incomplete version of this fourth factor, showed no difference at all between the various baseline/treatment combinations. The decision was then made to collapse each subject's trials within the various baseline and treatment periods and calculate one average score per period.

The rate of responding on the trainees' chosen feedback category was selected for the dependent measure employed in the analyses. This criterion measure was calculated by dividing the frequency of occurrence of the feedback category during any one session by the time of that session. This measure was selected because it was considered more stable than frequency or percentage of frequency, due to the varying length of some teaching sessions. The rate measure also provides a method of combining under one measure the two orthogonal components of frequency and time.

Table 7 contains the analysis of variance source tables for the feedback category as well as the second and third choice categories. F ratios for the feedback category indicate that there were no significant differences in the Group (G) and Treatment (T) main effects, while there was a significant finding in the Period (P) main effect ($p<.01$). An examination of the means in Table 8 and Figure 6 reveal that feedback trials were significantly higher than baseline trials. In addition, the Treatment by Period interaction (TP) was also significant ($p<.05$). All other interactions were statistically nonsignificant.

Due to the significant TP interaction, a qualification of the main effect of T and P is required. Figure 6 illustrates the TP interaction plot of the mean scores from Table 8 on the criterion feedback category. A simple main effects analysis (Kirk, 1968) was performed on T and P in order to qualify these main effects. The results of the simple effects analysis are also indicated in Table 8.

As expected, the qualification of the Treatment effect (T) indicated that the scope and printout feedback condition (T_2) and the printout—only feedback condition (T_1) were not significantly different during baseline trials (P_1), but did reach significance ($p<.05$) during the treatment trials (P_2). The Period effect (P) is qualified by showing that the printout—only treatment group (T_1) did not significantly increase their mean rate of responding between baseline (P_1) and treatment (P_2) trials, while the scope display-printout treatment group (T_2) increased significantly ($p<.01$) between baseline and treatment trials.

Table 7

Analysis of Variance Table for Feedback and Second and Third Choice Categories

SOURCE	FEEDBACK CATEGORY				SECOND CHOICE CATEGORY				THIRD CHOICE CATEGORY			
	SS	df	MS	F	SS	df	MS	F	SS	df	MS	F
Between Subjects	1.5082	23			7.9394	23			6.2206			
Groups (G)	.0711	2	.0356	<1	.7119	2	.3559	<1	1.0766	2	.5383	2.34
Treatments (T)	.1728	1	.1728	2.79	.2806	1	.2806	<1	.1055	1	.1055	<1
Btwn T at P_1	.0000	1	.0000	<1								
Btwn T at P_2	.3455	1	.3455	7.15*								
Within cell	.5796	12	.0483									
GT	.1476	2	.0738	1.19	.2739	2	.1369	<1	.8933	2	.4496	1.95
S w/ groups S(GT)	1.1167	18	.0620		6.6730	18	.3707		4.1452	18	.2303	
Within Subjects	1.6212	24			.5104	24			.2725			
Periods (P)	.7301	1	.7301	21.10**	.0054	1	.0054	<1	.0105	1	.0105	<1
Btwn P at T_1	.0962	1	.0962	2.78								
Btwn P at T_2	.8066	1	.8066	23.31**								
GP	.0852	2	.0426	1.23	.0024	2	.0012	<1	.0065	2	.0033	<1
TP	.1728	1	.1728	4.99*	.0609	1	.0609	2.67	.0137	1	.0137	1.02
GTP	.0096	2	.0048	<1	.0306	2	.0153	<1	.0014	2	.0007	<1
PxS w/ groups SP(GT)	.6235	18	.0346		.4111	18	.0228		.2404	18	.0134	
Total	3.1294	47			8.4498	47			6.4931	47		

* $P < .05$

** $P < .01$

Table 8
Mean Rate Performance on Categories and Treatment Phases

			P_1 BASELINE	P_2 TREATMENT	INCREASE FROM BASELINE
SELECTED FEEDBACK CATEGORY	Printout only	T_1	.1350	.2716	.1267
	Scope & printout	T_2	.1350	.5017	.3667
	$\bar{X}T_1$ & T_2		.1350	.3867	.2467
SECOND CHOICE	Printout only	T_1	.2675	.2175	-.0500
	Scope & printout	T_2	.3492	.4417	.0925
	$\bar{X}T_1$ & T_2		.3083	.3296	.0213
THIRD CHOICE	Printout only	T_1	.2908	.3542	.0634
	Scope & printout	T_2	.2308	.2267	-.0041
	$\bar{X}T_1$ & T_2		.2608	.2904	.0296

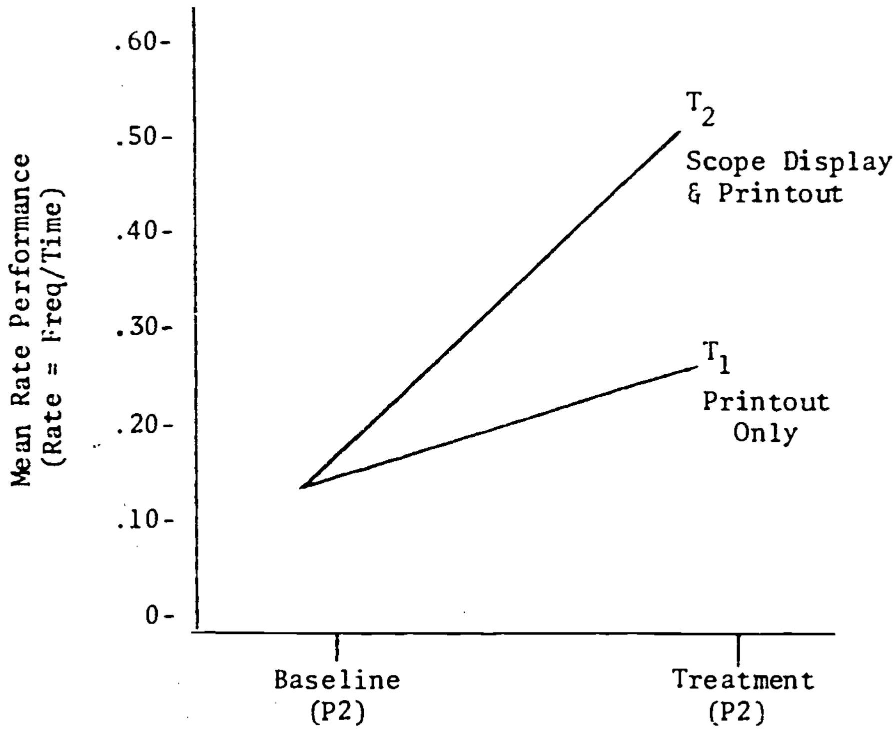


Figure 6. Mean rate performance by treatment across baseline and treatment trials (TP interaction).

In order to further support the simple main effects interpretation of the significant TP interaction, a Tukey post hoc analysis was performed on the means which were plotted in Figure 6. The same results were four. as reported above with the same alpha level of significance.

As indicated previously, prior to the trainees' selection of their criterion feedback category, they initially chose three cognitive teaching behaviors that they specified they would like to work on in the classroom. From this list of three, the final feedback category was selected. As a contrast to the selected feedback category, identical analyses were performed on the second and third choice behaviors that were not included in the concentrated CATTS feedback program. These second and third choice behaviors were not displayed on the scope instantaneous feedback display, but were included on the printout given to both treatment groups together with all of the other behavior categories coded in the classroom.

An examination of Table 7, Variance Components for Second and Third Choice Categories, revealed no significant main effect differences or interactions present in the data. Hence, only the criterion feedback category selected from the original chosen three categories was found to improve significantly from baseline to treatment. Figure 7 (a & b) graphically illustrates the mean rate increase from baseline and performance differences from baseline between the three feedback choice categories. Figure 8 further illustrates the differential effect of the two treatment feedback conditions, T_1 and T_2, across all three categories.

F max tests (Kirk, 1968) were performed on the error terms for all three categories found in Table 7 in order to test whether or not the assumption of homogeneity of the partitioned parts of the within-cell variation is tenable. F max ratios indicated that both the within-subjects and between-subjects variances error terms for the feedback category were found to be

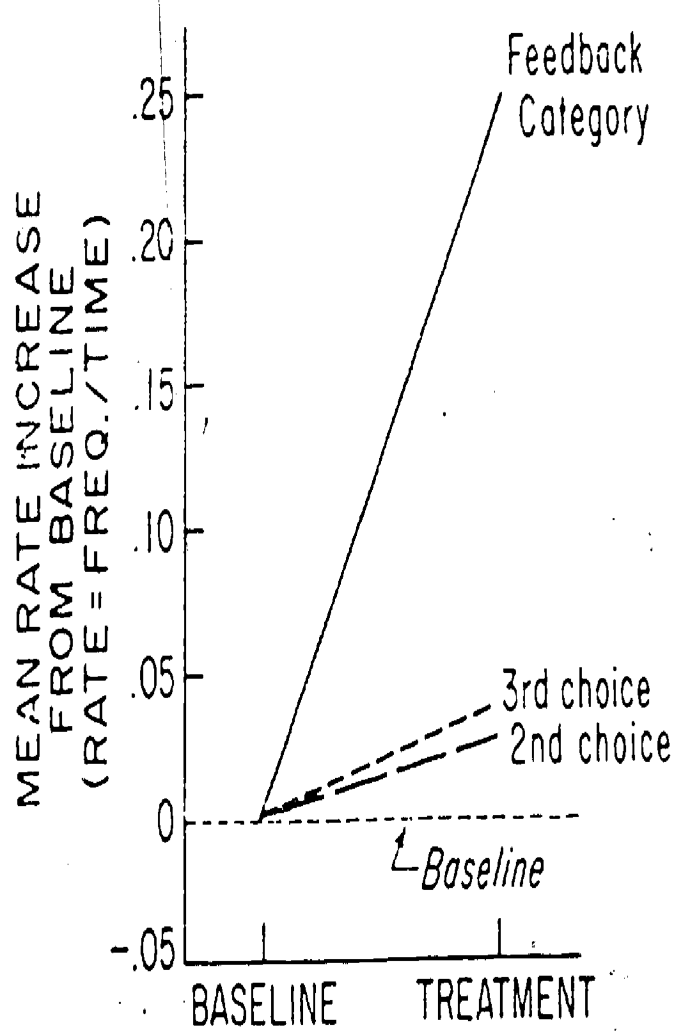


Figure 7a. Mean rate increase for feedback category and second and third choice categories.

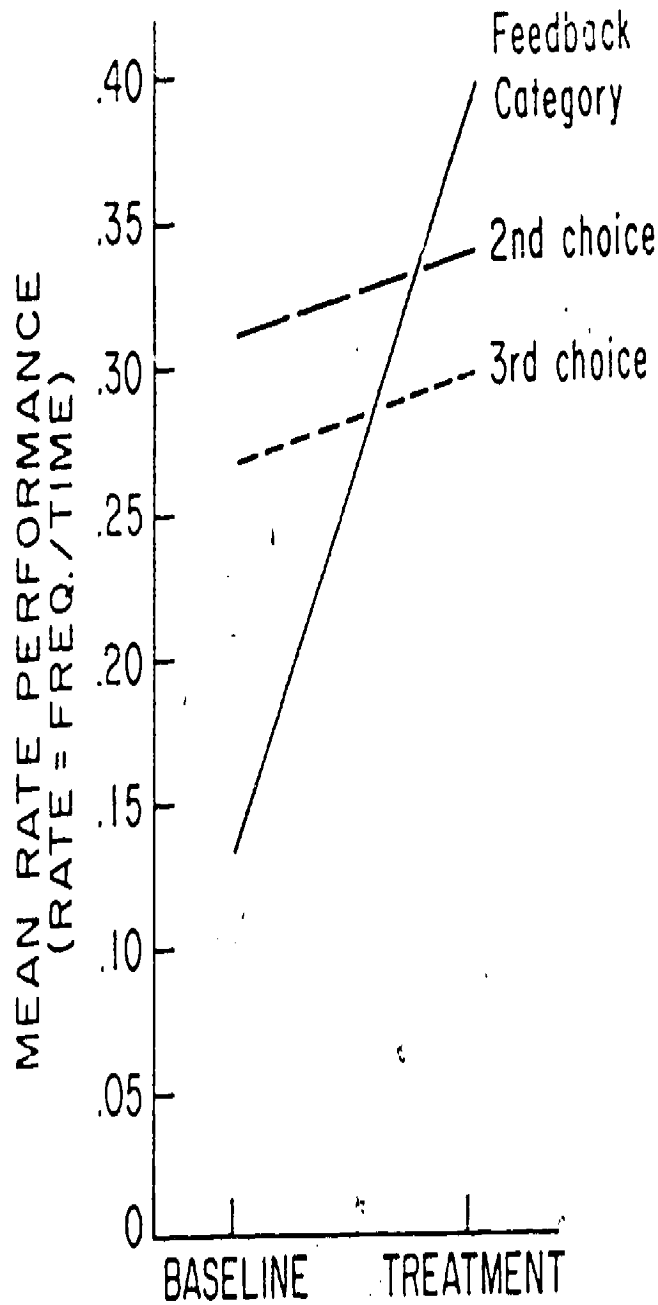


Figure 7b. Mean performance rate for feedback category and second and third choice categories.

