1976

The effects of rules and instructions, consultant feedback, and teacher self-monitoring on teacher and student behavior during consulting and non-consulting periods

Kenneth Slater Roach

THE EFFECTS OF RULES AND INSTRUCTIONS, CONSULTANT FEEDBACK,

AND TEACHER SELF-MONITORING

ON TEACHER AND STUDENT BEHAVIOR

DURING CONSULTING AND NON-CONSULTING PERIODS

BY

KENNETH SLATER ROACH

A THESIS
SUBMITTED TO THE GRADUATE FACULTY
OF THE UNIVERSITY OF RICHMOND
IN CANDIDACY
FOR THE DEGREE OF
MASTER OF ARTS
IN PSYCHOLOGY

AUGUST 1976

THE EFFECTS OF RULES AND INSTRUCTIONS, CONSULTANT FEEDBACK,

AND TEACHER SELF-MONITORING

ON TEACHER AND STUDENT BEHAVIOR

DURING CONSULTING AND NON-CONSULTING PERIODS

By Kenneth Slater Roach

Approved:

Edward H. Tilley
Chairman

Jean N. Dickinson

William Edwin Walker

ACKNOWLEDGEMENTS

Finally, Mrs. Katherine B. Roach, the writer's wife, is gratefully thanked for her time in checking computations and typing tables. More importantly, her frequent praise and encouragement and her total confidence are sincerely appreciated.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract

This study investigated the effects of rules and instructions, consultant feedback, and self-monitoring on teacher approval, disapproval, and student on-task behavior. Data was collected during a consulting period, measuring changes when the consultant was present, and during a non-consulting period, assessing whether similar changes occurred with the consultant absent. Three elementary teachers who exhibited more verbal disapproval than approval participated. On-task data was collected on three randomly selected students in each classroom. Following baseline, the teachers set classroom rules and were instructed to increase their approval and decrease disapproval. During the consulting period of the feedback phase, the consultant provided feedback every five minutes to the teacher on the frequency of her approvals and disapprovals. The teachers counted their approvals on a wrist counter during both periods of the self-monitoring phase and continued receiving feedback during the consulting period. Follow-up data was collected after the fourth phase.

A two by five (periods by phases) repeated measures analysis of variance revealed that the training program significantly increased teacher approval and student on-task behavior and decreased teacher disapprovals. Approvals increased significantly during each of the three training

phases. The nonsignificant interaction and significant phases effects showed that changes in teacher approval during the consulting period were not different from changes during the non-consulting period across phases. Throughout the study, though, the teachers approved significantly more during the consulting than non-consulting period. Follow-up data indicated the teachers were approving less than during the feedback and self-monitoring phases but more than during the rules phase.

CHAPTER I

Introduction

When a teacher is unable to change the inappropriate behavior of students in her classroom, she may ask a consultant to observe these students and make recommendations on how to manage them. The consultant often finds that the teacher is attending to what she does not want the students to do, through nagging or scolding, and that she seldom pays attention to desirable behaviors, such as working on assigned academic tasks. The consultant may recommend that the teacher ignore the students' inappropriate behaviors, since they may be misbehaving to obtain teacher attention. He may further recommend that the teacher praise or pay attention to appropriate student behaviors so as to increase their frequency.

Teachers have employed a variety of reinforcers to increase appropriate student behavior, including candy, free time, privileges, and many forms of teacher attention (Pinkston, Reese, LeBlanc, & Baer, 1973). Verbal praise, a form of teacher attention, has become the mainstay of many teachers' repertoire of reinforcers and a reinforcer recommended frequently by consultants. Verbal praise costs nothing, can be dispensed easily and immediately from many geographic locations throughout the classroom, has reinforcing effects for practically all students, and fits into a teacher's desire for a positive classroom atmosphere (Madsen & Madsen, 1974).

Studies in applied behavior analysis, the data-based systematic manipulation of antecedents and consequences of behavior in everyday environments, have demonstrated the effectiveness of contingent teacher praise in changing student behavior (Hall, Lund, & Jackson, 1968; Madsen, Becker, & Thomas, 1968). Hall et al. used contingent teacher praise to increase the frequency of student behavior from a baseline average of 25% to a treatment level of 79%. Madsen et al. (1968) compared rules, praise, and ignoring, and found praise to be the key to increasing appropriate student behavior.

Many teachers understand the effectiveness of praise and can quickly recite, "Ignore the bad; praise the good." Yet, according to Madsen and Madsen (1974), only 8% of classroom teachers employ more verbal approval than disapproval. When White (1975) investigated the praise rates of first through twelfth grade teachers, she also found that most teachers disapprove more than they approve. In her study only the first and second grade teachers used more verbal approval than disapproval. Praise is a necessary part of effective classroom management but, apparently, difficult to do. Programs have been developed recently to train teachers to increase their frequency of praising. Data from these training programs indicate that teachers can learn to praise effectively in relatively brief periods of time and increase appropriate student behavior to within acceptable levels (Horton, 1975).

Investigations of training teachers to praise may be placed into two major categories, based on whether the type of training occurs primarily outside or within the classroom. Studies in training outside the classroom have utilized inservice programs to provide teachers with management skills. Madsen, Madsen, Saudargas, Hammond, Smith, and Edgar (1970) presented behavioral principles to teachers during a two-week, summer inservice workshop, and later, when school began, provided feedback to the teachers on both teacher and student behaviors. The workshop stressed the RAID approach (rules, approval, ignoring, and disapproval) through lectures, discussion, role-playing, and video tape discrimination training. While approval included written words, physical expressions, closeness, activities, and things (e.g., food, tokens), verbal praise was emphasized. Teachers who completed the workshop exhibited over twice as much approval to appropriate student behavior than teachers who had not attended. Teachers who used four approvals to every disapproval (4:1 approval/disapproval ratio) had students with the highest rates of on-task behavior and also spent less class time approving and disapproving. In other words they praised more, yet talked less. In order to achieve a 4:1 approval/disapproval ratio, most teachers not only have to increase their rates of praising, but also decrease their disapproval rates.

Teachers also have been trained to increase their frequency of praising while they teach. The most frequent type of this on-the-job training has been to provide the teacher with feedback on the amount of praise she is using while teaching. Cossairt, Hall, and Hopkins (1973) and Parsonson, Baer, and Baer (1974), in contrast to Madsen et al. (1970), spent very little time discussing behavioral principles and definitions outside the classroom. Cossairt et al. (1973) investigated the effects of instructions, feedback at the end of a session, and feedback plus praise at the end of a session. Instructions and feedback produced inconclusive results, but the feedback phase had to be prematurely terminated, preventing a comparison of feedback with feedback plus praise. Feedback plus praise at the end of the sessions produced the highest rates of teacher praise. Parsonson et al. (1974) provided feedback more often than just at the end of the training session and without the praise used by Cossairt et al. (1970). After every 15 teacher attention responses, the consultant simply handed the teacher a slip of paper on which was written the percentage of attention responses given to appropriate and to inappropriate behavior. This feedback came every three to five minutes. After five to eight days of training the teachers were attending to appropriate student behavior over 80% of the time. Thus, providing a teacher with frequent feedback can be used to increase her

rate of praising. Furthermore, it is more effective than praising a teacher's performance.

From this information on teacher training programs it appears that training both outside and within the classroom has changed teacher behavior. While Madsen et al. (1970) primarily employed training outside the classroom, Cossairt et al. (1973) and Parsonson et al. (1974) demonstrated that feedback within the classroom alone would increase teacher approval.

The effectiveness of the training programs has been evaluated by changes in teacher and student behaviors during the training observation sessions. Most training programs consisting of consultant feedback have brief, daily contacts with the teacher over a set period of time or until a criterion has been met. However, it also is important to know what happens to both teacher and student behavior during the remainder of the day when the consultant is not present to provide the teacher with feedback. An effective training program would increase teacher praise throughout the day, enabling the teacher to manage student behavior at all times, not just in the presence of the consultant. Since consultant feedback has been an important variable in training teachers to praise, the presence of the consultant may prompt teacher praise and the absence may remove key praise eliciting stimuli. However, many stimuli remain the same regardless of the presence or absence of the consultant, including the teacher, the classroom, the students, and teaching techniques (e.g.,

reviewing classroom rules at the beginning of an academic period). If the teacher responds in a similar way (by praising) to situations other than the consulting one, stimulus generalization occurs. Generalization may take place when the teacher fails to discriminate between the praise eliciting stimuli of the consulting and non-consulting periods.

Few studies have collected daily measures during both a consulting and non-consulting time to see if behavior changes during the training are exhibited outside of the consulting session (Loeber & Weisman, 1975). These studies suggest that generalization outside of the consulting session may not occur without specific programming (Horton, 1975; O'Leary, Becker, Evans, & Saudargas, 1969; Wahler, 1969). Wahler and his associates worked with two children who showed similar inappropriate behavior both at home and at school. While contingency changes within the home successfully modified inappropriate behavior there, the children's behavior did not change in the unprogrammed school environment.

O'Leary et al. (1969) studied generalization between morning and afternoon sessions within the same classroom. In this study a teacher administered a token economy during the afternoon but did not administer it in the morning, even though the same students remained throughout the day. During the morning the teacher was requested to set the same rules, praise appropriate behavior, and

disapprove infrequently, as performed in the afternoon while administering the token economy. Increased appropriate student behavior during the afternoon token program failed to generalize to the non-token morning session. More important to teacher training, the teacher's behavior varied greatly between the morning and afternoon sessions, indicating teacher behavior also failed to generalize. For example, the teacher only responded to questions from students with raised hands during the afternoon (a classroom rule), but during the morning she answered questions from students who had not obtained permission to speak. This research demonstrated that both the teacher and the students acted differently in the afternoon token program than they did in the non-token morning sessions. It appeared that the teacher and the students discriminated between the stimuli in the two sessions.

One study (Horton, 1975) investigated generalization of teacher praise across academic subjects within the same classroom. However, Horton's training program was quite different from the research previously reviewed on teacher training. Horton used a video tape discrimination training technique outside the classroom to train teachers to discriminate instances of behavior specific praise from non-instances. The initial video tape contained examples of behavior specific praise taken from a reading class conducted by an anonymous teacher. After initial discrimination

training, the teachers increased their praise rates
during the reading period only, as assessed by a multiple
baseline design collecting data across all academic
subject areas. In order to generalize praise rates
across subjects, Horton had the teacher receive video
tape discrimination training with examples of behavior
specific praise from all academic subjects. After this
training, the teachers increased their praise rates
throughout all academic periods. Although this is an
effective training procedure, it has several drawbacks.
Not all school systems can purchase the necessary video
tape equipment, and the training requires much time out-
side of the classroom, an aversive event for many teachers.
Furthermore, if the teacher needs help in maintaining
praise rates, the teacher has to look outside the every-
day resources, back to the consultant, for video tape
retraining.

In summary the above data support the statement by
Baer, Risley, and Wolf (1968) that generalization must be
programmed in order to be certain it will occur. Loeber
and Weisman (1975), in their review on the training of
trainers, suggested self-control methods as a means of
programming generalization. Self-monitoring, a type of
self-control (Thoresen & Mahoney, 1974), was employed by
Thoresen, Hubbard, Hannum, Hendricks, and Shapiro (1973)
in a program training teachers to increase their praising.
These investigators trained teachers to self-monitor their
behavior within the classroom by having them count each

of their praise responses on a wrist counter. By simply
pressing a lever on the counter, the teacher kept a
cumulative count for feedback. The study compared
three types of self-monitoring training: 1) minimal
training consisting of asking the teacher to record each
praise response on the wrist counter; 2) daily feedback
on the teacher's accuracy of self-monitoring; 3) a
detailed modeling and video tape training session. While
the latter two types of training increased teacher accuracy
of self-monitoring, they produced relatively modest in-
creases in praise.

The purpose of this research was to investigate the
effects of a teacher training program on teacher and
student behavior. This study measured these behaviors
during a consulting and non-consulting period to see if
changes when the consultant was present (consulting period)
occurred when he was absent (non-consulting period).
Teacher approval was recorded when the teacher verbally
praised a student's appropriate behavior. Since many
studies have demonstrated that teacher approval for a
target student behavior increases that behavior, student
on-task behavior was monitored. On-task behavior con-
sisted of behaviors appropriate to the assignment or in-
structions of the teacher, including writing, reading,
listening to the teacher, raising a hand for permission
to speak, and following rules.

The training program was implemented in three cumulative phases during the consulting periods when the consultant was present. During the first phase (<u>rules</u> <u>and</u> <u>instructions</u>), the teacher implemented the RAID approach. It was hypothesized that teacher praise would increase during the consulting periods as compared with the baseline data. The predicted results for the study are presented in Figure 1. The consultant provided frequent feedback to the teacher within the classroom during the second training phase (<u>rules</u> <u>plus</u> <u>consultant</u> <u>feedback</u>). Teacher praise, it was predicted, would increase over the preceding phase. During the final training phase (<u>rules,</u> <u>consultant</u> <u>feedback</u> <u>plus</u> <u>self-monitoring</u>), the teacher counted her own frequency of praising on a wrist counter and continued setting rules and receiving consultant feedback. Since the teacher should have attained a proficient level of praising during the preceding phase, it was predicted that self-monitoring would not alter the rate of teacher approval.

