DOCUMENT RESUME

ED 244 482 EC 162 784

AUTHOR	Plienis, Anthony J.; Romanczyk, Raymond G.
TITLE	Teacher Versus Computer Delivered Instruction: An Analysis of Child Performance, Behavior, and Characteristics. Final Report, 5/15/82 to 11/14/82.
INSTITUTION	State Univ. of New York, Albany. Research Foundation.
SPONS AGENCY	Special Education Programs (ED/OSERS), Washington, DC.
PUB DATE	[Nov 82]
GRANT	G008200052
NOTE	116p.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC05 Plus Postage.
DESCRIPTORS	Behavior Disorders; Comparative Analysis; *Computer Assisted Instruction; Elementary Secondary Education; *Learning Disabilities; Reinforcement; *Student Teacher Relationship; *Teacher Behavior

ABSTRACT

A comparison between teacher and computer delivered instruction was conducted with 17 children (ages 4 to 14 years) with severe learning and behavior problems. The investigation was intended to analyze the stimulus functions of the presence of a teacher in the task-demand situation. The task employed was a two-choice discrimination that became progressively more difficult, analogous to educational tasks in general. Ss alternated sessions between conditions, but the criteria for advancement was independent between conditions. Measures of task performance and observational behavior data were collected. The results suggested that there was no overall difference in task performance between conditions but that the children as a group did exhibit more deviant behavior in the teacher condition. Correlational analyses on these dependent measures suggested that different patterns of relationship exist between behavior and performance in the two conditions. In an attempt to identify critical variables for the prediction of individual differences in performance and behavior, the results of regression analyses on diagnoses, developmental measures, and pretest compliance measures are presented. Results are discussed in the context of the task avoidance and inadvertent reinforcement conceptualizations and the implications for computer-assisted instruction in the education and treatment of severely disturbed children. (Author/CL)

Teacher Versus Computer Delivered Instruction: An Analysis of Child Performance Behavior, and Characteristics

Anthony J. Plienis

and

Raymond G. Romanczyk

Department of Psychology
State University of New York at Binghamton

Final report for U.S. Department of Education Grant #G008200052, Federal Assistance CFDA #84.023A

Abstract

A comparison between teacher and computer delivered instruction was conducted with 17 children (ages 4 to 14 years) with severe learning and behavior problems. The rationale for this comparison is based on the conceptualizations that adult attention is a source of inadvertent reinforcement for maladaptive behavior and adult directions set the occasion for avoidance responding. Therefore, this investigation provides an analysis of the stimulus functions of the presence of a teacher in the task-demand situation. The task employed was a two-choice discrimination that became progressively more difficult, analogous to educational tasks in general. Subjects alternated sessions between conditions, but the criteria for advancement was independent between conditions. Measures of task performance and observational behavior data were collected. The results suggest that there was no overall difference in task performance between conditions but that the children as a group did exhibit more deviant behavior in the teacher condition. Correlational analyses on these dependent measures suggest that different patterns of relationship exist between behavior and performance in the two conditions. In an attempt to identify critical variables for the prediction of individual differences in performance and behavior, the results of regression analyses on diagnoses, developmental measures, and pretest compliance measures are presented. The results of this investigation are discussed in the context of the task avoidance and inadvertent reinforcement conceptualizations and the implications for computer-assisted instruction in the education and treatment of severely disturbed children.

INTRODUCTION

With the advent of legally mandated educational services for all children as specified in P.L. 94-142, many specialized educational settings have emerged to serve children who typically would have been excluded from publicly funded education. These settings serve children with atypical development, severe behavior problems, and learning disorders. This includes children who exhibit aggression, self-injury, oppositional behavior, social withdrawal, and developmental deficits. Concern for the provision of services for such children has provided an impetus for research assessing the determinants of problematic behavior and learning deficiencies as well as the development of effective remedial strategies.

Adult Attention as a Reinforcer

From a behavioral developmental framework, adult attention is considered a fundamental generalized conditioned reinforcer for the behavior of children (cf. Bijou & Baer, 1961). Numerous investigations employing differential reinforcement, extinction, or time-out from positive reinforcement have provided support for the reinforcing role of adult attention on child behavior (cf. Leitenberg, 1976, Marholin, 1978, Ross, 1981). The social attention of adults has been demonstrated as a functional reinforcer even for severely maladaptive and disruptive behavior. Bucher and Lovaas (1968) present data on the self-injurious behavior of an 8 year old autistic boy who was typically maintained in physical restraints. For 90 minutes per day, the child was unrestrained but isolated from social attention. Initial high rates of self-injury gradually reduced over time and reached zero by the 8th session. In an analysis complementary to these results, Lovaas and Simmons (1969) demonstrated how social attention in the form of concern and comfort delivered contingent upon episodes of self-injury in an 11 year old autistic boy increased the rate of self-injury. Over a series of seven sessions, they demonstrated that this effect was reversible in that the withdrawal of attention contingent on self-injury resulted in a reduction of the behavior. Clearly, the intent of one's attention may have little to do with its functional effect.

Escape and Avoidance Responding to Teacher Demands

Adult attention in the form of performance demands can also set the occasion for maladaptive and disruptive behavior in severely disturbed children. For example, Carr, Newsom, and Binkoff (1976) demonstrated that self-injurious behavior in an 8 year old schizophrenic boy served to function as an escape response from teacher demands. In other words, when the teacher made task demands, the rate of self-injury increased; when the teacher ceased making demands, the rate decreased. Carr, Newsom and Binkoff (1980) also conducted a similar analysis of the escape-from-demand function of aggressive behavior in two retarded children. Disruptive behavior serving an escape function from teacher demands has also been demonstrated by Plummer, Baer, and Le Blanc (1977). With implementation of a procedural time-out contingency in the instructional situation, disruptive behavior increased in the autistic child studied. They demonstrated subsequent suppression of disruptive behavior by employing paced instructions without procedural time-out, i.e., teacher instructions were delivered at a set pace regardless of the child's behavior.

