THE EFFECTS OF VISUAL FEEDBACK FROM A CHART UPON THE RATE OF ACADEMIC PERFORMANCE OF JUNIOR AND SENIOR HIGH SCHOOL STUDENTS

By

PATRICK A. QUIGLEY, JR.

A DISSERTATION PRESENTED TO THE GRADUATE COUNCIL OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA
1972

To Charlotte

ACKNOWLEDGMENTS

The task of sorting out the contribution of various environmental changes with relationship to various behavioral changes is often a difficult task when precise data are not kept. When the target behavior is professional scholarship which has been shaped over a four-year period, then accounting for the contribution of each environmental change becomes even more obfuscated. There are a few contributors in the case of this dissertation who, however, have had an effect beyond all others:

First, to Dr. John Newell, who was the first to teach me learning theory and then to reinforce my beliefs in this area over a three-year period.

To Dr. Henry Pennypacker, who opened new vistas in my training as a learning theorist through precision teaching.

To Dr. Emile Jester, who both as my advisor for the master's degree and later as a committee member for the doctorate helped me realize the importance of psychology as the mother science for educational psychology.

To Dr. William Wolking, who helped me realize the applicability of the experimental analysis of animal behavior in the area of stimulus control to school learning. My

debt to him has not been paid in this dissertation and must wait future payment.

To Dr. William Ware, who taught me a good deal of all that I know in the areas of statistics and research design. His presence on the committee helped me realize the positive relationships between standard research methodology and the experimental analysis of behavior.

To Joseph and Mary Otchin, I owe a wealth of thanks for their support over the past few years.

Many thanks are due Pat Craig, Linda Howell, Pat Nobles, and Margaret Tomlinson for initiating this research and using their classrooms to further our understanding of junior and senior high school students.

Finally, and most important, to Charlotte, my wife, and to Leigh, Catie, and Shannon, my children, who, as the most significant features of my immediate environment, provided most the incentive for finishing this work.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO SYMBOLS OF ABBREVIATIONS

RCS - Rate Computation Sheet

SBC - Standard Behavior Chart

KOR - Knowledge of Results

Abstract of Dissertation Presented to the Graduate Council
of the University of Florida in Partial Fulfillment
of the Requirements for the Degree of Doctor of Philosophy

THE EFFECTS OF VISUAL FEEDBACK FROM
A CHART UPON THE RATE OF ACADEMIC
PERFORMANCE OF JUNIOR AND SENIOR
HIGH SCHOOL STUDENTS

By

Patrick A. Quigley, Jr.

December, 1972

Chairman: John M. Newell
Major Department: Foundations of Education

The purpose of the study was to produce self-control of academic behavior in ninth and tenth grade typing and eleventh grade accounting students by providing continuous, self-generated feedback through various devices and to assess the relative merits of each device. A review of prior self-monitoring studies suggested the appropriate controls used in the research design. The devices themselves proved no different from the controls. The accelerations in the past posttreatment baseline proved significantly different from all other phases for two sets of data. A third set of data only yielded a trend toward significant change during the last posttreatment baseline. A review of some individual data revealed directional changes in rate which were not measured by the analysis of variance technique. Results suggested that students demonstrated self-directed change of performance rates with the devices; terminating the last phase conjunctively with the end of the school year explains

the fixed interval scallop; statistical significance may not be sensitive enough to reflect meaningful changes in individual behavior; and weighting daily work more may reduce scalloping and assure the longevity of self-control through the end of the school year.

CHAPTER I

INTRODUCTION

Purpose

The purpose of this study was to examine the effects of self-charting on academic behavior in business education classes. The charts served as a vehicle whereby the student might generate his own feedback with respect to his daily classroom performance. Efforts were made to leave the ongoing procedures of the classes intact so that the effects of the chart alone might be clearly assessed. This research will have significance in two respects: first, the chart was studied as a device for shaping self-control of academic behaviors; and second, in terms of future work with self-generated charts, the experimenter will be informed as to the reactive effects of charting and thus the need to incorporate additional control groups or phases within an experiment to assess the effects of the independent variables.

Problem

Currently in the field of education there is a move afoot to make some of the goals of the educational process

behaviorally or concretely measurable. The recent clamor concerning such practices as performance contracting, accountability, contingency contracting, merit pay, and even behavioral objectives attests to this underlying movement toward the concrete statement of aims in those areas where such a thrust is technologically feasible. Generally speaking, those subject areas where the acquisition of a fundamental skill is involved are receiving most of the attention. These areas are encompassed in the traditional domains of reading, writing, and arithmetic. Other, more difficult to define areas of educational goals are not receiving as much pressure for reform. Under this heading may be lodged such goals as moral education, cooperation, and self-control, to name a few. Given the hurried atmosphere of most contemporary classrooms, the task of reform seems least likely where the teacher must try some types of innovative practices in addition to the handling of demands for skill acquisition.

Concerning the more ethereal goals of the process, the teacher at most may try to serve as a behavioral model, to issue exhortations, or to reprimand the students for undesirable conduct. Thus, while these goals may be of equal importance to those of skill learning with reference to the student as a future citizen, they receive little attention and must take second place until the skills are acquired or until a student misbehaves.

With the use of an experimental analysis of behavior, there remains the possibility of combining both the skill acquisition and the learning of self-control techniques. This is only possible because an experimental analysis of behavior requires the behavioral definition of any dependent variable under study. While school systems do not often state their goals in terms of experimental analogies, this language easily lends itself to the statement of outcomes or goals. Thus the teachers, curriculum, and peers of the student provide the major independent variables or environment in which the student's behavior is to be changed over a course of 12 or more years.

Self-control within the context of an experimental analysis of behavior entails two responses: the controlling response and the controlled response (Skinner, 1953). The controlling response is emitted by the behaver to manipulate those variables which are functionally related to the controlled response. Thus, the controlling response may determine the probability of occurrence and topography of the controlled response. In school learning situations, the controlled response may be identified as the student's academic behavior. The controlling response may be any behavior which the student emits to manipulate the occurrence and topography of his own academic performance.

The problem, then, of shaping self-control in classroom learning situations becomes one of identifying controlling

responses which the student may emit to manipulate or monitor his own academic behavior. Furthermore, in order to make such self-control feasible, the form of the controlling response should be one easily emitted by the student and involve a minimum of teacher influence. The former criterion should be tempered by the fact that at least initially the controlling responses of the students may involve some shaping to permit the students to emit the responses. However, violation of the latter criterion in any extended sense would by definition not permit self-control to be acquired.

Within the context of this research, the controlling response under study was the recording of academic behaviors on a standard behavior chart used in slightly different settings (Johnson, 1971; Starlin, 1971; Haughton, 1971; Duncan, 1971). After a small amount of instruction and practice in the use of the chart, the students were left alone to record their own daily academic behaviors. It was hoped that, by generating their own feedback for a period of days, the students would be able to manipulate the rate of their responses in a desirable direction. It was the thesis of this research that the visual display generated by the student on the chart would enable him to control his own academic behaviors.

In terms of research methodology, the study of self-control explores the reactivity of any given setting or

test. If the setting or testing process produces cues as to the acquisition of behavior which in itself will be detected or influence later testing, then the test or setting is said to produce a reactive effect (Campbell and Stanley, 1963). The same may be said of sampling procedures which change the members in any fashion. While such an artifact may be a problem to experimenters, it may be of benefit to educators searching for devices to shape self-control in students.

CHAPTER II

REVIEW OF LITERATURE

This chapter surveys the research and methodological literature pertinent to this study. For purposes of organization and clarity, this survey of the literautre is presented in a number of sections which include: self-control, self-monitoring methodology, self-monitoring in clinical settings; self-monitoring in school settings, and knowledge of results through behavior graphs. A final section presents a summary of the literature discussed in each of the sections previously mentioned.

Self-Control

Over the past half-century, behaviorists have worked persistently to detail a precise science of behavior. At times, this search for law and order led them to confuse objectivity with empirical verifiable facts. These behaviorists argued that unless a phenomenon were visible and could be reliably measured, it had no place in a science. Such limitations on the bounds of psychology caused confusion both within the field of behaviorism (Watson, 1913) and to a greater extent outside of behaviorism (Scriven, 1972) concerning those phenomena subject to empirical test. By

operating within the limited framework of accepting only visible and observable behavior, psychology would never adequately account for such subjective phenomena as moods, feelings, urges, insights, thoughts, and free will. The rejection of this last state resulted in little research on self-control. To grant the existence of self-control as a legitimate phenomena for study by psychologists would be tantamount to admitting the existence of free will and thus backsliding to a rationalist position (Bolles, 1967).

One notable exception in the arena of behaviorists was B. F. Skinner. Rather than admit the reliance of self-control upon any special, internal mechanism (such as free will), he described self-control as a behavioral phenomenon which influences an individual's subsequent behavior (1953). Thus the behaver emitted a response, called a controlling response, which alters the occurrence of another response, the controlled response within the repertoire. The controlling and controlled responses may be either public (outer or observable) or private (inner or unobservable to the public) events (Homme, 1965; Premack, 1971). Since the private event was said to share the same properties as the public event (Skinner, 1953), self-control may become a legitimate field of study for the behaviorist.

For behaviorists working within the Skinnerian framework, the definition of self-control was operationalized in terms of the context of research or therapy under

consideration. Goldiamond (1965) defined self-control as a functional relationship between behavior and the environment. Ferster, Nurnberger, and Levitt (1962) referred to self-control as self-selected change in the environment to produce changes in the individual's repertoire. For Duncan (1969), self-control involved the self-selection and application of behavior modification techniques. Later work in the area of self-control led Duncan to adopt a more euphemistic term (personal management) with relation to the control of inner and outer behaviors (1971). Berger, one of Duncan's students, has produced the most extensive examination of personal management techniques for inner (private) and outer (public) behaviors (1972). According to Blackwood (1970), self-control consisted of the emission and reinforcement of verbal behaviors to countercondition some temptation to misbehave.

All of the above definitions may be encompassed by Skinner's definition which emphasized the effect of self-control on subsequent behavior. Where research has been done in conjunction with the development of self-control, self-monitoring has been used as a data collection device (Goldiamond, 1965; Ferster et al., 1962; Duncan, 1969, 1971; Berger, 1972). Most of these studies use self-monitoring in conjunction with other independent variables to change behavior. However, self-monitoring may be studied as an independent variable by itself (McFall, 1970). Other

articles which elaborate the possible use of inner behaviors in conjunction with outer behaviors suggest the use of self-monitoring without demonstrating its uses (Homme, 1965; Premack, 1971). Cautela (1969) and other articles to be mentioned later in this chapter suggested possible uses of self-control procedures in clinical settings.

Methodological Considerations for the Use of Self-Monitoring Procedures

Before examining the self-monitoring research, a review of methodological considerations is necessary to assess the validity and generalizability of the data in question. Since self-monitoring as a data collection technique was at first used extensively only in clinical settings, the articles written with respect to validity and generalizability of results concern clinical data. These same methodological considerations seem applicable to educational settings with few exceptions and thus were applied to these data.

As result of studies by McFall (1970) and Marston and McFall (1971) where self-monitoring was used as both a data collection device and one of the major independent variables for research on smoking behavior, the following methodological discussions were prepared (Orne, 1970; Kanfer, 1970; Mausner, 1971; Lichtenstein, 1971).

Orne's questions concerning self-monitoring data centered upon the three major points which characterized the other critiques. First, are behavior changes in

self-monitoring settings due to unique treatments or may they be interpreted as a result of some demand characteristic of the experiment? If the noted behavior change may be attributed to some subtle cues from the experimenter or from the larger setting of just being in an experiment, then demand characteristics may account for the treatment effects. For example, if the subject either by volunteering or by learning the goal or nature of an experiment behaved accordingly, the validity of experimental effects must be questioned. Orne suggested the use of additional control groups to assess the effect of the major independent variable (or variables) beyond the demand characteristics of an experiment. Second, in those instances where a behavior is occurring during all wakeful hours, what is the generality of data taken in limited samples? This question is answered by either taking day-long samples or by being relatively certain that the behavior under study only occurs during the sampling time. Third, Orne questioned the reliability of the data gathered under self-monitored conditions. He raised the possibility of the subject reporting faked data. For example, when an observer of the subject's behavior is a colleague or relative, there could be some communication between the observer and experimental subject. Thus reliability measures taken to validate the self-monitored data should be independent of the subject's recording. Having an observer both personally or socially unknown to the

subject would constitute an ultimate answer. As will be noted later, this objective was often compromised in experimental studies which attempted to naturalize the setting and facilitate the collection of reliability data.

Kanfer (1970) raised many of the same questions and went further to state that the concurrent reliability and validity of self-reports were maintained when reliability of reports between subject and observer were high. The nature of the observer was given a broader definition to include behavioral products as evidence of a behavior occurring. Thus showing butts for smoking, weight loss for reduced consumption rates, longer fingernails for nail biting, fewer bruises on the wife's body for wife-beating urges may all serve as concurrent validations of self-monitored data. Kanfer also noted that the self-monitoring may serve as a controlling response and thus facilitate self-control. In addition, he viewed the results of self-recorded data as knowledge of results. Hence, the self-recording techniques used in this study were related to the general discussion of knowledge of results data later in this chapter. Finally, he proposed that self-monitored data was a poor substitute for control procedures prior to behavior modification. Thus some independent measure of responding or response rates should be gathered prior to any self-monitoring to capture the results of the recording alone.

Mausner (1971) added little to prior methodological considerations. However, as an alternative to recording public behaviors, he mentioned as a possibility the recording of a private event which represented a competing response for the public's behavior under study. For example, clicks on a wrist counter may be used to count the number of times resistance to smoking urges occurs. If there was a concurrent reduction in actual smoking rate, the data on the private event are valid (Homme, 1965; Premack, 1971).

Lichtenstein raised a further concern which was more relevant to meaningful therapy than to the research methodology of self-monitoring. In the process of doing research, the experimenter often categorized the treatments into various groups to obtain experimental control. With regard to the best outcomes for each individual in the experiment, any one subject may not receive the most promising treatment due to random assignment. One reply to this concern would be to use counterbalanced designs which permit each subject to receive each treatment or to base the construction of manipulation upon previously demonstrated treatments from small sample or single subject research.

Self-Monitoring in Clinical Settings

The clinical literature on self-monitoring dealt with a variety of human behaviors in many different settings.

References to self-monitoring as a self-control device were not always made. A few studies were reported in detail to illustrate the use of the criteria discussed in the last section. McFall (1970) used his own class to study the smoking behavior of college students. In this study, McFall asked two questions: first, does self-monitoring alter behavior; and second, does the nature of the target behavior yield differential effects in the self-monitoring situation? The validity of the data were judged by the reliability of observers covertly monitoring two treatment groups. The first group was asked to count the number and time spent smoking each cigarette in class and the second group was asked to count both resistance to urges and the total number of cigarettes smoked. Each of the sixteen smokers in the class was placed into one of the two treatment groups and sixteen non-smokers were assigned to monitor the smoking behavior of the smokers. A baseline period of nine days was followed by thirteen days of treatment and then an additional nine days were added as a return to baseline conditions. The smokers recorded only during the treatment phase while the observers recorded throughout the experiment. In order to facilitate smoking during class, the experimenter lit a cigarette and remarked to the students that the no smoking sign in the room only applied to afternoon classes. Furthermore, the experimenter expressed an interest in collecting data on smoking behavior from the students. Both treatment groups demonstrated changes

during the self-monitoring phase and these changes persisted during the rebase phase when self-monitoring was discontinued. Two of the eight pairs of subjects reported data for a day in which no class was held, and thus all of their data was suspect. For the other six pairs during the treatment phase the reliability between subjects and observers was .61 and for each pair taken separately the reliabilities ranged from -.05 to 1.00, with three of the six pairs above .75. McFall concluded that the self-recording was a reactive measure and thus should be assessed independent of any other treatment.