This research also investigated some of the stimulus conditions responsible for eliciting and supporting teacher praise in a non-consulting period when the consultant was absent. Generalization, based on a stimulus control model, occurs when a behavior supported by a set of stimuli in one situation is elicited by stimuli in another situation. The more similarity between the stimuli in the two different situations, the more likely the
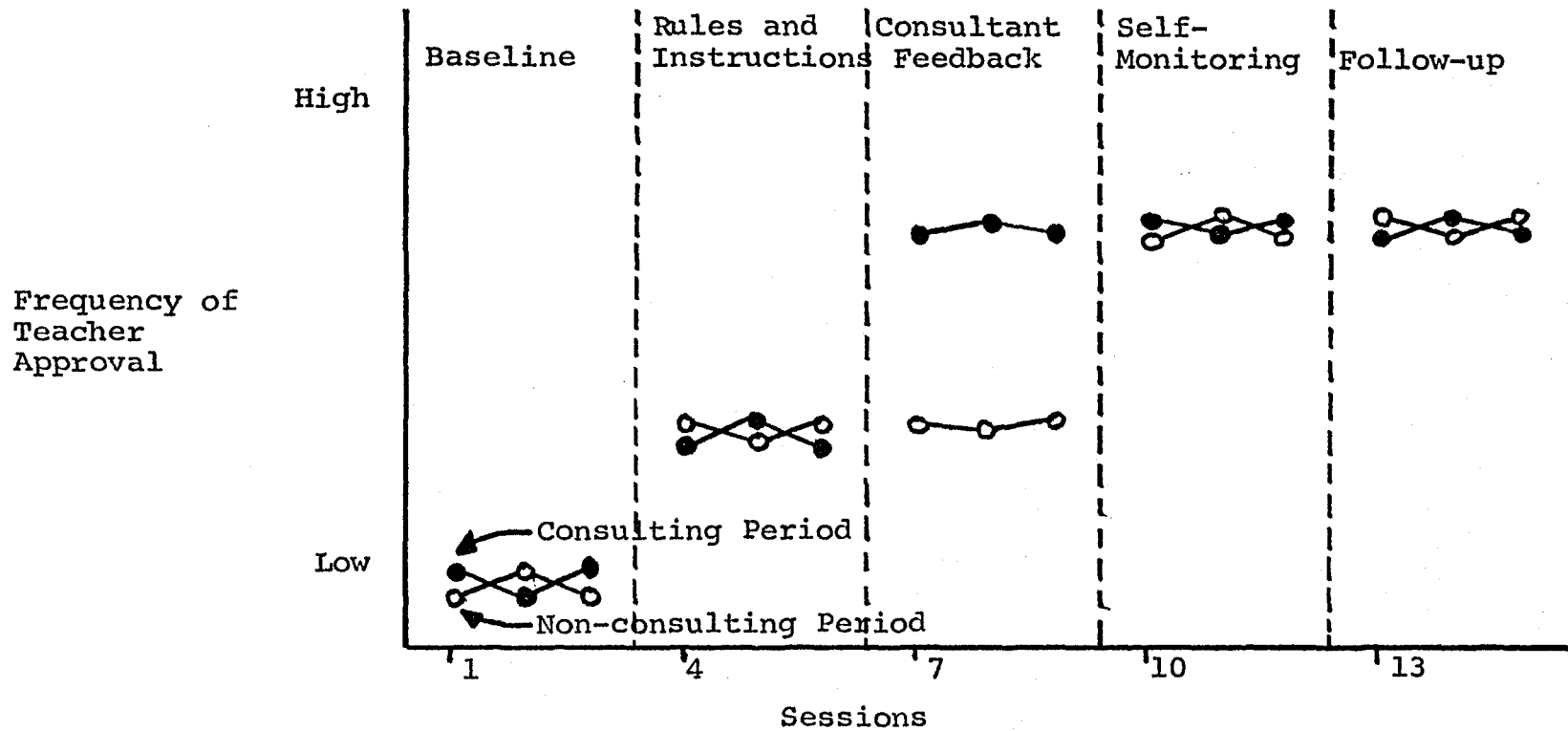
Figure 1. Predicted frequency of teacher approval during each session.

subject will behave similarly in both situations. Teacher praise will generalize outside of the consulting periods, according to this model, when praise eliciting stimuli present in the consulting period have been duplicated or programmed in the non-consulting period.

During the first training phase (<u>rules</u> <u>and</u> <u>instructions</u>), the teacher set rules both during consulting and non-consulting periods, so as to program rule setting across the two situations. It was predicted that the rate of teacher approval during the non-consulting periods of the rules phase would increase over the baseline rate and not differ from the approval rate during the consulting periods of this first training phase (Figure 1). No changes in the programming of teacher behavior took place in the non-consulting periods of the <u>rules</u> <u>plus</u> <u>consultant</u> <u>feedback</u> <u>phase</u>. It was hypothesized that teacher approval during the non-consulting sessions would not change in comparison with the previous phase. Furthermore, the frequency of teacher approval during the non-consulting periods of the feedback phase would be much less than during the consulting periods of the same phase.

The teacher self-monitored her praise frequency on a wrist counter during the non-consulting periods of the <u>rules</u>, <u>consultant</u> <u>feedback</u> <u>plus</u> <u>self-monitoring</u> phase, but without consultant feedback. Feedback has been a key variable in training programs, and self-monitoring duplicated feedback stimuli from the consulting to the

non-consulting period. It was predicted that teacher approval during the non-consulting periods would increase over approval in the same periods of the previous phase. It was further predicted that teacher approval during the non-consulting periods of the self-monitoring phase would not differ from approval during the consulting periods of this same phase.

Follow-up data was collected during the consulting and non-consulting periods after a single academic day break between the self-monitoring phase and this final one. Research has demonstrated that teachers maintain their praise rates after feedback has been discontinued (e.g., Greenwood, Hops, Delquadri, & Guild, 1974; Parsonson et al., 1974). It was predicted that teacher approval would remain at the same rate during follow-up as had been recorded during the self-monitoring phase.

In addition to data on teacher behaviors, data was collected on student on-task behavior during consulting and non-consulting periods throughout the study. Research has indicated that contingent teacher praise increases student on-task behavior (Hall et al., 1968; Madsen et al., 1968). It was hypothesized that on-task behavior would increase during each period as teacher behavior predictably increased.

CHAPTER II

Method

Subjects

Teachers. Three elementary classroom teachers from
a single school in suburban Chesterfield County, Virginia,
were selected for inclusion in this study according to
the following criteria: 1) the teacher volunteered to
participate in research on classroom management; 2) the
teacher was experiencing difficulties in classroom manage-
ment; 3) each teacher's verbal consequences to student
behavior contained at least 50% disapproval comments. The
three teachers had eight, nine, and twelve years of
experience.

Each teacher signed a written contract with the re-
searcher indicating she volunteered for the study, that
she had the right to withdraw from the experiment, and
that she would meet with the consultant for five, one-hour
consultation sessions, the final one being a de-briefing
meeting (Appendix A).

Students. Three students were selected randomly in
each classroom to be observed throughout the study. Other
students were added when target students were absent.

Dependent Teacher Behavior

Data was collected on teacher approval and disapproval.
The response definitions were adapted from Madsen and
Madsen (1974) but limited to verbal behavior only.

Approval was recorded when the teacher verbally praised appropriate academic or social student behavior. Approval for appropriate academic behavior occurred when the teacher stated that a student's academic response was correct. Examples: "Mark, you spelled that word correctly," or, "Yes, Joan, you are right." The definition excluded "OK" and "alright," except when clarified by further remarks from the teacher. Approval for social behavior indicated the teacher praised the social behavior of a student or group of students. Examples: "This reading group has been so quiet," or, "I like the way John is sitting," or, "Mary, you followed the rules, you may speak now."

Disapproval included teacher criticisms of both academic and social student behavior. Disapproval of academic behavior occurred when the teacher indicated verbally that a student's academic response was incorrect. Examples: "You spelled 'cat' wrong," or, "Jack, you know three plus two does not equal six." If the teacher criticized the inappropriate social behavior of student(s), a disapproval was recorded. Examples: "Be quiet!" or, "Sh-h-h," or, "Get back in your seat."

Dependent Student Behavior

Student on-task behavior consisted of following classroom rules and performing the assignments or instructions of the teacher. The student was following rules when complying with the specific classroom rules (e.g., talking

only after obtaining permission, or when raising a hand
to gain permission) and when obeying general classroom
rules (e.g., walking in the classroom as opposed to
running). The student was performing an assignment when
reading orally when asked by the teacher, looking at a
book, writing at the appropriate desk or table on paper,
or listening to the teacher. The student was considered
on-task when conforming to the teacher's instructions,
for example, by picking up the other students' papers.

Apparatus

Each observer was cued to observe and record from a
tape in a standard cassette tape recorder. The observer
listened through an earphone attachment to prevent the
students and teacher from hearing the tape. When inter-
observer reliability was taken, both observers listened
to the same tape recorder by separate earphones attached
to the single output connection on the recorder by a Y
couple. The earphones and Y couple had male miniature
plugs, while the couple had two phono jack female re-
ceivers. Adaptors were placed on the female ends so as
to accept both earphone plugs. An extension cord between
one earphone and the Y couple allowed the observers to
sit apart from one another.

The teachers counted their praise remarks on an Ajay
Scorekeeper (Par Golfer) by Ajay Enterprizes Corporation.
The counter, which resembles a wrist watch, is used
primarily by golfers to count their golf strokes.

Observational Procedures

Data was collected during the language arts and math periods, generally considered the most important academic periods. For one teacher, language arts was at 10:00 and math at 1:15. Language arts was at 9:30, with math at 12:10 for a second teacher, while the final teacher held language arts at 9:30 and math at 11:15. Three undergraduate, paid observers collected the data on both teacher and student behaviors throughout the study, with each observer assigned to a particular teacher according to which times of the day fit the observer's schedule best.

The observers used an interval recording method for recording teacher behavior and a time sampling method for student behavior. Both teacher and student behavior were observed within a 20 second period. An audio tape signaled the observers by an earphone to aid them in accurate observing and recording. The tape announced "observe" to start each 20 second interval, and the observer listened to the teacher for the first 10 seconds. At the conclusion of the 10 second interval, the tape told the observer to record on the record form (Appendix B) those behaviors which occurred. More than one category could be exhibited within the 10 second interval. In announcing the end of the 10 second interval, the tape signaled the observer to record the behaviors on the record form according to the particular minute of the observation session and to which of the three intervals

during that minute to mark. After allowing four seconds
for the observer to record teacher behavior, the tape
announced "count," and the observer quickly counted the
number of the three target students who were on-task,
counting from right to left. The observer had six
seconds to count the students and record the number on
the record form. A new interval started 20 seconds after
the previous one began, allowing three recordings per
minute of teacher and student behavior.

After each five minutes of observing and recording,
a one minute break allowed the observers to rest and the
consultant to provide feedback to the teacher in two
phases of the study. The one minute breaks were announced
on the tape during both math and language arts periods
throughout the study. Each observation session consisted
of four, five minute blocks of observation. Counting
the one minute breaks after the first three blocks of
observation, each session lasted 23 minutes.

Reliability. Each observer was trained to agreement
criteria before the study began and was checked by another
during each phase to ascertain that data was collected
accurately and objectively. The reliability of the data
of teacher and student behavior was computed separately.

The reliability of teacher behavior was computed five
different ways, all based on the following formula:

$$\frac{\text{agreements}}{\text{agreements} + \text{disagreements}} \times 100\% = \text{percentage of agreement.}$$

First, total interval (T-I) agreement was computed by
considering an agreement as an interval in which both
observers recorded both teacher behaviors similarly.
That is, both approval and disapproval had to be recorded
the same within an interval for an agreement to occur.
The observers were trained to an 85% agreement criterion
before the study began. For comparison with T-I data,
a more frequently employed reliability measure, category
agreement, was computed. Each category of teacher be-
havior was considered separately during each interval
in determining agreements or disagreements (category
approve, C-A, and category disapproval, C-D). The mean
reliabilities for these computations are presented in
Figure 2. T-I agreement was a more stringent measure of
reliability than category agreement.

Since interval data is affected by the rate the
particular behavior occurs, the reliability of each
teacher behavior was calculated separately by computing
the mean of the scored and unscored interval reliabilities
(Hawkins & Dotson, 1975). First, the scored interval (S-I)
reliability was computed by considering only those inter-
vals on which one or both of the observers scored the
particular behavior, disregarding all intervals not scored
for that behavior by both observers. Second, the reliability
of the unscored intervals (U-I) was calculated from those
intervals on which one or both observers did not score the
particular behavior, disregarding intervals on which both
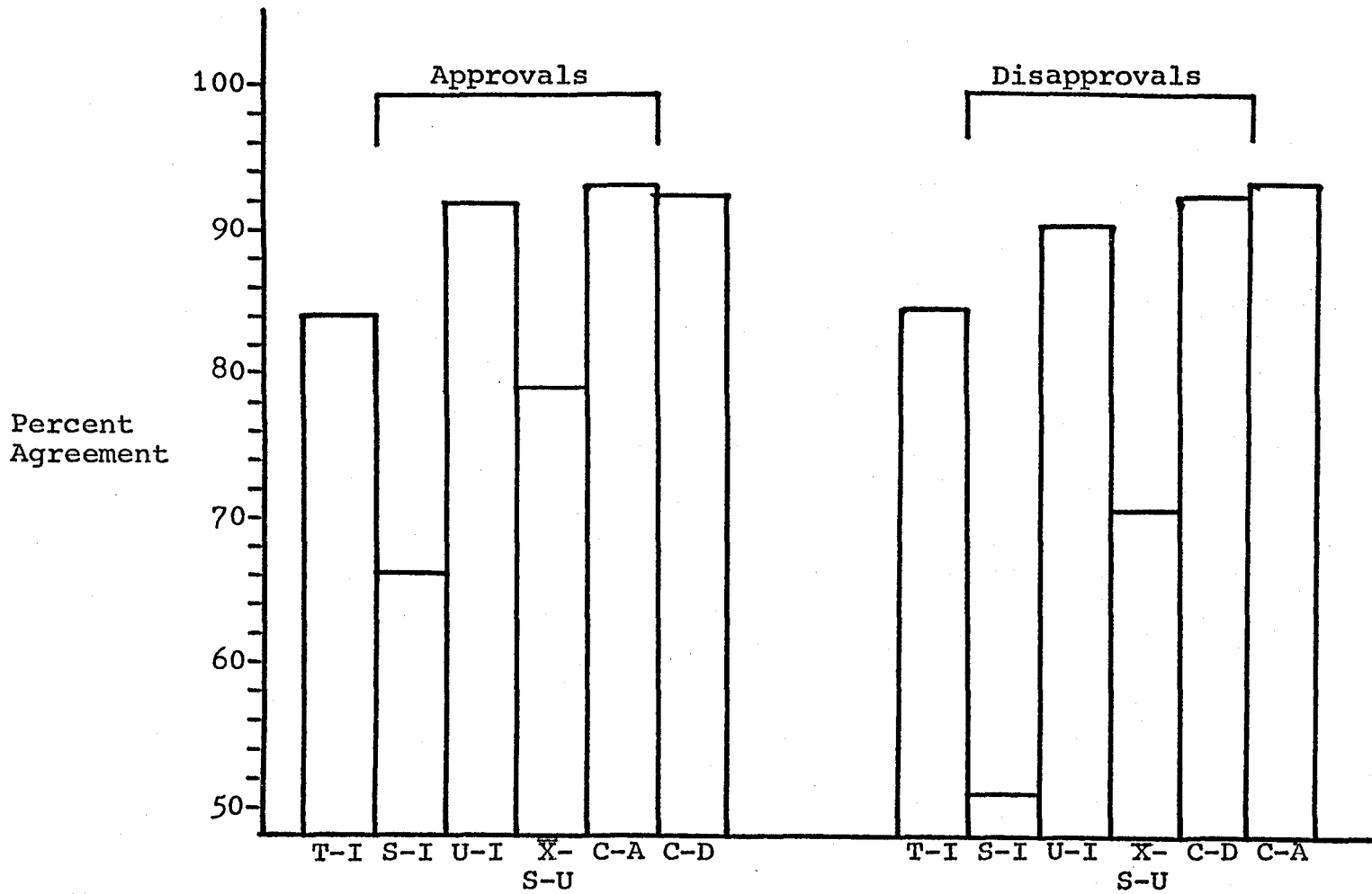
Figure 2. Reliability computations on total interval (T-I), scored interval (S-I), unscored interval (U-I), mean of the scored and unscored interval ($\overline{X}$-S-U), and category (C-A, C-D) agreement.