Instructional directions and adult demands may set the occasion for a variety of escape and avoidance behaviors. With severely disturbed, non-verbal children, topographies such as aggression, self-injury, autistic withdrawal (including gaze aversion), and self-stimulation may typically prevail. With verbal, disturbed children, different topographical task avoidant responses may likely occur, such as engaging in distracting conversation, sulking, inattentiveness, or outright verbal refusals. Often, these children are referred to as oppositional or noncompliant (Wahler, 1969; Forehand and King, 1977). Across levels of impairment, disturbed child populations may share a fundamental commonality in that adult instructions have poor stimulus control over desired child behavior. In addition, adult instructions and demands may often be discriminative for seriously disruptive and oppositional behavior. Further evidence indicating a critical relationship between noncompliance and other maladaptive behaviors is presented by Russo, Cataldo, and Cushing (1981). They provide data on three children suggesting that noncompliance covaries with other negative behaviors, e.g., aggression, self-injury, and tantrums. Their results show that when compliance was increased by contingent reinforcement, the negative behaviors decreased without the application of direct contingencies.

A notable hypothesis concerning the development of problematic behavior in response to instructional demands is a history of failure in past demand situations (Bijou & Baer, 1961; Bijou, 1966). Considerable potential exists for parents and teachers to make performance demands that are too far beyond the child's existing repertoire of discriminated responding. In addition, coercive procedures may be employed in attempting to generate correct performance. Under these conditions, it is likely that the instructional situation will acquire aversive properties and set the occasion for escape and avoidance responding. (cf. Skinner, 1953; Azrin and Holz, 1966).

Instructional Technology

Extensive research efforts have focused on the development of an effective teaching technology. Within this body of research, fundamental discrimination learning has received considerable emphasis. Extending the "errorless learning" investigations of Terrace (1966) to retarded children, Sidman and Stoddard (1966) trained a discrimination between a circle and a set of ellipses. In the early stages of stimulus presentation, the ellipses were narrow and easily distinguished from the circle. As the training program proceeded, the graduated series of presented ellipses were closer approximations to the circle. Touchette (1968) also applied an errorless training format to simple visual discrimination problems with severely retarded adolescents. The combined results of these investigations demonstrated that gradual stimulus fading procedures could establish discriminated responding with minimal errors, and stimulus control could be shifted or reversed in a similar manner. Both investigations provide some evidence in support of Terrace's earlier observation that errors produce more errors. That is, unreinforced responses to S– often resulted in subsequent multiple incorrect respponses, often accompanied by extreme emotional behavior.

However, the reports of collateral behavior were subjective observations (Sidman & Stoddard, 1966) rather than systematically measured.

Further support for the effectiveness of stimulus fading procedures has been provided by Koegel and Rincover (1976), Rincover (1978), and Schreibman (1975). These investigations provide evidence indicating that autistic children acquire visual discriminations more effectively with distinctive feature or within-stimulus prompting and fading rather than with extra-stimulus cue-fading procedures. This work was motivated by evidence suggesting that autistic children are particularly affected by the problem of stimulus overselectivity, that is, responding to an isolated non-relevant element of a stimulus complex rather than other relevant stimulus components (Lovaas, Koegel, Schreibman, and Rehm, 1971; cf. Lovaas, Koegel, and Schreibman, 1979). The data indicate that within-stimulus and distinctive feature prompt-fading procedures provide an effective strategy to overcome stimulus overselectivity in teaching certain types of discriminations. These procedures are highly consistent and similar to the errorless-learning stimulus fading procedures described by Sidman and Stoddard (1966), Terrace (1966), and Touchette (1968).

Automated Instruction

Although personalized systems of instruction and teaching machines had been popularized in the past two decades (Holland, 1960; Keller, 1968; Skinner, 1968), the utility of this technology has not been successfully demonstrated with severely disturbed children. In a comparison between teacher delivered and automated instruction with autistic children on a conditional discrimination task, Russo, Koegel, and Lovaas (1978) demonstrated better acquisition and performance with the teacher. However, a potentially significant procedural difference in the two conditions existed in that training trials presented by the teacher were typically contingent on the subjects exhibiting attending behavior. Similar response requirements and contingencies were not employed with the automated instruction for which trial presentations were paced. In the current study, this methodological problem was addressed by employing a standard temporal pacing of trials in both conditions. As previously noted, there is some evidence to suggest that paced instructions may be an effective teaching alternative with task avoidant children (Plummer et al., 1977).

In another recent investigation, Richmond (in press) compared automated and human instruction with developmentally retarded preschoolers on a visual discrimination task. This task employed an errorless learning format in which S- gradually increased in size until equivalent in size to S+, although different in form. Richmond reports an overall group difference in which human delivered instruction was superior. Again, however, the instructional interactions were quite different between conditions with the teacher employing verbal instruction and physical prompts. Contrary to these previous findings, Thorkildsen (1981) provides a preliminary report on comparisons between automated and teacher instruction with moderately and severely retarded children and severely retarded adults. The instructional tasks consist of functional daily living skills, e.g., time-telling and identification of coins. The automated instruction format employs a microcomputer operated videodisc interface which utilizes complex audio-visual displays in the program sequences. Thorkildsen reports that there were no significant differences between the teacher delivered and computer operated methods.

Purpose of Current Study

The purpose of this investigation is to assess task performance and collateral behaviors exhibited by severely disturbed children in two instructional conditions: teacher-delivered versus computer-delivered instruction. In both conditions, the instructional task was a basic discrimination learning task that becomes progressively more difficult, which is analagous to educational tasks in general. Although the training stimuli, teaching procedures, and performance contingencies were highly similar there are inherent differences that form the basis of the comparison. The presence of a teacher is associated with the child's past experience with teachers and other adults, and is likely to have differential discriminative properties for child behavior. In addition, the presence of a teacher provides a source of attention that can function as a reinforcer for child behavior.