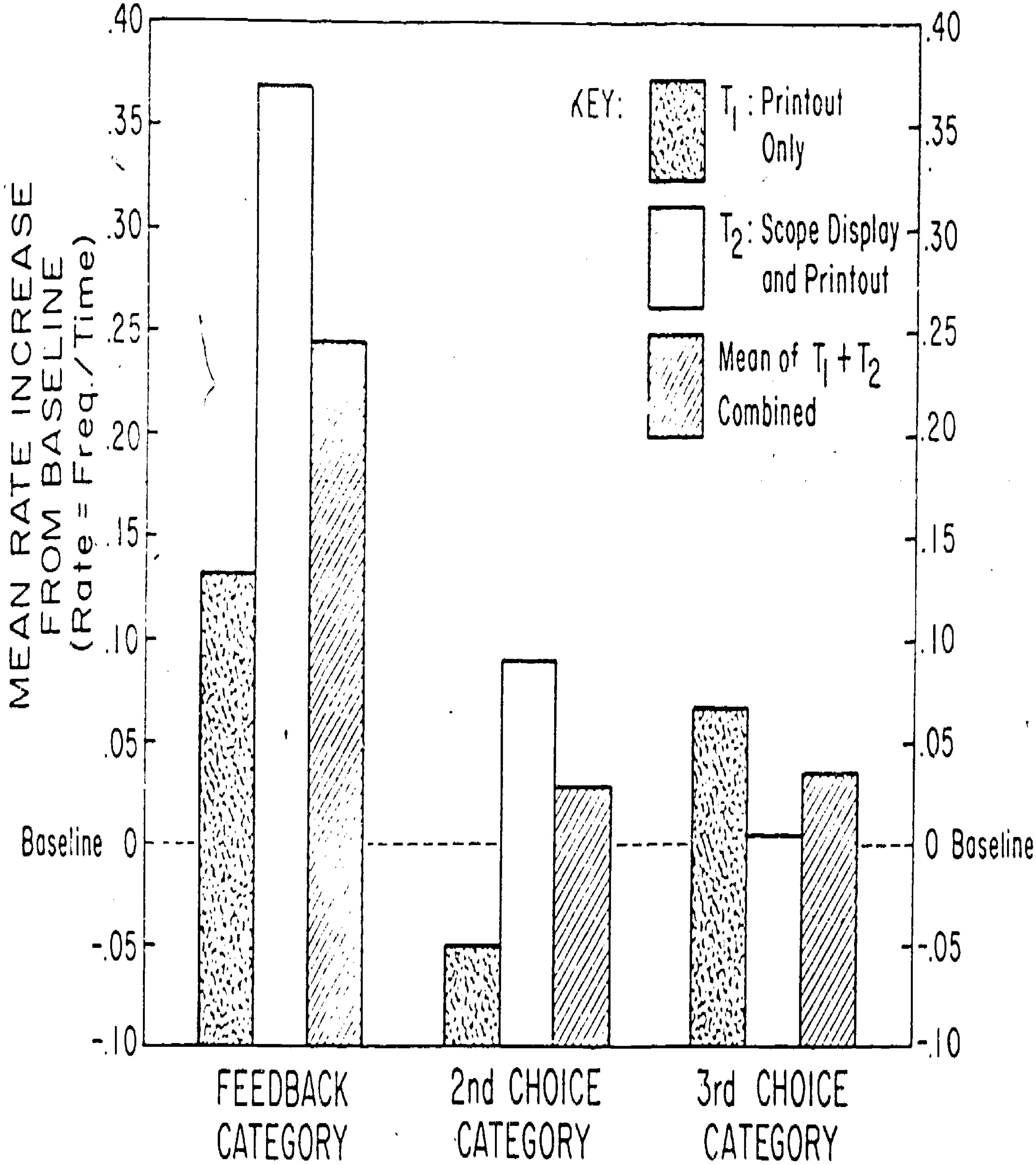


Figure 8. Mean treatment and category comparisons.

made up from homogeneous within-cell components. The within-subjects error term for the second choice category was also found to be homogeneous. The F max ratios for the remaining error terms rejected the homogeneity assumption. The application of this test qualifies the interpretation of the second and third choice categories, but lends support to the results attributed by CATTS to the chosen feedback category.

Figures 9 (a, b, c, & d) are plots of the groups of trainees who were included in the initial four types of baseline/treatment combinations. Most of the missing observations for the trainees occurred during the eighth, ninth, and tenth treatment trials. As a result, the projected points towards the end of the teaching periods reflect considerably reduced n's, and therefore are unstable. Missing observations also occurred during the initial baseline trials but were considerably fewer in number, so the stability of the baseline trials is fairly consistent with n's of 6 and 7.

Although 25 subjects were used in the analysis, there was data available from 27 trainees. Three of the trainees had identical baseline/treatment patterns to three other trainees, and the decision was made to combine those three pairs into three separate mean scores and provide equal cell sizes for analysis.

The results of the study revealed that all trainees in both treatment conditions significantly increased their rate of criterion performance by a ratio of 2.8:1 over the baseline, as a function of CATTS feedback. There were no differences between the two CATTS feedback groups during baseline on their individually chosen criterion measure. However, during treatment trainees in the CATTS scope and printout condition increased their criterion rate of responding to a significantly greater degree (2:1 ratio) than did trainees in the delayed printout-only condition.

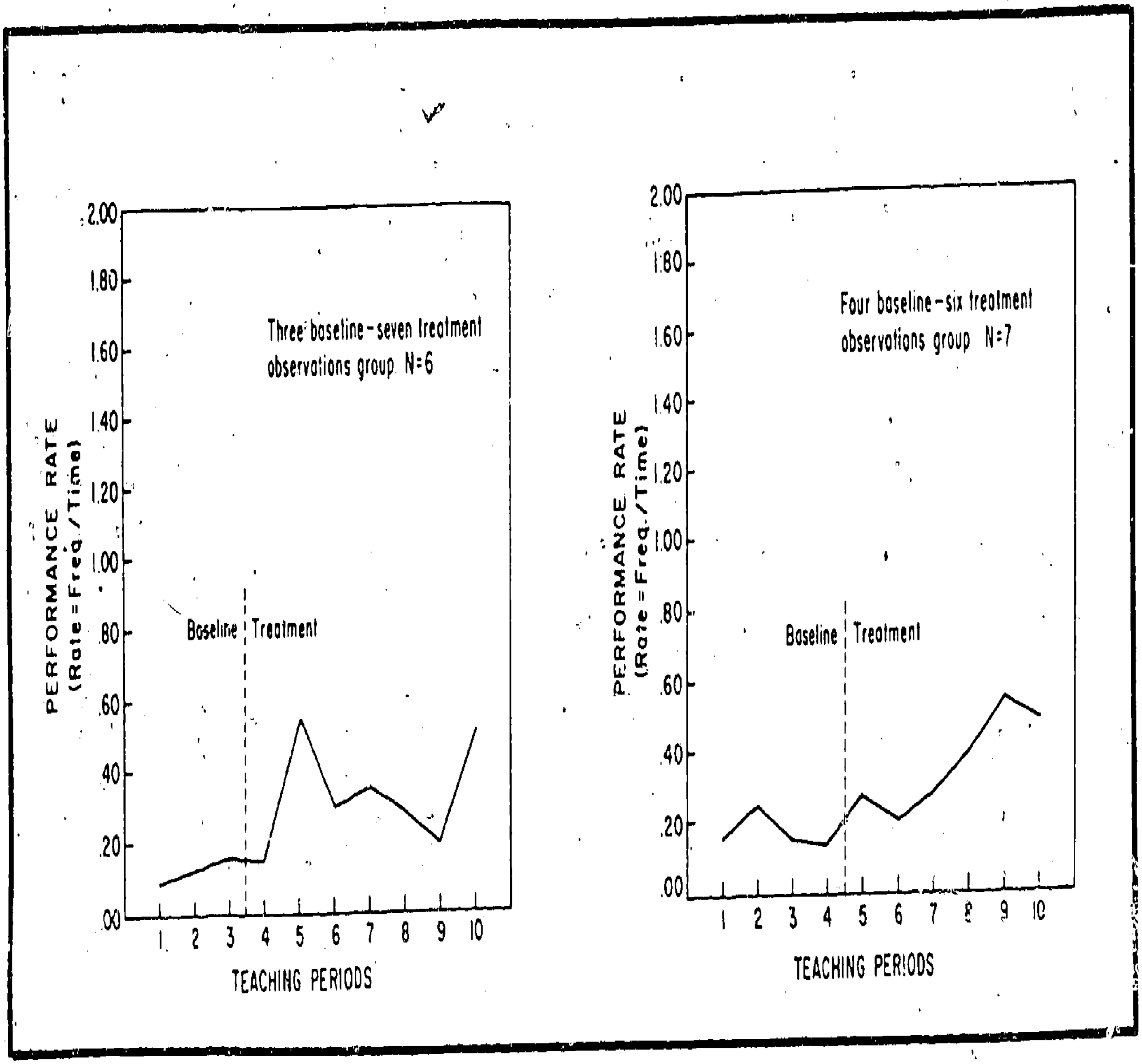
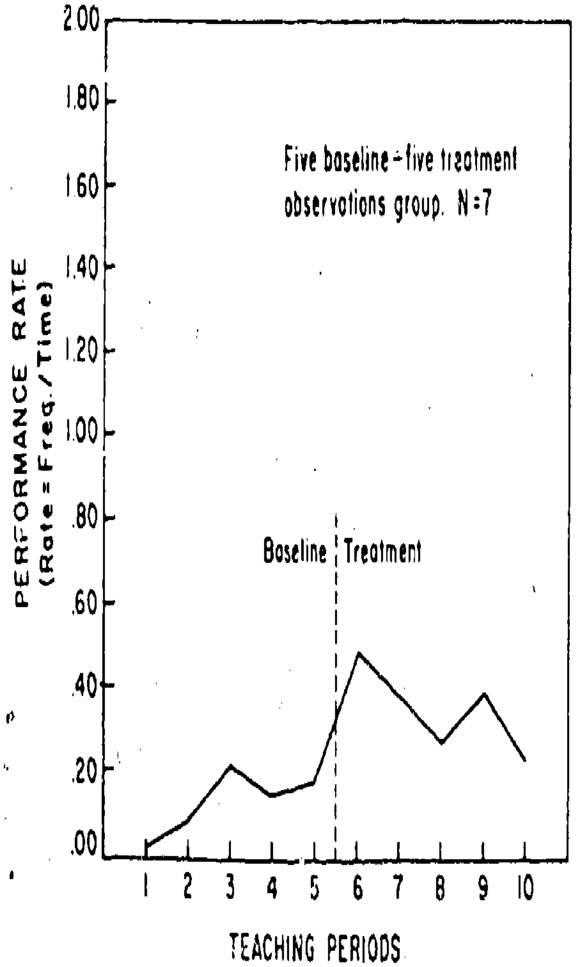


Figure 9a. Performance rate curve for the 3/7 group.

Figure 9b. Performance rate curve for the 4/6 group.



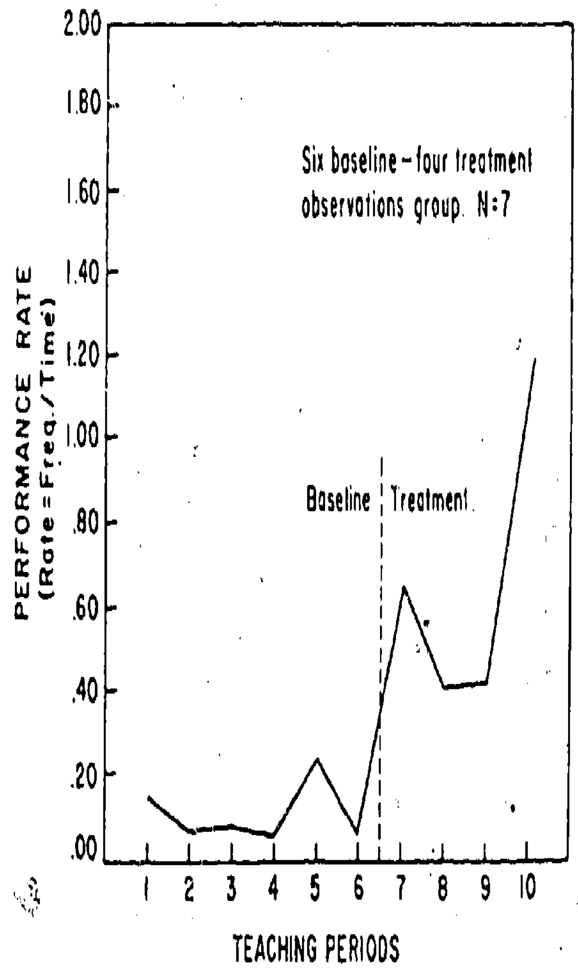


Figure 9c. Performance rate curve for the 5/5 group.