observers scored the behavior as occurring. Finally,
the mean of the scored and unscored interval ($\overline{X}$-S-U)
reliabilities was computed for each teacher behavior.
As presented in Figure 2, the data from the S-I, U-I,
and $\overline{X}$-S-U showed that the observers disagreed most often
when scoring disapprovals.

For student behavior the reliability was computed
for each observation interval by dividing the number of
agreements by the number of possible agreements (the
number of target students). For example, if both ob-
servers recorded all students on-task, they agreed on
all three students, yielding 100% reliability (three
agreements divided by three possible agreements). If
one observer recorded three students on-task and the other
observer counted only two on-task, the observers agreed
on two students and disagreed on one. Two agreements
divided by three possible agreements yields a 67% relia-
bility score. The reliability of the observation on
student behavior for an entire session was found by com-
puting the mean reliability of all the intervals. The
observers were trained to a 90% criterion before the
study began. The interobserver agreement for the study
was 90.6%.

Appendix C contains the mean reliability computed
by phases, the ranges of individual session reliabilities
for each behavior, and a further explanation of the T-I
data. Also, the raw data collected by the reliability
observers is presented graphically with the data collected
by the regular observer.

Training Procedures

The program to train the teachers to increase their frequency of praising was implemented in three cumulative phases: 1) rules and instructions; 2) rules plus consultant feedback; 3) rules, consultant feedback plus self-monitoring. The consultant for this study was the researcher, a male graduate student in the master's degree psychology program at the University of Richmond. He had 1½ years experience as a school psychologist.

Rules and Instructions. During a one-hour consultation session, the consultant asked the teacher to implement the RAID approach, consisting of setting rules, increasing approval and ignoring, and decreasing disapproval. The three teachers set the same rules for each of the two target periods and reviewed the rules at the beginning of each period. The following rules were set: 1) raise your hand and get permission to speak; 2) raise your hand and get permission to leave your seat; 3) work on your assignment. In discussing approval and disapproval, the consultant defined the two behaviors, modeled behavior specific remarks given to individual students and to groups, explained the advantages of the 4:1 approval/disapproval ratio, and asked each teacher to increase her praising by "catching the students being good." To help the teachers decrease their disapprovals, the consultant defined ignoring, suggested when and when not to ignore inappropriate student behavior, and asked the teachers to actively ignore by praising another student.

The teachers were told that the consultant would collect data on the two target teacher behaviors when he was present in the classroom. During this phase the consultant answered teacher questions but did not provide feedback on the amount of approval or disapproval exhibited.

Rules Plus Consultant Feedback. One consultation session was held on the school day immediately preceeding this phase and another session on the afternoon of the first day in this phase. During the first session, the consultant showed each teacher a graph of her behavior and that of the target students gathered during the first two phases. Next, the consultant discussed the implementation of the feedback procedures and related them to the teacher and student behavior on the graphs. On the afternoon of the first consultant feedback day, another consultation session was held with the teachers to discuss any difficulties. The teachers were primarily concerned with learning a variety of responses during this session. The consultant had the teachers share their most frequent and/or favorite responses, and the consultant shared his. Sample responses from Madsen and Madsen (1974) were read and discussed.

During this phase, the consultant provided feedback to the teacher after each five minutes of classroom observation. The consultant handed the teacher a slip of paper on which was written the number of approvals and disapprovals counted by the consultant (Appendix D).

A goal was set each day for the number of approvals and disapprovals to be given during each five minutes. The number of approvals was determined by adding two to the mean of the five minute intervals from the day before. The disapproval goal was one less than the previous day's mean. The goals were written in the boxes on the slip of paper that was handed to the teacher every five minutes. The frequency counts for the final five minute segment of the session were provided during a brief, post-session conference. The mean frequency of the five minute segments was computed before the conference, along with the percentage of students on-task. This information was graphed and shown to the teacher, allowing an opportunity for the consultant to praise desired changes in the teacher's behavior.

To aid the consultant in determining the beginning and ending of a teacher approval or disapproval, a new verbal remark began when the teacher addressed a different student, when three seconds had elapsed since the last remark, or when the teacher approved or disapproved a different behavior of the same student(s).

Rules, Consultant Feedback Plus Self-Monitoring. One consultation session was held on the school day prior to implementing the self-monitoring phase. When the teachers stated that the feedback had helped them change their behavior, one teacher asked what to do next year when the consultant was not present (the study was conducted at the end of the school year). The consultant suggested

the use of self-monitoring on a wrist counter, and he
asked each teacher to count her approvals during the
non-consulting period as well as during the consulting
period. The use of the wrist counter was explained.

Throughout this phase the consultant continued pro-
viding the frequency count of approvals and disapprovals
to the teacher during the consulting period. The teacher
self-monitored during the consulting period and compared
her approval count with that obtained by the consultant.
The teachers wrote their count on a slip of paper after
each five minutes during the consulting periods and
handed it to the consultant at the end of the period.
In order to assure the five-minute intervals of the
teacher and consultant coincided, the teacher self-
monitored when the consultant was sitting and had his
hands below the level of his head. The observer cued the
consultant when a five-minute segment began and ended.
For the non-consulting periods, the teacher was given a
goal based on the five-minute goal of the consulting
session, but adjusted for a count based on 23 minutes of
self-monitoring. The teacher began counting during the
non-consulting period after she had reviewed the rules
and stopped counting when the observer stood up to leave
the room. The count obtained during the non-consulting
period was for the teacher's benefit only and was not
given to the consultant.

Design for Analysis

A two factor (two periods by five phases) repeated measures analysis of variance design was employed (Bruning & Kintz, 1968, page 47). The study consisted of five phases: 1) baseline; 2) rules and instructions; 3) rules plus consultant feedback; 4) rules, consultant feedback plus self-monitoring; 5) follow-up. Each phase lasted three days with data collected daily in both a consulting period (consultant present) and a non-consulting period (consultant absent). Controlling for time of day and academic content, the consulting period occurred seven or eight times during language arts for each teacher throughout the study, with the remaining consulting periods occurring during math. The consulting period took place at least once in language arts and math for each teacher during each phase.

The teachers were told prior to baseline that the observers would collect data on the students, and that the consultant would gather information on the teacher. In actuality the observers collected data on both the teacher and the students. When the consultant was present he collected frequency data on the teacher, data that was presented to the teachers during feedback and self-monitoring phases.

Baseline. The first phase began after the observers had attained the 85% inter-observer agreement criteria for teacher and student behavior. The baseline phase was used to determine the effects of the presence of the

consultant on teacher and student behavior, and it served as a comparison with the other phases of the study.

Rules and Instructions. The first of the training procedures were implemented during the consulting and non-consulting periods, allowing a comparison of the effects of rules and instructions on teacher and student behavior when the consultant was present and absent (Fig. 3).

Rules Plus Consultant Feedback. During this phase, the effects of consultant feedback on both teacher and student behavior were determined in the presence of the consultant. The teacher continued setting rules and trying to change her behavior during the non-consulting period, when the consultant was not present to provide feedback.

Rules, Consultant Feedback Plus Self-Monitoring. The teacher self-monitored her behavior during both the consulting and non-consulting periods so as to determine the effects of self-monitoring on teacher and student behavior.

Follow-Up. Follow-up data was collected after only a one day break between the fourth and fifth phases due to the ending of school for the teachers and students. The teachers were asked to teach the way they wanted to teach. Data was collected during the consulting and non-consulting periods to determine the durability of the training.

	Baseline	Training 1	Training 2	Training 3	Follow-up
Consulting	consultant present	rules and instructions	rules plus consultant feedback	rules, con-sultant feedback plus self-monitoring	consultant present
Non-consulting	consultant absent	rules and instructions	rules and instructions	rules plus self-monitoring	consultant absent

Periods

Figure 3. Program changes in each period during each phase.

CHAPTER III

Results

Approvals

The mean approvals per phase during the consulting and non-consulting periods are depicted in Figure 4. The results of the two by five (periods by phases) repeated measures analysis of variance failed to yield a significant two-way interaction (Table 1). However, significant main effects of the periods factor ($F(1,8) = 395.738$, $p < .001$) and phases factor ($F(4,32) = 102.327$, p .001) were obtained. These results indicated that the teachers approved more frequently when the consultant was present than when he was absent, and that their rate of approving differed across phases of the study.

The Newman-Keuls test of multiple comparisons of means was performed on the data from the phases. As presented in Table 2, each mean differed significantly from every other mean, and the following ranking emerged: $\mu_{baseline} < \mu_{rules} < \mu_{follow-up} < \mu_{feedback} < \mu_{self-monitoring}$. This ranking showed that the teachers approved significantly more often during the rules phase than they did during baseline. Consultant feedback produced a significant increase in approvals during the third phase, as did self-monitoring in the fourth phase. Follow-up data indicated that approvals dropped below the feedback and

Figure 4. Mean number of approval intervals per session during the consulting and non-consulting periods of each phase.

TABLE 1

Analysis of Variance: Approvals

SOURCE	SS	df	ms	F
Between <u>Ss</u>				
Subjects	172.156	8		
Within <u>Ss</u>				
Periods	172.156	1	139.738	395.858*
Phases	3413.222	4	853.306	102.327*
Pds X Phs	59.844	4	14.961	0.338
Pds X <u>Ss</u>	2.822	8	0.353	
Phs X <u>Ss</u>	266.844	32	8.339	
Pds X Phs X <u>Ss</u>	1417.290	32	44.290	

*$p < .001$

TABLE 2

Newman-Keuls

Multiple Comparisons among Means: Approvals

	$\overline{X}_1$	$\overline{X}_2$	$\overline{X}_5$	$\overline{X}_3$	$\overline{X}_4$	
$\overline{X}_1=$ 9.1	---	12.6**	18.0**	24.2**	37.2**	$W_2=$ 2.559
$\overline{X}_2=$ 21.7		---	5.4**	11.6**	24.6**	$W_3=$ 2.913
$\overline{X}_5=$ 27.1			---	6.2**	19.2**	$W_4=$ 3.124
$\overline{X}_3=$ 33.3				---	13.0**	$W_5=$ 3.281
$\overline{X}_4=$ 46.3					---	

**$p<.01$

self-monitoring means, but remained significantly higher than the mean of the rules phase.

Disapprovals

Figure 5 graphically presents the mean disapprovals per phase from both the consulting and non-consulting periods, and Table 3 lists the results of the two by five analysis of variance. The main effects of phases on disapprovals yielded the only significant factor, $F(4,32) = 12.388$, $p < .001$.

A comparison of the phase means with the Newman-Keuls test produced the following ranking (Table 4):

$\mu_{feedback} < \mu_{rules} = \mu_{follow-up} = \mu_{self-monitoring} < \mu_{baseline}$.

The teachers disapproved significantly more during baseline than during any subsequent phase. The disapproval means of the rules, self-monitoring, and follow-up phases were not significantly different from one another. The teachers disapproved significantly less during the feedback phase than during the other phases of the study.