A significant component of this investigation is the assessment of compliance that was separately conducted and analyzed in relation to the behavior and performance measures in the two conditions. By definition, adult-delivered instructions and demands have poor stimulus control over the requested behavior of noncompliant or oppositional children. However, it is not known whether the avoidance or escape from adult demands is a function of the adult or the demand. The between condition comparison by pretest measures of compliance will provide some correlational information concerning these relationships.

This investigation is also concerned with the role of diagnostic classifications and developmental measures in the prediction of performance and behavior in the two conditions. Such relationships may assist in matching children to the appropriate instructional technology in the provision of educational services.

METHOD

Subjects

The subjects were 17 children, ages 4 years, 7 months to 14 years, 6 months (mean age = 8 years, 2 months) who were enrolled at the Children's Unit for Treatment and Evaluation. Subsequent to human subjects review, informed consent was obtained from the parents of all participating children. The generic descriptor of severely disturbed applies to these children in that they had exhausted the normal continuum of public education services as a function of the severity of their behavior and learning problems. As is typically the case in special settings that serve this population, these children represent a variety of diagnostic disorders and a range of functional levels.

Four diagnostic categories are represented in this sample: 6 children diagnosed as autistic, 5 as emotionally disturbed, 4 as neurologically impaired with significant emotional disturbance and 2 as psychotic. These diagnoses were obtained from the children's files, and determination of diagnoses was made prior to their participation.

In addition to chronological age, four other developmental measures are presented in order to not only provide further descriptive information on these children, but to be used in the data analyses of this study. These measures were obtained from administrations of the Peabody Picture Vocabulary Test-Revised (Dunn and Dunn, 1981) and the Vineland Social Maturity Scale (Doll, 1965). Members of the Children's Unit professional staff administered these assessments as part of the periodic evaluation of the children. All testing took place within three months of the children's participation in this investigation.

Three of the measures employed are the Vineland Social Quotient (VSQ), the Vineland Age Equivalent (VAE) and the PPVT-R Age Equivalent. The fourth measure was derived from the PPVT-R Standard Score Equivalents and is reported as subgroups of high (SSE > 70), medium (70 > SSE > 40) and low (SSE < 40). These subgroups were formed as a function of the limits of the standardization of the PPVT-R which subsequently prevents treating this measure as a continuous variable. Table 1 presents these diagnostic classifications and developmental measures for all subjects.

Setting and Apparatus

This investigation took place at the Children's Unit for Treatment and Evaluation, located in the Department of Psychology at The State University of New York at Binghamton. The Children's Unit for Treatment and Evaluation is a private nonprofit special education program approved by the New York State Special Education Department.

In one condition, an Apple II microcomputer was used for stimulus generation and presentation, response recording and timing and data summarization. The visual stimuli were presented to the subjects on a SONY Trinitron 12 inch color television set. The television is enclosed in a plywood case and the screen is faced by a two key plexiglass response panel which divides the television screen down the vertical midline. A Gerbrands M&M dispenser (Model Q5220) directly adjoins the response apparatus and is also adjoined by a speaker box operated by a Digi-Talker Micromouth speech synthesizer interface.

This response apparatus is located in a small (1.5 x 3 m.) sound attenuated experimental room, illuminated by two fluorescent bulbs (15 w.). This experimental room adjoins a control room containing the microcomputer and vidotape recording equipment. An identical experimental room provided the location for the teacher delivered instructional conditon. Ths room was equipped wth a child-sized desk with a portable lecturn securely fastened to the surface of the desk. The space between the desk top and slanted surface of the lecturn provided a location for stimulus materials and tangible reinforcers, while the topside of the lecturn's slanted surface held the teacher's data form for recording child performance. A plastic container attached to the side of the lecturn served as the receptacle for the delivery of the tangible reinforcers. The design of this structure allowed the teacher to have easy access to the instructional materials while obstructing the child's access.

In both conditions, taller children sat in a chair of standard height without armrests while shorter children sat in an elevated chair with armrests. These different height chairs allowed for the presentation of the task at eye-level in reference to a normal sitting position. Each child used the same chair in both conditions.

Pretest Compliance Observational Measures

Within two weeks prior to each child's participation in the experimental conditions, each child was observed for a minimum of three 10-minute samples while engaged in their regularly scheduled instructional activities. During these observations, an independent observer recorded compliance to task directions and compliance to behavior management directions according to the following criteria:

Compliance to task directions -

A task direction consists of an instructional statement to perform a task response. If the child initiates the task response within 3 seconds, compliance is scored regardless of response accuracy. If the child does not initiate the task response within 3 seconds or engages in alternative behavior requiring physical or verbal redirection, noncompliance to task direction is scored. If the task direction is restated or a physical prompt is delivered prior to 3 seconds elapsing, the preceding task direction is cancelled, i.e., neither compliance nor noncompliance is scored for that direction.

Compliance to behavior management directions -

Behavior management directions consist of instructions to engage in motor behavior, posturing and attending that are considered elements of the teaching situation, but are not part of the actual task performance, e.g., "Sit down", "Look at me", etc. The same criteria as described for compliance to task directions is employed with this class of directions with the exception of response accuracy, which is essentially synonymous to compliance in reference to behavior mangement directions.

Additional samples were collected in order to obtain a minimum of 10 occurrences of each type direction.