Orne (1970) and Kanfer (1970), commenting upon this study, suggested the possibility of demand characteristics explaining the treatment effects rather than the treatments themselves. By lighting a cigarette, the experimenter may have communicated his hypothesis that self-monitoring would change the smoking rates. The reduced smoking rate in the "resistance to urges" group may be due to instructions alone. Stated more specifically, the reviewers suggested that the same results might have been obtained by merely telling one group of students to smoke more and the other less because they were in an experiment. Since it may be safely assumed that the smokers in the study continued smoking all day, the generality of the data based upon the limited sample was questioned by the authors. In lieu of evidence to assure honest reports of those students not caught cheating, the authors raised the possibility of a

pact between smokers and observers. Given the opportunities for socializing in most college classrooms, this suspicion does not seem unreasonable.

Powell and Azrin (1968) studied the functional relationships between the contingent, self-application of electric shock and cigarette smoking. Three subjects volunteered for the study and two of them gave names of co-workers who would observe the smoking five to twelve times a day. The subjects gave a self-report of their smoking history and kept records of the count and time of the day smoking occurred. The major treatment was the use of a cigarette box which delivered an electric shock upon opening. The value of the shock was varied from 0.0 milliamperes to 2.0 milliamperes. The "participant-observers" served to insure the subject was using the device, collect smoking data, and check the working order of the cigarette box. The percentage of subject-observer agreement was considered acceptable for both subjects (98% and 88%). The main finding was that as the intensity of the shock increased, the smoking and the use of the box decelerated. Thus while the subjects found other means to obtain cigarettes, there was also a concurrent decrease in the total number of cigarettes smoked. Once the shock was removed, the smoking returned to its pretreatment level. The salient feature of this self-monitoring study is obvious; the participant-observer system insured reliability. While some suggestions

of a pact between the subject and the observer might be raised, this was thought unlikely since the subjects volunteered reports of not using the shocking device as according to instructions. It would have been interesting to have a covert observer gather additional data both when the participant-observer was present and at another time when the participant-observer was absent. These data would both settle the reliability question and perhaps reveal the extent to which the participant-observer may become a discriminative stimulus for non-smoking. If this third observer were paid, the social reasons for colluding would also be minimized. The same authors have also reported the use of the participant-observer system with similar reliability in a smoking study using discriminated extinction through a cigarette box which sounded an alarm, and unlocked after an elapsed time interval (Azrin and Powell, 1968) and in a study using response priming to improve prescribed self-medication (Azrin and Powell, 1969).

None of the studies reported hence have incorporated follow-up data collection to any significant extent to demonstrate potent treatment effects. With a behavior like smoking, long-term abstinence would be the only obvious demonstration of experimental control. In an attempt to compare several types of behavior modification techniques (anti-smoking pill, stimulus satiation, extinction, and hierarchical reduction), Marston and McFall

(1971) set up a smoking clinic to attract volunteers. Self-recording in a diary form was required of all subjects. The smoking behavior patterns were typical of other data using similar therapies. The only form of reliability data available was through the number of cigarettes bought at the clinic. Follow-up data revealed no differences due to treatment as all subjects had approximately recovered pre-experimental smoking rates.

After a review of previous stop-smoking literature when adequate follow-ups were done, McFall and Hammen (1941) agreed with Bernstein's (1969) conclusion that non-specific or secondary factors accounted for the reliable behavior patterns of temporary reduction and then long-term resumption of smoking. The non-specific factors isolated were: motivated subjects, structured participation in a program, and self-monitoring. In order to create a climate conducive to these factors, the authors advertised a clinic in the student newspaper, asked participants to postdate a twenty-five dollar ($25) check to insure against subject withdrawing during the study, and set up four treatment conditions. In the first condition, the subjects only handed in a daily report of smoking; in the second, the subjects used a wrist counter to monitor smoking and they said aloud, "I do not want to smoke" with each urge; in the third, the subjects awarded themselves twenty non-negotiable points to be counted on a wrist counter for each urge resisted;

and in the fourth, the subjects recorded resistances to temptations to smoke and award themselves one non-negotiable point for each occurrence on a wrist counter. Data collected six weeks and six months after treatment revealed no differences due to treatments. The authors concluded that the non-specific factors yielded as great a result as the more elaborate, theoretically derived, and presumably more potent procedures. In an earlier study comparing self-monitoring alone with other self-control treatments for smoking behavior, Rutner (1967) obtained the same results. As for the smoking therapies themselves, Bernstein (1970) has published a more extensive review of the literature than a review cited earlier.

Self-monitoring has also been used in other clinical settings. Leitenberg, Agras, Thompson, and Wright (1968) gave phobic subjects a stopwatch to self-monitor their own relaxation times in desensitization therapies. The self-monitoring was validated against automatic timing devices. Initially, the self-monitoring was confirmed and praised by the experimenter in conference with the subjects. Eventually, the experimenter feedback was faded and follow-up data revealed positive results of the therapy. Weight control studies have shown positive results with self-monitoring either by itself (Stollak, 1967) or in connection with other environmental changes (Ferster et al., 1962). Self-monitoring in diary or chart form with other

environmental changes (Goldiamond, 1965; Duncan, 1969) and chart form alone (Kolb, Winter, and Berlew, 1968) have proved successful in changing various social behaviors. In a study which revealed unreliable self-reporting and peer-reporting, Fixsen, Phillips, and Wolf (1972) demonstrated the use of various progressive ratio schedules of reinforcement to increase reliability among delinquents in an innovative therapy center.

Finally, Duncan (1971) and Berger (1972) demonstrated the use of self-monitoring for inner and outer behaviors. Both studies used the Standard Behavior Charts (Koenig, 1972) to record self-monitoring. Duncan had a three-and-a-half-year-old girl self-monitor selfish thoughts and acts and a twelve-year-old girl self-monitor feelings of anger and outbursts. Since the inner and outer behavior patterns were similar, it was assumed that the concurrent reports confirmed one another. Berger enlarged Duncan's strategies to include outer behaviors alone (twenty-three cases), inner behaviors alone (thirty-seven cases), inner and outer behaviors (three cases in which the urge or the act were recorded as the same response), and inner-outer behaviors (eight cases in which the movement cycle was in part outer and in part inner). These last behaviors included such pinpoints as eating when hungry and the occurrence of a meaningful conversation. The author offered as criteria for valid data such criteria as amount of "bounce" from

day to day on the chart, sequence of high and low rate days, correlation between charts from the same person, and length of charting period. These criteria were offered to make self-charted data valid in themselves without continuous observer checks. An example of the predictive validity of these criteria was given for an inner behavior of one subject. While the need for continual assessment of reliability remains an open issue, Berger's work made a substantial contribution in exploring new movement cycles and suggesting criteria for assessing the validity of self-charted behavior itself. It may prove that the use of movement cycles such as inner and outer behaviors and inner-outer behaviors will never satisfy the traditional methodologist and that criteria such as Berger listed will provide the only means of judging validity.

Self-Monitoring in School Settings

Self-monitoring has been used less frequently in school settings than clinical settings. Two studies of reliability under peer-monitoring conditions for peer-tutoring revealed 90 percent reliability between a fifth grader and his teacher for monitoring the academic behaviors of fourth graders (Surratt, Ulrich, and Hawkins, 1969), and 90 percent accuracy of plotting rates on the Standard Behavior Chart for a first grader tutoring classmates (Starlin, 1971). Both studies also pointed to the successful and timesaving features of peer-tutoring.

In an investigation of the relationship between non-verbal and verbal behaviors of preschool children, Risley and Hart (1968) found a low correspondence between the self-reports of play equipment use and the actual use of the materials. When food snacks were made contingent upon correspondence between the use of materials and the self-report, the use of the materials accelerated to the level of the self-reported use. This effect also generated to the use of other materials in which the correspondence was not made contingent upon food snacks.

Broden, Hall, and Mitts (1971) studied the effects of self-recording on the classroom behavior of two junior high school students. The first student went to the school counselor and volunteered to record the frequency with which she was on and off task during a history class. Two observers entered the class to record the percentage of ten-second intervals the girl was on or off task. Self-monitoring brought an increase in study time. The agreement between the two observers varied from 87 to 96 percent. Since the subject did not record on a fixed interval basis and even failed to record at times, the direct reliability between the observers' and the student's report was questionable even though the overall percentages were equivalent. Once teacher attention was increased, self-recording was no longer necessary to maintain study behavior. The second subject did not volunteer. He was told by the teacher to record on paper slips the number of talk outs for half of

the class session. Observers monitored the student as in the previous case (agreement ranged from 84 to 100 percent). There was little correspondence between the observer and the subject report. After the slips were removed, the talk-out rate accelerated. In this case, teacher praise was not manipulated with the fading of self-recording. The authors concluded that this fact accounted for accelerated talk-out rates during the last phase of the study.

The criteria listed in the methodology section of this chapter may be applied to the self-recording studies in school settings. A critique of Broden et al. (1971) suffices as a demonstration. Data based upon the self-recording of the first subject do not lend clear support to the positive effects of self-monitoring since the absence or presence of the recording slips alone may have accounted for the results. Broden et al. (1971) stated that physical presence of the paper slips alone may account for the results since the self-recording was often omitted by the subjects. To reiterate a previous question, could the same results be obtained by merely telling a subject at various times that an experiment is in progress? Since the data collected by the observers and the subject were gathered under different conditions (interval sampling as opposed to unspecified recording), the reliability must be questioned. As McFall and Hammen (1971) concluded, unspecific factors such as a formal setting or subject motivation may underlie other treatment effects. In this study, the formal setting and motivation (i.e., subject

initiated counseling) were confounded with the effects of self-recording. For the second subject, motivation due to volunteering was not a confounding feature. However, by having the teacher in control of the intervention, differential teacher attention may have accounted for some the effects. While the authors reported taking observations of teacher attention to both subjects, the data were missing for the second subject and thus the omission raises additional questions. Since the reliability for the second subject's recording was poor, this study represented, perhaps, more the effects of non-specific factors than of self-monitoring.

Knowledge of Results Through Behavior Charts

As Kanfer (1970) noted, information provided through self-monitored data qualified as knowledge of results (KOR). Studies demonstrating KOR through self-monitoring and self-charting have already been reviewed (eg., Duncan, 1971; Berger, 1972). There were a few examples in the literature of KOR through behavior charts which were not self-generated. The behavior rates were plotted on a chart by a supervisor or teacher and then shown to the behaver for evaluating his performance. Johnson (1971) used the Standard Behavior Chart to help one student in her first-grade class learn mathematics and another decelerate his rate of yawning. Many more examples of the effects of the Standard Behavior Chart as a decelerator or accelerator may be found in the handbook of Precise Behavior Facts (1971).

Jeus and Shores (1969) working with three trainable mentally retarded adolescents found that showing a chart of the previous day's work on a simple assembly task accelerated performance on the same task. Equivalent procedures did not yield similar results for ring assembly and packaging tasks.

With regards to the KOR literature itself, much work needs to be done to demonstrate the reinforcing or non-reinforcing value of KOR (Geis and Chapman, 1971). While pointing to self-monitored data collection as an example of KOR, the authors stated that the general reinforcing value of KOR in relation to schedule control, immediate versus delayed, and nature of the reinforcement remained an open question.

Summary

This chapter has reviewed the historical context of self-control with respect to the realm of behavioristic inquiry, the methodological considerations necessary for valid data collection in regards to one form of self-control--self-monitoring, and the behavioral literature of self-monitoring in clinical and educational settings. Self-control was defined as encompassing two responses--the controlling response and the controlled response. Self-monitoring of behavior as a controlling response was the focal point of the review. Criteria for valid data collection included controlling for demand characteristics of experiment so

that instructions alone would not account for treatment effects, reliability of self-monitored data with some unbiased source, assessment of the response rates or other measures of the behavior before self-monitoring began, control for any non-specific factors underlying the experimental setting (e.g., structured setting, and motivation of the subjects), and adequate sampling of the controlled response to insure generality of the findings.

All of the literature reviewed pointed to the positive effects of self-monitoring and thus qualified the self-monitoring response as a controlling response for the acquisition of self-control. The main types of self-monitoring responses were in diary form, note taking, and on behavioral charts. Furthermore, the behavior chart proved effective as a device for behavior change whether it was self-generated or not.

With the exception of the McFall and Hammen (1971) study, all of the studies failed to control for one or more of variables demanded by the criteria. Since the study by Berger (1972) both included some new behavioral definitions and cited new sets of criteria for the assessment of validity, the future course of scientific inquiry alone must judge the validity of the new criteria. The contemporary technology cannot adequately assess the reliability and validity of the movement cycles used.

As a result of this review, it is possible to restate the purpose of the research at hand and justify the methods

and design using the criteria cited. As the title of the present study states, this study was designed to assess the feedback through a chart. Not all of the non-specific features of the setting were controlled, rather a few were purposely left unmanipulated. Thus the structured setting for instruction in recording on unfamiliar forms was tested against a familiar form of response recording (i.e., posting responses and work times on manila envelopes). Two new types of recording, one of which was the Standard Behavior Chart, were studied in relationship to another, more familiar type. The control for demand characteristics through subjects guessing the hypotheses of the experiment by instructions was accomplished by exposing all subjects to the hypothesis or instruction that all self-monitoring improves performance. Short baselines were inserted between treatment phases to assess the persistence of experimental effects. In order to eliminate a bias in the results due to teacher commitment or differential teacher attention which might influence student performance, the treatments were counterbalanced in a Latin squares fashion with the use of the old recording type continuing while the new recording types were introduced. As it was clear from the literature, the self-monitoring is reactive for a number of reasons and thus serves as a poor control technique prior to manipulation. Therefore, an ability stratification relevant to prior classroom

performance was obtained to assess the approximate levels or rates of the individual performance before any type of self-monitoring was begun. The reliability of the self-recording was checked by each teacher during the term. The teacher made comments concerning the high reliability. One set of data was statistically examined to reinforce these comments. The observers and subjects were unknown to each other.

Finally, two points deserve further note. First, if certain variables are known to be reactive, it is only possible to examine the effects of one reactive variable against other reactive variables with the treatment of concern being hierarchically eliminated from other reactive variables as controls. Thus the use of Standard Behavior Chart was contrasted with the use of Rate Computation Sheet (RCS) not because of any special properties ascribed to the RCS, rather because it required recording responses. Second, the generality of findings based upon academic behavior sampled in classrooms is not as questionable as taking a small sample of behavior in a clinical setting. In a good number of cases, the academic behavior is under the stimulus control of classroom. The behaviors observed in the clinic can usually occur all day long. Thus the methods and design of this study proved adequate to assess the effects of visual feedback through a chart upon the rate of academic performance of junior and senior high school students.