Student On-Task Behavior

The results of the two by five analysis of variance showed that on-task behavior was affected significantly by the phases factor, $F(4,32) = 7.475$, $p < .001$ (Table 5). As graphically shown in Figure 6, student on-task behavior increased throughout the study. However, a comparison of the means with the Newman-Keuls, a less powerful test than the analysis of variance, failed to differentiate between the means (Table 6).
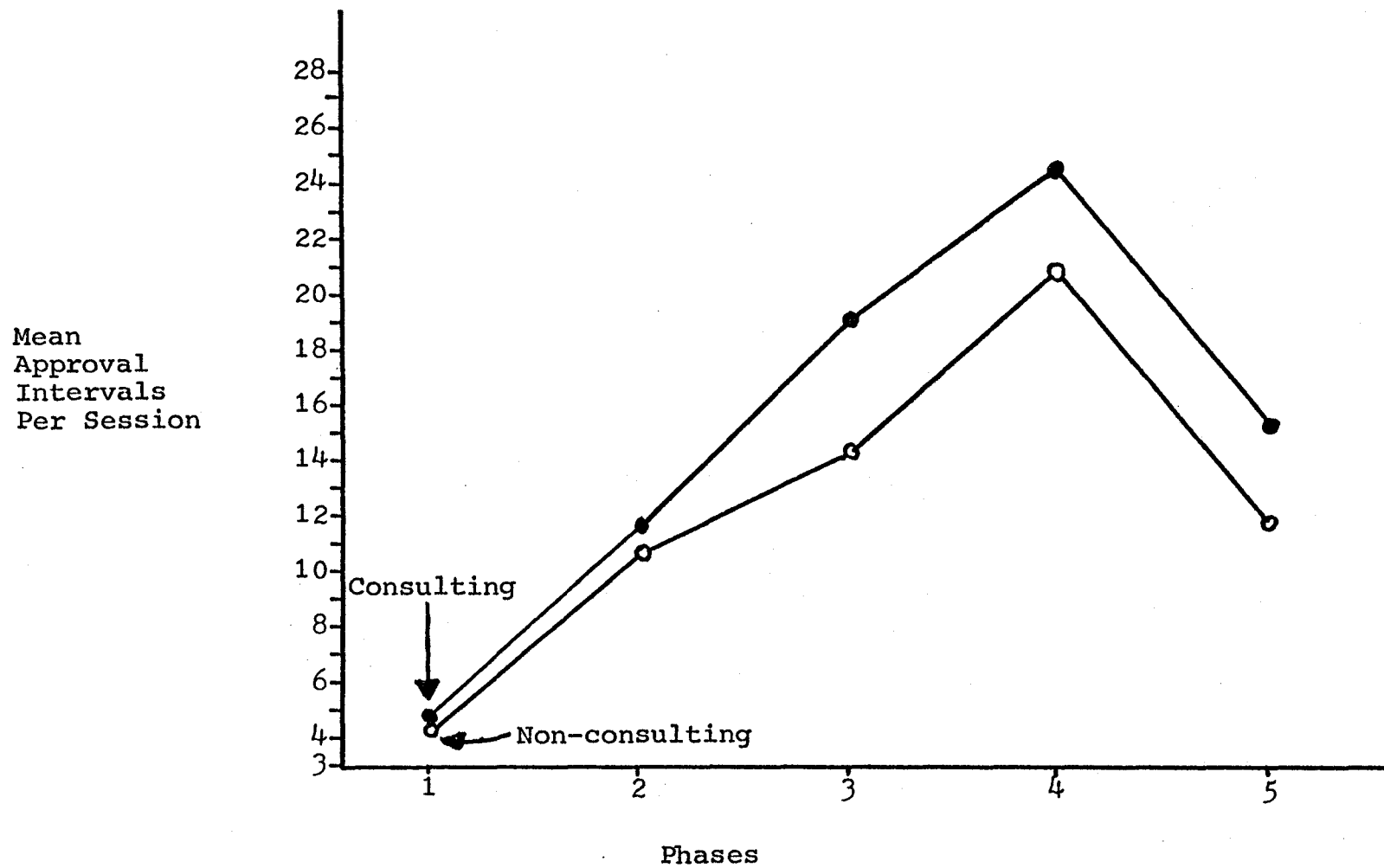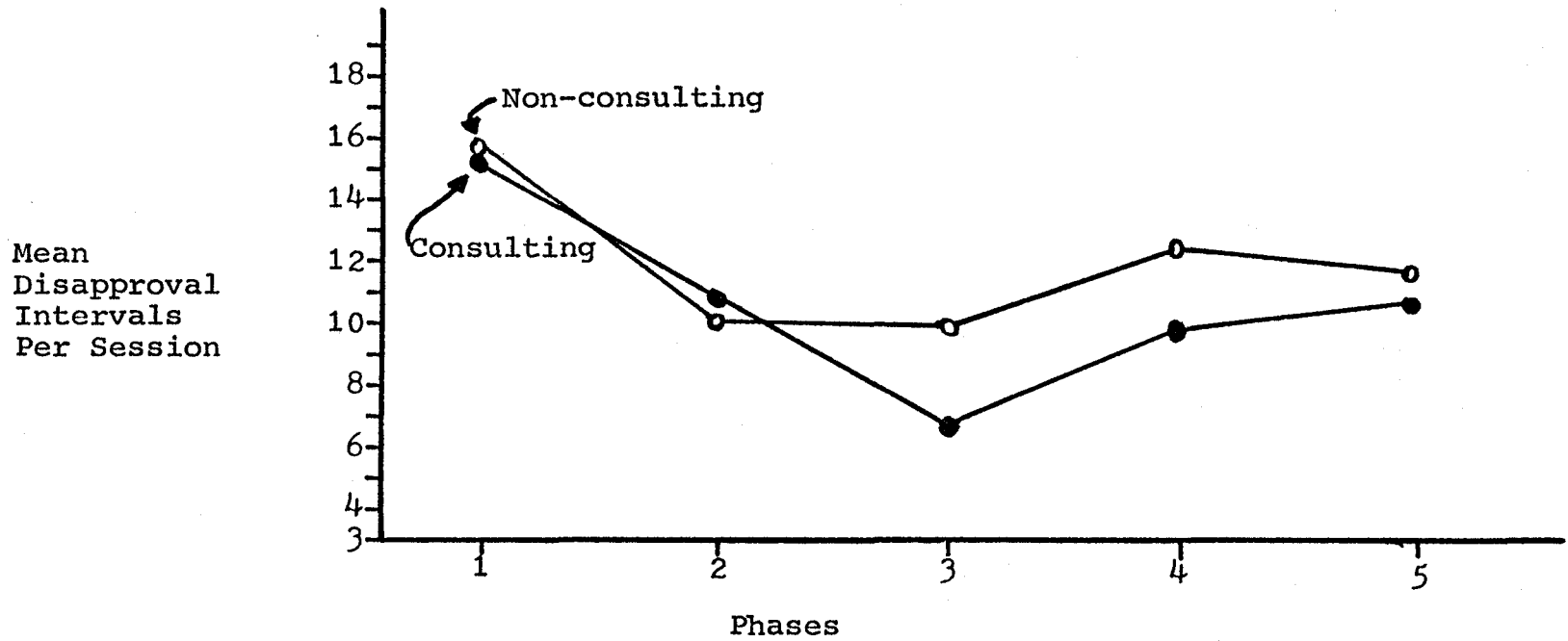
Figure 5. Mean disapproval intervals per phase during the consulting and non-consulting period.

TABLE 3

Analysis of Variance: Disapprovals

SOURCE	SS	df	ms	F
Between Ss				
Subjects	585.956	8		
Within Ss				
Periods	36.100	1	36.100	3.460
Phases	497.845	4	124.461	12.388*
Pds X Phs	47.844	4	11.961	0.320
Pds X Ss	83.467	8	10.433	
Phs X Ss	321.488	32	10.047	
Pds X Phs X Ss	1196.423	32	37.388	

*p<.001

TABLE 4

Newman-Keuls

Multiple Comparisons among Means: Disapprovals

	$\overline{X}_3$	$\overline{X}_2$	$\overline{X}_5$	$\overline{X}_4$	$\overline{X}_1$	
$\overline{X}_3 = 16.33$	----	4.34**	5.34**	5.67**	14.45**	$W_2 = 2.809$
$\overline{X}_2 = 20.67$		----	1.00	1.33	10.11**	$W_3 = 3.198$
$\overline{X}_5 = 21.67$			----	0.33	9.11**	$W_4 = 3.429$
$\overline{X}_4 = 22.00$				----	8.78**	$W_5 = 3.601$
$\overline{X}_1 = 30.78$					----	

**$\underline{p} < .01$

TABLE 5

Analysis of Variance: On-Task Behavior

SOURCE	SS	df	ms	F
Between Ss				
Subjects	28627.800	8		
Within Ss				
Periods	11.378	1	11.378	0.125
Phases	2826.556	4	706.639	7.475*
Pds X Phs	99.166	4	24.792	0.086
Pds X Ss	726.022	8	90.753	
Phs X Ss	3024.977	32	94.531	
Pds X Phs X Ss	9266.101	32	389.566	

*$p < .001$
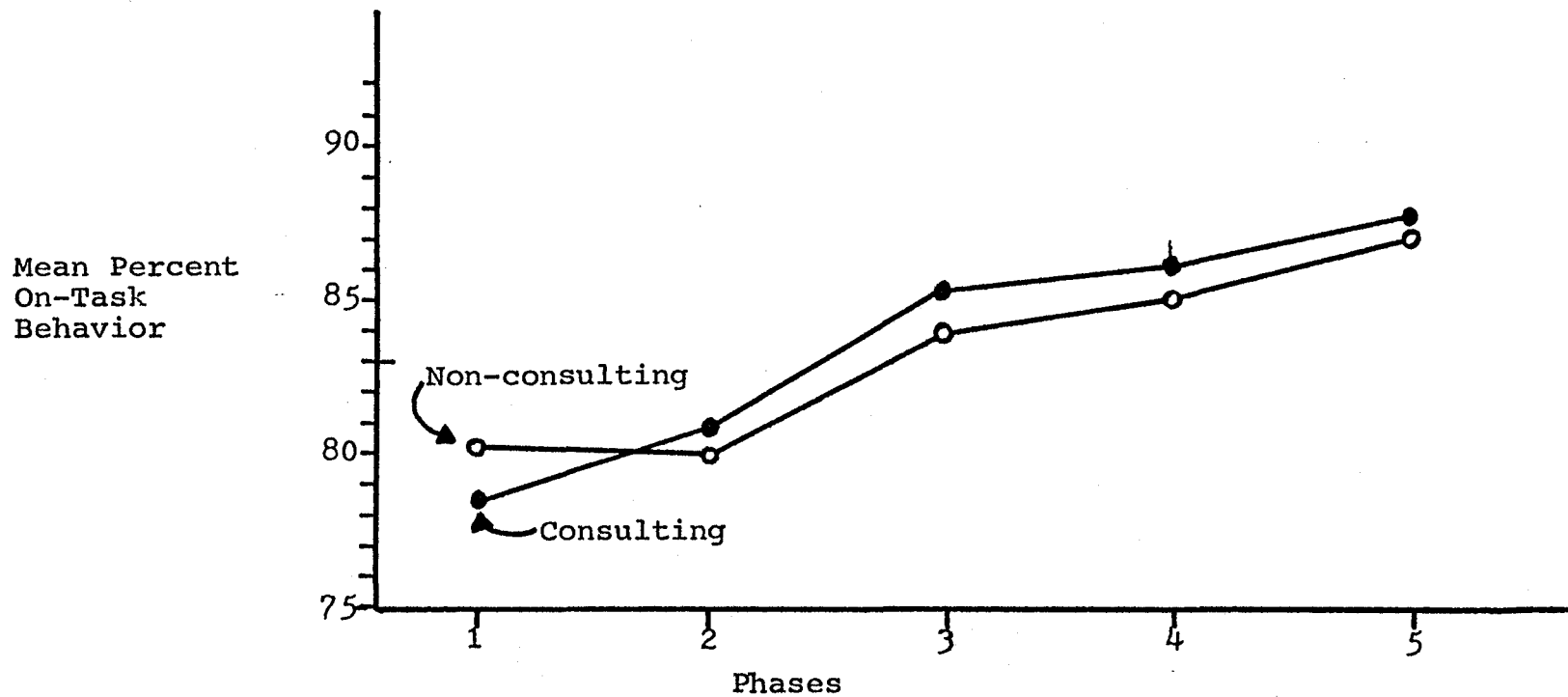
Figure 6. Mean percent on-task behavior per session during the consulting and non-consulting periods of each phase.

TABLE 6

Newman-Keuls

Multiple Comparisons among Means: On-Task Behavior

	$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$	$\bar{x}_4$	$\bar{x}_5$	
$\bar{x}_1 = 79.4$	-----	1.0	5.2	6.3	8.1	$W_2 = 6.485*$
$\bar{x}_2 = 80.4$		-----	4.2	5.3	7.1	$W_3 = 7.792*$
$\bar{x}_3 = 84.6$			-----	1.1	2.9	$W_4 = 8.571*$
$\bar{x}_4 = 85.7$				-----	1.8	$W_5 = 9.121*$
$\bar{x}_5 = 85.5$					-----	

$*p < .05$

CHAPTER IV

Discussion

This study investigated the effects of a three

phase training program (rules and instructions, feed-

back, and self-monitoring) on teacher approval, disap-

proval, and student on-task behavior. Data was collected

during a consulting period, assessing behavior changes

when the consultant was present, and during a non-con-

sultant period, recording whether similar changes occurred

with the consultant absent. The research investigated

a predicted interaction between periods and phases, that

is, a difference in the differences between consulting

and non-consulting periods across phases. Specifically,

it was predicted that the difference in approvals between

the consulting and non-consulting periods would be greater

during the feedback phase than during any other phase

of the study.