Reliability training was conducted over a six week period with eight undergraduate research assistants who served as observers. Rotating assignments to observer pairs were employed to reduce problems with observer drift. Reliability training occurred under identical conditions to the collection of pretest data. At the end of this period, seven of the eight observers had reached a training criteria of

80% across three samples calculated by the following formula: reliability coefficient = (lower observed frequency/higher frequency) x 100. These seven observers commenced collection of the pretest data with frequent reliability checks scheduled such that the data for each child had at least one reliability check. Assignment to pairs for reliability checks was rotated. The mean reliability on compliance to task directions is 87% with a range of 0-100%. Removal of the low frequency sample with 0% agreement provides a range of 73% - 100% for the remaining subjects. The mean reliability on compliance to behavior directions is 85% with a range of 0 - 100%. Again, the removal of the low frequency sample with 0% agreement provides a considerably improved range of 64 to 100%. Table 2 provides the obtained pretest compliance measures for each subject. Compliance to task directions has a mean of 92% with a range of 63 - 100%. Compliance to behavior management directions has a mean of 73% with a range of 40 - 100%.

Teacher Ratings on Noncompliance

Teacher ratings on noncompliance to task directions and noncompliance to behavior management directions were collected to serve as an additional measure to analyze with the results of the experimental manipulations and to compare with the observational compliance measures. The professional staff of the Children's Unit served as raters. The classroom teachers responsible for supervising the educational program for each child served as the primary raters for the children. Assistant teachers who also had considerable direct contact with the children served as reliability raters. The participation of these raters was voluntary. Instructions in completing the rating scales were explicit in requesting independence of ratings. (See Appendix for a copy of the rating form used.) For each of the two types of noncompliance, the rating forms employed a scale of 1 (almost never) to 7 (almost always). A correlation matrix of the ratings of the two types of noncompliance by the two sets of raters is presented in Table 3. The correlation coefficients on each type of noncompliance by set of raters provide measures of reliability. For noncompliance to task directions, $r=.78$ $(p<.01)$. For noncompliance to behavior directions, $r=.68$ $(p<.01)$. These reliability estimates reflect a moderate level of concordance between raters. Correlations within raters and between noncompliance categories are .90 for primary raters and .84 for reliability raters. This high correlation between noncompliance categories will be addressed in the discussion section. The ratings by primary observers for each subject are presented in Table 2 along with the observational compliance measures.

Experimental Design

All subjects participated in both experimental condtions:

Condition A - Teacher Delivered Instruction

Condition B - Computer Delivered Instruction

Subjects were given the same task in both conditions. The computer was programmed to present visual and auditory stimuli in a manner similar to the teacher. Sequential sessions alternated between the two conditons. Subjects were nonsystematically assigned sequential numbers determining their order of participation. Odd number subjects had odd number sessions in Condition A, and even numbered subjects had even numbered sessions in Condition A. Subjects participated in 2 to 6 sessions per day, wth a minimum interval of 15 minutes between sessions. Criteria for advancement through the instructional task were independent between conditions.

Instructional Task

The two instructional conditions employed the identical visual discrimination task in which S+ is a 6 sq. cm. figure and S- progressed through 8 stages of closer approximations to S+. An errorless learning format (Terrace, 1966; Sidman and Stoddard, 1966) was employed with intrastimulus fading procedures (Koegel and Rincover, 1976; Rincover, 1978; Schreibman, 1975). Four different randomized sequences for the left or right position presentation of S+ were predetermined and subjects were rotated through these sequences according to the following criteria: no two consecutive sessions within or between conditions employed the same randomization sequence. The exception to these left/right sequences was a position preference algorithm: if the child made seven consecutive responses to one position, S+ was presented in the other position until a correct response was made. For Stage 1 presentations, S+ alone was presented. At Stage 2, S- was a 6 cm. vertical line. At Stage 3, S- was a 6 cm. vertical bar. For Stages 4 through 7, S- was a progressively wider rectangular shape. At Stage 8, the horizontal dimension of S- was 4 mm. less than S+. (See Appendix for reproductions of the stimuli presented.) In the teacher condition, stimuli were presented on cards the size of the response keys in the computer condition. During stimulus presentation, the teachers positioned and held the cards against a plexiglass shield (23 x 30 cm.) to both protect the cards and to more closely replicate the tactile contact of a response in the computer condition.

During a session, 10 consecutive correct responses on the current stage was the criterion for advancing to the next stage. Three consecutive errors in a session was the criterion for moving back to the previous stage. Failure to respond trials were not equivalent to incorrect response trials and were not evaluated as part of a consecutive response sequence. Therefore, failure to respond trials were not included in the criteria for changes in stage or reinforcement schedule. There were 40 trials per session. Each trial lasted 20 sec. and had three component parts. A 5 sec. warning interval was initiated by the statement "Ready". At the end of this interval, the verbal direction "Please point" was delivered simultaneous to the onset of the task stimulus presentation. Offset of stimulus presentation occurred after a task response or a maximum of 5 sec. An intertrial interval of 10 sec. followed the 5 sec. maximum stimulus presentation interval.

A praise statement was delivered contingent upon each correct response. (The computer was programmed to variably select one of six praise statements, e.g., "That is right"). No programmed events occurred following failure to respond trials. Each incorrect task response was followed by a negative feedback statement, "No, try again". A small food reward was dispensed for correct responses according to the schedule of correct respondng and stage of stimulus presentation depicted in Table 4. All but two subjects received M&M's as the food reward. Carob chips were used for two subjects with dietary restrictions. Criteria for termination of sessions were 10 consecutive correct responses at Stage 8 in both conditions or the completion of 10 sessions in each condition.