Figure 9d. Performance rate curve for the 6/4 group.

V. SUMMARY AND CONCLUSIONS

This project attempted to demonstrate the relative effectiveness of CATTS instantaneous and delayed feedback in generating specific teacher behaviors in a preservice teaching environment. The project built on previous developmental work at Indiana University involving the CATTS system. Specifically, the purpose of the project was to demonstrate the effect of CATTS immediate visual and delayed (printout) feedback on increasing various cognitive and management behaviors of teacher-trainees in a controlled laboratory classroom setting. The investigation was carried out in conjunction with the practicum phase of a junior level course in special education at Indiana University.

There were 27 majors in special education who participated in the study. In all, the study entailed three major phases: first, discrimination training in which trainees learned an observation system; second, collection of a baseline of trainees' cognitive demands and behavioral control strategies in the classroom; and third, measurement of trainees' cognitive demands and behavioral control strategies under CATTS instantaneous and/or delayed feedback conditions.

In the first phase of the project trainees learned two category observation systems entitled COG-STRAT and MAN-STRAT. Categories on the COG-STRAT included teacher and pupil cognitive styles of interaction, while MAN-STRAT categories focused on pupil on- and off-task behavior and strategies of teaching management of pupil behavior. The behaviors chosen to be included in these systems were those which were deemed important during classroom interaction with handicapped children. All trainees were trained as coders on the COG-STRAT system. CITH personnel were trained as coders on the MAN-STRAT system. Acquisition of observation-coding skills was facilitated by a computer-aided training device called DITRMA. The total time for trainee acquisition of the COG-STRAT system varied from 8 to 10 hours, while MAN-STRAT training took slightly longer.

Following observer training, both criterion-related and intracoder measures were obtained respectively for COG-STRAT and MAN-STRAT observers via video-taped criterion tests. All trainees achieved initial criterion-related agreement scores $\geq$ .85. In addition to initial criterion-related agreement, initial intra-observer agreement measures were obtained for both systems. All trainees exceeded the preset standard of .75 for both pupil and category identification. In fact, almost all exceeded a .85 standard. Maintenance checks for criterion-related agreement were similarly performed about one-third of the way through the study, except different video tapes were used. In addition, "live" maintenance checks were performed throughout the study on both observation systems in order to monitor the observers' agreement with expert coders during actual trainee classroom lessons. The majority of the trainees exceeded the .85 standard on the maintenance check.

In each of the two teaching phases (i.e., Baseline and Feedback), trainees taught language arts lessons they had prepared. Lesson objectives that trainees used for lesson plan preparation were stated in behavioral terms. Expected pupil terminal behaviors, indicating the pupils attainment of a concept or ability to perform a task, were the basis for a given objective. The trainees were required to submit a written lesson plan for each lesson they taught one week prior to teaching, for comment and evaluation. Each practicum student was randomly assigned to teach the same group of children, once each week for the duration of the study. There were four to five children in each group and there were three such groups. These groups were drawn from a class of EMR children whose regular classroom was located in the Teacher Education Laboratory (TEL) at CITH. The I.Q. of the children in this class ranged from 59 to 84, and they were grouped for instruction according to the similarity of instructional needs and social maturity. Each group of pupils received a total of one hour of supplementary reading instruction by CATTS trainees each day. The same group of children was taught by one trainee for half an hour and then by another trainee for the next half hour.

In addition, each trainee coded another trainee in his/her group at least once a week on the COG-STRAT system. The scheduling of COG-STRAT coders depended upon the trainees' teaching assignment. For example, if a trainee taught during the first half hour, s/he would code during the second half hour and vice versa. Coders were randomly assigned to teachers and scheduled so that each teacher was observed by a different coder on successive lessons, and assignment of trainees to day of teaching was also random. Assignment of trainees to teaching room rotated between the three classrooms in the TEL facility. Trainees observed and coded classroom interaction on button boxes which were linked to a PDP-12 computer located in the TEL. Coding for each lesson continued for 30 minutes after which the computer automatically shut off the coding terminals. The three instructional groups were taught simultaneously in three different classrooms.

A multiple baseline design was selected for the two teaching phases of the project. In the baseline condition, trainees taught the lessons they had prepared without receiving feedback of any kind. The number of lessons taught without feedback varied from trainee to trainee. The baseline teaching always included at least the first three lessons and varied up to the first six lessons taught. Assignment of trainees to variable baseline teaching trials was random. The second controlled condition was teaching after the trainee received computer printout information or feedback on his/her baseline teaching performance and a set of instructions in modular form on how to read the computer printout. This feedback was based on both COG-STRAT and MAN-STRAT categories.

The trainees were also instructed in procedures for using the printout information to graph the percent frequency of occurrence of behaviors for each of the observation system categories. They were also assigned the task of graphing all categories in the baseline period. At this stage the trainees were also required to evaluate their baseline teaching on the basis of their graphs. They then selected and listed in order of personal preference three of the observation system categories from COG-STRAT and MAN-STRAT that they judged as most important to their teaching and wished to work on for improving their teaching performance. Within a week of receiving baseline information, the trainees were required to select a single behavioral category for subsequent improvement. The final selection was arrived at in consultation with a staff member. This choice became, in most cases, the category that was targeted for improvement in subsequent lessons. The trainee also had to indicate whether she/he wanted to accelerate or deccelerate the percent frequency of occurrence of the target behavioral category in subsequent lessons. In most cases the trainees' first choice was an acceptable behavioral category indicative of a relatively stable baseline performance on that category during baseline. All of these activities were carried out after the baseline period and before the trainees taught with feedback.

In addition to these procedures, the trainees were randomly assigned to feedback conditions. Half the trainees were assigned to a printout-only feedback group, and half to a scope and printout feedback group. The scope feedback group was presented with a self-instructional module and instructions from a staff member on how to interpret CATTS video scope feedback. This step was performed prior to the trainees teaching with feedback. During the feedback phase, the weekly lessons continued as during baseline and, in addition, all trainees received printouts on their latest lesson within two hours after teaching. The group receiving instantaneous video or scope feedback always had a video monitor in their classroom which showed the cumulative percent frequency of occurrence of the targeted behavioral category that the trainee had chosen for teaching behavior change. Hence, this group obtained immediate feedback on the criterion teaching behavior while continuing to teach. The printout-only group was required to graph the percent frequency for the targeted category after the printout was received. This data was recorded on a cumulative individual graph for the category and included the trainees baseline percentages. In addition, all trainees kept a running record of the percent frequency in category for each successive lesson by entering the information for the previous lesson on a graph in the laboratory.

The instantaneous video feedback group was required to fill out a post-teaching questionnaire after each lesson they taught. It was expected that the questionnaire would serve to reinforce the trainees attention to the scope during teaching. Among the items in the questionnaire was one in which the trainee was required to estimate the number of times he looked at the video monitor. Another question required the trainee to reconstruct the video feedback display for the entire 30-minute lesson. The number of lessons taught with feedback varied from trainee to trainee. The feedback teaching included at least four lessons and varied up to seven lessons taught, depending on the number of lessons taught during the baseline phase. The teaching phases of the project covered a total period of ten weeks.

The results of the study revealed that all trainees in both treatment conditions significantly increased their baseline criterion rate of performance by a ratio of 2.8:1 as a function of CATTS feedback. There were no differences between the two CATTS feedback groups during baseline on their individually chosen criterion measure. However, during treatment, trainees in the CATTS scope and printout condition increased their criterion rate of responding to a significantly greater degree (2:1 ratio) than did trainees in the delayed printout-only condition.

Conclusions. The results of the present CATTS project have demonstrated the effectiveness of CATTS immediate visual and delayed (printout) feedback on increasing critical cognitive and management behaviors of special education teacher-trainees in a controlled laboratory setting. In this project, all trainees in both CATTS treatment conditions were able to significantly increase their baseline rate of teaching performance as a function of CATTS feedback. Moreover, trainees who received the CATTS instantaneous visual feedback were able to increase their criterion rate of responding to a relatively greater degree than did trainees who received CATTS delayed or printout feedback following their teaching. Hence, these findings once again verify the importance of feedback and, in particular, immediate feedback or knowledge of results in human learning and skill acquisitic . Feedback which is transmitted within seconds of the occurrence of the behavior to be modified appears to facilitate greater teacher behavior change than feedback which is delivered after the lesson. In our view, successful acquisition of teaching skills in <u>any</u> competency or competency-based training program in special education is critically dependent upon (a) specification of "appropriate" teaching behaviors, patterns, and environments, (b) reliable and valid feedback of teaching performance during practice or acquisition trials, (c) <u>immediate</u> availability of feedback information to the trainee during the course of an ongoing lesson, and (d) access to information on previous performance for post-teaching analysis and review. Hence, the CATTS system can serve as a prime vehicle for the discrimination, generation, and evaluation of specific teaching competencies in existing teacher education programs in special and/or regular education. In essence, CATTS is a versatile and comprehensive system which can be applied in numerous ways within the teacher-training field.

REFERENCES

1. Allen, D. W., & Ryan, K. Microteaching. Palo Alto, Calif.: Addison-Wesley, 1969.

2. Anderson, H. H. Domination and social integration in behavior of kindergarten children and teachers. Genetic Psychology Monograph, 1939, 3, 285-385.

3. Borg, W. R., Kelly, M. L., Langer, P., & Call, M. The minicourse: A microteaching approach to teacher education. Beverly Hills, Calif.: MacMilian, 1970.

4. Fink, A. H., & Semmel, M. I. Indiana Behavior Management System II. Bloomington: Indiana University, Center for Innovation in Teaching the Handicapped, 1971.

5. Flanders, N. A. Estimating reliability. In Amidon, E. J., and Hough, J. B. (Eds.), Interaction analysis: Theory, research, and application. Reading, Mass.: Addison-Wesley, 1967, 161-166.

6. Flanders, N. A. Analyzing teaching behavior. Reading, Mass.: Addison-Wesley, 1970.

7. Frick, T., & Semmel, M. I. Observational records: Observer agreement and reliabilities. Bloomington: Indiana University, Center for Innovation in Teaching the Handicapped, 1974.

8. Fuller, F. F., Baker, H. P., & Manning, B. A. Self-confrontation counseling: A selective review to 1972. Paper presented at AERA, Chicago, 1972.

9. Kirk, R. E. Experimental design: Procedures for the behavioral sciences. Belmont, Calif.: Wadsworth, 1968.

10. Kreider, J. M. The effect of computer-assisted teacher training system feedback on increasing teacher use of pupil ideas with EMR children. Unpublished doctoral dissertation, University of Michigan, 1969.

11. McGaw, B., Wardrop, J. L., & Bunda, M. A. Classroom observation schemes: Where are the errors? American Educational Research Journal, 1972, 9(11), 13-27.

12. Medley, D. M., & Mitzel, H. E. Measuring classroom behavior by systematic observation. In N. L. Gage (Ed.), Handbook of research on teaching. Chicago: Rand McNally & Co., 1963, 247-328.

13. Medley, D. M., & Norton. The concept of reliability as it applies to behavior records. Prepared for the 1971 meeting of APA in Washington, D.C.

14. Lynch, W. W., & Ames, C. Individual cognitive demand schedule. Bloomington: Indiana University, Center for Innovation in Teaching the Handicapped, 1971.

15. Medley, D. M., & Soar, R. S. Assessment and research in teacher education: Focus on CBTE. Washington, D.C.: American Association of Colleges for Teacher Education, PBTE Series: No. 17, 1975.

16. Peck, R. F., & Tucker, J.A. Research on teacher education. In R. M. Travis (Ed.), Second Handbook of Research on Teaching, Chicago: Rand McNally, 1971.

17. Schmitt, J. S. Modifying questioning behavior of prospective teachers of mentally retarded children through a computer-assisted teacher training system (CATTS). Unpublished doctoral dissertation, University of Michigan, Ann Arbor, 1969.

18. Semmel, M. I., Semmel, D. S., & Morrissey, P. A. Competency-based teacher education in special education. Bloomington: Indiana University, Center for Innovation in Teaching the Handicapped, 1976.

19. Semmel, M. I. Application of systematic classroom observation to the study and modification of pupil-teacher interaction in special education. In Weinberg, R. A., & Wood, F. H. (Eds.), Observation of pupils and teachers in mainstream and special education settings: Alternative strategies. Minneapolis: Leadership Training Institute/Special Education, University of Minnesota, 1975, 231-264.

20. Semmel, M. I., & Olson, J. L. CATTS Technical Reports: Bloomington, Indiana University, Center for Innovation in Teaching the Handicapped, 1977.