CHAPTER III

METHODS, PROCEDURES, AND DESIGN

Subjects and Settings

As a result of a workshop given to business education teachers, three classrooms were secured for research purposes. Two of the classes were introductory typing at the ninth- and tenth-grade levels while the third class was an introductory accounting class at the eleventh-grade level. The academic behaviors for charting the two typing classes consisted of three-minute, timed typing of a sample from the texts.

In the introductory accounting class, the curriculum consisted of workbook exercises and bookkeeping activities. The students in the two typing classes were accustomed to identifying the errors made during the three-minute typing sample and making note of these along with the total number of words or words per minute on envelopes which the students used to file each exercise. This recording on envelopes was introduced well in advance of the first baseline for all three classes. For the ninth-grade class in which the students posted the words per minute, the teacher had student-aides proofread the

exercises at a later time in order to identify any additional errors overlooked by the typist. The other class, which just posted the total number of words, exchanged their papers with fellow students who then proofread the paper. In both cases, number of additional errors identified by the proofreader was posted on the typist's envelope. The essential differences then, between the typing classes, consisted of posting the typing rate for three minutes and the delayed feedback of a day on the number of additional errors found by a student-aide in the ninth-grade class as opposed to posting the total number of words typed for three minutes and the immediate feedback from a fellow student concerning additional errors identified in the tenth-grade class. Students in both typing classes used the vertical scale in the texts to compute the total number of words typed. The sample for typing was changed each day during the experiment.

The ninth-grade class was composed of 29 students of which 26 were females and 3 were males. Of this number, 24 females and 3 males were included in the study for analysis purposes. Of those not included for statistical analysis of this class and the other two classes, some were discarded for reasons of prominent absenteeism and the rest were discarded by the randomization procedures employed in the selection of subjects. These students did, however, undergo the same treatments as the rest.

The tenth-grade class was composed of 25 students of which 15 were females and 10 were males. Of this number 10 females and 8 males were included in the study for analysis purposes. The remaining students underwent the same treatments but their data were not included in the statistical analysis.

The third class consisted of eleventh-grade, introductory accounting students. These students worked at their own pace through workbook exercises of fill-in-the-blank questions and longer problems from the text of bookkeeping activities. The sequence of classroom activity involved reading the introductory material in a unit of the test; answering short questions concerning the reading in the workbook; and finally completing the bookkeeping problems from the text in the workbook. The teacher made herself available during class time on an individual basis to help students and to provide answer sheets. The longer problems were self-correcting because columns would not balance if numbers were entered incorrectly or if there were mistakes in addition or subtraction. By providing answer sheets, the teacher made attempts to have errors corrected as soon as they occurred. These students were accustomed to posting responses on envelopes before experimentation began. The response unit used for analysis consisted of any fill-ins, column entries, and mathematical computations accomplished during the total class time. Due to the modular scheduling

at this school, the duration of class time varied from 50 minutes on three days of the week to 33 minutes for the remaining two days.

In the Accounting class, a total of 21 students were available for the study. Sixteen were female and 5 were male. Of the students selected for the experiment, there were 13 females and 5 males.

Procedure

In addition to the daily classwork performance by all students in each of the three groups, each group of students was assigned three different methods for recording the results of their classwork for the day. The first method of recording was "no method." This recording procedure consisted of employing the usual method of recording classwork performance that had been used in the classes before the treatment phase of this study began. The second method was the use of Rate Computation Sheets (RCS) only. A rate computation sheet is designed to facilitate the computation of behavior rates. Spaces are provided for entering the amount of time recording occurs and the number of movements observed during that time. Each division panel is keyed to the day of the week. Appendix 1 provides a sample frame for a week's recording on the RCS. The RCS forms were used by the students. The third method of recording consisted of the Standard Behavior Chart (SBC)

which is designed to provide a visual display of rates in graphic form. The ordinate axis allows for recording behavioral rates occurring as slow as once in a thousand minutes or as fast as a thousand per minute. This axis is arranged in a six-cycle, semilogarithmic fashion. The abscissa permits recording by the calendar day. Samples of the SBC are provided in Figure 4, 5, 6, and 7 (pp. 67, 68, 72, and 73) without the grid usually given to facilitate recording.

During a given time period of the treatment phase of this study, each group of students employed only one of the three methods described above but over the entire treatment phase each group employed all three of the methods of recording. The method of assigning each recording method to the three groups is described more fully in the design section of this chapter.

Instruction in the use of the rate computation sheets was given during the first two days of each treatment phase. During the first 20 minutes of each day, the group using the rate computation sheet (RCS) alone and the group using the RCS and the chart were instructed in the use of the RCS together. At the end of the 20 minutes, the group using the RCS and the chart together remained with the experimenter to practice using the chart. Generally, this further instruction took more than an additional 15 minutes and thus another day of instruction was planned. A second day was used for review of procedures for both groups. Each

student demonstrated his proficiency with both forms by logging the previous week's data in addition to the data from the first two treatment days. The group using the chart was also given instructions concerning the means of interpreting this visual display of academic behavior. This instruction involved brief statements concerning distances between the daily rates and the record floor (i.e., the reciprocal of the amount of time behavior is observed for charting) and the general trend suggested by any sequence of rates.

At the end of the second day of instruction, the students were asked to use the particular type of recording procedure to which they were assigned for a duration of 10 school days. The teachers were also asked to check the students' envelopes occasionally during the treatment phases to insure the daily recording. Without the recording, the effect of unused charts would be difficult to interpret and one which the present design was not constructed to handle. Each treatment phase was terminated by the teacher removing all forms from the students' envelopes.

Design

In order to assess the effects of the standard behavior charts as a device for affecting response levels, the treatment phases incorporate the use of groups not keeping any new forms, another group using the RCS alone, and a group using the RCS and the standard behavior charts.

If no significant differences in increase in response levels occurred for those groups not using the standard behavior chart, when compared to the group using the chart, the value of employing the standard behavior chart as a device for effecting levels of classroom performance would be questioned. Since all three classes were subjected to special attention, novelty in the classroom as well as the presence of a stranger in the classroom, it is assumed that the effect of these variables on changes in classroom levels of performance would be about equal for all three classes during the treatment phase of this study. The treatment phases are counterbalanced in a Latin squares fashion to permit all three groups to receive one of the three experiences during the course of the three treatment phases. There are also incorporated into the design four baseline phases interspersed before, between, and after each treatment phase. During these baselines, the students merely did their classwork and posted the responses on envelopes as mentioned in an earlier section of this chapter. The actual assignment of the three treatment methods as well as the overall experimental design of this study are summarized in Table 1.

During the initial planning of this research, the teachers expressed concern for the possible significance of interactions of treatment with ability level of the students. In order to study this possible interaction, it was agreed to stratify the samples into three ability levels (high,

average, and low). Since there were no standardized measures available to permit an objective and replicable stratification, the teachers merely used each student's previous class performances and test scores to assign the students to one of three ability levels.

Once three ability groups were identified, the students in each ability group were randomly assigned to one of three treatment sequences. As stated above, each student received each treatment once during the course of the experiment. Thus, there were equal numbers of students to each strata in each treatment sequence. For the tenth-grade typing class and the eleventh-grade accounting class there were two students per each of the nine cells. The ninth-grade class had three students per each of the nine cells.

The three treatment sequences were randomly chosen from a possible set of six treatment combinations with the exception that any given treatment appeared only once in a given column or row. The first combination was assigned to the first ability clustering, the second combination to the second clustering, and so on for the last combination to the last ability clustering.

The overall sequence of treatments and subject groupings are presented in Table 1. This design held for all classes with the exception that the ninth-grade typing class had three subjects nested in each ability by treatment order

Table 1

SEQUENCE OF EXPERIMENTAL PHASES

		B1	T1	B2	T2	B3	T3	B4
C1 C2 C3	S1 S2 S3 S4 S5 S6	CW	CW	CW	CW, RCS, Chart	CW	CW, RCS	CW
C1 C2 C3	S7 S8 S9 S10 S11 S12	CW	CW, RCS, Chart	CW	CW, RCS	CW	CW	CW
C1 C2 C3	S13 S14 S15 S16 S17 S18	CW	CW, RCS	CW	CW	CW	CW, RCS Chart	CW
		1..5,	6..15,	16..20,	21..30,	31..35,	36..45,	46..50

Days

CW....Classwork
RCS...Rate Computation Sheet
B.....Baseline
T.....Treatment
C1....High ability
C2....Average ability
C3....Low ability

cluster. Each baseline lasted for five days and each treatment phase for ten days. Whenever variations in the number of days for each treatment phase occurred due to difficulties in scheduling class time for instructing students in the use of the Rate Computation chart or the Standard Behavior Chart, the onset of the treatment phase began at the earliest possible date. Hence, the baselines prior to such treatment phases were extended. Where this occurred, the phases were lengthened rather than protracted.

Research Questions

How data from the various phases and treatment sequences were reorganized for purposes of statistical analysis has been summarized in Table 2. According to this design, an analysis of variance permitted the following questions to be asked:

1. What are the contributions of the main effects (Treatments, Order, and Ability) to the population variance? The term treatments is used here to mean any one of the seven phases, whether experimental manipulation phase or baseline. These, of course, must show some differences in order for additional and more pertinent questions both to be asked and answered succinctly. If the treatments are significant, there is justification for proceeding with post hoc comparisons of means to identify the sources of this variance. The treatment order was not a central question to this study since the order variable was

Table 2

DESIGN FOR ANALYSIS

			b1	b2	b3	b4	b5	b6	b7
	C1	S1							
		S2							
.1	C2	S3							
		S4							
	C3	S5							
		S6							
	C1	S7							
		S8							
.2	C2	S9							
		S10							
	C3	S11							
		S12							
	C1	S13							
		S14							
3	C2	S15							
		S16							
	C3	S17							
		S18							

A....Order of treatments
C....Ability levels
b1...baseline for classwork alone
b2...classwork alone
b3...baseline for classwork + RCS
b4...classwork + RCS
b5...baseline for classwork + RCS + Chart
b6...classwork + RCS + Chart
b7...final baseline

originally used as a device for counterbalancing three treatment combinations for each treatment phase and for assuring each subject's exposure to each treatment. Traditionally, a design controlling for order effects is used when there is some support for the belief that the effect of the independent variable or variables on the dependent variable is irreversible. Such is not the case with the treatments used in this study. Ability in itself is not a prime concern to this study. Rather, in relationship to previous behavioral studies (Rosenfeld, 1972), there is a concern for the Treatment X Ability interaction. Should the main effect of ability prove statistically significant, this would indicate that initial differences in ability among students contributed to differences in levels of performance but would not contribute to the major question being examined in this study, the effect of recording procedures on levels of performance.

2. Are there any differences among the interactions of the main effects? Of concern here were the Treatment X Order, Treatment X Ability, and Ability X Order interactions. A significant Treatment X Order interaction would yield valuable information with regard to the primacy or recency of the effects or, by association, the effect of other naturally occurring phenomena within the school setting (i.e., grading periods or termination of the school year). Interest in the Treatment X Ability has

already been suggested by past behavioral literature. A significant Order X Ability interaction would be difficult to explain in terms of past behavioral literature but would nonetheless yield information with regards to the work patterns of the various ability levels over the duration of the school semester.

3. Are there differences between the seven treatments themselves? If the treatment effects are significant, then mean comparisons by the Scheffé's technique will be done to identify the sources of variability. This question will also permit a comparison of baselines (both pre- and post-) and experimental manipulation phases. If the data for the manipulation phases prove better than their pretreatment baselines, there would be some evidence to suggest the effectiveness of the manipulations and thus self-control on the part of the students. If the manipulations prove better than their posttreatment baselines, then a reversal shall have occurred and self-control on the part of the students thus diminished. If the opposite occurs, self-control will have been retained, unless some alternative explanation is possible because of the occurrence of events beyond experimental control. Such events might be an upcoming grading period in which a given amount of work needs completion or possibly the completion of the school year and the attendant rush to finish incomplete work. An inspection of the means for experimental manipulation phases will suggest any differences among the

three treatment combinations. If the recording in count form on envelopes alone proves better than the Rate Computation Sheet (RCS) alone and the RCS with the chart, then there would be reason to suspect the overall novelty of a visitor, intraclass communication, and the fact that the instruction in the type was given by teacher and not by the experimenter account for the difference. If the RCS recording alone proved superior to recording on envelopes and charts, then just having a novel form or practice using novel forms (the RCS alone is given after the RCS and chart manipulations for the first two treatment orders), or perhaps any difficulty the students might have had using the chart with the RCS might explain the difference. If the chart alone proved better than the other two manipulations, then there is justification for believing the usual display through the chart accounts for the difference. Should there be no differences among all three manipulations, then each must take credit as controlling responses for varying reasons. The posting of a response count would serve as a controlling response; however much this recording can be explained on the basis of intraclass communication or novelty of a visitor to a class. The use of the RCS alone would also qualify as a controlling response even though its use was originally intended by the experimenter as a control for the novelty effects of introducing a new form to the students. While it has been argued earlier in this study that the use of the chart, when under

tight experimental control, would alone represent a controlling response less confounded with extraneous sources the self-control acquired through the use of the chart may be no different in terms of ends of performance than the other two controlling responses in contrast.

If any one of the sets of data is so contaminated with procedural difficulties that the data collected by the students is rendered suspect, an accounting of these difficulties will be made anecdotally and quantified to show the extent of the damage done to the original experimental strategies. Possible sources for such difficulties may include: the experimenter failing to instruct the students in the use of forms adequately enough to be used by the students alone; the experimenter or teacher failing to terminate experimental phases; teachers changing the classroom procedures after the experiment has begun; teachers not giving students the opportunity to use the forms; students not being able to perform the simple, mathematical computations necessary to use the forms correctly; or students failing to use the forms for long periods of time. While most or all of these difficulties could be remedied as they occurred, this would entail a radical change in the classroom environment and thus render clear demonstrations of self-control impossible.

Finally, the last features of data to be analyzed will be those pieces of individual data which either hold clinical

significance with regard to the control of academic behavior or suggest meaningful significance beyond the significance or non-significance indicated by the particular statistical analysis used in this study. Since the import of this study lies in the test of possible techniques to be used in classrooms, any data which may hold some application value is considered worthy of mention. Furthermore, such individual data may suggest tentative hypotheses for further classroom research or intervention.

If there are missing data, there are two solutions which may be used. Should there be missing data and there appears to be no accounting for the loss due to the nature of the treatments, an unweighted means solution will be employed. If there are missing data and there is some reason to believe that the nature of the treatments in some way accounts for the missing data, then a least squares solution must be employed (Kirk, 1968, p. 204). Thus, should a student complain of his difficulty with the forms and subsequently refuse to perform, there would be justification for using the least squares solution. Alpha levels for all statistical analyses were set at the ($p < .05$) level.