Teacher Behaviors

Since no periods by phases interaction occurred,

approval and disapproval data was collapsed within each

category of behavior from both consulting and non-

consulting periods (Figure 7). Changes in the training

program across phases significantly affected teacher

approval and disapproval. Teacher behavior is discussed
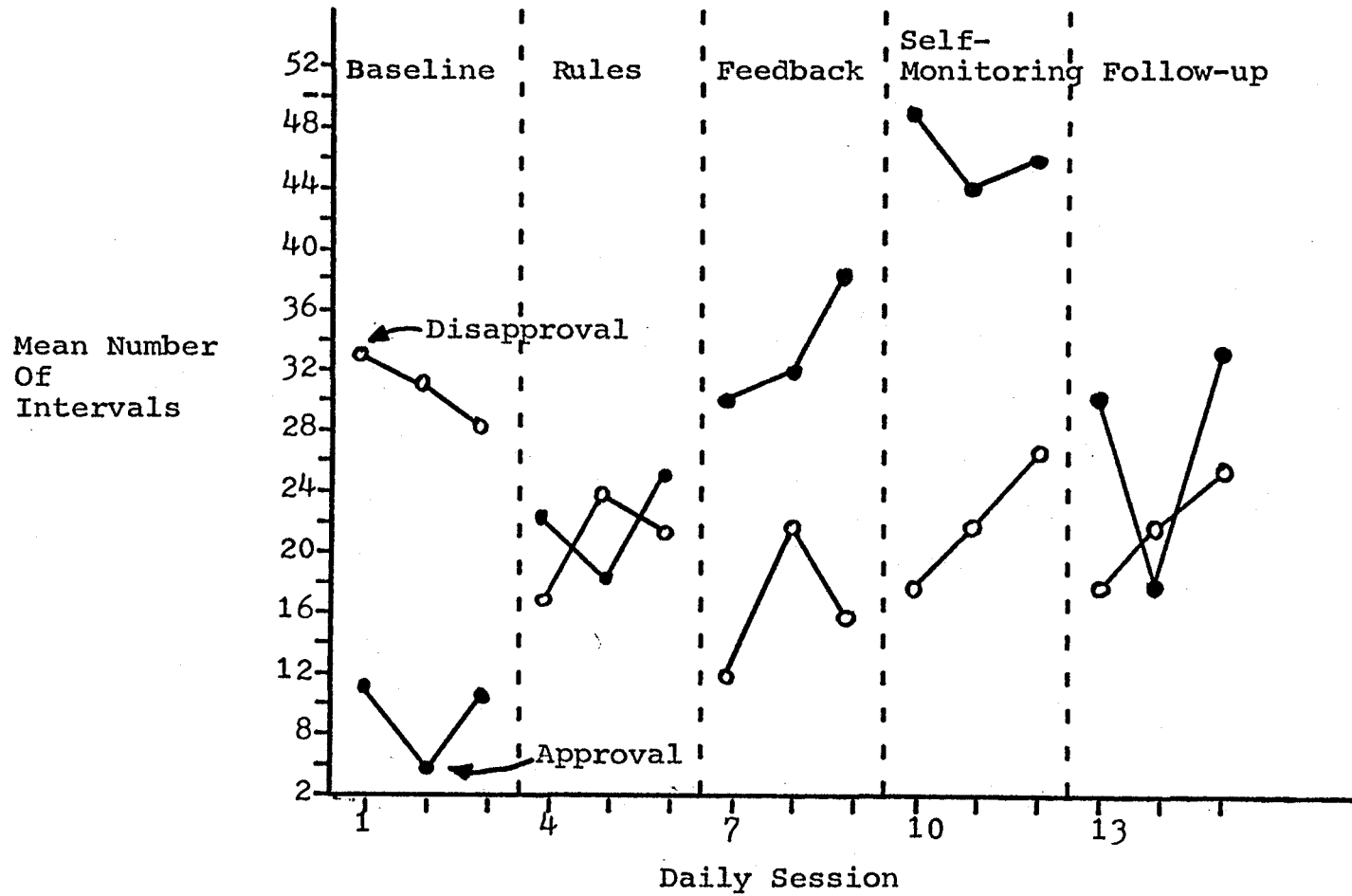
Figure 7. Mean approval and disapproval intervals given per teacher during each daily session (considering both consulting and non-consulting periods).

first by changes across phases due to the training procedures, followed by a discussion of the generalization of these behavior changes.

Training procedures. The baseline established the rate of approval and disapproval before training began. The three teachers disapproved three times as much as they approved, exhibiting 1.56 disapprovals and .46 approvals per minute. Since rate figures (responses per minute) were based on interval data when only one response could occur within a ten second interval, they represent conservative estimates of true rates based on a frequency count of every response. That is, only one response could be recorded per interval, while several responses may have occurred. These approval and disapproval rates are not unlike those of the average elementary teacher (Madsen & Madsen, 1974).

Instructions to set the three classroom rules, increase approvals and ignoring, and decrease disapprovals (RAID) immediately and significantly affected teacher behavior as expected during the second phase. The teachers approved and disapproved once each minute, doubling the baseline approval rate and decreasing disapprovals by one-third. While all teachers changed their behavior, two teachers were affected more than a third one.

Some studies have reported that instructions have no effects on most teachers (Cossairt et al., 1973; Parsonson et al., 1974). The explicit rule setting and review at

the beginning of each period may have been a key difference between those studies and the current one.

The positive effects of consultant feedback on teacher behavior supported the findings of Cossairt et al. (1973) and Parsonson et al. (1974). Consultant feedback provided every five minutes increased the rate of approvals to 1.67 per minute, three times the baseline rate. While the daily session means appeared to indicate a steady increase in approvals throughout the feedback phase, the data from only one teacher fit that curve. The approval rates of the other two teachers were affected more abruptly and less consistently. Disapprovals decreased to .82 per minute, significantly lower than during the rules phase. It appeared to the consultant that academic disapproval or corrections were more frequent than disapproval to social or off-task behavior, in comparison with the previous phases.

During the feedback and self-monitoring phases the consultant walked to the teacher, if she was sitting, and handed her the slip of paper containing the feedback. If she was standing, the paper was placed on the teacher's desk or on a convenient table for her to pick up. This method of providing feedback was an easy technique for communicating to the teacher her frequency of approval and disapproval. Neither the students nor teacher appeared to be disturbed by the consultant's walking around the room.

Approvals increased significantly to 2.32 responses per minute during the self-monitoring phase. These results contrasted with the slight increase in approval reported by Thoresen et al. (1973). However, the teachers in this study increased their approvals during the feedback phase and then used self-monitoring during the next training phase. Learning to praise and self-monitor at the same time may be a more difficult task than first learning to praise and then to self-monitor.

In comparison with the feedback phase, disapprovals increased significantly to 1.10 responses per minute during the self-monitoring phase, a change not consistent with the predictions on approval data. Several factors may have influenced disapprovals. First, counting on the wrist counter focused the teachers' attention on that category of behavior. Even though the teachers continued receiving feedback from the consultant on approvals and disapprovals, they may have been primarily concerned with counting approvals in agreement with the consultant. Second, the teachers possibly created more opportunities to disapprove than previously existed. In order for the teachers to attain the high approval rates, they increased their walking among the students in order to praise them as they worked, as opposed to staying at the teacher's desk. The teachers were actualizing behaviorally the saying, "One teacher on her feet is worth two in their seats" (Madsen & Madsen, Note 1).

But in this correct attempt to increase their approvals,
the teachers frequently looked at the students' papers,
noticed mistakes, and verbally clued the students to
the errors (disapprovals).

In conjunction with the second explanation, the
teachers may have been less concerned about consultant
feedback on disapprovals. The consultant pointed out
that disapprovals to academic behavior would probably
increase because the teachers were walking to the students'
desks more often. Since avoiding academic disapprovals
might be difficult and unwarranted, the teachers were
asked to minimize disapprovals to social behavior in
order to maintain or achieve the 4:1 approval/disapproval
ratio.

The teachers experienced varying success in self-
monitoring. The agreement of one teacher's count with
that of the consultant was low throughout this phase
(mean 54.4% agreement). She stated she would forget to
press the lever on the counter. This teacher taught in
a trailer and the counter noise was noticeable (it
was not noticeable in the regular classrooms). Pressing
a quieter counter may have been a more pleasant experience
for her. Another teacher could not keep track of her
record forms and each five minute block during the first
consulting period of the phase, making it impossible to
compute her agreement according to each five minute interval.
The third teacher had relatively good agreement (mean

76.1%), but awkwardly kept both her hand and the counter
in a pocket at practically all times. Self-monitoring
was not the most pleasant experience of the study for
any of the teachers and was the least liked part by one
of them.

The teachers approved significantly less during
follow-up than during the feedback and self-monitoring
phases, despite only a single academic day between the
fourth and fifth phases. No change in the approval
rate had been predicted. However, the teachers were
approving 1.4 times per minute during the final phase,
a rate significantly higher than during the rules phase.
Several factors may have decreased approvals. Obviously,
the lack of feedback may have affected teacher behavior.
Also, the teachers still may have been acquiring skills
and were not ready to maintain them without feedback.
This six day training program involving feedback was
much briefer than the 14 days or more of training in
studies reporting no changes in follow-up (Cossairt et
al., 1973; Parsonson et al., 1974). Furthermore, the
fact that two of the three teachers could decrease approvals
and not affect student on-task behavior might also account
for decreased approval. Since student on-task behavior
was maintained during this phase in two of three class-
rooms, the teachers may not have found it worthwhile to
praise two or three times per minute, a goal set by Madsen
and Madsen (Note 1) and achieved by the teachers during
the self-monitoring phase.

The teachers disapproved 1.09 times per minute during follow-up, a rate commensurate with the rules and self-monitoring phases. Since disapprovals were maintained (1.10 during self-monitoring) and approvals were not (2.32 during self-monitoring), the stimuli eliciting each teacher behavior must have been differentially changed during follow-up. Specifically, approval eliciting stimuli decreased during follow-up, while disapproval eliciting stimuli remained the same. Disapproval eliciting stimuli (off-task behavior and incorrect academic work) were present during follow-up as they had been during the self-monitoring phase. Approval eliciting stimuli included consultant feedback and self-monitoring, both of which were discontinued during follow-up.

Generalization. This research produced evidence for and against the generalization of teacher approval from the consulting to non-consulting periods. Factors supporting generalization include the nonsignificant periods by phases interaction and the significant changes in approvals across phases. The nonsignificant interaction indicated the differences between the consulting and non-consulting periods failed to vary significantly across the five phases of the study. Any change occurring during the consulting period was also recorded during the non-consulting period. The significant phases effects showed that approvals increased during each of the first four phases. Therefore, in light of the nonsignificant interaction, increases across phases were similar during both consulting and non-consulting periods.

Factors within and outside of the classroom may
have facilitated generalization. Within the classroom,
the setting and reviewing of classroom rules during
each period may have programmed approval eliciting
stimuli in both periods of the rules and instructions
phase. Furthermore, reviewing the rules before each
period of the subsequent phases may have helped to
support generalization in those phases. Self-monitoring
may have served a similar purpose by programming approval
eliciting stimuli in the non-consulting period of the
fourth phase. Striving to meet a goal based on approval
during the consulting period may have increased teacher
praise during the non-consulting period of the self-
monitoring phase, too.

The presence of the observer may have affected
teacher behavior during the non-consulting periods, also.
Since both the observer and consultant were present
during the consulting period, the observer's presence
may have elicited teacher approval when the consultant
was absent.

Generalization may have resulted also from two
factors within the classroom due to the design of the
study. First, one method of promoting generalization
from one environment to a second is to perform some of
the training in each setting. For example, a therapist
might program the generalization of a client's newly
learned behavior from the office to the client's home by

performing some of the therapy in the home setting.
Varying the consulting period between math and language
arts may have served a similar purpose. Also, simply
learning to approve math and language arts subject
matter may have influenced generalization. As mentioned
in the introduction, Horton (1975) had teachers view
video tapes containing approvals from either one or
several academic areas. Viewing approval in one sub-
ject increased approval in the classroom during that
subject only, while viewing several areas led to changes
in an equal number of subjects. In this study the
teachers learned to praise math and language arts responses
of the students. Learning to praise both academic sub-
jects may have served the same purpose as viewing several
academic subjects on video tape, as done by the teachers
in Horton's study who showed generalized praise.

Training outside of the classroom may have affected
generalization. The one hour session on the RAID approach,
the discussions preceding and following the first day of
feedback, plus the meeting prior to the self-monitoring
phase may have elicited equal changes within periods
across the training phases.

The significant periods effects indicated that the
teachers approved more during the consulting than non-
consulting period throughout the study, showing that
teacher approval did not fully generalize from one period
to the other. The differences between the two periods

were more pronounced after feedback began (Figure 4,
page 30). Perhaps, the consultant, the consultant's
behavior, the teacher's interactions with the consultant
and/or the teacher's behavior during the consulting
period served as discriminative stimuli for teacher
approval.

Student On-Task Behavior

The training program improved the teachers' class-
room management skills, as indicated by a significant
increase in student on-task behavior (Figure 8). The
mean percentage of students on-task increased from 79.4%
during baseline to 87.5% during follow-up. This base-
line rate appears high when compared with the data from
Madsen et al. (1970) and the guidelines in Madsen and
Madsen (1974). However, the Madsen research recorded
off-task behavior with a 10 second interval recording
technique, while this research recorded on-task behavior
by a time sampling method. The probability of showing
a class of students to be unmanageable would be greater
when students are considered off-task if that behavior
occurs at any point in a 10 second interval, in comparison
with an instantaneous glance of a time sampling technique.