21. Simon, A., & Boyer, E. G., (Eds.). Mirrors for behavior II: An Anthology of observation instruments. (Vol. A & B). Philadelphia, Pa.: Research for Better Schools, 1970.

22. Smith, B. O., & Meux, M. A study of the logic of teaching. USOE Cooperative Research Project No. 258. Urbana, Ill.: University of Illinois, 1962.

23. Tiedt, I. M., & Tiedt, S. W. Contemporary English in the elementary school. Prentice Hall Inc., Englewood Cliffs, N.J., 1967.

24. Van Allen, R., & Van Allen, C. Language experiences in reading. Encyclopedia Britannica Press, Chicago, 1970.

25. Weaver, P. Effects of a computer-assisted teacher training system and teacher expectancies on teacher-pupil verbal interactions with EMR children. Unpublished doctoral dissertation, University of Michigan, 1969.

APPENDIX I

TABLES

TABLE 1.2

CATTS OBSERVATION SYSTEMS CODING SCHEDULE

Time	January										February			
	22	23	24	25	26	27	28	29	30	31	1	2	3	4
	M	T	W	R	F	S	S	M	T	W	R	F	S	S
00 - 8:30	G-1	G-6	G-5	G-4				G-3	G-2	G-1	G-6			
30 - 9:00	T-1	T-3	T-2	T-3				T-1	T-3	T=2	T-2			
00 - 9:30				G-5							G-1			
30 - 10:00	G-[illegible]							[illegible]			T-2			
00 - 10:30	T-1					open practice coding sessions		T-1			G-2		open practice coding sessions	
30 - 11:00				[illegible]-4							T-3			
00 - 11:30	G-3	G-8		G-1				G-5	G-3		G=3			
30 - 12:00	T-2	T-3		T-1				T-2	T-3		T-3			
00 - 12:30		G-2		G-2					G-4		G=4			
30 - 1:00	G-4	T-1		T-1				G=6	T-1		T-1			
00 - 1:30	T-2	G-3	G-6	G-3				T-2	G-5	G-2	G-5			
30 - 2:00		T-1	T=2	T-2					T-1	T-2	T-1			
00 - 2:30	G=5	G=4	G-1					G=1	G-6	G-3				
30 - 3:00	T=3	T=2	T-3					T=3	T-2	T=3				
00 - 3:30		G-5	G-2						G=1	G-4				
30 - 4:00	G-6	T-2	T-3					G-2	T=2	T-3				
00 - 4:30	T-3		G-3					T=3		G=5				
30 - 5:00			T-1							T-1				
00 - 5:30			G-4							G-6				
30 - 6:00			T-1							T-1				
00 - 6:30														
30 - 7:00														
00 - 7:30														
30 - 8:00														
00 - 8:30														
30 - 9:00														
00 - 9:30														
30 - 10:00														
00 - 10:30														
0:30 - 11:00														
1:00 - 11:30														

= COG-STRAT Training

= MAN-STRAT Training

= Trainee in-class periods

G = COG-STRAT Training Group

T = COG-STRAT Trainer

TABLE 1.3 TENTATIVE CATTS TRAINING SCHEDULE

Week #1

Session #1 (1½ hours) Monday	1.1	Go over any questions on written script (10 min.)
	1.2	Intro. to DITRMA: (30 min.) 1.2.1 Call out pupils & categories for button box and mental warm-up drill 1.2.2 Code handed-in examples 1.2.3 Break (5 min.)
	1.3	Code COG-STRAT Training Tape Seg's 1-40 (40 min.)
	1.4	Assignment #2: Prepare a 5-min. Lesson relevant and appropriate to Steakley's class. Due on Session #2 (5 min.). (Be sure to include all categories.)
Session #2 (1 hour) Tuesday	2.1	Warm-up: Code seg's 1-40 again (30 min.)
	2.2	Role-playing (30 min.) 2.2.1 1 teacher 2.2.2 2 pupils 2.2.3 1 or 2 coders & trainer/coder
Session #3 (1 hour) Wednesday	3.1	Code continuous segments on COG-STRAT Training Tape (1 hour) 3.1.1 Start by coding 5, 6, 7 only, using pupil code 1, 2, 9, 0 3.1.2 Replay and code pupils only, using above codes 3.1.3 Replay and code everything 3.1.4 Continue with seg. #2 (Ibid)
Session #4 (1 hour) Thursday	4.1	Continue coding continuous segments on training tape (30 min.)
	4.2	Role-playing (30 min.)
Session #5 (1½ hours) Monday	5.1	Criterion Tape (Segments 1-54) (45 min.)
	5.2	Read printout and determine who passes and who does not (answer questions)
	5.3	Replay seg's 1-54 for people who do not pass
	5.4	Read printout, etc.

TABLE 1.3 TENTATIVE CATTS TRAINING SCHEDULE
(Continued)

Session #6 (1 hour) Tuesday	6.1	Criterion Tape (Zuck's lesson): to be coded twice (45 min.)
	6.2	Read printout
Session #7 (1 hour) Wednesday	7.1	Make-up sessions
	or 7.2	Retraining
	or 7.3	Retake criterion
Session #8 (1 hour) Thursday	8.1	Make-up or retrain or retake criterion

TABLE 1.4: OBSERVER AGREEMENT CHECKS

System	Number of Observers Involved	Type of Check	Agreement Coefficient(s) Used
COG-STRAT	27	Initial Criterion-Related (#1)	P_o; P_o'
COG-STRAT	27	Initial Criterion-Related (#2)	P_o; P_o'
MAN-STRAT	9	Initial Criterion-Related (#1)	P_o; P_o'
MAN-STRAT	9	Initial Criterion-Related (2)	P_o; P_o'
COG-STRAT	27	Initial Intra-Observer	P_{o_f}; π_f
MAN-STRAT	9	Initial Intra-Observer	P_{o_f}; π_f
MAN-STRAT	9	Maintenance Criterion-Related	P_o'
MAN-STRAT	9	Maintenance Intraobserver	P_{o_f}; π_f
COG-STRAT	26	Maintenance Criterion-Related	P_o'
MAN-STRAT		Live Maintenance	P_{o_f}; π_f
COG-STRAT		Live Maintenance	P_{o_f}; π_f

TABLE 1.5 INITIAL CRITERION-RELATED OBSERVER AGREEMENT: COG-STRAT

OBSERVER		PUPIL IDENTIFICATION					
		PUPIL #0		PUPIL #1		PUPIL #2	
		P_o^*	P'_o	P_o	P'_o	P_o	P'_o
1 Jo Ann	A.	.93	.94	1.00	1.00	1.00	.98
2 Laura	S.	1.00	1.00	.95	.96	.87	.91
3 Mary Kay	Y.	1.00	1.00	.95	.99	1.00	1.00
4 Kate	Z.	.97	.98	.90	.92	1.00	1.00
5 Joyce	B.	.92	.95	.95	.95	.97	1.00
6 Janette	S.	1.00	1.00	1.00	1.00	.97	.99
7 DeLynn	T.	1.00	1.00	.95	.97	1.00	1.00
8 Teri	H.	1.00	1.00	.95	.97	1.00	1.00
9 Janet	R.	1.00	1.00	1.00	1.00	.93	.97
10 Paula	S.	1.00	1.00	1.00	1.00	.90	.95
11 Bev	R.	1.00	1.00	.86	.88	.87	.89
12 Terri	P.	1.00	1.00	.95	.97	.93	.97
13 Barb	W.	.92	.95	.95	.95	1.00	.98
14 Chris	T.	.92	.91	.90	.93	.97	.99
15 Shelly	L.	.89	.92	.90	.90	.90	.96
16 Susie	A.	.92	.96	.90	.88	.93	1.00
17 Rueben	F.	.95	1.00	1.00	.96	.87	.96
18 Marcy	S.	1.00	1.00	1.00	1.00	.93	.96
19 Chris	N.	.95	.96	.90	.90	.93	.99
20 Marilyn	O.	1.00	1.00	.86	.88	.93	.97
21 Carol	M.	1.00	1.00	.95	.96	.93	.98
22 Jim	F.	1.00	.99	.95	.98	1.00	1.00
23 Diane	V.	1.00	1.00	1.00	1.00	.97	.98
24 Jo	S.	1.00	1.00	1.00	1.00	.93	.99
25 Kathy	B.	.95	.99	.95	.93	.87	.94
26 Diana	M.	.86	.84	1.00	1.00	.83	.84
27 Arlene	G.	.97	1.00	1.00	.93	.77	.88
	MEAN	.969	.977	.95[illegible]	.956	.933	.966

*P_o = Percent agreement against criterion based on raw frequencies

P'_o = Percent agreement against criterion based on percent frequencies

TABLE 1.5 INITIAL CRITERION-RELATED OBSERVER AGREEMENT: COG-STRAT

OBSERVER	Pupil Identification #3 P_o^*	Pupil Identification #3 P_o'	Pupil Identification #4 P_o	Pupil Identification #4 P_o'
1	.95	.96	.96	.95
2	.95	.95	.96	1.00
3	.95	.99	.96	.98
4	.90	.97	1.00	1.00
5	1.00	.98	.96	1.00
6	.95	.99	1.00	1.00
7	.95	.96	.92	.96
8	.95	1.00	1.00	1.00
9	1.00	1.00	.96	.97
10	.90	.98	.96	.96
11	.90	.96	.93	.94
12	1.00	1.00	.96	1.00
13	.86	.94	.96	.95
14	1.00	1.00	.90	.92
15	.95	.93	1.00	.99
16	.90	1.00	.96	1.00
17	.95	.98	.92	1.00
18	.91	.93	.92	.95
19	1.00	.98	.96	1.00
20	.95	1.00	.96	1.00
21	.76	.82	.93	.93
22	1.00	1.00	1.00	1.00
23	.95	1.00	.92	.96
24	.95	1.00	.96	1.00
25	.95	1.00	1.00	.97
26	.95	.99	.77	.77
27	.90	1.00	.88	1.00
MEAN	.940	.974	.949	.970

*P_o = Percent agreement against criterion based on <u>raw</u> frequencies

P_o' = Percent agreement against criterion based on <u>percent</u> frequencies

TABLE 1.5 INITIAL CRITERION-RELATED OBSERVER AGREEMENT: COG-STRAT

OBSERVER	Category Identification					
	Category #1		Category #2		Category #3	
	P_o*	P'_o	P_o	P'_o	P_o	P'_o
1	1.00	1.00	1.00	1.00	1.00	1.00
2	1.00	.99	.83	.99	.90	.36
3	.90	.93	.90	.93	.77	.79
4	.90	.96	.91	.91	.91	.91
5	1.00	.98	.90	.97	.83	.81
6	.83	.85	.83	.85	1.00	1.00
7	.83	.84	.71	.72	.70	.74
8	.90	.95	1.00	1.00	.91	.92
9	.90	.96	.91	.91	.90	.96
10	.91	.90	.91	.90	.80	.86
11	.91	.92	.83	.84	.90	.95
12	.83	.83	.91	.91	.80	.85
13	.90	.97	.83	.81	.80	.88
14	.91	.92	1.00	1.00	1.00	1.00
15	.91	.89	.91	.89	.50	.54
16	.91	.83	1.00	.97	.90	.99
17	1.00	.94	.90	1.00	1.00	.94
18	.91	.92	1.00	1.00	.90	.95
19	.83	.82	1.00	.98	1.00	.98
20	.91	.91	1.00	1.00	.83	.83
21	1.00	.99	.91	.90	.90	.96
22	.91	.92	.91	.92	1.00	1.00
23	.90	.95	.70	.74	.77	.78
24	1.00	.98	1.00	.98	1.00	.98
25	1.00	.97	1.00	.97	1.00	.97
26	.83	.87	.60	.62	.77	.80
27	.83	.76	.9[illegible]	1.00	.70	.82
MEAN	.913	.919	.900	.915	.870	.895

*P_o = Percent agreement against criterion based on <u>raw</u> frequencies

P'_o = Percent agreement against criterion based on <u>percent</u> frequencies

TABLE 1.5 INITIAL CRITERION-RELATED OBSERVER AGREEMENT:

COG-STRAT

OBSERVER	CATEGORY IDENTIFICATION					
	Category #4		Category #5		Category #6	
	P_o*	P'_o	P_o	P'_o	P_o	P'_o
1	.90	.90	1.00	1.00	1.00	.94
2	1.00	.99	1.00	1.00	1.00	1.00
3	.90	.93	1.00	1.00	.88	.86
4	1.00	1.00	.85	.90	.92	.71
5	1.00	.98	.85	.92	1.00	.90
6	1.00	1.00	1.00	1.00	1.00	1.00
7	1.00	1.00	1.00	.95	1.00	1.00
8	1.00	1.00	1.00	1.00	1.00	1.00
9	1.00	1.00	1.00	1.00	.83	.79
10	1.00	.98	1.00	1.00	1.00	.98
11	.80	.85	1.00	1.00	1.00	1.00
12	1.00	1.00	1.00	.94	.91	.86
13	1.00	.98	1.00	1.00	.95	.88
14	1.00	1.00	1.00	.96	1.00	.96
15	.90	.97	1.00	.98	1.00	1.00
16	1.00	.97	.92	1.00	1.00	1.00
17	1.00	.94	1.00	.94	1.00	1.00
18	.90	.95	1.00	1.00	.87	.83
19	1.00	.98	1.00	.99	1.00	.93
20	1.00	1.00	.92	.98	1.00	.99
21	1.00	.99	1.00	.94	.95	.89
22	.90	.95	1.00	1.00	.88	.87
23	1.00	1.00	.85	.89	.94	.93
24	1.00	.98	1.00	.99	.88	.89
25	1.00	.97	.62	.67	1.00	1.00
26	1.00	1.00	.76	.75	1.00	.96
27	.80	.93	.85	.98	1.00	.95
MEAN	.967	.972	.949	.955	.963	.927

*P_o = Percent agreement against criterion based on raw frequencies

P'_o = Percent agreement against criterion based on percent frequencies

TABLE 1.5 INITIAL CRITERION-RELATED OBSERVER AGREEMENT:

COG-STRAT

OBSERVER	Category Identification					
	Category #7		Category #8		Category #9	
	P_o^*	P'_o	P_o	P'_o	P_o	P'_o
1	1.00	1.00	.91	.91	.91	.97
2	1.00	1.00	1.00	.99	.91	.90
3	.85	.82	1.00	1.00	.91	.93
4	.65	.61	1.00	1.00	.91	.91
5	.73	.68	.92	.89	.91	.89
6	1.00	1.00	1.00	1.00	.80	.83
7	1.00	1.00	.92	.92	1.00	1.00
8	1.00	1.00	1.00	1.00	1.00	1.00
9	1.00	1.00	1.00	1.00	1.00	1.00
10	1.00	1.00	1.00	.99	1.00	.98
11	1.00	1.00	.92	.92	.90	.95
12	.75	.75	1.00	1.00	1.00	1.00
13	1.00	1.00	1.00	.98	.91	.89
14	1.00	.98	1.00	1.00	1.00	1.00
15	1.00	1.00	.92	.89	.83	.81
16	1.00	.91	1.00	.96	.91	.88
17	1.00	.98	.91	1.00	1.00	.94
18	.63	.62	1.00	1.00	.91	.92
19	.88	.89	.91	.99	1.00	.98
20	1.00	1.00	1.00	1.00	1.00	1.00
21	.88	.89	1.00	.99	1.00	.99
22	.85	.81	1.00	1.00	.91	.92
23	.73	.69	1.00	1.00	.91	.92
24	.85	.78	1.00	.99	.91	.90
25	.79	.71	.91	1.00	1.00	.97
26	1.00	1.00	.92	.95	.90	.92
27	.88	.97	.82	.96	.90	1.00
MEAN	.906	.892	.965	.975	.939	.941

$^*P_o$ = Percent agreement against criterion based on raw frequencies

P'_o = Percent agreement against criterion based on percent frequencies

TABLE 1.5 INITIAL CRITERION-RELATED OBSERVER AGREEMENT:

COG-STRAT

OBSERVER	Identification					
	Category #10		Category #11		Category #12	
	P_o^*	P'_o	P_o	P'_o	P_o	P'_o
1	.92	.98	1.00	1.00	1.00	1.00
2	.64	.68	1.00	1.00	1.00	.93
3	1.00	1.00	.88	.85	.85	.82
4	.91	.96	.88	.87	1.00	.94
5	.55	.59	.88	.90	1.00	1.00
6	.91	.95	1.00	.98	.80	.83
7	.64	.67	.88	.87	.90	.95
8	1.00	1.00	1.00	1.00	1.00	1.00
9	.55	.58	.88	.87	.92	.86
10	.73	.78	.88	.89	1.00	.93
11	.82	.86	1.00	1.00	1.00	1.00
12	.64	.67	1.00	1.00	.90	.96
13	.55	.59	1.00	1.00	.90	.97
14	.91	.95	.88	.87	1.00	1.00
15	.82	.88	.63	.64	1.00	.92
16	.73	.80	.75	.79	1.00	1.00
17	.46	.52	.75	.80	1.00	.89
18	.82	.86	1.00	1.00	1.00	1.00
19	.82	.88	1.00	1.00	.80	.86
20	.82	.87	1.00	1.00	.90	.96
21	.73	.78	1.00	1.00	.90	.96
22	.91	.95	1.00	.98	.90	.95
23	.91	.96	1.00	1.00	1.00	1.00
24	.46	.49	1.00	1.00	1.00	1.00
25	.91	1.00	1.00	1.00	.90	.99
26	.18	.19	.92	.90	1.00	.98
27	.82	.96	1.00	.85	.80	.93
MEAN	.746	.793	.934	.928	.943	.914

$^*P_o$ = Percent agreement against criterion based on raw frequencies

P'_o = Percent agreement against criterion based on percent frequencies

TABLE 1.5 INITIAL CRITERION-RELATED OBSERVER AGREEMENT:

COG-STR...

OBSERVER		PUPIL IDENTIFICATION					
		PUPIL #0		PUPIL #1		PUPIL #2	
		P_o^*	P_o'	P_o	P_o'	P_o	P_o'
Jo Ann	A.	.93	.94	1.00	1.00	1.00	.98
Laura	S.	1.00	1.00	.95	.96	.87	.91
Mary Kay	Y.	1.00	1.00	.95	.99	1.00	1.00
Kate	Z.	.97	.98	.90	.92	1.00	1.00
Joyce	B.	.92	.95	.95	.95	.97	1.00
Janette	S.	1.00	1.00	1.00	1.00	.97	.99
DeLynn	T.	1.00	1.00	.95	.97	1.00	1.00
Teri	H.	1.00	1.00	.95	.97	1.00	1.00
Janet	R.	1.00	1.00	1.00	1.00	.93	.97
Paula	S.	1.00	1.00	1.00	1.00	.90	.95
Bev	R.	1.00	1.00	.86	.88	.87	.89
Terri	P.	1.00	1.00	.95	.97	.93	.97
Barb	W.	.92	.95	.95	.95	1.00	.98
Chris	T.	.92	.91	.90	.93	.97	.99
Shelly	L.	.89	.92	.90	.90	.90	.96
Susie	A.	.92	.96	.90	.88	.93	1.00
Rueben	F.	.95	1.00	1.00	.96	.87	.96
Marcy	S.	1.00	1.00	1.00	1.00	.93	.96
Chris	N.	.95	.96	.90	.90	.93	.99
Marilyn	O.	1.00	1.00	.86	.88	.93	.97
Carol	M.	1.00	1.00	.95	.96	.93	.98
Jim	F.	1.00	.99	.95	.98	1.00	1.00
Diane	V.	1.00	1.00	1.00	1.00	.97	.98
Jo	S.	1.00	1.00	1.00	1.00	.93	.99
Kathy	B.	.95	.99	.95	.93	.87	.94
Diana	M.	.86	.84	1.00	1.00	.83	.84
Arlene	G.	.97	1.00	1.00	.93	.77	.88
	MEAN	.969	.977	.953	.956	.933	.966

$^*P_o$ = Percent agreement against criterion based on raw frequencies

P_o' = Percent agreement against criterion based on percent frequencies

TABLE 1.6 INITIAL CRITERION-RELATED OBSERVER AGREEMENT:

MAN-STRAT

OBSERVER	Pupil #0 Po*	Pupil #0 P_o'	Pupil #1 Po	Pupil #1 P_o'	Pupil #2 Po	Pupil #2 P_o'
1	1.00	.96	.88	.99	.92	.98
2	1.00	1.00	1.00	1.00	1.00	1.00
3	.99	1.00	.97	1.00	1.00	.97
4	.98	.99	1.00	.99	1.00	1.00
5	.92	.98	.97	.97	1.00	.92
6	.89	.96	.84	.95	1.00	.93
7	.89	1.00	.81	.96	.84	.95
8	.95	1.00	.88	.99	1.00	.91
9	.80	.92	.91	.96	1.00	.87
Mean	.936	.979	.918	.979	.973	.948

OBSERVER	Pupil #3 Po*	Pupil #3 P_o'	Pupil #4 Po	Pupil #4 P_o'		
1	.87	.93	.88	.99		
2	1.00	.98	.94	.97		
3	.95	.98	.88	.95		
4	1.00	.97	1.00	.99		
5	1.00	.92	.77	.86		
6	.93	.82	.79	.90		
7	1.00	.89	.88	1.00		
8	1.00	.93	.68	.77		
9	.92	.94	.82	1.00		
Mean	.963	.929	.849	.937		

*Po=Percent agreement against criterion based on raw frequencies

P_o^1=Percent agreement against criterion based on percent frequencies

TABLE 1.6 INITIAL CRITERION-RELATED OBSERVER AGREEMENT:

MAN-STRAT

OBSERVER	Category 1 Po	Category 1 P'_o	Category 2 Po*	Category 2 P'_o	Category 3 Po	Category 3 P'_o
1	.99	.89	1.00	.94	.90	.96
2	1.00	.99	1.00	1.00	.96	.95
3	1.00	.94	1.00	.97	.97	.99
4	.99	1.00	1.00	1.00	.99	.99
5	.93	1.00	1.00	.94	.98	.97
6	.81	.88	1.00	.94	1.00	.94
7	.76	.87	1.00	.88	.93	.97
8	.92	.99	1.00	.94	.99	.88
9	.74	.86	1.00	.88	.93	.95
Mean	.904	.936	1.00	.943	.961	.956

OBSERVER	Category 4 Po	Category 4 P'_o	Category 5 Po	Category 5 P'_o	Category 6 Po	Category 6 P'_o
1	.89	.99	1.00	.90	1.00	.93
2	.89	.91	1.00	.96	1.00	1.00
3	.78	.80	.88	.90	.75	.73
4	1.00	1.00	.90	.86	.86	.86
5	.89	.99	.88	.93	.86	.80
6	1.00	.93	1.00	1.00	.75	.71
7	.90	.80	.89	.83	1.00	.89
8	.56	.61	.90	.80	.50	.57
9	.44	.53	1.00	[illegible]	.86	.75
Mean	.817	.829	.939	.890	.842	.804

*Po = Percent agreement against criterion based on raw frequencies
P'_o = Percent agreement against criterion based on percent frequencies

TABLE 1.6 INITIAL CRITERION-RELATED OBSERVER AGREEMENT:

MAN-STRAT

OBSERVER	CATEGORY IDENTIFICATION					
	Category 7		Category 8		Category 9	
	Po	P_o'	Po*	P_o'	Po	P_o'
1	1.00	.93	1.00	.94	.71	.75
2	.86	.89	1.00	1.00	.77	.75
3	1.00	.97	1.00	.97	.71	.72
4	1.00	1.00	1.00	1.00	1.00	1.00
5	.86	.80	.86	.98	.77	.80
6	.67	.62	1.00	.94	.77	.82
7	.67	.59	.86	1.00	.94	1.00
8	.75	.71	.86	.98	.59	.63
9	.75	.65	.71	.87	.94	1.00
Mean	.840	.796	.921	.964	.800	.830

OBSERVER	Category 10		Category 11		Category 12	
	Po	P_o'	Po	P_o'	Po	P_o'
1	1.00	.94	.77	.82	.33	.33
2	1.00	1.00	.85	.82	1.00	1.00
3	1.00	.97	.77	.78	.67	.67
4	1.00	1.00	.77	.76	1.00	1.00
5	1.00	.94	.77	.82	.67	.75
6	1.00	.94	.69	.74	1.00	1.00
7	.43	.53	.92	1.00	.67	.75
8	1.00	.94	.62	.65	.67	.75
9	.86	1.00	.92	1.00	1.00	1.00
Mean	.921	.918	.787	.821	.768	.806