In summary then, three classes were secured for research at the junior and senior high school levels. Two were in typing and the remaining was in accounting. All three classrooms were accustomed to posting the number of responses on envelopes before experimentation began. Later

this response was to be contrasted with two other forms of controlling responses: recording responses in rate form on a sheet which lacked visual, graphic display of the academic behavior and recording responses in rate form on a chart which provided visual, graphic display of the academic behaviors. The response units for the typing class were words per minute and error rate for a three-minute, once daily, timed typing. The response units for the accounting class was any fill-in, row entry, column entry, or mathematical computation occurring during the class period. Counting responses and recording them in one of three forms occurred once daily during the five-day school week. Manipulations were counterbalanced in a Latin squares fashion to insure equal exposure of all manipulations to each student. Baselines were inserted before, between, and after each manipulation phase to provide a contrast for each manipulation phase. The students were stratified into three ability levels and randomly assigned one of three treatment sequences. The treatment sequences themselves were randomly chosen from a possible set of six combinations. The major research questions were listed and possible answers were discussed with reference to the various results which analyses may yield. Possible procedural difficulties were suggested and the means of handling these data were mentioned. The possibility of presenting individual data was indicated and the reasons for this presentation

were given. Lastly, the techniques and rationale for handling missing data were discussed.

CHAPTER IV

ANALYSES

This chapter is divided into four major sections: analyses of group data, analyses of selected individual data, reliability, and summary. There is a summary after the first two subheadings to facilitate the reading. Before proceeding to review the analyses of group data, the reader should refer back to Table 2 and review the treatment designations since cross-reference will be made to these designations in the body of the current chapter.

Group Data Analyses

Three analyses were run: these were on the accounting work rates for the eleventh-grade class, words per minute and error rate for the ninth-grade typing class. The data from the tenth-grade class in typing were not treated in a similar fashion for reasons to be given later. A log (base 10) transformation was performed initially on all rates because the effects of one of the major independent variables, the Standard Behavior Chart, contained a $\log_{10}$ transformation. Due to absenteeism and to one student refusing to perform when asked to use the recording forms, there were missing data and these were generated by using

the least squares estimates provided by biomedical computer program X-63 (Dixon, 1971). This program performed regression analysis on the data. Using those parameters from the regression analysis, an estimate of the missing values could be obtained which would not radically effect the total variance. A check on the validity of these estimates was obtained by comparing the sums of squares for error using the X-63 program and the sums of squares obtained from the analysis of variance table.

If the estimates are satisfactory there should be very little difference between the two error terms. The differences were felt to be negligible as indicated in Table 3.

The analysis on logs of accounting work rates is presented in Table 4. The treatments were significant ($p>.05$), yet there were no significant differences between the manipulation phases (B_2, B_4, B_6) and their pretreatment baselines (B_1, B_3, B_5) using the Scheffé's multiple comparison technique. There was no difference between the manipulation phases and their posttreatment phases (B_3, B_5, B_7). Ranking the means of the treatments and making all pair-wise comparisons showed that B_2 was significantly smaller and B_7 was significantly larger than B_1, B_4, B_6, B_5, and B_3 using Scheffé's criterion. The only two treatments that differ are classwork alone in a manipulation phase (B_2) and the last posttreatment baseline for all treatment orders (B_7).

Table 3

A COMPARISON OF PROGRAM X-63 AND THE ANALYSIS OF VARIANCE OF SUM OF SQUARES FOR ERROR

Data Source	Sum of Squares for Error		
	X-63	ANOVA	Difference
Accounting	.7010	.6998	.0012
Words per minute	.0639	.0640	-.0001
Error rate	1.5048	1.5703	-.0655

Table 4

SUMMARY TABLE OF ANALYSIS OF VARIANCE
ON LOGS OF ACCOUNTING WORK RATES

Source	Sum of Squares	Degress of Freedom	Mean Square	Error Term	*F*
Order	.0751	2	.0375	Subj/OrderXAbility	.160
Ability	2.455	2	1.228	Subj/OXA	5.227*
OrderXAbility	.3345	4	.0836	Subj/OXA	.356
Subj/OrderXAbility	2.1138	9	.2349		
Treatment	.3869	6	.0645	TrtXSubj/OXA	3.59*
TrtXOrder	.7371	12	.0614	TrtXSubj/OXA	3.42*
TrtXAbility	.3251	12	.0271	TrtXSubj/OXA	1.51
TrtXOrderXAbility	.6146	24	.0256	TrtXSubj/OXA	1.43
TrtXSubj/OrderXAbility	.6998	39	.0179		

*Significant at .05 level

The ability levels also differed significantly with respect to means ($p < .05$). Their means rank in the expected order of high ability highest, average ability next, and low ability last. The Treatment X Ability and Order X Ability interactions were not significant.

The Treatment X Order interaction was also significant ($p < .05$). An inspection of the cell means in Table 5 for this interaction shows no significant patterns with the exception being the last posttreatment baselines which for the first and third orders are higher than all other cell means. The cell means for classwork-alone group in a manipulation phase showed lower work rates than their pretreatment baselines for the first two orders. Since according to the original temporal sequence of the experiment (see Table 1) the cell mean for the classwork alone in a manipulation phase for the second order came in the next-to-last phase of the experiment and since the rates were higher for the last posttreatment baseline in comparison to this treatment, even the second order tends to support the conclusion that the high work rates during the last post-treatment baseline contributed to the variance to this interaction.

With regard to self-control, it cannot be said that the manipulations were effective as accelerators for performance. Further, it appears that for the group continuing to work as usual during the introduction of forms to other

Table 5

CELL MEANS FOR $\log_{10}$ ACCOUNTING WORK RATES BY ORIGINAL TEMPORAL SEQUENCE OF THE EXPERIMENT WITH TREATMENT DESIGNATIONS SUBSCRIPTED FOR TREATMENT X ORDER INTERACTION WITH RAW SCORE MEANS IN BRACKETS

Parameter	Treatment Designation						
	1st	2nd	3rd	4th	5th	6th	7th
Treatment means	.2709 [1.87]	.2124 [1.63]	.3310 [2.14]	.2717 [1.87]	.2862 [1.93]	.2836 [1.92]	.4048 [2.54]
	Temporal Sequence of the Experiment						
Treatment order							
First	.3330 (1) [2.15]	.2485 (2) [1.77]	.3638 (5) [2.31]	.3947 (6) [2.48]	.1493 (3) [1.41]	.3436 (4) [2.20]	.4691 (7) [2.95]
Second	.3078 (5) [2.03]	.21517 (6) [1.64]	.39717 (3) [2.50]	.2605 (4) [1.82]	.2755 (1) [1.89]	.1718 (2) [1.49]	.3060 (7) [2.02]
Third	.1718 (3) [1.49]	.1736 (4) [1.49]	.2320 (1) [1.71]	.1598 (2) [1.44]	.4338 (5) [2.71]	.3353 (6) [2.16]	.4392 (7) [2.75]

groups, their exclusion served as a decelerator for performance rates (except for the third treatment order and here the cell means of .1718 for the first baseline and of .1736 for the classwork alone during a treatment phase were hardly different). With respect to the continuous feedback through almost immediate knowledge of results provided by the teacher and answer sheets, the self-generated feedback did not implement accelerations significantly beyond those occurring through other available sources. The superiority of the occurrence of final examinations and termination of the school year to serve as accelerators was far more significant than any source of feedback in this class environment.

The analyses of logs for the words per minute data from the ninth-grade class showed treatments to be significant ($p < .05$). Using Scheffé's multiple comparison technique, it was found that the three manipulations (B_2, B_4, B_6) differed significantly from their pretreatment baselines (B_1, B_3, B_5). There were no differences between the manipulations and their posttreatment baselines (B_3, B_5, B_7). Ranking the means and making all possible pair-wise comparisons showed that the last baseline is significantly higher than all of the other treatments while none of the other means differ.

Ability levels differed significantly ($p < .05$) and they were in the expected order with high ability highest, the average ability next, and the low ability last.

The Treatment X Order interaction differed significantly ($p < .05$). Table 6 sheds some light on this interaction. There is a positive acceleration of cell means for the logs of the data with only two exceptions occurring as noted. The rank order is almost precisely the same as the temporal pacing of phases as in Table 1. This trend was shown less pointedly in the accounting work rate data. In the typing data, however, there are no reversals occurring as a result of doing regular classroom work while some students are receiving instruction in special forms. Using the data from Table 6 and the mean comparisons of Treatment effects, it can be shown that the data points are positively accelerating with the only exceptions occurring where one of the three manipulations is in effect. Thus the treatments are functionally related to major upward movements over baselines with the accelerations slowing during each posttreatment baseline (with the exception of the seventh phase for the second treatment). Without regard to treatment order, however, the last posttreatment baseline represents an acceleration over the highest manipulation phase mean.

The words per minute data indicate the occurrence of self-control during manipulation phases for all three devices in comparison to pretreatment baselines. Further, there are no data to suggest that any one of the manipulations was any better than the others with regard to its properties for

Table 6

SUMMARY TABLE OF ANALYSIS OF VARIANCE
ON LOGS OF WORDS PER MINUTE TYPING

	Sum of Squares	Degress of Freedom	Mean Square	Error Term	*F*
Order	.1799	2	.0900	Subj/OXA	1.982
Ability	1.184	2	.5918	Subj/OXA	13.035*
OrderXAbility	.3396	4	.0849	Subj/OXA	1.870
Subj/OrderXAbility	.8172	18	.0454		
Treatment	.0549	6	.00915	TrtXSubj/OXA	13.578*
TrtXOrder	.1731	12	.0144	TrtXSubj/OXA	21.405*
TrtXAbility	.0109	12	.00091	TrtXSubj/OXA	1.343
TrtXOrderXAbility	.0262	24	.0011	TrtXSubj/OXA	1.617
TrtXSubj/OXA	.0640	95	.00067		

*Significant at .05 level

Table 7

RANK ORDER FROM LOWEST TO HIGHEST OF CELL MEANS FOR $\log_{10}$ WORDS PER MINUTE DATA BY TREATMENT ORDER WITH RAW SCORE MEANS IN BRACKETS

ORDER	1st	2nd	3rd	4th	5th	6th	7th
First	1.4317 [27.01]	1.4775 [30.02]	1.5073 [32.16]	1.5295 [33.85]	1.5343 [34.22]	1.5515* [35.60]	1.5523* [35.67]
Second	1.5225 [33.30]	1.5481 [35.33]	1.5754 [37.62]	1.5814 [38.14]	1.6035 [40.13]	1.6085 [40.60]	1.6171 [41.41]
Third	1.4506 [28.22]	1.4950 [31.26]	1.5222 [33.28]	1.5296 [33.85]	1.5310 [33.96]	1.5423** [34.86]	1.5433** [34.94]

*These two phases are the only two in reverse order from their original temporal sequence for the first treatment order.

**These two phases are the only two in reverse order from their original temporal sequence for the third treatment order.

helping students demonstrate self-control. Thus this self-control must be explained as being confounded with novelty, intraclass communication, and perhaps differential teacher enthusiasm occurring at the beginning of manipulation phases. The multiple comparisons data suggest that self-control was lost after manipulation phases and the naturally occurring events as final examination period and final exams possess much more control over academic behaviors than the students themselves possess.

The analyses of variance on the logs of error rate for the ninth-grade typing class as indicated in Table 8 showed only the Treatment X Order interaction to be significant ($p < .05$). The data do not lend themselves to clear interpretation since the treatment means do not progress in any orderly fashion nor in the original temporal sequence of phases as in the words per minute data. By inspecting the magnitude of treatment means there appears to be an acceleration of error rates for the last posttreatment baseline. Thus it appears that as the words per minute increase, there is an increase in error rate but not to a commensurate extent. The use of the chart alone seems to facilitate control of error rate; however, its use is not better than the precharting baseline. Even this conclusion must be modified by adding that the precharting baseline is only better than the charting phase in the first treatment order. Without even statistical support, many of

Table 8

SUMMARY TABLE OF ANALYSIS OF VARIANCE
ON LOGS OF ERROR RATE FOR TYPING

Source	Sum of Squares	Degrees of Freedom	Mean Square	Error Term	F
Order	.3126	2	.1563	Subj/OXA	.5827
Ability	.7256	2	.3628	Subj/OXA	1.3525
OrderXAbility	1.8036	4	.4509	Subj/OXA	1.6810
Subj/OXA	4.8281	18	.2682		
Treatment	.0539	6	.00898	TrtXSubj/OXA	.5433
TrtXOrder	.5336	12	.0445	TrtXSubj/OXA	2.692*
TrtXAbility	.1757	12	.0146	TrtXSubj/OXA	.8832
TrtXOrderXAbility	.3403	24	.0142	TrtXSubj/OXA	.8590
TrtXSubj/OXA	1.5703	95	.01653		

*Significant at .05 level

the above conclusions are at best tentative for this group of students as a class. There is support in the typewriting literature to anticipate increased error rates with increased overall typing speed (West, 1969).

With regards to self-control, these group data offer little support for the presence of controlling responses. The effects of self-generated feedback through the chart were only suggestive in two of the three treatment orders. Furthermore, with the highest mean occurring in the last posttreatment phase, the data share some of the same conclusion from the other data sets--that the occurrence of events such as final grading periods and termination of the school year possess more control over academic behavior than the students themselves possess.

The data for the tenth-grade typing class was so contaminated with procedural errors that the analyses performed for the other data sets was not done. A summary of the difficulties collecting data representative of what the original procedures demanded includes poor communication between the experimenter and the teacher, teacher forgetting to have the timed typing daily, rushing the students through the recording of rates on the forms, starting the timed typing before each student was ready with a piece of paper in the carriage, and not checking to see that the forms were being used daily. There was also an example of a few students calculating the rates one way during the manipulation phases

and in another fashion during the baseline phases. Related to all of these faults which must be shared by both experimenter and teacher, the students themselves failed to use the forms as frequently as demanded by the procedures of this study.

Figure 1 presents the percentage of form usage by the entire class and by temporal order of introduction to the RCS and the chart. These figures give some justification for believing the loss of data for the chart is not solely due to construction of the chart itself. There was some expectation on the part of the experimenter and all three teachers prior to introducing it to the classes that the students would experience more difficulty with the chart than the RCS due to the semi-logarithmic scale on the ordinate of the chart. This, however, was not the case. The greatest losses temporally occurred in the second order. This may be explained by a four-day holiday occurring a few days after instruction in the use of the forms and still another four days before timed-typing exercises resumed. The phase was thus extended to compensate for the holiday and lost typing practice.