One class of students began this study with fewer
students on-task than the other two. Furthermore, on-
task behavior fluctuated within this classroom during the
self-monitoring phase and decreased during follow-up,
while the other two classes showed a continual increase
during these phases. The types of disapprovals given by

90 –

85 –

80 –

75 –

70 –

Baseline | Rules | Feedback | Self-Monitoring | Follow-up

Mean
Percent
On-Task
Behavior

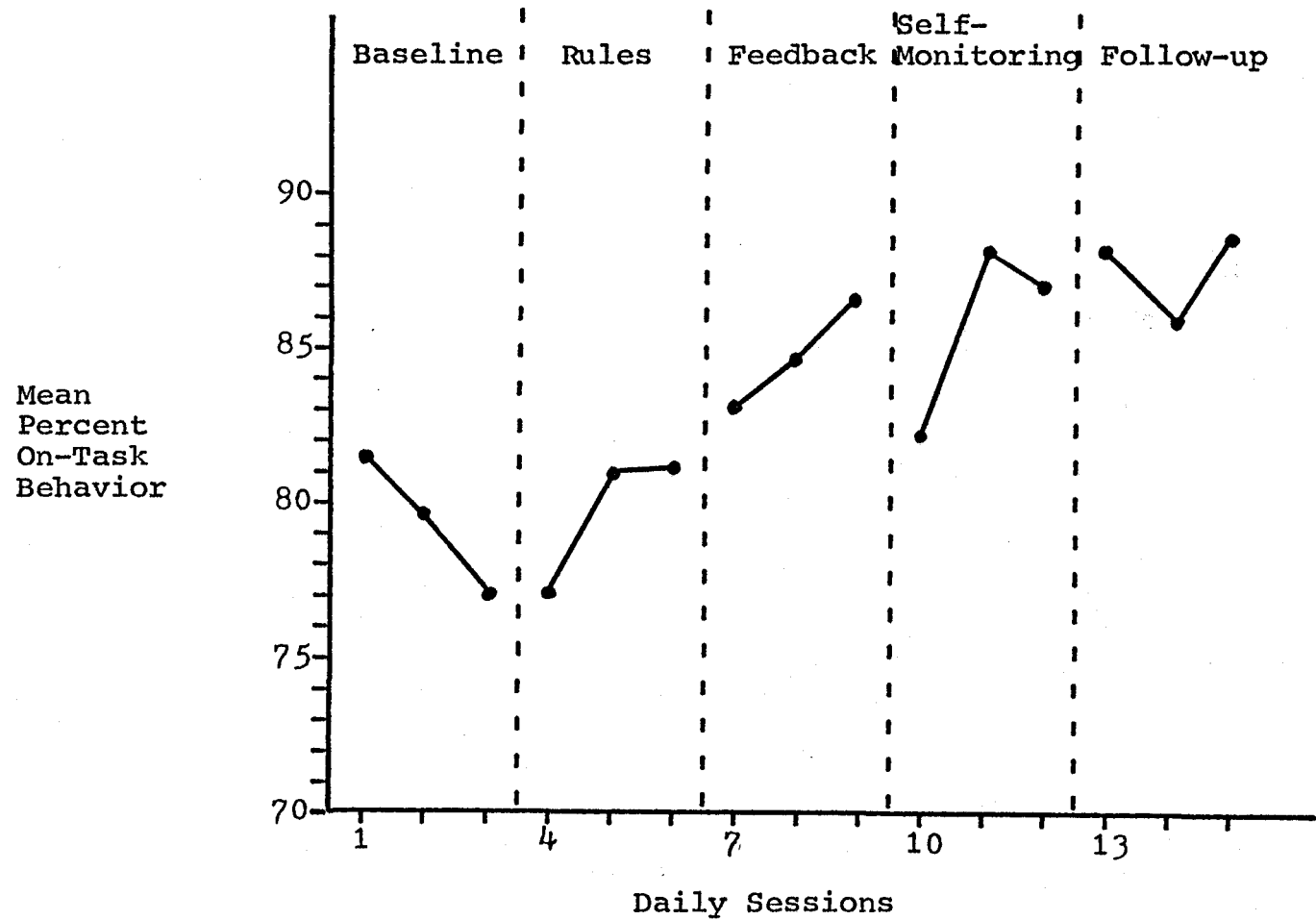1   4   7   10   13

Daily Sessions

Figure 8.   Mean percent on-task behavior per classroom during each daily session
(considering both consulting and non-consulting periods).

the teacher of the deviant class during the fourth and
fifth phases may account for the decrease in student
on-task behavior. It was the consultant's impression
that this teacher increased her disapproval to social
behavior during the final two phases, while the other
two teachers increased their disapproval to academic
behavior during these phases.

Implications for Future Research

This research indicated that rules and instructions,
feedback, and self-monitoring increased teacher approval
and student on-task behavior, and decreased teacher
disapproval. The training procedures were more effective
during the consulting as opposed to non-consulting period;
that is, the teacher approved more often when the con-
sultant was present than when absent. However, the in-
creases in approvals across phases during the non-consult-
ing period were not different from changes across phases
during the consulting sessions. Despite the positive
effects of the training procedures, the teachers unexpect-
edly decreased their rate of approval during follow-up.

In view of these results, the factors affecting
the maintenance of teacher approval should be investigated.
The number of training days may be an important variable
based on the positive results of studies using longer
training programs (Cossairt et al., 1973; Parsonson et al.,
1974). Research could indicate whether one or both of
the feedback or self-monitoring phases should be lengthened.

However, reaching a criterion based on the number of responses per minute for several consecutive days may be a more important factor than simply the number of training days. The data collecting of a longer training program based on attaining a criterion could be eased by lenghtening the self-monitoring phase or by training paraprofessionals, such as classroom aides, to perform the task of providing feedback.

Providing feedback on an increasing variable interval schedule might affect maintenance, also. This feedback might be given by the consultant, a classroom aide, or by self-monitoring. A rotating schedule among all three people might be beneficial.

Maintenance might be enhanced by combining the current program with parts of other programs. For example, both video tape discrimination training (Horton, 1975) and role-playing techniques (Jones & Eimers, 1975) have increased approval, and changes have been maintained during follow-up. Sessions from each of these programs might be added after the teachers have had initial success due to feedback.

Hopefully, research on these variables will lead to effective, brief training programs. Shorter programs would mean less of an intrusion on a teacher's time inside and outside of the classroom. Also, school systems would be more likely to support a brief training program that would involve less consultant, teacher, and possibly classroom aide time. With programs frequently offered, more teachers will learn to be effective classroom managers.

REFERENCES

Baer, D. M., Wolf, M. M., & Risley, T. R. Some current

dimensions of applied behavior analysis. Journal

of Applied Behavior Analysis, 1968, 1, 91-97.

Bruning, J. L., & Kintz, B. L. Computational handbook

of statistics. Glenview, Ill.: Scott, Foresman, 1968.

Cossairt, A., Hall, R. V., & Hopkins, B. L. The effects

of experimenter's instructions, feedback, and praise

on teacher praise and student attending behavior.

Journal of Applied Behavior Analysis, 1973, 6, 89-100.

Greenwood, C. R., Hops, H., Delquadri, J., & Guild, J.

Group contingencies for group consequences in class-

room management: A further analysis. Journal of

Applied Behavior Analysis, 1974, 7, 413-425.

Hall, R. V., Lund, D., & Jackson, D. Effects of teacher

attention on study behavior. Journal of Applied

Behavior Analysis, 1968, 1, 1-12.

Hawkins, W. P., & Dotson, V. A. Reliability scores

that delude: An alice in wonderland trip through

the misleading characteristics of interobserver agree-

ment scores in interval recording. In F. Ramp &

G. Semb (Eds.), Behavior analysis, areas of research

and application. Englewood Cliffs, N. J.: Prentice-

Hall, 1975.

Horton, G. O. Generalization of teacher behavior as a function of subject matter specific discrimination training. <u>Journal of Applied Behavior Analysis</u>, 1975, <u>8</u>, 311-319.

Jones, F. H., & Eimers, R. C. Role playing to train elementary teachers to use a classroom management "skill package." <u>Journal of Applied Behavioral Analysis</u>, 1975, <u>8</u>, 421-433.

Loeber, R. & Weisman, R. G. Contingencies of therapist and trainer performance: A review. <u>Psychological Bulletin</u>, 1975, <u>82</u>, 660-688.

Madsen, C. H., Jr., Becker, W. C., & Thomas, D. R. Rules, praise, and ignoring: Elements of elementary classroom control. <u>Journal of Applied Behavior Analysis</u>, 1968, <u>1</u>, 139-150.

Madsen, C. H., Jr., & Madsen, C. K. <u>Teaching/discipline: A positive approach for educational development, expanded second edition for professionals</u>. Boston: Allyn and Bacon, 1974.

Madsen, C. H., Jr., Madsen, C. K., Saudargas, R. A., Hammond, W. R., Smith, J. B., & Edgar, D. E. Classroom raid (rules, approval, ignore, disapproval): A cooperative approach for professionals and volunteers. <u>Journal of School Psychology</u>, 1970, <u>8</u>, 180-185.

O'Leary, K. D., Becker, W. C., Evans, M. B., & Saudargas, R. A. A token reinforcement program in a public school: A replication and systematic analysis. Journal of Applied Behavior Analysis, 1969, 2, 3-13.

Parsonson, B. S., Baer, A. M., & Baer, D. M. The application of generalized correct social contingencies for group consequences in classroom management: An evaluation of a training program. Journal of Applied Behavior Analysis, 1974, 7, 427-437.

Pinkston, E. M., Reese, N. M., LeBlanc, J. M., & Baer, D. M. Independent control of a preschool child's agression and peer interaction by contingent teacher attention. Journal of Applied Behavior Analysis, 1973, 6, 115-124.

Thoresen, C. E., Hubbard, D. R., Jr., Hannum, J. W., Hendricks, C. G., & Shapiro, D. Behavioral self-observation training with a nursery school teacher. Stanford, Calif.: Stanford Center for Research and Development in Teaching, 1973. (ERIC Document Reproduction Service No. ED 081 508)

Thorensen, C. E., & Mahoney, M. J. Behavioral self-control. New York: Holt, Rhinehart and Winston, 1974.

Wahler, R. G. Setting generality: Some specific and general effects of child behavior therapy. Journal of Applied Behavior Analysis, 1969, 2, 239-246.

White, M. A. Natural rates of teacher approval and
disapproval in the classroom. <u>Journal of Applied
Behavior Analysis</u>, 1975, <u>8</u>, 367-372.

REFERENCE NOTES

1. Madsen, C. H., Jr., & Madsen, C. K. What is behavior
    modification? --- "it is a process of learning."
    Presentation at Drive-In Conference on Behavior
    Modification, Longwood College, Farmville, Va.,
    May 3, 1975.

APPENDIX A

Teacher-Researcher Contract

Teacher - Researcher Contract

Teacher

1. I, _____, volunteer as a teacher to participate in classroom management research.

2. This teacher will try diligently to implement the requests of the researcher.

3. This teacher agrees to meet for four, one hour sessions with the researcher, and also to attend a de-briefing session at the conclusion of the study.

4. The degree of anonymity of the teacher will conform to the desires of each teacher, as set individually in the de-briefing session.

5. This teacher agrees not to discuss this study with other faculty members or participants in this research until the study has been completed.

6. This teacher has the right to withdraw from this research if so desired.

7. Any difficulties with the classroom observers will be reported to the researcher.

Researcher

1. I, Kenneth Roach, agree that the above teacher volunteered for the study.

2. After the observers have been trained, the study will last fifteen data days, to be completed as quickly as possible, but allowing for a brief break between the twelth and fifteenth data days.

3. The researcher agrees that the five sessions described above will be the only requests made for the teachers' time outside of the classroom.

4. This researcher agrees to abide by the requests of the teachers during the de-briefing conference.

_____
(Teacher)

_____
(Researcher)

APPENDIX B

Record Form

TEACHER OBSERVATION FORM

Observer_____   Teacher_____
Reliability Observer_____   Date_____
Names of Target Students_____

Intervals

Min		a		b		c

I

Min		a		b		c
1	OBSERVE	A D 1 2 3	OBSERVE	A D 1 2 3	OBSERVE	A D 1 2 3
2		A D 1 2 3		A D 1 2 3		A D 1 2 3
3		A D 1 2 3		A D 1 2 3		A D 1 2 3
4		A D 1 2 3		A D 1 2 3		A D 1 2 3
5		A D 1 2 3		A D 1 2 3		A D 1 2 3

BREAK   BREAK   BREAK

II

Min		a		b		c
1	OBSERVE	A D 1 2 3	OBSERVE	A D 1 2 3	OBSERVE	A D 1 2 3
2		A D 1 2 3		A D 1 2 3		A D 1 2 3
3		A D 1 2 3		A D 1 2 3		A D 1 2 3
4		A D 1 2 3		A D 1 2 3		A D 1 2 3
5		A D 1 2 3		A D 1 2 3		A D 1 2 3

BREAK   BREAK   BREAK

III

Min		a		b		c
1	OBSERVE	A D 1 2 3	OBSERVE	A D 1 2 3	OBSERVE	A D 1 2 3
2		A D 1 2 3		A D 1 2 3		A D 1 2 3
3		A D 1 2 3		A D 1 2 3		A D 1 2 3
4		A D 1 2 3		A D 1 2 3		A D 1 2 3
5		A D 1 2 3		A D 1 2 3		A D 1 2 3

BREAK   BREAK   BREAK

IV

Min		a		b		c
1	OBSERVE	A D 1 2 3	OBSERVE	A D 1 2 3	OBSERVE	A D 1 2 3
2		A D 1 2 3		A D 1 2 3		A D 1 2 3
3		A D 1 2 3		A D 1 2 3		A D 1 2 3
4		A D 1 2 3		A D 1 2 3		A D 1 2 3
5		A D 1 2 3		A D 1 2 3		A D 1 2 3

APPENDIX C

Reliability

Tables A and B contain the mean and range per phase of the five computations of reliability.

T-I agreement dropped to its lowest level in the self-monitoring phase. This decrease may have been a function of several variables. Interval agreement is affected by the rate of the observed behavior (Hawkins & Dotson, 1975). A behavior recorded in 50% of the intervals in a session has the greatest probability of having the lowest observer agreement. As the frequency deviates from 50% occurrence, the probability of agreement increases. Teacher responses (approvals and disapprovals) were closest to 50% during this phase. Also, the number of intervals in which one or both observers scored both an approval and disapproval increased during the fourth phase. The observers agreed infrequently when both behaviors were recorded in the same interval. It may have been that the observer began thinking about recording the response as soon as it occurred and paid less attention to teacher behavior during the remaining part of the interval.

A change in the type of teacher responses may have affected observer agreement during the self-monitoring phase. The teachers began giving more subtle responses that made it more difficult to discriminate scorable from

TABLE A

Reliability of Approval Data

	Baseline		Rules		Feedback		Self-Monitoring		Follow-Up	
	Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range
T-I	87.8	78.3-95.0	81.7	71.7-91.7	82.2	80.0-85.0	77.2	68.3-88.3	91.1	86.7-93.3
S-I	51.2	14.3-72.7	54.4	37.5-85.7	70.7	52.6-84.6	71.5	60.7-80.0	81.7	75.0-87.5
U-I	94.1	89.8-98.3	90.7	81.5-95.8	89.9	82.0-95.9	84.8	74.4-90.9	94.9	90.2-98.1
$\overline{X}$-S-U	72.7	52.1-83.5	72.6	59.5-90.8	80.3	67.3-90.3	78.2	67.6-85.5	88.3	85.7-92.8
C-A	94.4	90.0-98.3	91.7	83.3-96.7	91.7	85.0-96.7	88.9	81.7-93.3	96.1	93.3-98.3

Table A. Approval data, mean and range per phase of total interval (T-I), scored interval (S-I), unscored interval (U-I), mean of the scored and unscored intervals ($\overline{X}$-S-U), and category (C-A) agreements.