Procedural Consistency Between Conditions

In the computer condition, the procedural algorithms and contingencies were simply a matter of programming. In the teacher condition, procedural accuracy was a more salient concern. Eight advanced undergraduate research assistants served as teachers in the teacher condition. All had been previously enrolled in an undergraduate psychology practicum course which included the implementation of systematic teaching programs with severely disturbed children. In order to minimize the potential effects of individual differences in the teachers, the following two strategies were employed: 1) teacher assignments in the teacher condition were rotated such that children were not assiigned to the same teacher for two consecutive sessions and maximum variability in teacher assignments was attempted within the constraints of the availability of the teachers; and 2) in addition to a four week period of intensive training in the implementation of the instructional procedures, several audio-visual aides were provided to facilitate accurate and consistent implementation. Four sets of performance record sheets which corresponded to the four sets of predetermined left/right position sequences were produced. In addition to these visual cues for the position in which S+ was to be presented, these performance record forms provided matrices for tracking consecutive sequences of correct and incorrect responses, consecutive sequences of responses to the right or left position and scheduled delivery of reinforcement. (See Appendix for example of performance record form.) The teachers also wore a miniature ear phone connected to an audio cassette tape player which provided prompts for the position of

S+ presentation and for the pacing of the trials.

After the completion of the study, several measures were collected to assess the accuracy of the im lementation of the instructional procedures in the teacher condition, i.e., procedural reliability (Peterson, Homer, and Wonderlich, 1982). First, an independent observer viewed videotapes of a random selection of one sesson conducted by each of the eight teachers. This observer scored the accuracy of the children's responses in these sessions. Out of 320 total trials, one disagreement was recorded providing a reliability measure of 99.7%. Another independent observer reviewed the performance record forms for all 119 sessions conducted in the teacher condition. This review was conducted to assess the accuracy of the implementation of the stage change criteria. Errors were associated with 7 out of 134 total stage changes, or a reliability estimate of 94.8%. See Table 5 for a detailed presentation of these errors.

Prior to running subjects in the computer condition, multiple testing sessions were conducted to calibrate the timing of the component intervals of trial presentation and to validate the accurate implementation of the designated algorithms and contingencies.

Procedural Adaptation Training

In order to establish the component skills that are prerequisite to the implementation of the instructional task, all subjects received procedural adaptation training in both conditions. This consisted of presentations of S+ as a white surface area defined in size by the dimensions of S+ (6 cm. sq.) and S- as a charcoal gray surface defined in size by the dimension of a response panel key. Trials adhered to the same procedural specifications as the instructional task, with one exception. This procedural difference was the use of faded manual guidance to train the pointing response if the child was not exhibiting this behavior. A continuous schedule for the delivery of an edible treat contingent on correct responses was employed during this procedural adaptation training. The completion criteria for this pre-experimental phase was 10 consecutive independent correct responses. Two children did not reach this criterion level within 5 sessions in this phase and were not included in the study. The seventeen children for whom data is reported did meet this criteria.

Measures From Experimental Conditions

Performance

The following measures of task performance were collected: percent correct, attempts correct, difficulty level, percent incorrect and no response (see Table 6 for descriptions of these measures). Since the criteria for advancement through the stages of the task were independent between conditions, it was necessary to develop a measure that accounted for accuracy of performance and stage of difficulty while reflecting the rate of progress through the task sequence. The General Performance Index (GPI) serves these purposes and allows for comparisons across children and between conditions. See Figure 1 for the formula for the calculation of the GPI. This formula takes into account the number of trials and percent correct at each stage of the task while weighting these components by the stage number, i.e., the level of difficulty. The subsequent values for each stage are summed and then divided by the total number of sessions, since this could vary from 3 to 10 sessions.

Collateral Behavior

All experimental sessions were videotaped. In the teacher condition, the audiotapes used to pace the presentation of the task were dubbed on to the videotape recordings of the sessions. The overlay of these audiotapes provided signals for each of the four 5 sec. components of a trial which were then used for continuous 5 sec. interval observational data collection. In the computer condition, the programmed "Ready" and "Please point" statements signalled the beginning of the first two 5 sec. intervals of a trial and the computer also generated two brief beeping tone sequences to signal the beginning of the third and fourth intervals of a trial. These beeping tones were dubbed on to the videotape recordings of the sessions. In this manner, the videotapes of the computer condition sessions were similarly designed for continuous 5 sec. interval observational data collecton (Gelfand and Hartmann, 1975).

The following categories of child collateral behavior were scored: disruptive behavior, out of seat and self-stimulation (see Table 7 for operational definitions). In addition to these child behaviors, observational data were also collected on the occurrence of redirective behavior on the part of the teacher or monitoring adult. In both experimental conditions, the children were given considerable freedom to do as they pleased without adult intervention or redirection. However, two major factors provided the criteria for the necessity to redirect a child's behavior: 1) the safety of the child and 2) interference with the presentation of the instructional task. Redirection consisted of verbal instruction and, if necessary, physically guiding the child to sit down in the designated chair. Examples of behavior requiring redirection due to safety included leaving the room, pulling on heavy fixtures and climbing precariously on furnitire or fixtures. Examples of behavior requiring redirection as a functon of interfering with task presentation included attempts to manipulate the task stimulus materials and the analogy in the computer condition of attempting to manipulate the wiring in the rear of the plywood case (which also overlapped with safety concerns). In the teacher condition, the primary agent of redirection was the teacher. Although an adult was not present in the room with the child in the computer condition, the children were continuously monitored via closed circuit television and a one-way window. A session supervisor would enter the experimental room to redirect the child according to the above criteria.

Observational data were collected on a keypad interfaced with an IMSAI 8080 microcomputer, with data stored on minidisks upon completion of scoring each session. The keypad was set up such that four keys on the lefthand panel coded type of interval and four keys on the righthand panel coded the occurrence of specific target behaviors. Therefore, observers would enter type of interval as each new interval was signalled on the videotape and a subsequent entry of the occurrence of a target behavior would be scored in that interval. Two complete scorings of the tapes took place. On the first round of scoring, three observers scored out of seat, disruptive behavior and redirection. On the second round of scoring, five observers participated in the scoring of self-stimulation and the rescoring of disruptive behavior. Disruptive behavior was rescored as a function of the problematic level of reliability obtained on several children on the first round of scoring.