It is clear from Figure 2 that the greatest loss of data by ability occurred in the low-ability groups. Even here it may be noted that the chart fared no worse than the RCS. Rather, by reviewing the forms of all three groups, the experimenter found a number of division mistakes for

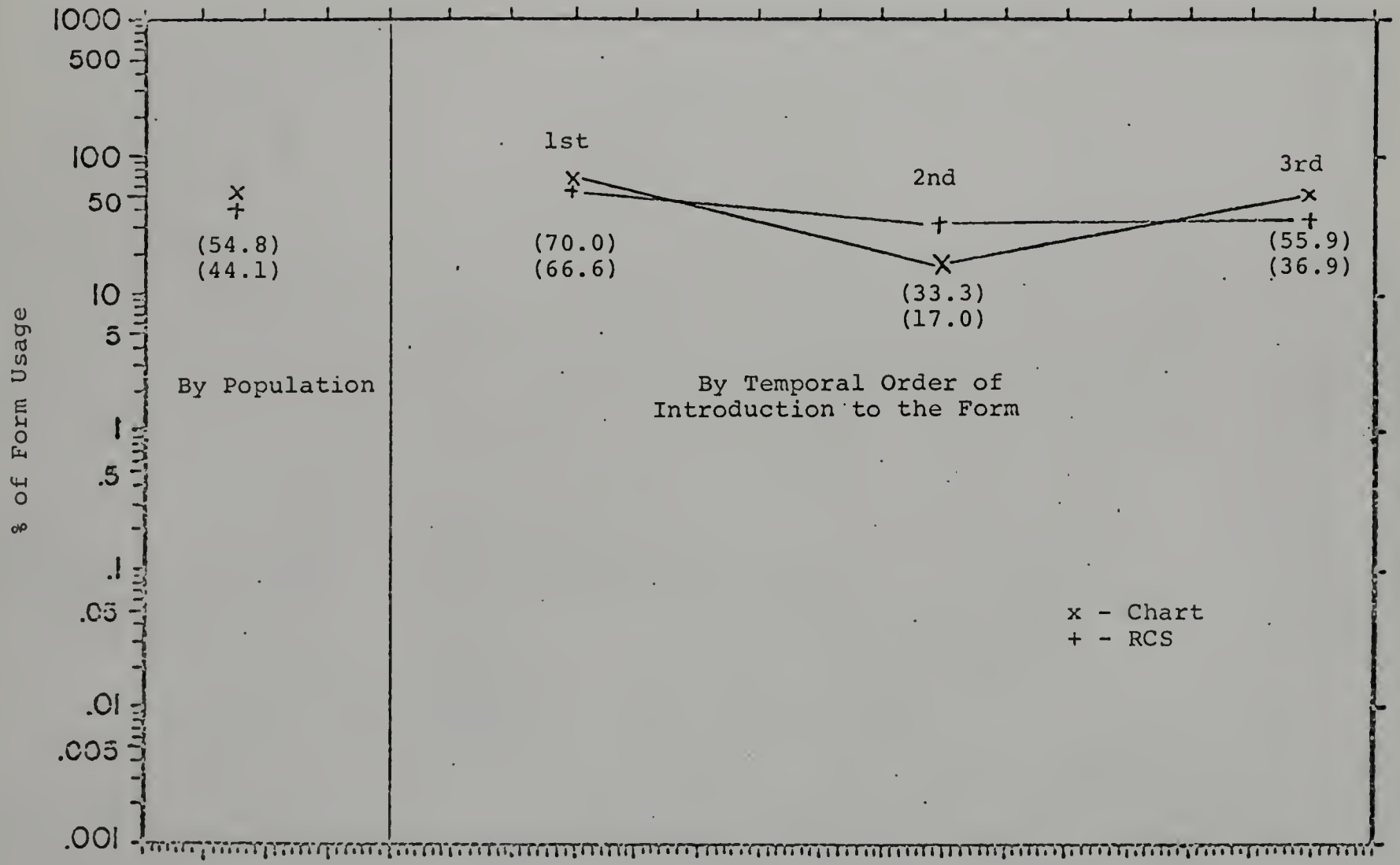

Figure 1. PERCENTAGE OF FORM USAGE IN TENTH-GRADE TYPING CLASS BY POPULATION AND BY TEMPORAL ORDER

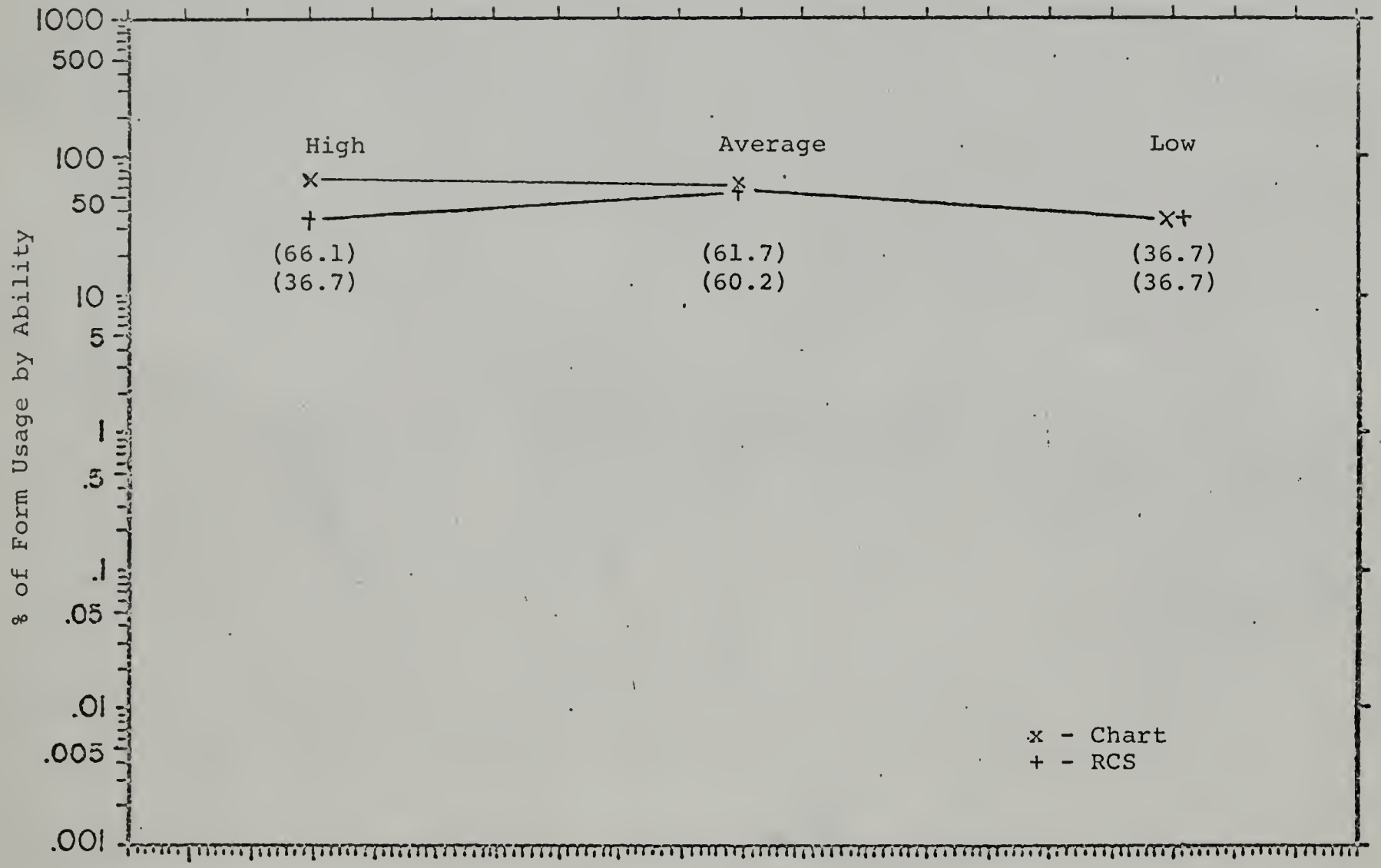

Figure 2. PERCENTAGE OF FORM USAGE IN TENTH-GRADE TYPING CLASS BY ABILITY LEVELS

the error rate data. The inability to do division when not supervised by the experimenter perhaps accounts for the greatest loss by the low-ability group with both forms.

Figure 3 presents the percentage of data loss for the Treatment X Ability X Order interaction. Overall, the lines restate the representations and explanations given for the Treatment X Order interaction (Figure 1) and Treatment X Ability interaction (Figure 2). The notable exceptions occurred for the high- and low-ability groups in the second and third order, respectively, using the RCS. There does not appear to be any plausible explanation for handling these data which would also explain the more frequent use of the forms by the average-ability RCS users.

In summary, then, an analysis of variance was performed on work rates for the eleventh-grade typing class and on the words per minute and error rate data for the ninth-grade typing class. Data in the form of percentage of form usage alone was presented for the tenth-grade typing class because the student performance data did not come in adequate contact with the procedures for the study. The analysis of accounting data revealed significant Treatment, Ability, and Treatment X Order interaction effects. A Scheffé's mean comparison revealed the last posttreatment baseline to be significantly different from all other treatments. Furthermore the comparison indicated that for individuals continuing to do regular classwork while others received instruction in the

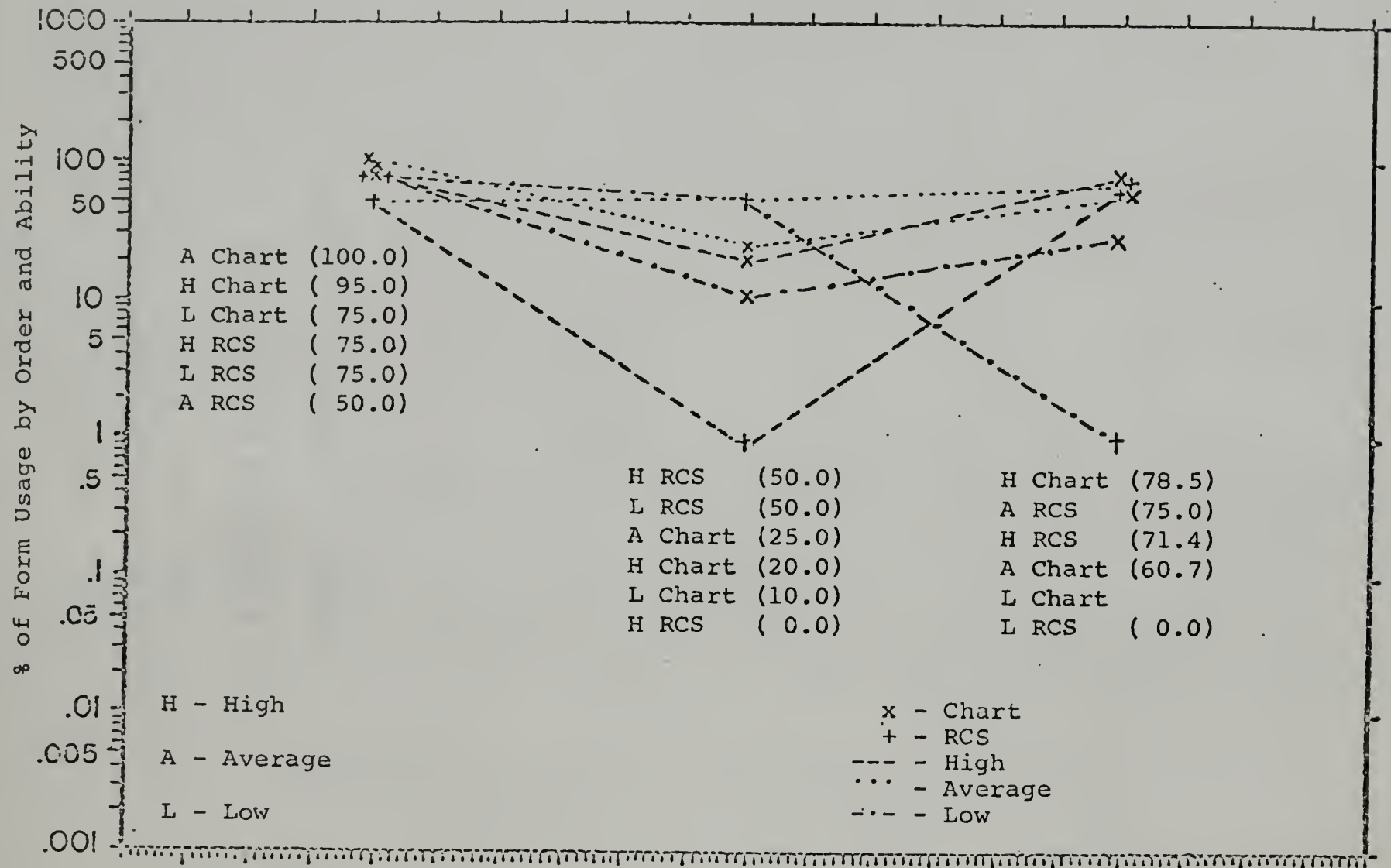

Figure 3. PERCENTAGE OF FORM USAGE IN TENTH-GRADE TYPING BY ORDER AND ABILITY

use of novel forms there was a decelerating effect. It was concluded that no self-control wad demonstrated but rather that the control of academic behavior was under the control of final examinations and the approaching end of the school year.

An analysis of variance on the words per minute for the ninth-grade typing class showed significant Treatment, Ability, and Treatment X Order interaction effects. A Scheffé's mean comparison showed all three manipulations to be different from their pretreatment baselines but not different from their posttreatment baselines. It was concluded that all three manipulations, and not any one in particular, served as accelerators. An overall mean comparison indicated the difference between the last posttreatment baseline and all other phases. Thus the self-control that was demonstrated during manipulation phases was confounded with a long-term, fixed interval scallop peaking with the occurrence of final examinations and the end of the school year. These data did not reveal any decelerating effect of doing classwork alone during manipulation phases.

The analysis of variance for the error rate data of the ninth-grade typing class was less revealing and only suggested the usefulness of the chart in reducing error rate and in reconfirming the fixed interval scallop with the impending examinations and end of the school year.

Both sets of typing data tended to reaffirm the finding of increased error rates with increased typing speeds but not to any demonstrable extent.

Because of procedural difficulties which did not occur in the eleventh- and ninth-grade classrooms, the tenth-grade class was analyzed only in terms of the percentage of form usage. The data for the population did not indicate that the students had any more difficulty using the charts than the RCS. Looking at the loss by temporal order confirms the procedural interruptions and other weaknesses as cited. The extreme loss of data with low-ability students may, at least, in part be explained by the inability of these students to readily perform the required division. Inspection of the loss by Treatment X Order X Ability interaction reaffirm the conclusions drawn for Treatment X Order and Treatment X Ability interactions. Where exceptions were noted, no explanation was thought plausible.

The main conclusion to be drawn from all of data analyzed was that where a demonstration of self-control existed it was confounded by fixed scallops (a positively accelerating increase in rate). This finding was clearly indicated in the accounting and words per minute data but only suggested by the error rate data. Suggestions concerning the future prevention of data loss and the flattening of the fixed interval scallop will be discussed in the fifth chapter.

Selected Individual Data

The individual data are presented to suggest, first, that the group statistical analysis used in this study may not represent meaningful significance with regard to effects of the chart upon the rates of academic performance and, second, that for at least one set of comparisons of individual data there is demonstrated clinical value for the inspection of individual charts. The first set of data came from two individual subjects in the eleventh-grade accounting class.

Both charts contain the baselines prior to the charting phase for each student. In both cases the obtained rates are contrasted with the record floor (i.e., reciprocal for the amount of time the behavior was occurring for charting purposes). Most noticeable in both sets of data are the downward trends during baseline periods and the reversal of these trends during the charting phase.