*Po = Percent agreement against criterion based on raw frequencies

P_o' = Percent agreement against criterion based on percent frequencies

TABLE 1.7 COG-STRAT INITIAL INTRA-OBSERVER AGREEMENT

	PUPIL IDENTIFICATION			CATEGORIES		
OBSERVER	P_{e_f}	P_{o_f}	π_f	P_{e_f}	P_{o_f}	π_f
1 Jo Ann A.	.26	.91	.88	.14	.86	.84
2 Laura S.	.24	.88	.84	.16	.87	.85
3 Mary X. Y.	.23	.95	.94	.15	.85	.82
4 Kate Z.	.21	.96	.95	.16	.86	.83
5 Joyce B.	.23	.88	.84	.20	.90	.88
6 Jannette S.	.24	.84	.80	.16	.87	.85
7 DeLynn T.	.24	.88	.84	.16	.87	.85
8 Teri H.	.25	.92	.89	.18	.90	.88
9 Janet R.	.23	.88	.84	.17	.88	.86
10 Paula S.	.22	.94	.92	.17	.94	.93
11 Bev R.	.28	.87	.82	.17	.93	.92
12 Terri P.	.24	.94	.92	.19	.88	.85
13 Barb W.	.21	.90	.87	.16	.86	.83
14 Chris T.	.21	.96	.95	.22	.92	.90
15 Shelly L.	.22	.90	.87	.17	.85	.82
16 Susie A.	.25	.94	.92	.16	.85	.82
17 Rueben F.	.22	.92	.90	.16	.90	.88
18 Marcy S.	.25	.95	.93	.18	.90	.88
19 Chris N.	.22	.94	.92	.17	.86	.83
20 Marilyn O.	.25	.88	.84	.17	.90	.88
21 Carol M.	.23	.87	.83	.15	.87	.85
22 Jim F.	.22	.96	.95	.19	.80	.75
23 Diane V.	.24	.86	.82	.19	.89	.86
24 Jo S.	.23	.95	.94	.19	.93	.91
25 Kathy B.	.26	.95	.93	.15	.88	.86
26 Diana M.	.23	.86	.82	.14	.77	.73
27 Arlene G.	.23	.95	.94	.17	.88	.86

TABLE 1.8 INITIAL INTRA-OBSERVER AGREEMENT:

MAN-STRAT

OBSERVER	Pupil I.D.			Categories		
	P_{e_f}	P_{o_f}	π_f	P_{e_f}	P_{o_f}	π_f
1	.20	.83	.79	.19	.82	.78
2	.21	.79	.73	.16	.84	.81
3	.20	.84	.80	.17	.83	.80
4	.21	.87	.84	.21	.73	.66
5	.22	.79	.73	.18	.91	.89
6	.21	.78	.72	.14	.79	.76
7	.21	.85	.81	.15	.81	.78
8	.21	.89	.86	.16	.77	.73
9	.20	.92	.90	.17	.80	.76

TABLE 1.9 MAINTENANCE INTRA-OBSERVER AGREEMENT:

MAN-STRAT

	Pupil I.D.			Categories		
OBSERVER	P_{e_f}	P_{o_f}	π	P_{e_f}	P_{o_f}	π
1	.37	.78	.65	.21	.82	.77
2	.35	.92	.88	.15	.89	.87
3	.30	.79	.70	.15	.80	.77
4	.41	.65	.41	.13	.71	.66
5	.36	.79	.67	.18	.63	.55
6	.38	.87	.79	.15	.82	.79
7	.37	.65	.44	.17	.79	.75
8	.39	.79	.66	.17	.82	.78
9	.49	.90	.80	.20	.89	.86

TABLE 1.10 MAINTENANCE CRITERION-RELATED OBSERVER AGREEMENT:

MAN-STRAT

	PUPIL-IDENTIFICATION					
	0		2		3	
OBSERVER	Po*	P_o'	Po	P_o'	Po	P_o'
1	---	---	---	---	---	---
2	1.00	.99	1.00	1.00	1.00	.99
3	.64	.71	.91	1.00	.99	.89
4	---	---	---	---	---	---
5	.88	.90	1.00	.97	1.00	.98
6	.57	.69	.82	.98	.97	.85
7	.93	.98	.82	.87	.95	1.00
8	---	---	---	---	---	---
9	.45	.54	.82	.98	.98	.82
Mean	.745	.802	.895	.967	.982	.922

	4	
OBSERVER	Po*	P_o'
1	---	---
2	1.00	.98
3	.75	.60
4	---	---
5	.73	.71
6	.63	.76
7	..80	.76
8	---	---
9	.80	.67
Mean	.785	.747

*Po = Percent agreement against criterion based on raw frequencies
P_o' = Percent agreement against criterion based on percent frequencies

TABLE 1.10 MAINTENANCE CRITERION-RELATED OBSERVER AGREEMENT:

MAN-STRAT

	CATEGORY IDENTIFICATION					
	1		2		3	
OBSERVER	Po*	P'o	Po	P'o	Po	P'o
1						
2	1.00	.99	----	----	1.00	.99
3	.64	.72	----	----	.98	.88
4						
5	.88	.90	----	----	.98	.96
6	.56	.67	----	----	.93	.89
7	.91	.96	----	----	.91	.97
8						
9	.44	.53	----	----	.99	.84
MEAN	.738	.795			.965	.922

	4		5		6	
OBSERVER	Po*	P'o	Po	P'o	Po	P'o
1						
2	1.00	1.00	1.00	1.00	1.00	.99
3	1.00	1.00	.88	.77	.70	.79
4						
5	1.00	1.00	.86	.88	.67	.65
6	1.00	1.00	.88	.73	1.00	.82
7	1.00	1.00	.86	.92	.90	.96
8						
9	1.00	1.00	.86	.96	.71	.59
Mean	1.00	1.00	.890	.877	.830	.800

*Po = Percent agreement against criterion based on raw frequencies
P'o = Percent agreement against criterion based on percent frequencies

TABLE 1.10 MAINTENANCE CRITERION-RELATED OBSERVER AGREEMENT:

MAN-STRAT

OBSERVER	CATEGORY IDENTIFICATION 7 Po*	7 P_o'	8 Po	8 P_o'	9 Po	9 P_o'
1						
2	1.00	1.00	1.00	.99	1.00	.99
3	1.00	.89	1.00	.88	.89	.81
4						
5	.95	.92	1.00	.98	.73	.70
6	.95	.87	.89	.74	.88	.95
7	.91	.86	1.00	.95	.89	.95
8						
9	.84	.99	1.00	.83	.80	.67
MEAN	.942	.922	.982	.895	.865	.845

OBSERVER	10 Po*	10 P_o'	11 Po	11 P_o'	12 Po	12 P_o'
1						
2	1.00	1.00	1.00	1.00	1.00	.99
3	1.00	.89	.92	.98	.75	.60
4						
5	1.00	.96	.75	.78	1.00	.98
6	1.00	.83	.83	.99	.89	.74
7	1.00	.96	.89	.84	.50	.51
8						
9	1.00	.83	.83	.69	.50	.62
Mean	1.00	.912	.870	.860	.773	.740

*Po = Percent agreement against criterion based on raw frequencies
P_o' = Percent agreement against criterion based on percent frequencies

TABLE 1.11 MAINTENANCE CRITERION-RELATED OBSERVER AGREEMENT:

COG-STRAT

OBSERVER	PUPIL IDENTIFICATION 0 P'_o *	1 P'_o	2 P'_o	3 P'_o	9 P'_o
1	1.00	.97	.89	.98	.99
2	.89	.89	.95	.84	.92
3	1.00	.97	.99	.93	1.00
4	.88	.96	1.00	1.00	.99
5	.83	.93	.89	1.00	.96
6	.80	.93	.85	.90	.92
7	.92	.94	.98	.92	.97
8	.87	.97	.97	1.00	.99
9	1.00	1.00	.96	1.00	.99
10	.93	.90	1.00	.94	.94
11	.89	.95	.89	.97	.98
12	1.00	1.00	.88	.95	.95
13	.98	1.00	.93	1.00	.97
14	.70	.87	.87	.97	.97
15	----	----	----	----	----
16	.81	.93	.95	1.00	.98
17	.92	.84	.92	.65	1.00
18	.99	.90	.96	.88	.97
19	.94	.98	.96	.91	.96
20	.90	.93	.92	1.00	.99
21	.88	.90	.89	.98	1.00
22	.74	.92	.94	.95	.95
23	.94	1.00	.86	1.00	.98
24	.90	.87	.81	.74	1.00
25	.95	1.00	.97	1.00	.97
26	.79	.92	.84	1.00	.99
27	.80	.86	.95	1.00	1.00
MEAN	.894	.936	.924	.943	.974

*P'_o = Percent agreement against criterion based on percent frequencies

TABLE 1.11 MAINTENANCE CRITERION-RELATED OBSERVER AGREEMENT:

COG-STRAT

OBSERVER	CATEGORY IDENTIFICATION 1 P'_o *	2 P'_o	3 P'_o	4 P'_o	5 P'_o	6 P'_o
1	.90	.65	.85	.99	.87	.69
2	.94	.73	.93	.99	.77	.84
3	.95	.94	.92	1.00	.86	.98
4	1.00	.90	.92	.95	.52	.83
5	.99	.91	.74	.85	.70	.98
6	1.00	.90	.87	.24	.79	.81
7	.99	.84	.71	.61	.86	[illegible]
8	1.00	.84	.84	.95	.78	1.00
9	1.00	.86	.71	[illegible]	.81	[illegible]
10	.93	.91	1.00	.95	.85	.95
11	.75	.81	.84	.97	.67	1.00
12	1.00	.80	.75	.72	.93	.80
13	.95	.76	.47	.72	.94	.94
14	.76	.84	.83	.82	.83	1.00
15	----	----	----	----	----	----
16	.99	.72	.90	.58	.74	[illegible]
17	1.00	.92	.93	.64	.99	.95
18	1.00	.94	.88	.95	.94	1.00
19	.94	.84	.67	.61	1.00	.87
20	1.00	.90	.97	.97	1.00	1.00
21	1.00	.94	.85	.60	.81	.95
22	1.00	.87	.75	.60	.81	1.00
23	1.00	.79	.65	.85	.70	.90
24	1.00	.72	.63	.85	.95	1.00
25	.69	.98	1.00	.97	1.00	1.00
26	.96	.84	.81	.97	.93	.68
27	1.00	.85	.87	.92	.59	.98
MEAN	.952	.844	.819	.816	.832	.916

P'_o = Percent agreement against criterion based on percent frequencies

TABLE 1.11 MAINTENANCE CRITERION-RELATED OBSERVER AGREEMENT:

COG-STRAT

OBSERVER	CATEGORY IDENTIFICATION			
	7	8	9	10
	P'_o *	P'_o	P'_o	P'_o
1	.74	.98	1.00	.77
2	.90	.95	.89	.79
3	.87	.87	.94	.54
4	.78	.78	.97	.99
5	1.00	.95	1.00	.91
6	.79	.93	.94	.90
7	.71	.82	1.00	.89
8	.98	.82	.96	.74
9	.90	.96	.99	.40
10	.89	.91	.98	.83
11	.94	.92	.95	.98
12	.95	.99	.91	.55
13	.90	.95	.87	.54
14	.91	.98	1.00	.99
15	----	----	----	----
16	.76	1.00	1.00	.88
17	1.00	.89	1.00	.70
18	1.00	.92	1.00	.85
19	.70	.79	.99	.71
20	.92	.94	1.00	.70
21	.90	.99	.99	.78
22	.97	.97	.95	.85
23	.86	.90	1.00	.97
24	1.00	.93	.96	.72
25	1.00	.99	.95	.77
26	.80	.93	1.00	1.00
27	.95	.88	.97	1.00
MEAN	.889	.921	.970	.796

*P'_o = Percent agreement against criterion based on percent frequencies

APPENDIX II

PILOT STUDIES

PILOT STUDY I

The Effect of CATTS Feedback on the Question Behavior of Teacher-Trainees in a Real Class Instructional Setting.

October 26, 1973

PILOT STUDY I

The Effect of CATTS Feedback on the Questioning Behavior of Teacher-Trainees in a Real Class Instructional Setting.

Introduction

This pilot study of the effect of CATTS feedback was carried out in conjunction with the practicum phase of a course in methods of teaching educable mentally retarded children. The students enrolled in the course were mainly I. U. juniors, majoring in the education of the mentally retarded.

The aim of the study was to assess the baseline frequency of use of specific teaching categories of behavior by trainees in a real classroom situation, and then measure changes in the use of these same categories of behavior under various feedback conditions.

Utilizing trained observers to record the cognitive demands (levels of questions) and behavioral control strategies of the teacher-trainees, a baseline of performance was obtained for each trainee. During the feedback phase, the observers similarly recorded the teaching behaviors of the trainees engaged in classroom teaching. The purpose of the study was to demonstrate and evaluate the effect of CATTS feedback (instant feedback) on increasing the number of high-level questions asked by the teacher during a given lesson.

In order to sensitize teacher-trainees to an awareness and understanding of different levels of questioning that can be generated in an instructional interchange with pupils, the trainees underwent a ten-hour training program in the Individual Cognitive Demand Schedule (ICDS) observational system. The system was developed by Lynch and Ames (1971), and is based in part on the theory of instruction proposed by Gagné (1967). It is a hierarchical system that categorizes questions asked by teachers into various high and low levels. The low-level categories include simple verbal demands that require, for example, habitual responding, observing, recitation of previously learned materials and remembering. Higher level categories include questions that require explaining, defining, classifying, applying and comparing, and inferential and problem solving responses. Also included in the higher level classifications are questions requiring value judgment and "make-believe" responses.