TABLE B

Reliability of Disapproval Data

	Baseline		Rules		Feedback		Self-Monitoring		Follow-Up	
	Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range
T-I	87.8	78.3-95.0	81.7	71.7-91.7	82.2	80.0-85.0	77.2	68.3-88.3	91.1	86.7-93.3
S-I	69.9	41.7-86.7	41.4	25.0-52.9	38.1	28.6-44.4	40.3	8.3-81.8	62.6	42.9-75.0
U-I	92.2	87.3-95.8	88.7	84.3-94.9	87.9	81.1-91.4	87.1	81.4-96.1	94.5	93.1-96.3
$\overline{X}$-S-U	81.0	64.5-91.2	65.1	60.0-68.6	63.0	60.0-67.8	63.8	44.9-89.0	78.6	67.9-85.7
C-D	93.3	88.3-96.7	92.8	86.7-96.7	88.9	83.3-91.7	87.8	81.7-96.7	95.0	93.3-96.7

Table B. Disapproval data, mean and range per phase of total interval (T-I), scored interval (S-I), unscored interval (U-I), mean of the scored and unscored intervals ($\overline{X}$-S-U), and category (C-D) agreements.

unscorable comments. Instead of the teacher saying,
"Jimmy, that's a good boy for raising your hand," she
might ask, "Who has their hand raised? Jimmy, tell
me . . ." An increase in the blending of approval and
disapproval also produced subtle discriminations. For
example, a teacher would say, "I wish table two would
see how quietly table one is sitting," or "That's
close but try again."

A disapproval definition problem caused agreement
difficulties during the fourth phase, also. The relia-
bility observer heard comments which were in fact
disapprovals (academic corrections), but were not
recorded. The reliability observer was enured to the
types of comments made by the teacher she regularly
observed, and this different teacher used more subtle
academic corrections.

Figures 9, 10, and 11 present the teacher data
collected during the non-consulting periods by the
regular observer. Also shown is the data gathered by
the reliability observers.

Table C contains the interobserver agreement on
student on-task behavior. The data collected by the
regular observer and reliability observer is presented
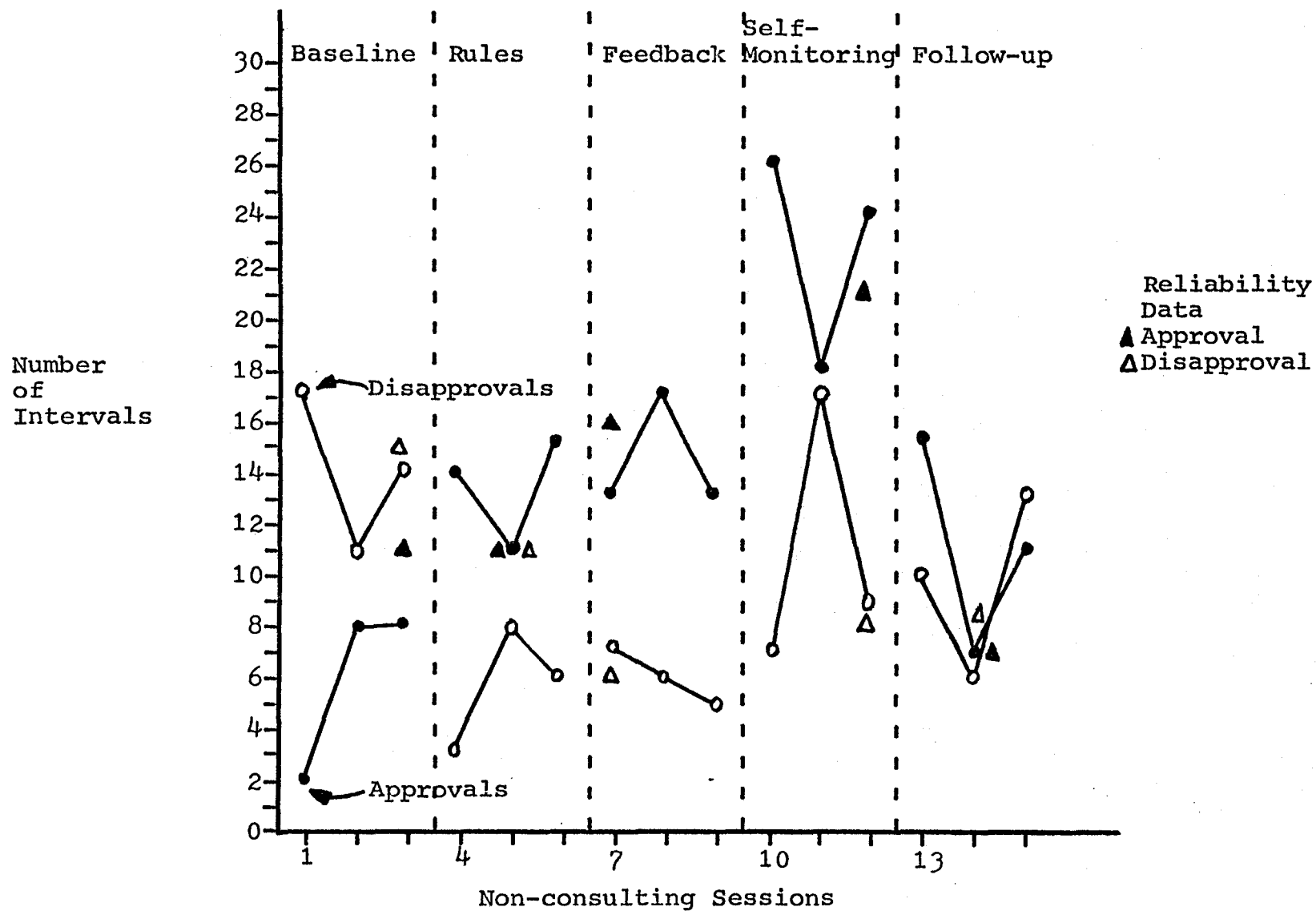graphically in Figures 12, 13, and 14.

Figure 9. Number of approval and disapproval intervals per non-consulting session as recorded by the regular observer. The data collected by the reliability observer during each phase is shown. Teacher 1.
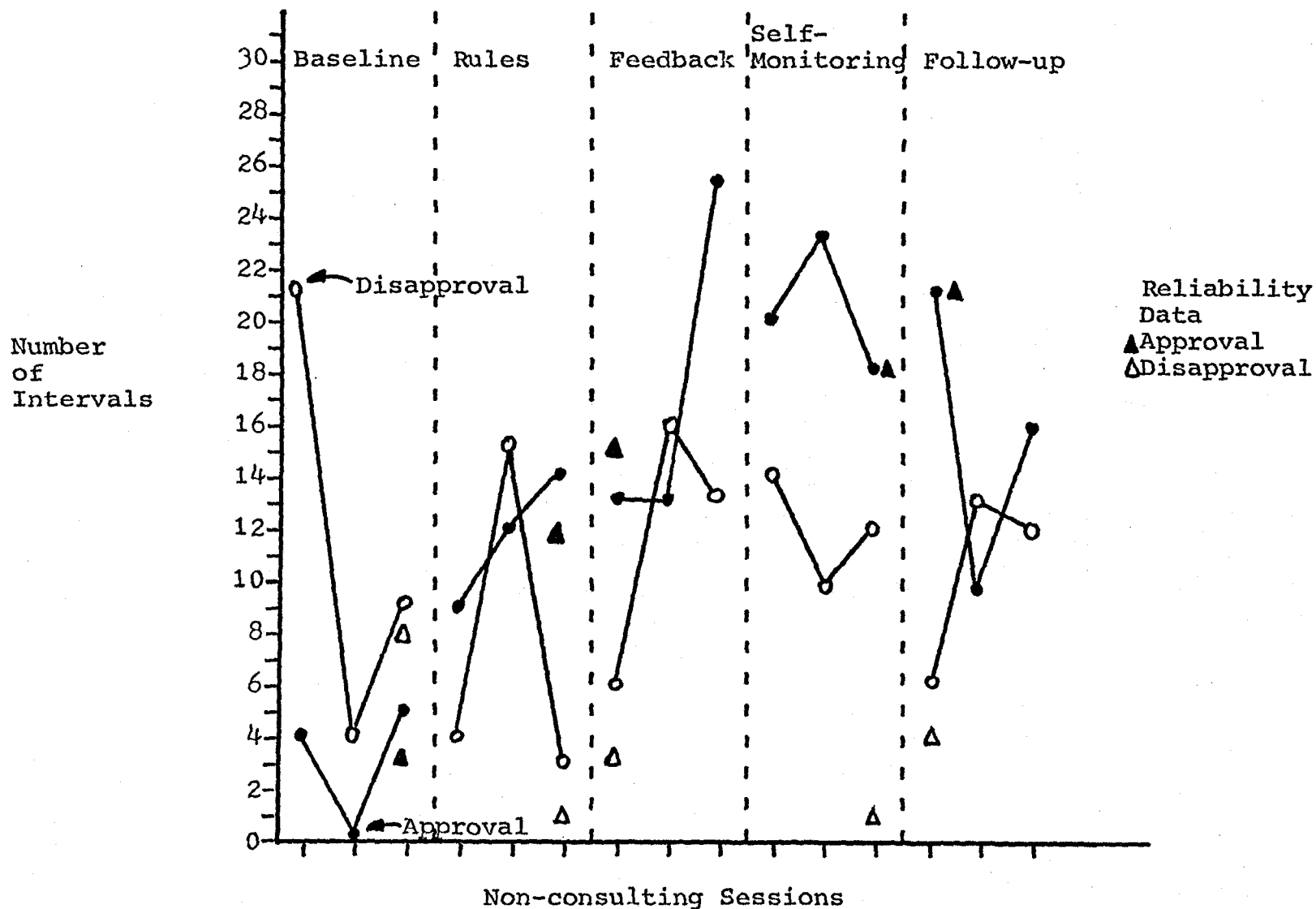
Figure 10. Number of approval and disapproval intervals per non-consulting session as recorded by the regular observer. The data collected by the reliability observer during each phase is shown. Teacher 2.
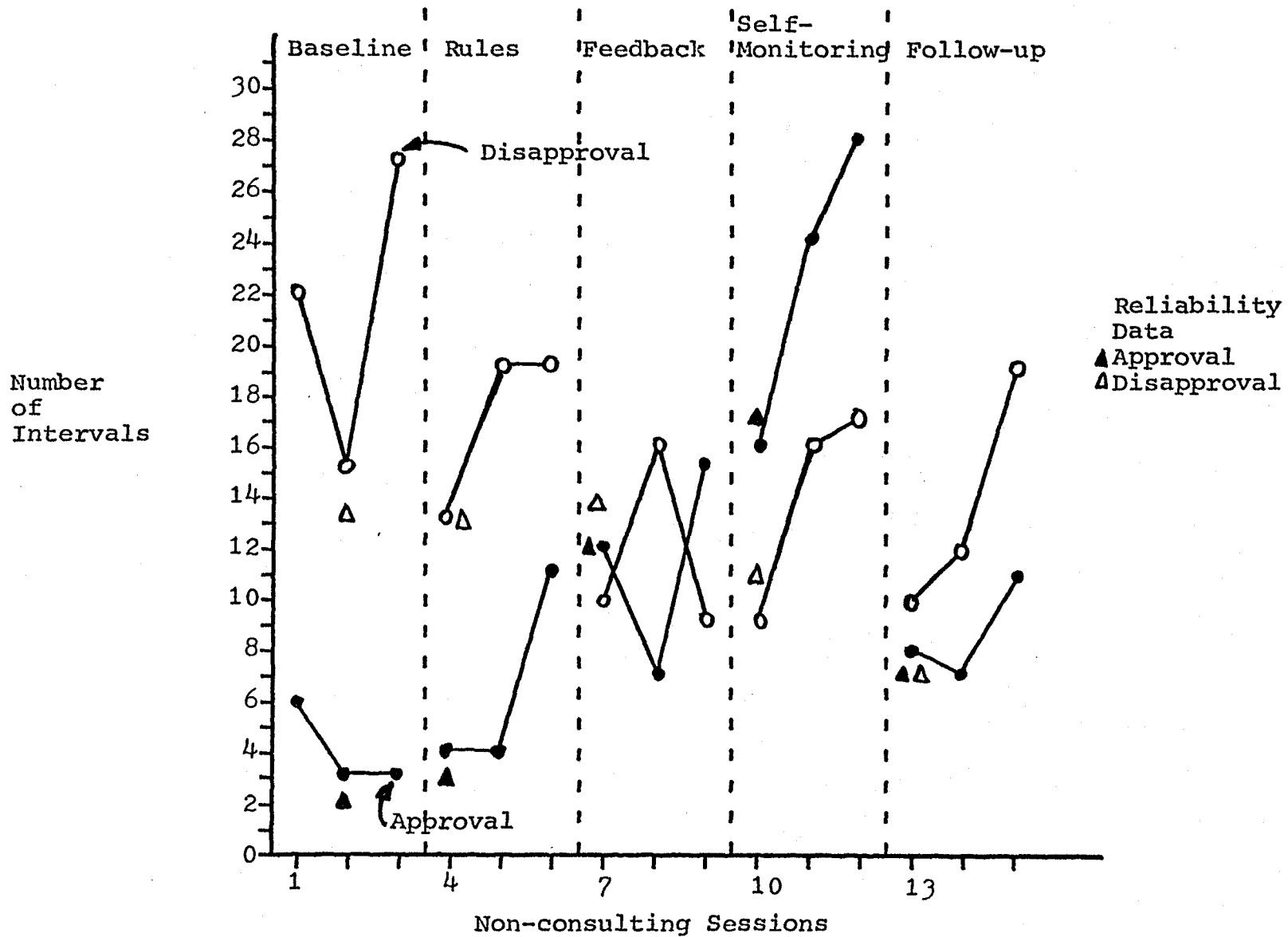
Figure 11. Number of approval and disapproval intervals per non-consulting session as recorded by the regular observer. The data collected by the reliability observer during each phase is shown. Teacher 3.