As may be seen in Table 9 for the mean comparisons (Glass and Stanley, 1970) there is no difference between the means of the two phases for each subject. The significance then which seems meaningful is the directional change of the trends and further this change is not statistically visible. It seems appropriate then to recommend the reliance upon visual inspection of celerations for ascertaining the significance of effects in those cases where the means are similar and the variances are small.

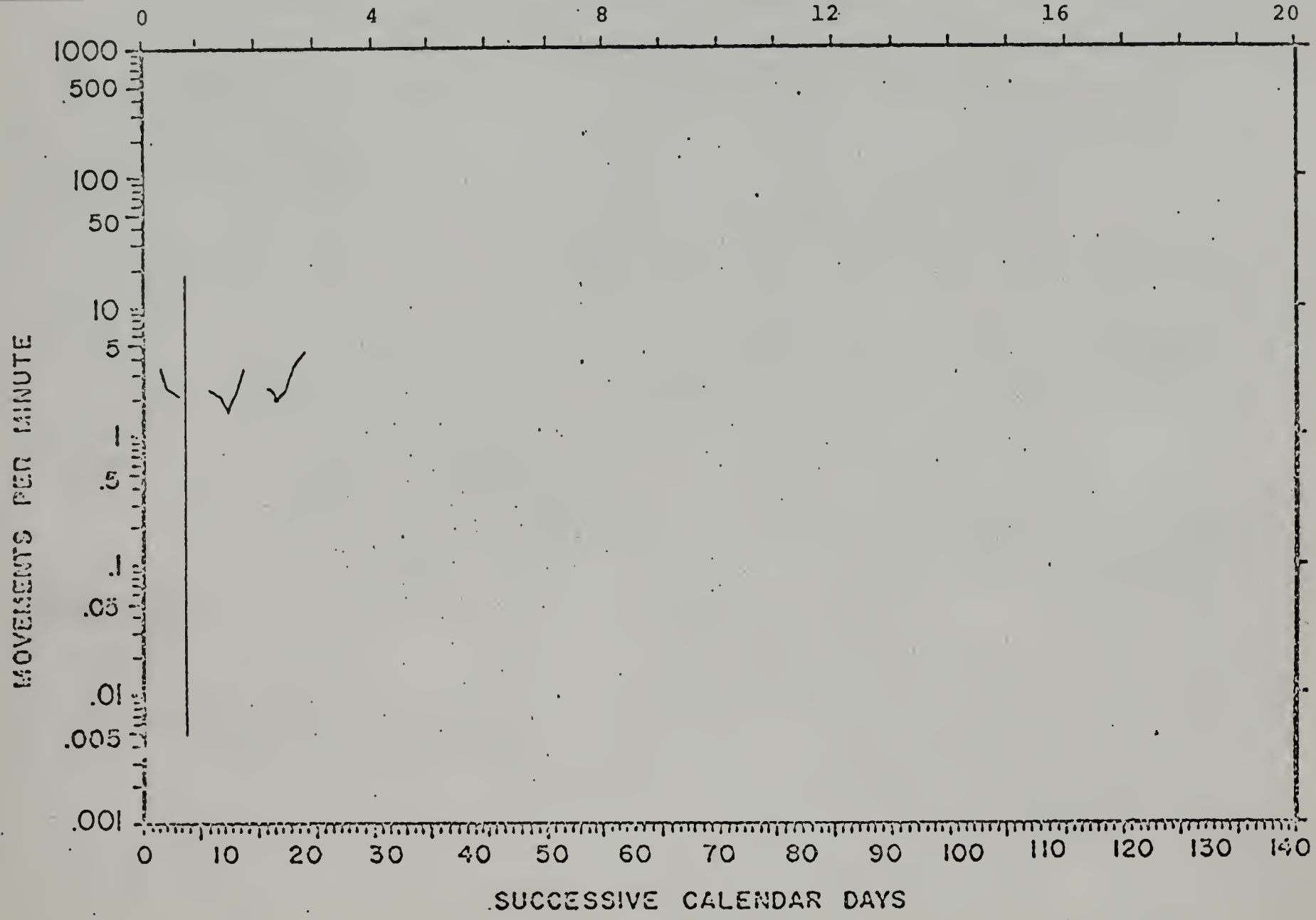

Figure 4. COMPARISON OF WORK RATES IN ACCOUNTING--FIRST SUBJECT

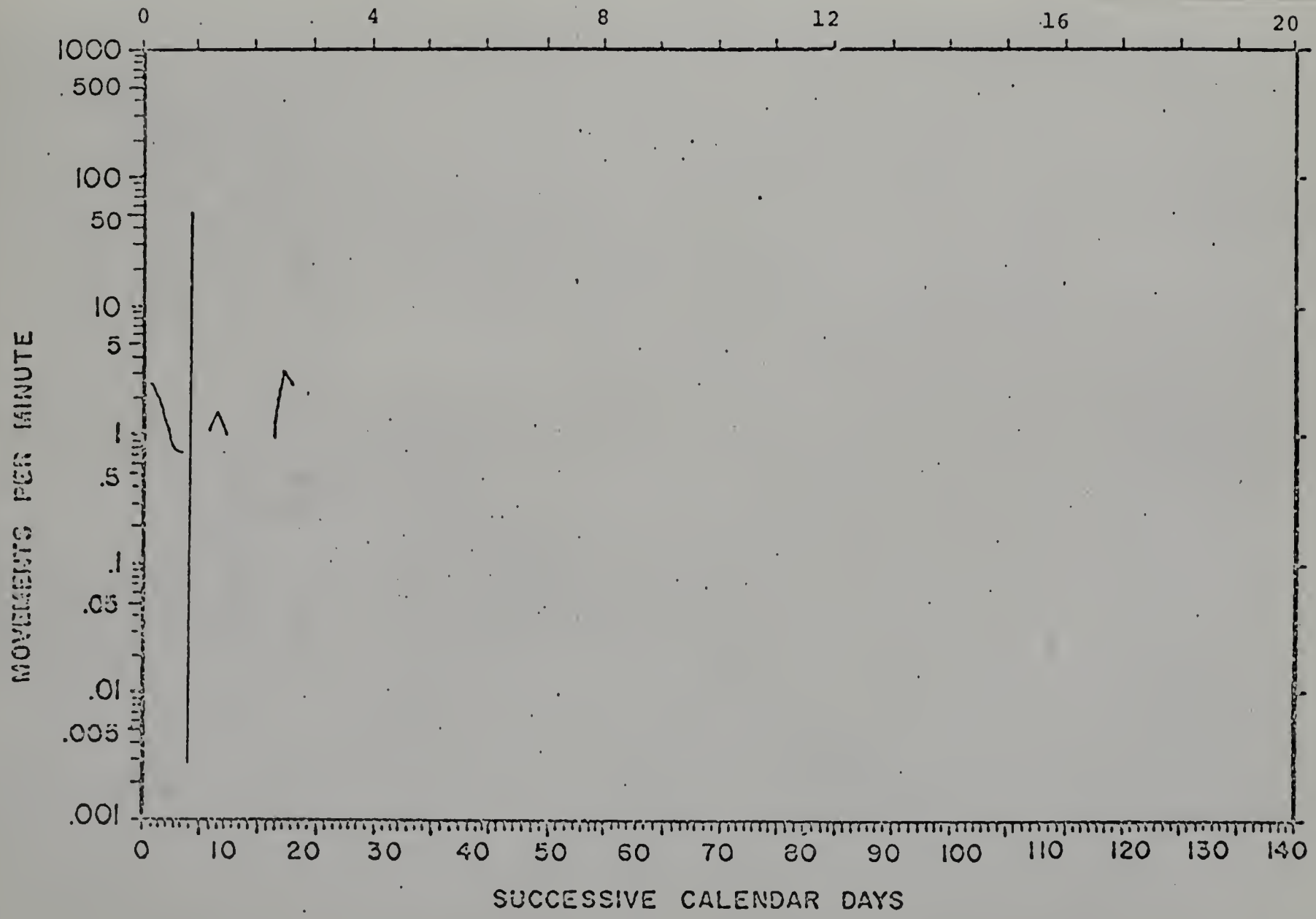

Figure 5. COMPARISON OF WORK RATES IN ACCOUNTING--SECOND SUBJECT

Table 9

RANK ORDER FROM LOWEST TO HIGHEST OF CELL MEANS FOR $\log_{10}$ ERROR RATES WITH TREATMENT DESIGNATIONS SUBSCRIPTED WITH RAW SCORE MEANS IN BRACKETS

Parameter	Rank Order						
	1st	2nd	3rd	4th	5th	6th	7th
Treatment means	.4741 (5) [2.98]	.4748 (6) [2.98]	.4985 (1) [3.15]	.5062 (3) [3.20]	.5091 (4) [3.23]	.5160 (2) [3.28]	.5176 (7) [3.29]
Treatment order							
First	.3597 (5) [2.29]	.4318 (1) [2.70]	.4652 (7) [2.92]	.4654 (6) [2.92]	.4742 (3) [2.98]	.4920 (4) [3.10]	.5342 (2) [3.42]
Second	.4370 (3) [2.74]	.5244 (6) [3.35]	.5430 (4) [3.49]	.5597 (1) [3.63]	.5793 (2) [3.80]	.6118 (5) [4.09]	.6337 (7) [4.30]
Third	.4346 (2) [2.72]	.4347 (6) [2.72]	.4506 (5) [2.82]	.4540 (7) [2.84]	.4924 (4) [3.11]	.5041 (1) [3.19]	.6075 (3) [4.05]

Table 10

TESTS OF DIFFERENCES FOR $\log_{10}$ RATES
BETWEEN BASELINE AND CHARTING PHASES
FOR TWO SUBJECTS WITH RAW SCORE MEANS IN BRACKETS

Phase	Mean	Variance	N	Test of Significance
		Subject 1		
Baseline	.4048 [2.54]	.0098	3	
				t = .0802
Charting phase	.4027 [2.53]	.0160	10	
		Subject 2		
Baseline	.1920 [1.56]	.0218	5	
				t = .0036
Charting phase	.1924 [1.56]	.0411	6	

Critical value for Subject 1: $t_{.05,\ 11} = 1.796$
Critical value for Subject 2: $t_{.05,\ 9} = 1.833$

Figure 6 is a chart of a student who represents an absentee problem to the school. Besides the scattered instances of missing classes, there was one gross example occurring from the middle of the third week lasting until the end of the eighth week. While it is clear that this student does not attend school frequently, it is also evident that there is little decleration between the last few work rates before the break and the work rates occurring after the break. When this student was in school, he worked quite consistently. The clinician's task would then become largely one of keeping this student coming to class and not one of further devising techniques to motivate the student to work while there.

Figure 7 represents an entirely different case. This student, as the chart depicts, dropped out of school. While there are instances of absenteeism as with the previous student, this student's behavior demonstrates the work of punitive contingencies even while in school. Not only did the student dislike coming to school, he also found the work aversive. This conclusion was also supported by the presence of various scatological comments peppering his work envelope. Thus the clinician's task here would be to reinforce both attendance and higher work rates.

The differences between the two students whose data are represented on Figure 6 and 7 are only noticeable because of the calendar-day comparisons of rates provided by the