Modifications in the operational definition for disruptive behavior and a lengthier training period preceded the second round of scoring. All reported measures of disruptive behavior are from this second round of scoring.

Assessment of Reliability

Three methods were used to evaluate the reliability of the observational data. The first method will be referred to as direct interval reliability, in which agreements and disagreements are based on comparisons between the directly corresponding intervals from two independent observations of the same session. The calculational formula for direct interval reliability is the number of agreements of occurrence divided by the sum of the agreements plus disagreements. This measure provides a conservative estimate of reliability for low to moderate frequency data.

The second method for assessing reliability will be referred to as adjacent interval reliability. This method is a variation on direct interval reliability in that each interval of recorded occurrences of the primary observer is compared with the directly corresponding interval of the reliability observer. In addition, the intervals in the primary record are also compared with the intervals that immediately precede and immediately follow the directly corresponding interval in the reliability record. The presence of a scored occurrence in any of these three intervals constitutes an agreement and the absence constitutes a disagreement. In addition, the reliability record is further reviewed for any additional scored occurrences that had not yet been accounted for in the previous step. These occurrences constitute additional disagreements. The same calculational formula as used with the direct interval reliability is then applied to this determination of agreements and disagreements. This more liberal reliability estimate was designed to assess the extent of disagreement on the direct interval reliability that could be accounted for by reaction time differences in the entry of interval change-overs and target behaviors on the keypad system, as well as differential discriminations of response carry-over between two intervals. However, when this form of reliability is viewed alone, it has the potential for considerabe distortions leading to inflated estimates of the level of agreement. An example of the extreme case would entail a primary record in which an occurrence is scored in every alternate interval and a reliability record in which an occurrence is scored in every interval. This will produce an adjacent interval reliability estimate of 100% which extremely misrepresents the proportion of intervals scored in the two records, 50 and 100% respectively. Therefore, a third method of reliability assessment was employed to provide further information on the interpretation of the adjacent interval reliability. This method is called session totals and simply consists of dividing the number of scored occurrence intervals in the lower frequency sample by the number of occurrences scored in the higher frequency sample. High levels of agreement on this form of reliablity reduce the probability of artifactual inflation of high levels of agreement on the adjacent interval reliability measure.

In the first round of data collection, the observers were two graduate students (including the project supervisor) and a paid technician. After reaching a training criteria at the 80% level on direct reliability for a subset of training tapes, the three observers were assigned to two pairs: A + B and A + C. Two graduate students who were not associated with this study made random assignments of sessions for these observer pairs with a scheduled overlap of 33% of the total sessions. All observers were assigned sessions in both conditions and were blind as to which assigned sessions constituted reliability comparisons. For each of three forms of reliability assessment, Table 8 presents the overall total reliability measures for out of seat and redirection by condition. The total number of sessions and intervals observed in each condition are also presented. For the direct interval and adjacent interval reliability calculations, Table 8 also presents the sum total of agreements plus disagreements of occurrence. For the calculation of the overall reliability on the sessions totals measure, the sum of the individual session totals for observer A constitutes one term in the fraction while the pooled individual session total of observers B and C constitute the other term. All measures of direct reliability are within an acceptable range with respect to the relative frequency of occurrence. A comparison of the three forms of reliability suggests that a notable proportion of the disagreements reflected in the direct reliability measures are accounted for by issues pertaining to immediate interval of occurrence. This is particularly striking on the measures for redirection. With an overall proportion of less than 3% occurrence in the teacher condition, the reliability estimate increases from 76% direct reliability to 87% adjacent interval, with session totals reliability of 91%. With an overall proportion of less than 1% occurrence in the computer condition, the reliability estimate for redirection increases from 77% direct reliability to 95% adjacent interval with a session totals reliability of 90%. See Tables A, C and

D in the Appendix for reliability measures by subject and by observer pair.

The five observers on the second round of scoring were undergraduate research assistants. Reliability training continued over a period of seven weeks with a criteria of 70% on direct interval reliability and 80% on adjacent interval reliability for a subset of training tapes. All observers were assigned to score sessions from both conditions and pairings for reliability observations were rotated. A 40% overlap for reliability observations was assigned with the observers remaining blind as to which assigned sessions constituted reliability samples. For each of the three forms of reliability assessment, Table 9 presents the overall total reliability measures by condition for disruptive behavior and self-stimulation. This data is presented in the same format as used in Table 8. The total number of sessions for one behavior in a condition varies slightly from the number of sessions on the other behavior. This is a function of rescoring all sessions for some children for a single behavior when the reliability levels were unacceptable. These rescorings were preceded by brief periods of retraining.

Comparisons of the different forms of reliability produce trends that are similar to those discussed for Table 8. This is particularly important in viewing the reliability estimates for self-stimulation in the computer condition. With less than 13% of the intervals in these samples having occurrences of self-stimulation scored, the reliability estimate increases from 66% for direct interval to 85% for adjacent interval with 84% agreement on session totals. Based on the previous discussion of these comparisons, it is argued that this data on self-stimulation is acceptably reliable at a level that is comparable to the other behaviors and conditions. For a molecular presentation of reliabilities by observer and by subject, refer to Tables B, E and F respectively, in the Appendix. From each pair of reliability observations, one observation was randomly selected for summation and analyses of data.

RESULTS

Analysis by Counterbalance Sequence

Although the sequence of experimental sessions for each subject alternated between the computer and teacher conditions, each subject participated in one of two sequences of participation, i.e., first experimental session with the computer or with the teacher. Subjects were nonsystematically assigned to one of the two sequences with 9 subjects receiving their first session with the teacher (Sequence 1) and 8 subjects receiving their first session with the computer (Sequence 2). In order to assess for possible effects on performance by sequence, an analysis of variance of the General Performance Index (GPI) difference scores by sequence was conducted. The results of this analysis (see Table 10) indicate that although the mean GPI difference scores are 11.6 for Sequence 1 and -6.0 for Sequence 2, this difference was not significant, $F(1,15) = 2.57$, $p = .13$. The preceding counterbalance sequence analysis was conducted to assess for a possible directional generalization effect on performance between conditions. A similar analysis on collateral behavior by sequence was not conducted since no analagous conceptual basis exists for this type of effect on deviant behavior.