TABLE C

Reliability of Student On-Task Behavior

	Phases				
	1	2	3	4	5
Classroom 1	94.4	88.3	83.3	93.9	96.1
Classroom 2	91.7	90.6	92.8	92.2	95.6
Classroom 3	88.9	87.8	83.9	88.9	90.6

Table C. Reliability of student on-task behavior for each
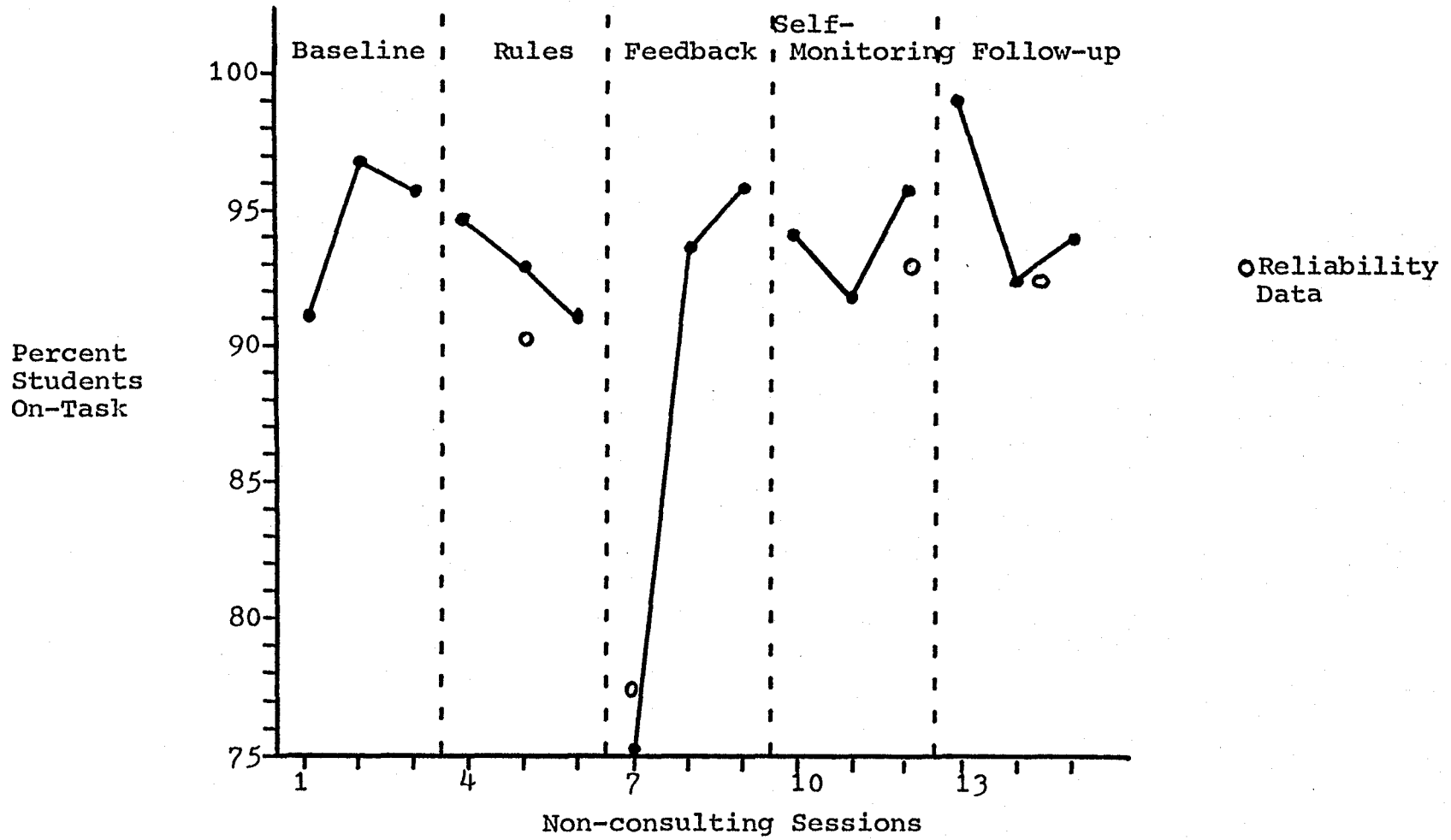classroom during each reliability session per phase.

Figure 12. Percent of target students on-task in the
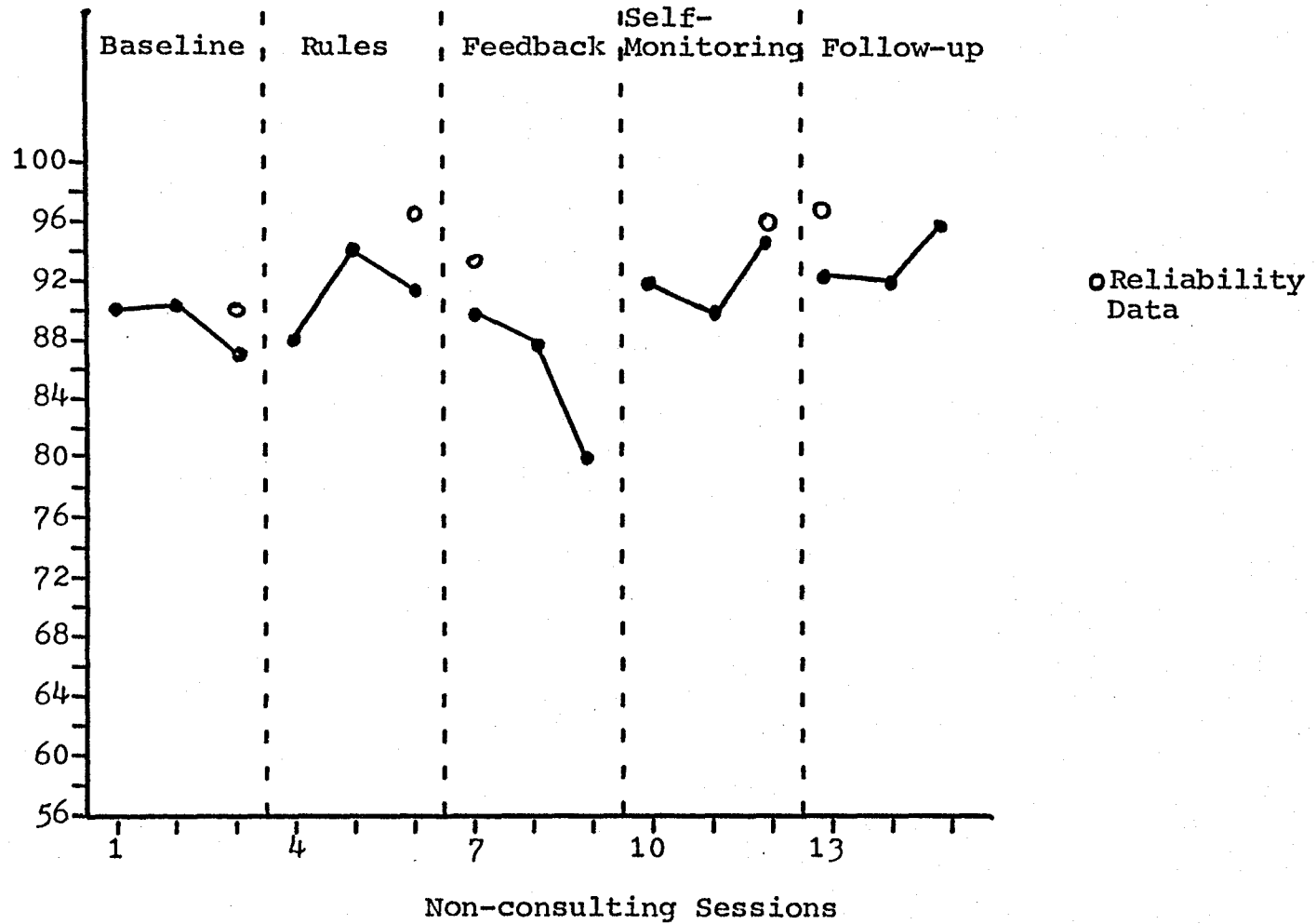classroom of Teacher 1 during the non-consulting sessions.

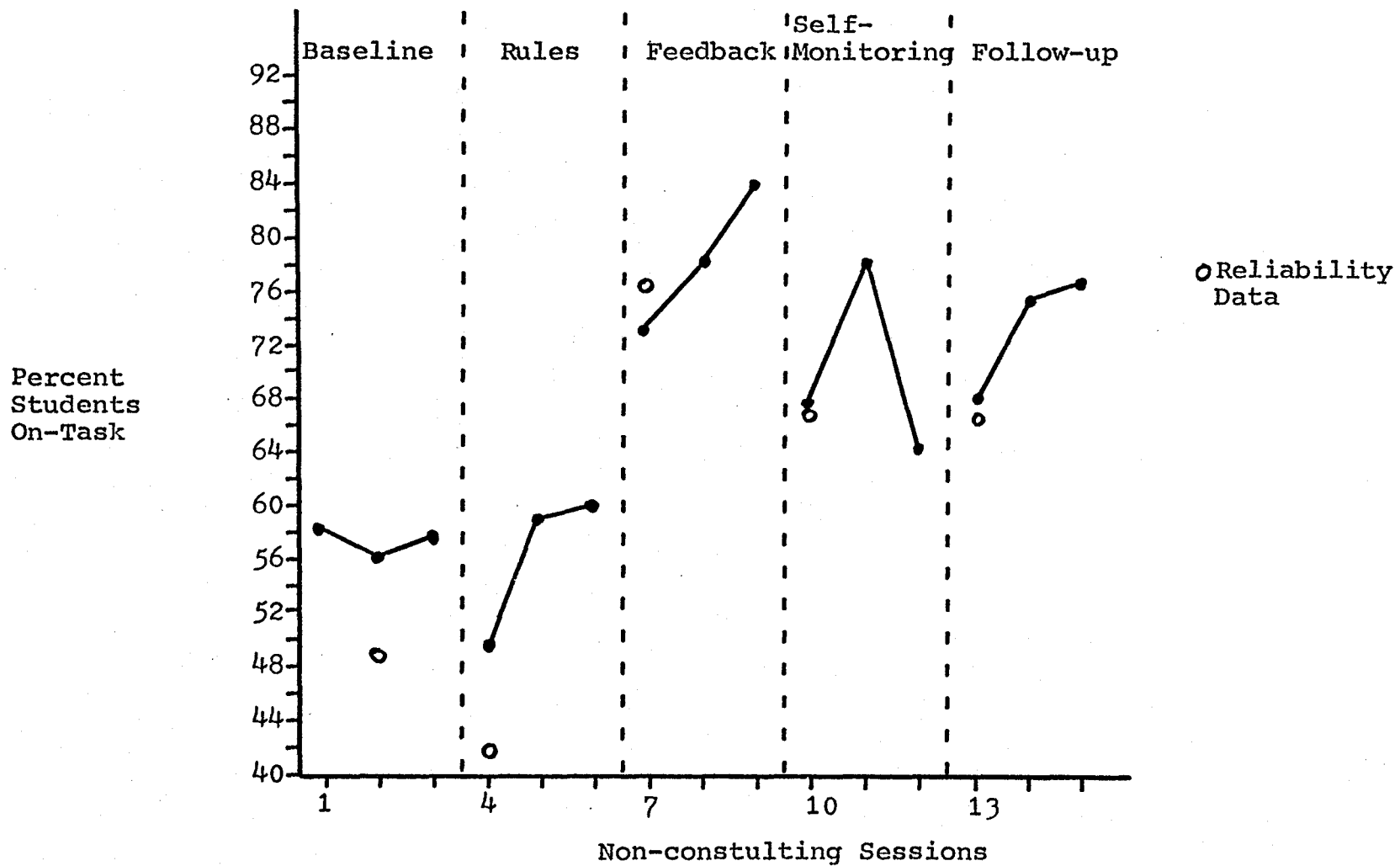Figure 13. Percent of target students on-task in the class-room of Teacher 2 during the non-consulting sessions.

Figure 14. Percent of target students on-task in the class-
room of Teacher 3 during the non-consulting sessions.

APPENDIX D

Feedback Form

```
┌─────────────────────────────────────┐
│ APPROVALS                            │
│         ┌───────┐                    │
│   Goal  │       │                    │
│         │       │         _____     │
│         └───────┘                    │
├─────────────────────────────────────┤
│ DISAPPROVALS                         │
│         ┌───────┐                    │
│   Goal  │       │                    │
│         │       │         _____     │
│         └───────┘                    │
└─────────────────────────────────────┘
```

VITA

Kenneth Slater Roach was born on October 5, 1948, in Radford, Virginia, and reared in that city until he was almost 16 years old, when his family moved to Nelson County, Virginia. He received his high school diploma from Virginia Episcopal School, Lynchburg, Virginia, in 1967, and he earned a B.A. in psychology from the University of Virginia in 1971. From September, 1972, through August, 1973, he gained experience as a Psychologist Assistant at Lynchburg Training School and Hospital, Lynchburg, Virginia.

Mr. Roach entered the graduate psychology program at the University of Richmond in August, 1973. After completing the courses required for the M.A. degree, he worked as a school psychologist for Title III and then Title I in the Chesterfield County Public Schools. Mr. Roach completed an internship in school psychology and is currently finishing the requirements for certification as a school psychologist in Virginia.

Mr. Roach expects to receive his M.A. in psychology from the University of Richmond in August, 1976.