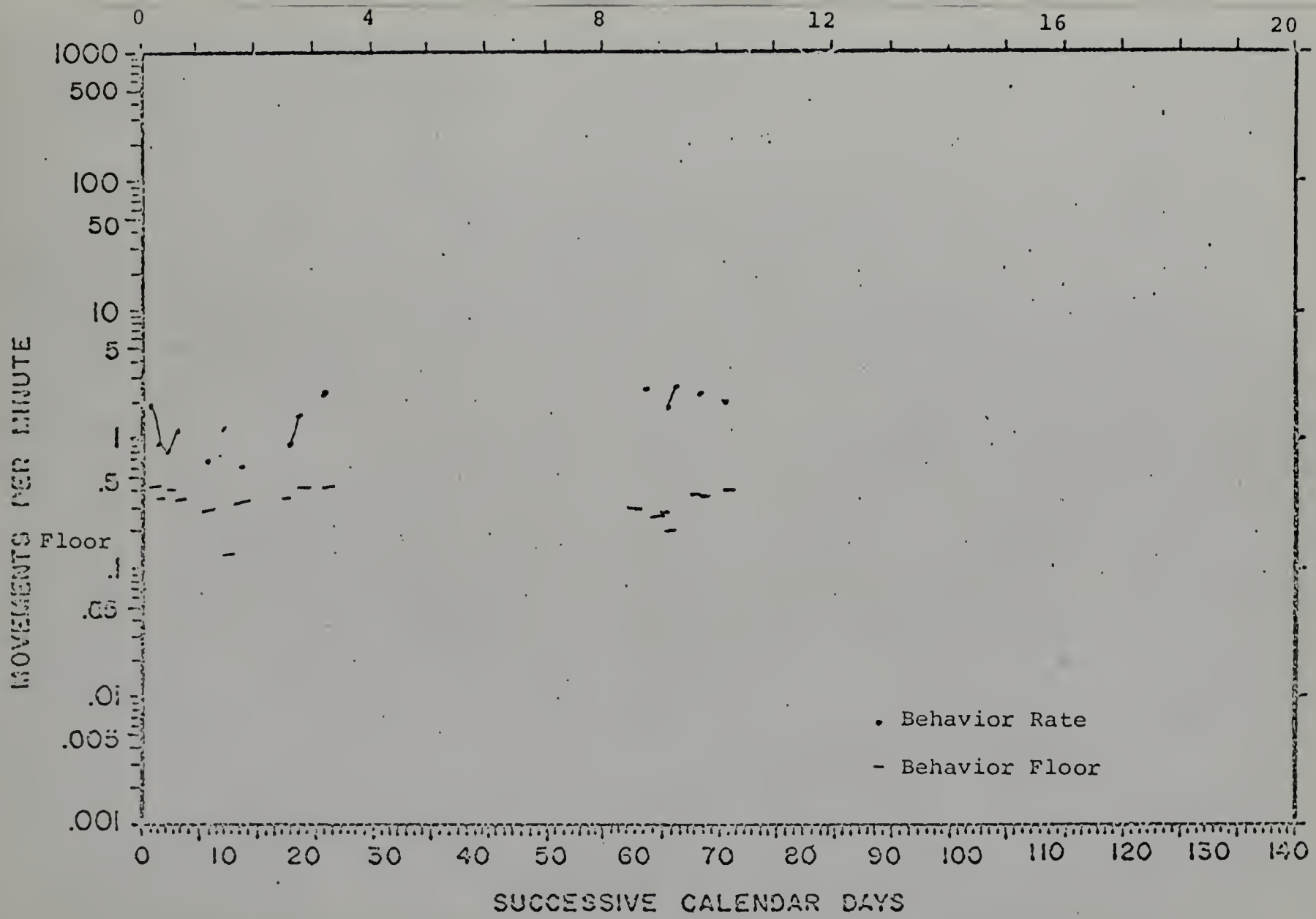

Figure 6. ACCOUNTING WORK RATES OF A HIGH ABSENTEE STUDENT

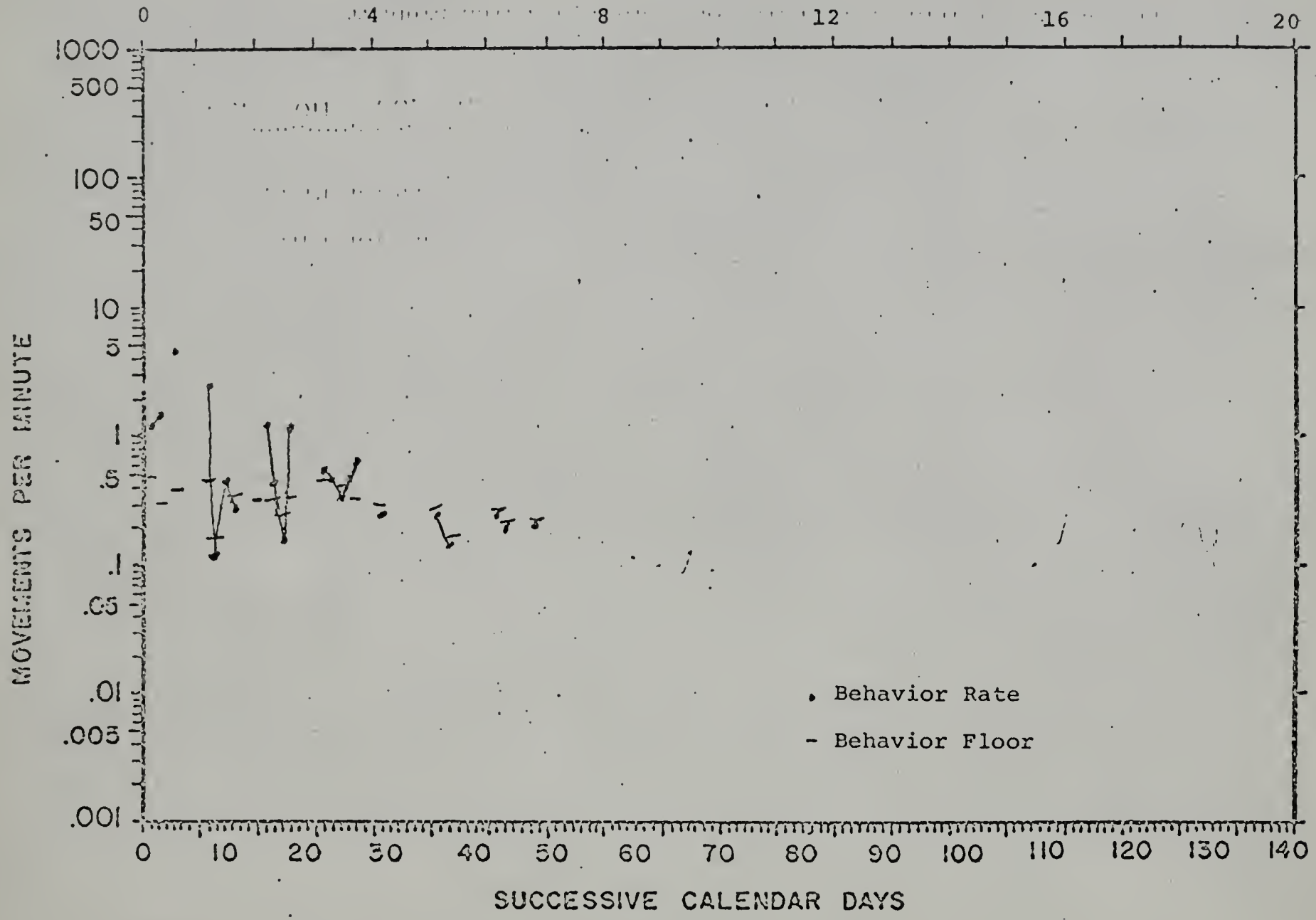

Figure 7. ACCOUNTING WORK RATES OF A SCHOOL DROP-OUT

Standard Behavior Chart. Without this information, it might have been quite easy to mistake the two, given the frequent occurrences of absenteeism. As may be noted by the various recommendations, these students are different and the types of intervention required would differ in each case as suggested by the charts.

In summary for this section, it has been noted that the reliance upon the visual inspection of individual data is necessary at times to assess the meaningful significance of manipulations and to suggest appropriate interventions where they can be attempted. For those cases where treatment means are similar and the variances are small, a statistic relying upon mean comparison is not as useful as a directional or trend change. Further, it was suggested that the calendar-day comparisons of rates on the Standard Behavior Chart are not only suggestive of diagnosis but also of prognosis for intervention with academic behavior problems.

Reliability

Due to the need for reliability with self-monitored data, the teachers were asked to check the students' envelopes on an intermittent basis. The teachers spent additional class time checking the recording on the envelopes daily upon initial introduction and thereafter only checked on a random basis after class time. All teachers reported accurate recording.

As a demonstration of this reliability, the error rate components (i.e., number of errors found by the typist and the number of errors found by the proofreader) were analyzed for the ninth-grade typing class and the aides who rotated the responsibility of checking the papers. A technique suggested by Winer (1962, pp. 124-128) was used to compare frequency of student error and the frequency of proofreader-identified error. Five days were randomly sampled and the analysis yielded the following coefficients for those days: .94 for the first day, .90 for the tenth day, .96 for the twenty-seventh day, .82 for the fifty-second day, and .96 for the sixty-third day. Summary tables for the analysis of variance used to compute the reliabilities are located in Appendix 2. These coefficients were computed by subtracting the quotient of the mean squares within people divided by the mean squares between people from 1. These reliabilities were felt high enough to demonstrate accurate recording.

Summary

Both group analysis and individual data were reviewed. The group analysis of the eleventh-grade accounting class suggests above all else the significance of the occurrence of final examinations and the termination of the school year as reinforcers for academic behavior. The analysis of the ninth-grade words per minute typing data confirms this conclusion significantly but go on to lend support for the

experimental manipulations as accelerating the fixed interval scallop at intermediate points along the curve. The analysis of the error rate data for the ninth-grade typing class only tends to show support for the fixed interval curve and to suggest that the Standard Behavior Chart may have helped the students gain control over their error rates. An inspection of form usage by the tenth-grade typing class indicated that the chart is no more difficult for adolescent students to use than the RCS and that low-ability students may need remedial instruction in simple division before these forms may be used without the aid of some prosthetic device (i.e., calculator or frequency finder).

The individual data were presented to indicate a difference between meaningful direction change of rates as opposed to statistical change for the chart users. Two data sets were then added to indicate the usefulness of the chart in depicting individual differences with regard to academic performance over time.

Self-control is said to have been demonstrated in a statistically significant manner during the manipulation phases for all recording responses in the words per minute data of the ninth-grade typing class. The presence of self-control was only suggestive for chart users within the error rate recording for the ninth-grade typing class. Two individual cases were presented for the accounting class in which significant self-control was demonstrated meaningfully, but not statistically.

Accurate self-reports of work rates were insured by teacher checks throughout the semester. A statistical comparison was performed for reliabilities between the student- and proofreader-identified errors for five days selected at random. The coefficients were high enough to demonstrate accurate recording.

CHAPTER V

CONCLUSIONS

This final chapter contains three major sections: discussion, implications, and summary. The discussion section provides a review of the results as related to prior literature and how this study has added to that body of knowledge. The implications section reviews the results and presents recommendations for future research and classroom practices. Finally, the summary provides a precis of the entire text.

Discussion

The literature pertinent to this discussion falls under three headings: self-monitoring, self-control, and knowledge of results. Previous research has indicated that self-monitoring may serve as a device for behavior change (Stollak, 1967; Rutner, 1967; Leitenberg *et al.*, 1968; Kolb, 1968; McFall, 1970; McFall and Hammen, 1971), that self-monitoring must be validated by an observer or observing device (McFall, 1970; Orne, 1970; Kanfer, 1970), that self-monitoring may be included as one of many non-specific factors such as structured situation and subject

motivation accounting for behavior changes in research settings using elaborate treatments (Orne, 1970; Kanfer, 1970; McFall and Hammen, 1971), and that the effects of self-monitoring are persistent for a period after self-monitoring is terminated (Leitenberg et al., 1968; McFall, 1970).

The analysis of the group data for the words per minute data and for two of the pieces of individual data indicated that self-monitoring during a treatment phase served to change behavior in a desired direction. Besides replicating previous research, this finding provided information with regard to the measure of the dependent variable. Many of the prior studies use only gross number or percentage data. The present study employed rate measures and obtained many of the same findings reported by others. Explanations for the non-replication in the accounting and error rate data are presented under the discussion of self-monitoring and non-specific factors.

Although steps were taken by the teachers (through spot checks) in all three classes to insure reliable data recording, there was only one set of data available for a statistical description of the reliability. This set of data was obtained from a comparison of self-recorded and teacher-aide-recorded errors in the ninth-grade typing class. Since the student and the aide were unknown to one another, the reported high reliabilities were taken as

honest measures. It was unfortunate that the tenth-grade typing class data could not be analyzed. The proofreading technique employed in this class was a variant of the "participant-observer" method used in clinical settings (Powell and Azrin, 1968; Azrin and Powell, 1968; Azrin and Powell, 1969). Thus the data would have served as a systematic replication of this method for classroom settings. As for the suggestion that self-recording serves as a poor control device prior to behavior modification (Kanfer, 1970), the self-monitored data yielded the same ability stratification as that given by teachers prior to the study. Thus the self-recording served as a good control.

Just as McFall and Hammen (1971) used a sophisticated design to isolate non-specific factors in smoking reduction, the present study employed a design which would isolate non-specific factors in self-recording through charts. This design permitted the most conservative test of self-monitoring data up to the present since the recording on the chart was contrasted to self-monitoring employed with relatively non-reactive procedures (student recording on envelopes after instruction by the teacher before the experiment began) and to self-monitoring with reactive procedures (students using the RCS after special instruction from the experimenter). Thus the design recommends itself for further use whenever behavioral devices are tested. A simpler version of this design without the within-subjects replication may be found in Azrin and Powell (1969).

The non-replication of self-monitoring effects for the accounting work rate and error rate data was probably due to the power of the design. Whereas prior research combined self-monitoring with other reactive variables, this design controlled for them. The results of mean comparisons for the accounting work rates, i.e., the negative effects, further demonstrate an instance of self-monitoring divorced from the reactive variables. Therefore, it may be concluded that the effects of self-monitoring are functionally related to the immediate environmental context in which this recording is accomplished.

As for the finding of persistent self-recording effects lasting after this recording is terminated (McFall, 1970), both the data which did not reveal effects different from baselines (the accounting work rate and error rate) and the data which did reveal effects that were confounded with other reactive variables (words per minute data) indicated no lasting effects. In relationship to the administratively scheduled events of exams and the termination of the school year, the effects of self-monitoring were shown to be evaporable.

The analysis of the group and individual data revealed only a few instances of self-control. Where it was demonstrated, the self-control was confounded with other variables such as instructions and novelty. This finding indicated that the controlling responses employed were

functionally related to the immediate environmental context. Thus more powerful contingencies must be used and then later faded for controlling responses if self-control is to remain intact. The self-control acquired through the recording responses used in this study was not sufficiently buffered against the effects of examinations and the end of the school year.

The use of the RCS and SBC as feedback devices for knowledge of results showed only qualified results. Thus the question of KOR as a reinforcer remains open. The error rate data, which was gathered under delay conditions of a day, pointed to the superiority of immediate feedback for simple skill learning. A clear resolution to this question may have to start in a more tightly controlled, laboratory setting before clear generalizations to the school setting may be drawn.

Implications

The implications of this research are twofold first, with regard to future educational research and, second, with regard to classroom practices. The impetus for this study came from and was in part planned and conducted by four high school teachers. Their active participation during workshops and in the initial planning insured a maximum of benefits for each student during the course of the study and at the same time permitted the conduct of this research from escaping the ubiquitous demands of running a classroom

on a daily basis. While communications problems still existed, many difficulties listed in an earlier chapter could have been remedied either by working with fewer classrooms or by soliciting colleagues of the experimenter to instruct the classes in the use of the recording forms. Research conducted in this fashion has value for those most involved in the educational process--teachers and students. All of the teachers involved in this study noticed positive changes in their students and thus planned to continue self-monitoring in their future classes. Furthermore, the teachers have expressed interest in doing further research in their classes. Thus their participation will prove mutually rewarding to both researcher and teacher.

Two additional remarks with regard to educational research seem noteworthy because of events which failed to occur. First, West (1969, P. 286) after a review of much research on the point concluded that intensive practices of timed typing incurs high error rates. This result was not replicated under daily timed typings over a period of twelve weeks. Perhaps the simultaneous self-monitoring of both words per minute and error rates accounted for the results.

Prior research has limited the effectiveness of behavior modification to low ability and special education classrooms (Rosenfeld, 1972). The author reported a study in which money was employed as a reinforcement for sixth-grade students having a high I.Q. (above 110 on the Lorge

Thorndike) and money reinforcement plus stars on a publicly located paper for average and high I.Q.'s (106 or better) showed significant achievement gains. Rosenfeld's findings, which ran counter to prior research, are replicated here because no significant Treatment X Ability interaction occurred.

The implications for classroom practices are directly related to the analyses of the data. Even when the treatments proved effective over pretreatment baselines, the effects of these treatments did not effect the acceleration of rates obtained at the end of the school year. While a qualified demonstration of self-control was achieved, this control passed to the occurrence of administratively scheduled events. This finding suggests that, when possible, testing or the deadline for large projects should be intermittently scheduled to insure the longevity of student self-control. Weighting daily work more and terminal work rates less might have the same effects.

Another finding which holds implications was the fact that rate measures may be used to achieve ability grouping. While this measure has been used for such purposes in typing class prior to this research, the replication of this finding with a pinpoint as hetergeneously mixed as the accounting work rate lends generality to the practice. Further study with other types of curricula is necessary to insure this generalization.

Finally, as a result of inspecting some of individual data, it was suggested that the satisfaction of criteria by statistical techniques may not always be consistent with significant behavior changes actually occurring. The statistical techniques may prove too time consuming, expensive, and not immediate enough to satisfy the daily needs of teachers and students. Data collection in the form of rates does not suffer these drawbacks. Furthermore, when rates are plotted on the SBC, long-range trends may be assessed at a moment's glance and communication time is commensurately reduced. These features alone urged a number of students to remark that "really bad days" were not taken as seriously as they were before the SBC was introduced.

Summary

The purpose of this study was to produce self-control of academic behavior in high school business education students by providing continuous, self-generated feedback through various feedback devices and to assess the relative merits of each device. This study was done to enlarge the settings and vary the parameters under which the effects of self-generated feedback have been found to be significant. By using Skinner's (1953) definition of self-control, the academic response rates were used as a controlled response and the generation of feedback responses through the various devices was used as the controlling response. For purposes

of this study, self-control is said to exist when celerations (acceleration or deceleration depending upon the nature of the pinpointed responses) are better under treatment conditions than under pretreatment baseline conditions.

The research was done in three classrooms. The first was a ninth-grade, second semester typing class with twenty-nine students. The second class was an eleventh-grade, second semester accounting class with twenty-one students. The last classroom was a tenth-grade, second semester typing class with twenty-five students.