Differences Between Conditions on Within Subject Dependent Measures

In order to assess differences between conditions on the within subject dependent measures, t-tests on the difference scores from paired observations were conducted (see Table 10). The difference scores for each subject were calculated by subtracting the overall obtained value in the computer condition from the overall obtained value in the teacher condition. No significant differences between the teacher and computer condition were found on five of the six performance measures, including the General Performance Index (GPI), which takes into account the number of trials and percent correct at each stage addressed. A marginal difference on the difficulty level was detected with a mean of 5.13 in the teacher condition and 4.55 in the computer condition, $t(16) = 1.40$, $p<.10$. However, a comparison of this result with the nonsignificant difference on the GPI would suggest that subjects may have had a tendency to advance through the earlier stages of the instructional program more rapidly with the teacher than with the computer, but that the performance difference across subjects dropped out at the higher stages which are more heavily weighted in the calculation of the GPI. Therefore, the overall results of these analyses on the performance measures demonstrate that the subjects as a group performed equivalently on the instructional task, whether delivered by the teacher or by the computer.

The analyses on the measures of collateral behavior reveal that the teacher condition produced significantly higher Total Deviance Scores than the computer condition, $t(16)=1.80$, $p<.05$. Comparisons on the measures of the component behaviors of the Total Deviance Scores (TDS) indicate a marginally significant difference in the same direction on disruptive behavior, $t(16)=1.41$, $p<.10$, and no significant differences on out of seat and self-stimulatory behavior. However, the difference scores on out of seat and self-stimulatory behavior were also in the direction of greater proportions of occurrence in the teacher condition. The measures on redirection, which are not included in the calculation of the TDS, show a large contrast with a mean of 2.55 in the teacher condition and a mean of 0.44 in the computer condition, $t(16)=3.$, $p<.005$. These measures indicate that about five times as many interventions in the form of verbal and/or physical redirection occurred in the teacher condition in comparison to the computer condition.

Assessment of Effects by Teacher

In order to assess whether any systematic relationships developed between subject performance and assignment to teachers, a correlational analysis was conducted on teacher assignment by subject, percent correct and difficulty level for all 119 sessions in the teacher condition (see Table 11). The results of this analysis suggest that there were no significant relationships between teacher assignment and subject performance. Percent correct and difficulty level are positively correlated ($r=.26$, $p<.05$), but these measures are unrelated to teacher assignment.

Interrelationships Between Dependent Measures

Table 12 presents correlation matrices on the dependent measures for the teacher and computer conditions respectively. With the exception of GPI and TDS, the measures used in the calculation of these correlations are based on cumulative total proportions across sessions for each subject. These matrices portray relationships between the dependent variables within each condition as well as providing the opportunity to inspect possible differences in these relationships between conditions.

Correlations between the GPI and the unidimensional performance measures are substantial and significant on all measures in both conditions with the exception of difficulty level in the teacher condition. This exception may be indicative of a possible trend of more rapid advancement through the early stages in the teacher condition, which would be less heavily weighted than performance at the higher stages in calculating the GPI. Overall, the magnitude of these correlations provides further validity for the GPI as a multidimensional measure of performance on this task. The correlations between the TDS and the component behaviors show strong and similar relationships with both out of seat and disruptive behavior in both conditions. However, the measures of self-stimulatory behavior do not correlate with the TDS in either condition. This does not negate the importance of self-stimulation as a component of TDS in that this lack of correlation is probably a function of the extremely low rates of self-stimulation for many of the subjects, regardless of their rates of out of seat or disruptive behavior. The measures on redirection were significantly correlated ($r=.56$, $p<.05$) with TDS in the teacher condition, but more modestly correlated with TDS in the computer condition ($r=.30$, n.s.). However, redirection is significantly correlated with disruptive behavior ($r=.63$, $p<.01$) in the computer condition.

In comparing the relationship between performance and collateral behavior measures, TDS exhibits a substantial inverse relationship with GPI in the teacher condition ($r=-.58$, $p<.05$) and a more modest inverse relationship in the computer condition ($r=-.40$, n.s.). The proportion of no response trials is positively correlated with TDS in the teacher ($r=.68$, $p<.01$) and the computer conditions ($r=.43$, $p<.10$). The overall trend in the results of the correlations between TDS and the unidimensional performance measures does suggest that the TDS is measuring behavior that is counterproductive to good performance. However, this inverse relationship is more clearly evident in the teacher condition, but only suggestive in the computer condition.

In reviewing the correlations between the component measures of TDS and the performance measures, disruptive behavior demonstrates a relatively consistent moderate inverse relationship with GPI and percent correct in both conditions ($.42 < r < .51$). However, disruptive behavior does not appear to be related to accuracy of responding in the teacher condition as measured by attempts correct ($r=-.02$) and percent incorrect ($r=-.07$), but is moderately correlated to response accuracy in the computer condition with $r=-.42$ ($p<.10$) on attempts correct and $r=.38$ (n.s.) on the proportion of percent incorrect trials. Measures of disruptive behavior also demonstrate a strong relationship with the proportion of no response trials in the teacher condition ($r=.88$, $p<.001$), while only exhibiting a moderate correlation with no response trials in the computer condition ($r=.47$, $p<.10$). The correlations with disruptive behavior indicate a similar inverse relationship with overall performance in both conditions as measured by GPI and percent correct. However, the comparison of the patterns of correlations suggest that different aspects of performance are contributing to this inverse relationship. In the computer condition, disruptive behavior seems to be more strongly related to response accuracy, while in the teacher condition, the detrimental relationship between disruptive behavior and overall performance is more strongly related to no response trials.