In all, the study entailed four major phases: First, participation by trainees in observation system training; second, measurement of trainees' cognitive demands in the classroom; third, measurement of trainees' cognitive demands in the classroom under CATTS instant, delayed, or no-feedback conditions; fourth, measurement of trainees' cognitive demands with no feedback, to determine the maintenance of behaviors obtained under feedback.

1. Observation System Training.

The teaching behaviors seen as critical to the establishment of a successful teaching environment by the teacher fall into two broad categories: behavioral management strategies and pedagogical strategies manifested by cognitive demands initiated by the teacher. Observation-coding systems are probably the most reliable technique for objective measurement of those classroom behaviors deemed important in any given approach to instructional method. Rosenshine & Furst (1973), and Flanders (1970) have extensively discussed the advantages of observation systems as a reliable, low-inference measure of classroom behaviors.

Two observation instruments were used to record the classroom behaviors of trainees and pupils: The Indiana Behavior Management Scale (IBMS) was used to record a wide range of off-task behaviors manifested by pupils, and the teachers' response to these behaviors. The IBMS, developed by Fink and Semmel (1971), categorizes pupils' behaviors into classifications such as types of verbal or physical manifestations of off-task behavior, self-involvement, and whether the behaviors are aggressive or interactive in nature. Teacher-control responses are categorized into eleven classifications of types of control statements.

The data derived from the IBMS was expected to result in a profile of individual and group patterns of pupil off-task behavior, including percentage of off-task behavior in the total instructional period, frequency of off-task behaviors that occur during instruction, and frequency of types of control measures invoked by teacher-trainees in response to pupil off-task behaviors.

The pedagogical strategies displayed by teacher-trainees were observed and coded in terms of the types of questions asked of pupils. The ICDS (Lynch and Ames, 1971) was used to record and describe classroom behavior in the cognitive domain. The teacher-pupil categories that are subsumed in the system were described earlier in the report (page 97). There are four low and seven high-level cognitive demands, as well as two miscellaneous and four feedback categories. The data derived from this system should yield a profile of frequency and percentage of occurrence of any category over the entire instructional period.

Coder training. Training procedures for each of the observation systems were developed independently and involved different personnel. A pool of approximately ten trained IBMS observers were available at CITH, and it consisted of center personnel who had previously mastered the system. Each coder had undergone intensive training sessions, and most of the IBMS coders were also experienced coder-trainers. All had criterion test scores that ranged from .79 up to .93. A criterion test was administered during the fourth week of the project as a check of maintenance of coder reliability. During the last two weeks of the project, three students from the K580 methods course who had received IBMS training in connection with other methods course activity, and who had criterion test scores of .80 and above, were selected as coders for the project.

There were fewer trained ICDS coders available at CITH, so that three CITH staff members were specially trained as ICDS coders during the week prior to initial data collection. Through consensus-training workshop sessions, all coder trainees were able to achieve criterion test scores of .80 and above. There were six different ICDS coders used throughout the data collection period. A reliability maintenance test was administered twice during data collection and once immediately following the completion of the study.

In addition to mastering the system to a reliability score of above .80, each ICDS coder underwent training to criterion in "button box" coding.

Both systems used in the present study are normally arranged so that coders pencil-check letter symbols representing each category on data sheets. For automation of the ICDS coding, an on-line button box (4 x 4 with ten numbered buttons, resembling a TOUCH-TONE telephone face) was used. Through transfer of training from letter symbols to numbers, each ICDS coder was able to observe and button-code each observation. The coded data generated was transmitted by telephone cable to the TEL computer and processed.

All 33 practicum students participating in the classroom phase of the CATTS feedback study were arbitrarily divided into three groups for observation system training. There were four two-and-one half to three-hour training sessions over a two-week period.

The instruction was centered around a pre-instructional package developed at the Center for Innovation in Teaching the Handicapped (CITH). The materials for observation training included a programmed training manual, a workbook and several audio and video tapes for instruction and practice. Each session was conducted by an instructor who had previously been through full ICDS training and who had also participated in two retraining workshop* sessions, during which an abridged version was used for instruction of teacher-trainees in the ICDS system.

Training in ICDS included introductory remarks by course instructors on the nature of observation systems and their relevance to teaching skills, a review of all ICDS teacher categories, role-playing, and practice coding from audio- and video-taped materials. Criterion tests were administered

---

*The workshop sessions were organized and conducted by Karen Greenough, whose efforts are much appreciated.

during the last hour of the third session and the beginning of the fourth session. Although coder training was not the purpose of these observation system training sessions, the criterion tests were used to assess the degree of accuracy with which the student trainees were able to discriminate the various categories of classroom questions.

II Measuring baseline teaching behaviors.

Subjects and class organization. The 33 methods class students who participated in the ICDS training were scheduled into a classroom teaching session to conduct a fifteen-minute lesson. The classes utilized for teaching were those sponsored by the University Affiliated Facility (UAF), a residential/educational unit for EMR children, ages 7 - 12. There are three UAF classes located in Building 8 of the I.U. University School complex. The CATTS installation of the CITH Teacher Education Laboratory (TEL) is located in Building 7 of the same University School facility.

For the duration of the project, all lessons were taught between 10:30 and 11:30 A.M. Each day, pupils were randomly assigned to one of the three classrooms.

Three consecutive fifteen-minute lessons were taught each day, with a five-minute break between lessons. Each class contained a maximum of seven pupils.

Lesson assignments. Each teacher-trainee received a lesson topic outline twenty-four hours before the scheduled teaching period. The lesson topic assignments were developed in cooperation with the head teacher at the UAF, and were chosen to augment the ongoing language arts program. An effort was made to establish topical continuity in the lessons taught consecutively.

The lesson topic outline included a listing of primary objectives, the purpose of the lesson, suggested approaches and terminal goals (desired outcome of lesson). Teacher-trainees were encouraged to be creative in lesson planning. It was noted that the approaches suggested in the outline were for assistance in planning, but were not to be interpreted too rigidly. Among topics included were such items as vocabulary building, and size and spatial relationships associated with three concepts. Also included were compound words, prefixes, classification and logical grouping and vowel sounds.

III. Experimental phase.

Overview of study. The first two weeks of the project were used to cycle all thirty-three student trainees through ten hours of instruction in discriminating the categories of the Individual Cognitive Demand Schedule (ICDS). During the third week these same student trainees were scheduled to teach one fifteen-minute lesson. The data gathered from this initial teaching contact was used to describe the baseline of teaching performance obtained prior to any experimental intervention. Following completion of the baseline data phase, each trainee taught three more lessons, under one of three feedback conditions. Twenty-seven of the thirty-three trainees were randomly assigned to one of three groups. Each group received a different feedback condition: instant f.b. (CATTS monitor), delayed f.b. (computer printout after the lesson), and no feedback. The remaining students participated in the baseline and maintenance teaching session, but did not teach the three lessons on which feedback was controlled.

The major objective of the study was to demonstrate the effect of feedback on the trainees' ability to increase specific types of questioning behavior. All of the trainees had as their goal the increase of the number of high-level questions asked during the lessons they taught. The categories of questions which they were asked to increase were inferring, problem-solving, defining-classifying and applying-comparing. The instant feedback displayed on the video monitor, therefore, showed the current cummulative frequency of questions in each of the four cognitive categories.

There was no feedback given during the fifth lesson taught by trainees. The data gathered during this last instructional session was used to determine the maintenance of the teaching behaviors in each of the feedback groups.

Behavior management workshop. There were several factors affecting the decision to interpose a workshop in classroom behavior management techniques between the initial teaching week (baseline) and the first experimental week. First, the educational program at the UAF relies heavily on a program of behavior modification. The teacher-trainees in the project had no previous experience with the technique and many had little knowledge of it. During the first week of teaching, the trainees were briefed on the Behavior Modification technique prior to their entering the classroom. Project assistants described only the rudiments of the program, e.g., the mechanics of appropriately marking "in control" or "out of control" checks on "control" sheets taped to the children's desk, and the role of positive or negative reinforcement in management of behavior. Secondly, the teacher-trainees in the present program were mostly inexperienced in the classroom or in group instruction. For many, this was their first experience ever in a classroom situation, and some encountered behavior management problems in their first attempt at teaching. In addition, the problems arising out of lack of management experience or skills were exacerbated by the extremely heterogeneous nature of the learning and behavioral characteristics of the UAF children. According to the head teacher at the UAF, several of the pupils were persistently hyperactive and volatile, a few had impaired hearing or vision, and a few were markedly withdrawn. In all, about two-thirds of the pupils had serious behavioral problems.

Data-gathering procedures and organization of the teaching practicum under controlled feedback conditions. A substantial effort was required to coordinate the operation of many separate systems that were necessary for the realization of the study. For example, each lesson taught by a trainee was prescheduled, and lesson topics were assigned and ready twenty-four hours prior to the lesson. Lists of the daily randomization of pupils were made available to the head teacher of the UAF, so that random assignment of pupils to classrooms was completed before the first teacher-trainee arrived. Six trained observation system coders were required for each session. The coders were briefed daily on names of teachers, pupils, nature of lesson, etc., prior to their assignment to coding stations. The teacher-trainees were briefed prior to their entering the classroom to teach. Primary technical systems (e.g., CATTS feedback monitor, transmission of coded data to computer, printout for delay f.b. group) described elsewhere in this report, were in operation during each teaching session. Back-up technical systems, such as auxiliary audio and video taping of the sessions, were also incorporated in the total operation of the data-gathering procedure. Contingency procedures were also established in anticipation of a possible breakdown of any of the above systems.

Coding procedures. At least fifteen minutes prior to each instructional session, six coders were present at the Teacher Education Laboratory (TEL) and were briefed on daily procedures. Each coder was assigned (by prescheduled randomization) to a classroom. There was one ICDS and one IBMS coder for each classroom. They were given sufficient data sheets for three consecutive fifteen-minute observations, and also received copies of the lesson outline and the names of pupils in each room for that day. After synchronizing watches, coders entered the observation rooms.

The IBMS coders had the additional task of signaling the waiting teacher-trainee to enter the classroom. A flashing red light on the "button box" indicated that coding transmission would begin in two minutes. After two minutes, a red light again flashed on the "coding device" and both ICDS and IBMS coders began recording observations; the ICDS observer on the button box, the IBMS observer using data sheets. [illegible] data sheets were also available to the coder, in case of system failure. After fifteen minutes the IBMS coder entered the classroom to signal the teacher that the session was over. A five-minute break followed, and then the second session began.

The procedures for the second and third instructional sessions were the same as for the first, with a [illegible] interval between sessions.

Coders returned to [illegible] upon completion of the last lesson, and they checked and completed data sheets which were filed along with the computer tape output for each lesson.

Teacher-trainee role. The student trainees were briefed on the responsibilities and requirements for conducting their lessons during a group meeting and through written communications.

In addition to information concerning scheduling and procedural routines, students were apprised of the nature of the laboratory school (UAF), the behavior modification program, and the names and characteristics of children with special problems.

Before their first lesson, trainees in the immediate f.b. group were told that the frequency of occurrence of any of the four target question categories (Defining-Classifying, Applying-Comparing, Inferring, Problem-Solving) would be flashed on the TV monitor in their classroom. The teacher entered the classroom upon signal from the IBMS coder, spent from two to a maximum of five minutes in acclimatization and introductory matters, and then commenced with the prepared lesson.

Feedback procedures. Each trainee was asked to set as a goal the increase in the number of high-level questions asked during a lesson. They were asked to try to generate increases in four specific categories: Defining-Classifying, Applying-Comparing, Inferring, and Problem-Solving.

The immediate f.b. group taught in a classroom equipped with a TV monitor. The frequency of occurrence of any one of the four target categories was flashed on the screen. The number flashed represented the cummulative frequency from the time the lesson began. The teacher could, therefore, glance up at the monitor and see how many questions each of the four types she had this far asked. Figure 5 shows the video display configuration print as the classroom monitor.

The students in the delay feedback group had a computer printout of frequencies and percentages of all ICDS categories generated during instruction available to them within an hour after teaching. During a group meeting of all students in the delay f.b. group, prior to the experimental phase of the project, an assistant explained the symbols and information available on the printout sheet. In addition, a reference sheet for guidance in print-out interpretation was given to each student in the delay group. The print-out sheet listed all ICDS categories that occurred within a five-minute interval. There were three such five minute intervals. A summary of the frequencies and percentages obtained during the full fifteen-minute period was also included in the printout.

The no-feedback group was encouraged to try to increase the number of high-level questions but did not receive any information about the frequency of occurrence of the four high-level question types they were working toward increasing. The printouts were made available to this group upon completion of the last teaching session.

Maintenance of teaching behaviors. The fifth and last teaching session in the project was conducted in the same manner as the first session. There was no feedback of any type for any of the trainees. The purpose of this session was to see whether any of the increases expected in the groups receiving feedback would be maintained in the absence of feedback of information.

PILOT STUDY II

The Effects of CATTS Feedback on the Question Behavior of Teacher-Trainees in a Simulation Classroom Experience

In design, format and procedures employed, pilot study II was essentially similar to pilot study I. The critical difference however, was that in the present study, student trainees role-played EMR children as a substitute for real children during the experimental (CATTS Feedback) phase.