The response unit for the accounting class was any fill-in-the-blank item, row or column entry, or any mathematical computation. The dependent variable for this class was work rate alone without respect to correct or incorrect responding. Since the students had answer sheets and the teacher available during the class for individual help, it was felt that work rate alone was a sufficient pinpoint. The response units for the typing classes were gross words per minute and error rate for a three minute, once daily, timed typing. The sample for typing was changed every day during the experiment. Each student computed the total number of words typed by using the vertical scale in his text for each sample. The students then proofread their own papers for errors. Later, a student-aide or a neighbor reread the papers to sift out any additional errors. Thus, during the treatment phases, error rates were computed a

day after the typing had been done for the class using the aides and immediately for the class using the neighbor.

In all classes prior to the experiment, the students were accustomed to counting the number of responses for either the entire class period in the accounting class or for the three minutes in the typing class and posting this count on a manilla envelope used to hold all of their work material. This type of self-recording was continued throughout the experiment and contrasted with recording on a rate computation form which lacked visual, graphic display of the rates and Standard Behavior Chart which lent visual, graphic display of the rates. During the first two days of each treatment phase, the experimenter visited the classes to teach the students the use of the new forms. At this time, one-third of the students continued working and recording as usual while the other two-thirds received instruction on the rate computation sheet or the chart. This was to control for any celerations due to novelty due to the presence of the experimenter, intraclass communication, or teacher enthusiasm. After instruction in the forms, the students were asked to use the forms daily for two weeks. There was no differential reward given by the teachers for using the forms. Since most of the students in these classes were doing their daily work assignments, the reward was thought superfluous and even contradictory to the goal of self-control. Each treatment phase was interspersed with a week's baseline condition before, between, and after the

three treatment-order combinations. Each treatment phase was terminated by the teacher taking the forms from the students' envelopes.

The three sets of data (accounting work rate for the eleventh grade, typing rate, and error rate for the ninth grade) were run independently through an analysis of variance after appropriate transformations of the data were performed. Since the tenth-grade typing class did not experience the procedures as contended, their data were not analyzed in this fashion. Due to the data loss by subject absenteeism, drop-out, and even one case of a student complaining about her difficulty performing the operations necessary to use the feedback devices, a least squares solution was used to predict the missing scores. The students were put into three ability levels according to past class work and then randomly assigned to one of three treatment orders. There were two subjects nested in each Treatment X Order X Ability crossing for the accounting class and three students in the nesting for the typing class.

The main effects of treatments and ability were significant for the accounting class ($p < .05$ and $p < .05$, respectively) and typing rate data ($p < .05$ and $p < .05$ respectively). These effects were not significant for the error rate data. The Treatment X Order interaction was significant ($p < .05$) for all sets of data. Using Scheffé comparisons, there was a significant difference between the first six phases and the last posttreatment baseline for all treatment orders

($p<.05$) in the accounting and typing rate data. A comparison of the first three pretreatment baselines and their treatments reveals a difference in favor of the treatments ($p<.05$) for the typing rate data. There were no significant differences between the treatments themselves and their posttreatment baselines (with the exceptions as noted above) for the typing rate data. The reliability of the self-monitored data was found to be acceptable for the error rate data sampled.

The main conclusion to be drawn from the study is that the students did evidence self-control for some of the treatment conditions but not with regard to the last post-treatment baseline. The Standard Behavior Chart while serving as an accelerator of performance rates did no better than the other devices. In the accounting class, due to the presence of answer sheets and individual teacher help, the devices alone were no better than these other modes of feedback. The error rate data, which was gathered under delay conditions of one day before feedback, confirms the superiority of immediate feedback for simple skill learning found by other researchers.

None of the treatment or baseline means were any better than the last posttreatment baseline. In graphic form, this behavior would when plotted cumulatively represent a fixed interval scallop. Given the fact that the daily classwork played little part in grade determination and that this phase ended simultaneously with the end of the school

year, this finding is in no way surprising. For educators who would make more of daily classwork, the grades for daily work should be weighted in proportion to a final examination with more frequent, terminal projects receiving equal weight as the final exam. Hopefully, this procedure would eliminate the accelerated performance prior to exam periods and the end of the school year. Finally, by using the individual data from the charts, both teachers and experimenters could achieve an understanding of the students which is mutually helpful in evaluating student progress and the effects of some manipulation. The individual data revealed differences between drop-outs and high absentee students. Research making use of individual data will hopefully aid teachers to plan instruction according to individual needs and, at the same time, advance the science of education.

APPENDICES

APPENDIX 1

SAMPLE OF A FRAME FROM THE RATE COMPUTATION SHEET

PROJECT NAMES: TRAINER ADVISER MANAGER PROTEGE MOVEMENT TARGET

USE No. 2 LEAD PENCIL

RATE COMPUTATION INSTRUCT FRAME
WEEK ○ (RCF-4)

Behavior Research Co.
Box 3331 Kansas City, Kansas 66103

TIME CLOCK READING Hrs:Min	TIME CLOCK Hrs x 60	TIME IN MINUTES	RATE COMPUTATION	DAY	COMMENT
FINISH START Minutes of Counting			Rate Movements Counted	LAST MON	INVALID IGNORED
FINISH START Minutes of Counting			Rate Movements Counted	LAST TUE	INVALID IGNORED
FINISH START Minutes of Counting			Rate Movements Counted	LAST WED	INVALID IGNORED
FINISH START Minutes of Counting			Rate Movements Counted	LAST THU	INVALID IGNORED
FINISH START Minutes of Counting			Rate Movements Counted	LAST FRI	INVALID IGNORED
FINISH START Minutes of Counting			Rate Movements Counted	LAST SAT	INVALID IGNORED
FINISH START Minutes of Counting			Rate Movements Counted	LAST SUN DAY NO	INVALID IGNORED

APPENDIX 2

SUMMARY TABLES OF ANALYSIS OF VARIANCE FOR RATER RELIABILITY DATA

Summary Table of Analysis of Variance
for Rater Reliability on Day 1

Source	SS	df	MS
Between people	3268.48	26	125.71
Within people	206.00	27	7.63
Between judges	65.51	1	65.52
Residual	140.48	26	5.40
Total	3474.48	53	

Summary Table of Analysis of Variance
for Rater Reliability on Day 10

Source	SS	df	MS
Between people	3008.92	26	115.732
Within people	325.0	27	12.03
Between judges	90.7408	1	90.74
Residual	243.25	26	9.36
Total	3333.93	53	

Summary Table of Analysis of Variance
for Rater Reliability on Day 27

Source	SS	df	MS
Between people	3003.94	25	120.15
Within people	127.50	26	4.90
Between judges	35.55	1	35.55
Residual	91.94	25	3.67
Total	3131.44	51	

Summary Table of Analysis of Variance
for Rater Reliability on Day 52

Source	SS	df	MS
Between people	600.25	23	26.09
Within people	114.00	24	4.75
Between judges	39.45	1	39.45
Residual	74.54	23	3.24
Total	714.25	47	

Summary Table of Analysis of Variance
for Rater Reliability on Day 63

Source	SS	df	MS
Between people	1763.97	23	76.69
Within people	66.50	24	2.77
Between judges	38.52	1	38.52
Residual	27.97	23	1.21
Total	1830.47	47	

REFERENCES

Azrin, N., and Powell, J. Behavioral engineering: The reduction of smoking behavior by a conditioning apparatus and procedure. Journal of Applied Behavior Analysis, 1968, 1, 193-200.

Azrin, N., and Powell, J. Behavioral engineering: The use of response priming to improve prescribed self-medication. Journal of Applied Behavior Analysis, 1969, 2, 39-42.

Berger, K. Adolescent self-awareness and change. Doctoral dissertation. Ferkauf Graduate School of Education, Yeshiva University, 1972.

Bernstein, D. A. Modification of smoking behavior: An evaluative review. Psychological Bulletin, 1969, 71, 418-440.

Bernstein, D. A. Modification of smoking behavior: An evaluative review. In W. A. Hunt (Ed.) Learning mechanisms in smoking. Chicago: Aldine Publishing Company, 1970.

Blackwood, R. O. The operant conditioning of verbally mediated self-control in the classroom. Journal of School Psychology, 1970, 8, 251-258.

Bolles, R. C. Theory of motivation. New York: Harper and Row, 1967.

Broden, M., Hall, R. V., and Mitts, B. The effects of self-recording on the classroom behavior of two eighth-grade students. Journal of Applied Behavior Analysis, 1971, 4, 191-199.

Campbell, D. T., and Stanley, J. C. Experimental and quasi-experimental designs for research. In N. L. Gage (Ed.) Handbooks of research on teaching. Chicago: Rand McNally, 1963.

Cautela, J. R. Behavior therapy and self-control: Techniques and implications. In C. M. Franks (Ed.) Behavior therapy: Appraisal and status. New York: McGraw-Hill, 1969.

Dixon, W. J. (Ed.). BMD Biomedical computer programs. Berkeley: University of California Press, 1971.

Duncan, A. D. Self-application of behavior modification techniques by teen-agers. Adolescence, 1969, 4, 541-556.

Duncan, A. D. The view from the inner eye: Personal management of inner and outer behaviors. Teaching Exceptional Children, 1971, 3, 152-156.

Ferster, C. B., Nurnberger, J. I., and Levitt, E. B. The control of reacting. Journal of Mathematics, 1962, 1, 95-97.

Fixsen, D. L., Phillips, E. L., and Wolf, M. M. Achievement Place: The reliability of self-reporting and peer-reporting and their effects on behavior. Journal of Applied Behavior Analysis, 1972, 5, 19-30.

Geis, G., and Chapman, R. Knowledge of results and other possible reinforcers in self-instructional systems. Educational Technology, 1971, 11, 38-51.

Glass, G. V., and Stanley, J. C. Statistical methods in education and psychology. New Jersey: Prentice-Hall, Inc., 1970.

Goldiamond, I. Self-control procedures in personal behavior problems. Psychological Reports, 1965, 17, 851-868.

Haughton, E. Great gains for small starts. Teaching Exceptional Children, 1971, 3, 141-146.

Homme, L. Perspectives in psychology: XXIV. Control of coverants, the operants of the mind. Psychological Record, 1965, 15, 501-511.

Jeus, K. G., and Shores, R. E. Behavioral graphs as reinforcers for work behavior of mentally retarded adolescents. Education and Training of the Mentally Retarded, 1969, 4, 21-27.

Johnson, E. C. Precision teaching helps children learn. Teaching Exceptional Children, 1971, 3, 106-110.

Kanfer, F. H. Self-monitoring: Methodological limitations and clinical applications. Journal of Consulting and Clinical Psychology, 1970, 35, 148-152.

Kirk, R. Experimental design: Procedures for the behavioral sciences. Belmont: Brooks-Cole Publishing Company, 1968.

Koenig, C. Charting the future course of behavior. Kansas City: Precision Media, 1972.

Kolb, D. A., Winter, S. K., and Berlew, D. E. Self-directed change: Two studies. Journal of Applied Behavioral Science, 1968, 4, 453-471.

Leitenberg, H., Agras, W. S., Thompson, L. E., and Wright, D. E. Feedback in behavior modification: An experimental analysis in two phobic cases. Journal of Applied Behavior Analysis, 1968, 1, 131-137.

Lichtenstein, E. Modification of smoking behavior: Good designs--ineffective treatments. Journal of Consulting and Clinical Psychology, 1971, 36, 163-166.

Marston, A. R., and McFall, R. M. Comparison of behavior modification approaches to smoking reduction. Journal of Consulting and Clinical Psychology, 1971, 36, 153-162.

Mausner, B. Some comments on the failure of behavior therapy as a technique for modifying cigarette smoking. Journal of Consulting and Clinical Psychology, 1971, 36, 167-170.

McFall, R. M. Effects of self-monitoring on normal smoking behavior. Journal of Consulting and Clinical Psychology, 1970, 35, 135-142.

McFall, R. M., and Hammen, C. L. Motivation, structure, and self-monitoring: The role of nonspecific factors in smoking reduction. Journal of Consulting and Clinical Psychology, 1971, 37, 80-86.

Orne, M. T. From the subject's point of view, when is behavior private and when is it public: Problems of inference. Journal of Consulting and Clinical Psychology, 1970, 35, 143-147.

Powell, J., and Azrin, N. The effect of shock as a punisher for cigarette smoking. Journal of Applied Behavior Analysis, 1968, 1, 63-71.

Precise behavior facts. Kansas City: Precision Media, 1971, 1.

Premack, D. Mechanisms of self-control. In W. Hunt (Ed.) Learning mechanisms in smoking. New York: Aldine, 1971.

Risley, T. R., and Hart, B. Developing correspondence between the non-verbal and verbal behavior of preschool children. Journal of Applied Behavior Analysis, 1968, 1, 267-281.

Rosenfeld, G. W. Some effects of reinforcement on achievement and behavior in a regular classroom. Journal of Educational Psychology, 1972, 63, 189-193.

Rutner, I. T. "The modification of smoking behavior through techniques of self-control." Master's thesis, Wichita State University, 1967.

Scriven, M. Freedom beyond Beyond Freedom. Paper presented at the annual meeting of the American Educational Research Association, Chicago, April 1972.

Skinner, B. F. Science and human behavior. New York: Macmillan Company, 1953.

Starlin, C. Peers and precision. Teaching Exceptional Children, 1971, 3, 129-132, 137-140.

Stollak, G. E. Weight loss obtained with different experimental procedures. Psychotherapy: Theory, Research, and Practice, 1967, 4, 61-64.

Surratt, P. R., Ulrich, R. E., and Hawkins, R. P. An elementary student as a behavioral engineer. Journal of Applied Behavior Analysis, 1969,], 85-92.

Watson, J. B. Psychology as the behaviorist views it. Psychological Review, 1913, 20, 158-177.

West, L. Acquisition of typewriting skills, methods of research in teaching typing. New York: Pitman Publishing Company, 1969.

Winer, B. Statistical principles in experimental design. New York: McGraw-Hill, 1962.

BIOGRAPHICAL SKETCH

Patrick Andrew Quigley, Jr., was born November 30, 1944, in the Bronx, New York. In June of 1962, he graduated from Ryken High School in Leonardtown, Maryland. In August, 1968, he graduated from Brooklyn College of the City University of New York with a Bachelor of Arts and Sciences in history. The following month, he entered the Graduate School of Education at the University of Florida. While working on his advanced degrees, he worked as a graduate research assistant and graduate teaching assistant. In June of 1970, he received an M.Ed. degree and in December of 1972, he received his Ph.D. in the psychological foundations of education. In the fall of 1972, he began working as an assistant professor of psychology at the University of Southern Mississippi.

Mr. Quigley is married to the former Charlotte Otchin Moyer and they have three children: Elizabeth Leigh, Cathleen Erin, and Shannon Maura.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

John Newell

John Newell, Chairman
Professor of Education

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Wm D Wolking

William Wolking
Associate Professor of Education

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Robert E. Jester

Robert E. Jester
Associate Professor of Education

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

William B. Ware

William B. Ware
Assistant Professor of Education

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

H. S. Pennypacker
Professor of Psychology

This dissertation was submitted to the Dean of the College of Education and to the Graduate Council, and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

December, 1972

Dean, College of Education

Dean, Graduate School