The purpose of this study was to determine whether CATTS was an effective training method when trainees instruct peers instead of children. The creation of a simulated classroom through role-playing was expected to provide teaching experiences and an opportunity to develop teaching skills when children and real classroom settings are unavailable to the trainee.

Overview

All of the trainees in this study were juniors majoring in special education and members of an EMR methods class. They were randomly selected from the class to participate in the study.

The design of the study provided that all trainees would teach five, fifteen-minute lessons. The trainees were divided into three experimental groups; an immediate feedback group, a delayed feedback group (in the form of a computer printout), and the no-feedback group. The initial teaching session provided the baseline data, the next three teaching sessions were treatment phases, and a final session was held to determine the maintenance of behaviors.

The study utilized a simulated setting in which trainees role-played educable mentally retarded children. The trainees assumed the roles prescribed for them in a written script. These scripts supplied information on the emotional, social, physical, and academic characteristics of an EMR child. Trainee peers were asked to portray the roles as they perceived them from the information on the scripts. (See sample script, Appendix). In the baseline phase of the simulation study, the children in the University Affiliated Facility (UAF), a residential educational center for primary EMR and intermediate children, were taught by the teacher-trainees. During the three treatment lessons, teaching took place in the simulated EMR classroom setting. In the maintenance phase, EMR classes at the UAF were used again.

Trained observers coding the ICDS and IBMS systems recorded the cognitive demands and behavioral control strategies of all the trainees in each stage of the study.

The specific purpose of the study was to demonstrate the effects of CATTS immediate visual feedback on increasing the level of the trainee's cognitive demands (questions) in a fifteen-minute teaching session. It was hypothesized that those trainees receiving the immediate visual feedback (CATTS) would produce more high-level cognitive demands than trainees receiving either delayed feedback (printout), or trainees receiving no feedback.

A workshop was instituted to better acquaint the trainees with the conceptual as well as the practical aspects of the Individual Cognitive Demand System. During the workshop the theoretical basis of the ICDS was presented and explained in detail. The trainees were given several opportunities to conceptualize the different levels of demands (high and low) by improvising high- or low-level cognitive demands. Role-playing was utilized to give the trainees practice in using the ICDS. Each trainee was asked to assume the role of teacher, coder, and student.

In sum, there were four phases of the CATTS simulation study; first, the training of teacher-trainees on the ICDS observation system; second a baseline ICDS measurement of the trainees questioning behavior with the EMR children, thirdly, the experimental phase in which CATTS immediate visual feedback, delayed feedback, and no feedback were provided in a simulated class setting, and fourthly, the maintenance phase (with EMR children), in which no feedback was provided in order to determine whether the levels of the questioning behavior obtained under the various feedback conditions were maintained.

Observation System Training

All the teacher-trainees participated in a one-day observation system training workshop. The workshop was an intensive session designed to acquaint the practicum trainees with the Individual Cognitive Demand System. The workshop began with all participants reading an abridged version of the ICDS training manual, containing only the definitions and a brief explanation of each category. The trainees were instructed to read the abridged training manual until they felt they had an understanding of the ICDS categories. They were then required to write definitions for each ICDS category and their abbreviations. Each trainee was subsequently given a slip of paper with an ICDS category on it and instructed to generate the questions or demands indicated. This activity was continued until the trainees could identify the question types by category. It was expected that constant repetition of ICDS categories in simulated situations (role-playing) would be the most effective means of category retention. Thus, the teacher-trainees role-played for several hours. Each trainee assumed the role of teacher, student, and observer (coder). After each role-playing session a discussion was held. Role-playing continued until trainees felt that they could generate examples of specific cognitive demands given by either the trainer or their peers. Trainees were then asked to simulate a specific classroom within a given setting and with a given lesson topic.

Included in the lesson were specific cognitive demands to be made in the five-minute lesson simulation. Every trainee participated in the simulated class experience. Discussions were held after these sessions and definitional problems discussed. At the end of the training sessions oral quizzes were administered in order to measure the degree of accuracy of discrimination of the various categories of the ICDS. The trainers of the teacher-trainees were skilled observers who had participated in several weeks of observer training on the ICDS. All had obtained above .80 reliability as ICDS coders.

Baseline Measurement of Teaching Behavior:

Simulated Subjects and Classroom Organization

The teacher-trainees in the baseline phase of the study taught in the University Affiliated Facility primary and EMR intermediate classrooms. There were three classrooms, with a total of twenty-one children, located in Building 8 of the University Elementary School complex. In the three treatment phases the classrooms used for simulations were located in the Teacher Education Laboratory. Fourteen different EMR roles were assigned to the role-playing trainees each day. The roles were in script form and gave detailed descriptions of the physical, social and emotional, and academic characteristics of a ficticious EMR child. The trainees were asked to portray the roles as they perceived them. Simulators role-played for an hour each day for four weeks enabling the teacher-trainees to complete the three treatment phases.

Behavior Management Workshop

The teacher-trainees participated in a behavior management workshop prior to the baseline teaching session. Since most of the trainees did not have any previous experience with educable mentally retarded children, it was anticipated that the trainee would need some guidelines in classroom management.

The purpose of the workshop was to give the trainees specific techniques on behavior management. Problem-solving techniques for specific behavior problems were presented by a task analysis method. The trainees were urged to carefully consider any behavioral problem situation and to try to give more positive feedback. Allowing for individual differences in children was strongly emphasized.

In order to give trainees an overview of the types of behavioral controls used by the UAF teachers, the two classroom teachers spoke at the workshop. They explained the behavior modification system used and provided examples of when to use positive or negative controls. Video tapes of the EMR children in their usual classroom environment were shown. The teachers described the "in control"/"out of control" sheets and procedures to be used by the teacher-trainees in the classroom. In addition to the information on behavior controls, the UAF teachers gave a brief description of the childrens' home background and academic problems.

The last phase of the workshop concentrated on specific behavioral controls that could be used with the children. One of the trainers who had five years experience with EMR children guided this discussion. It was pointed out that voice changes as a controlling measure are effective in instituting "on-task" behaviors in the deviant children. Emphasis was placed on the positive aspects of controlling off-task behaviors. Trainees were urged to be more probing and emphatic-sympathetic in dealing with off-task behaviors.

Lesson Plan Assignments

The curriculum topics that were used in this study were arrived at by discussions held with the head teacher of UAF. The lesson plans were a continuation of the curriculum areas that were stressed in UAF classrooms. Lesson plans were given to the teacher-trainee 24 hours before the assigned teaching period. The lesson plans format was behaviorally based. The format began with behavioral objectives to be stressed in the lesson. The purpose of the lesson, various possible approaches, and terminal behavioral goals were included in each lesson. Topics that were covered in the curriculum were science, perceptual skills, oral and written language and listening skills. The lessons were constructed so that there was a conceptual framework from which the trainee could proceed. The lesson ssignments encouraged individual creativity in planning by the teacher-trainee.

Experimental Phase

The twenty-four trainees participating in the experimental phase were randomly placed in three equal groups. One group received immediate visual feedback. The second group received delayed feedback in the form of a computer printout, and the third group received no feedback of any type. The objective of the experimental phase was to demonstrate the effectiveness of visual feedback, delayed feedback, and no feedback on the teacher-trainees' ability to increase specific types of cognitive demands. The goal for all trainees was to increase the frequency of four ICDS categories: value-judging, inferring, defining-classifying and applying-comparing.

Those trainees in the immediate visual feedback group viewed the TV monitor display which showed the frequency tally of the four ICDS categories that were to be increased by the trainees. The trainees in the delayed feedback group received feedback in the form of a computer printout, within 2 hours of their assigned teaching session. The printout showed the frequency tally of all thirteen ICDS categories and a summary of the four specific categories to be increased. The third group of trainees received no feedback of any kind.

Coding Procedures

Coders were randomly assigned on a daily basis to classroom observation rooms. There were three classrooms being observed and two observers coding in each. There was one ICDS and one IBMS coder in each observation area. Coder scheduling was randomized so that no two observers would code consistently together.

All coders met 15 minutes prior to the initial scheduled teaching session. They were then given coding packets which consisted of a daily class list, lesson plans to be taught by trainee, trainees names, paper coding booklets and room number assignment. Stop watches were given to all observers as well as the computer operator, and these were synchronized in order to assure that everyone began coding at the same time.

Observers then went into observation rooms and started an audio as well as a video recorder machine. The computer operator then flashed a one-minute warning signal on a button box. This reminded the IBMS observer to go to the classroom and tell the trainee to begin the teaching session. Then the IBMS and ICDS observers began coding. After fifteen minutes of continuous coding by the IBMS and ICDS coder, the teaching session terminated. The IBMS coder then went to the classroom and signaled the teacher-trainee to terminate the teaching session. The coders then had five minutes between sessions to check all technical hardware (audio and video tape recorder, button boxes). The second and third teaching sessions employed the same coding procedures as the first teaching session.

Teacher-Trainee Role

Trainees arrived at the TEL fifteen minutes before their assigned teaching time. They were given a packet which contained a class list of the children in their classroom, room number, and behavioral control sheets for the behavior modification system normally used by the classroom teachers. Trainees were then briefed on any changes in the classroom, i.e., if a child was sick or if there had been any emotional upsetting experience earlier that morning for any child. Trainees were also informed of options available if children went out of control and needed to be removed from the classroom. They were also told that they would have behavior control support from a crisis teacher.

Trainees in the immediate visual feedback group were reminded to look at the CRT during the experimental phase of the study. Trainees in the delayed feedback group were reminded that they could pick up their delayed feedback forty-eight minutes after they taught their lesson.

The teacher-trainees went to their assigned classrooms five minutes before their teaching time to get acquainted with the children. The IBMS coder then signaled the trainee to start the teaching session. The teaching session terminated when the IBMS observer opened the classroom door.

Feedback Procedures

The ICDS observers collected data by entering observations on the button box connected to the computer.

Each trainee's goal was to increase the frequency of occurrence of the four ICDS categories (VJ, AC, IN, DC) during the fifteen-minute teaching session. Thus, the teacher-trainee concentrated upon generation of high-level cognitive demands.

As the teacher-trainee taught the lesson during the experimental phase, a coder entered the observation by entering numbers on the button box. The real-time data processing continuously updated the CRT display. Thus, each time a trainee generated a demand in either of four specified categories, it was displayed on the CRT unit. The trainee then was able to see the frequency of high-level demands he made while teaching. Trainees who received delayed feedback also had as their goal the increase of high-level demands, as did trainees who received no feedback. Within two hours after the termination of the teaching session these trainees obtained printouts of the codes of their lesson. The delayed feedback, in a form of a computer printout, showed frequencies and percentages of ICDS categories generated by the trainee during the fifteen-minute teaching period. Trainees in this group received a reference sheet to the printout which explained the number tallies and category summaries. The printout had the data summarized in three five-minute segments. All thirteen categories of the ICDS used by the trainee were summarized by frequency and percentage.

Those trainees in the no-feedback group were encouraged to generate higher level demands during their teaching session, but never received any information on the frequency of occurrence of the higher level demands that they generated during their teaching sessions. However, this group had an opportunity to review video tapes of their teaching experiences at the end of the study.

Maintenance of Teaching Behaviors

The maintenance lesson was the fifth time that the trainees taught and was identical to the baseline or first lesson. The same procedure, lesson plans and coding techniques were implemented. The trainees did not receive any type of feedback but were encouraged to generate higher levels demands (DC, AC, VJ, IN). The purpose of the baseline lesson was to see if any of the increases expected in the groups receiving feedback would be maintained in the absence of feedback.

January 25, 1973

CATTS V Schedule: Teaching Times for Groups X, Y, Z

	Session	
Monday:	A	9:30-10:00
	B	10:30-11:00
Tuesday:	C	1:00-1:30
	D	2:00-2:30
Wednesday:	E	1:00-1:30
	F	2:00-2:30
Thursday:	G	9:30-10:00
	H	10:30-11:00
Friday:	I	12:45-1:15

PILOT STUDY REFERENCES:

References

1. Fink, A. H., & Semmel, M. I. Indiana Behavior Management System II. Bloomington: Indiana University, Center for Innovation in Teaching the Handicapped, 1971.

2. Flanders, N. A. Analyzing teaching behavior. Reading, Mass.: Addison-Wesely, 1970.

3. Gagnè, R. M. The conditions of learning. New York: Holt, Rinehart & Winston, 1967.

4. Lynch, W. W., & Ames, C. Individual Cognitive Demand Schedule. Bloomington: Indiana University, Center for Innovation in Teaching the Handicapped, 1971.

5. Rosenshine, B., & Furst, N. The use of direct observation to study teaching. In Travers, R. M. (Ed.), Second handbook of research on teaching. Chicago: Rand-McNally, 1973.