The correlations with the performance measures and out of seat are highly similar in both conditions. The strongest apparent relationship is between out of seat and the proportion of no response trials with $r=.44$ ($p<.10$) in the teacher condition and $r=.43$ ($p<.10$) in the computer condition. The out of seat measures are not related to the response accuracy measures of attempts correct and percent incorrect in either condition ($-.14 < r < .12$), but the correlations with GPI and percent correct suggest a weak inverse relationship ($-.37 < r < -.29$, n.s.). Again, the relationship between no response trials and out of seat may account for this possible detrimental relationship with overall performance. In examining the correlations between self-stimulation and performance, a striking contrast is present between conditions. The response accuracy measures of percent incorrect and attempts correct are substantially related to the amount of self-stimulation observed in the teacher condition with $r=.68$ and $-.66$ ($p<.01$) respectively. Overall percent correct trials are also significantly inversely related to self-stimulation in the teacher condition ($r=-.48$, $p<.05$). This negative relationship between performance and self-stimulation is not significantly evidenced by the correlations on these three measures in the computer condition with r ranging from $-.20$ to $.17$. In addition, the teacher condition contains a strong inverse relationship between self-stimulation and overall difficulty level ($r=-.71$, $p<.01$). This result suggests that the children who exhibited the higher levels of self-stimulation moved more slowly through the stages of the task in the teacher condition. This relationship is not clearly demonstrated in the computer condition with $r=-.31$, n.s. In both conditions, the correlations between the no response trials and self-stimulation are near zero. Despite this evidence that the rate of self-stimulation has a demonstrable negative relationship with response accuracy in the teacher condition, but not in the computer condition, the relationship between self-stimulation and GPI exhibits a modest nonsignificant relationshp with r=-.34 with the teacher and r=-.28 with the computer. However, this does not negate the significant relationships with response accuracy in the teacher condition. Given the relatively low rates of self-stimulation for many of the subjects, several of the children who performed poorly, e.g., low GPI, may not have exhibited a substantial amount of self-stimulation, but may have exhibited higher levels of disruptive behavior.

Despite the relatively low proportion of intervals in which redirection occurred (particularly in the computer condition), this measure demonstrated substantial relationships with the performance measures in both conditions. The amount of redirection in the teacher condition is inversely related to GPI, percent correct, attempts correct and mean difficulty level with r ranging from -.45 ($p<.10$) to -.68 ($p<.01$). Redirection is also positively correlated with the proportion of incorrect responses and no response trials in the teacher condition (r=.51 and .47 respectively, $p<.05$). Redirection in the computer condition also shows substantial inverse relationships with GPI, percent correct, attempts correct and difficulty level with r ranging from -.63 to -.72 ($p<.01$). Positive correlations of r=.65 ($p<.01$) with percent incorrect and a more modest r=.39 (n.s.) with percent of no response trials also exist in the computer condition. Overall, these results suggest that children who engaged in higher rates of the criterion behaviors for being redirected performed poorly in both conditions, despite the fact that five times as much redirection occurred in the teacher condition.

Regression Analyses on General Performance Index and Total Deviance Scores

The previously presented results indicate that there is no significant difference between conditions on performance for these subjects. However, Table 10 and Figure 2 demonstrate that there is considerable variability across subjects on their performance both within and between conditions. Therefore, an important part of this investigation is an attempt to identify those subject characteristics which correlate with performance within each condition as well as to discriminate between those subjects who perform better in one condition than the other. In the area of maladaptive behavior, a significant difference between conditions was found with the teacher condition producing higher TDS than the computer condition. Again, however, considerable variability exists across subjects (see Table 10 and Figure 3). Therefore, the same question applies to the subject characteristics that correlate with the amount of deviant behavior within condition and the differences between conditions.

The dependent variables chosen for these regression analyses were: 1) the GPI in each condition and the GPI difference scores and 2) the TDS in each condition and the TDS difference scores. The previously presented correlational analyses provide data to support GPI as an overall measure of performance. Similarly, the GPI difference scores should reflect relative differences in performance between conditions for each subject. Although identification of the child characteristics that correlate with the specific categorical components of the TDS would be of interest, the TDS alone was selected with the following considerations in mind: 1) conceptually, the component behaviors of the

TDS are considered as maladaptive, 2) the previously presented correlational analyses provide some empirical support for the interrelationship between the two more prevalently observed categories of deviant behavior, out of seat and disruptive behavior, 3) the previous analyses also provide some support for the counterproductive relationship with performance although this relationship is complex and varies between conditions, and 4) given the restricted size of this sample, the TDS may provide the opportunity to identify a more generalized model for the prediction of deviant behavior under these conditions.

The predictor variables employed in these regression analyses fall into three major subgroups: 1) diagnostic classification, 2) developmental measures, and 3) compliance measures. These measures were previously discussed in the Method section. To assist in the interpretation of the regression analyses, Table 13 provides a correlation matrix of the predictor variables and dependent variables employed in the regression analyses.

The following regression analyses were conducted using the maximum R-squared improvement technique, a variation on stepwise multiple regression (SAS Institute Inc., 1979). These analyses were conducted in an attempt to identify the critical variables for the prediction of effects. However, an attempt is also made to present models that have conceptual integrity, therefore, not reporting some models with significant high R-squared values defined by complex interactions with many variables. Such models lose their predictive validity by increased potential for spurious results or overestimation of the derived R-squared values (Tatsuoka, 1969).

Table 14 presents the results of the multiple regression analyses on the GPI in both conditions and the GPI difference scores. The best one variable model in the teacher condition contains the Vineland Age Equivalent (VAE) scores and accounts for 45% of the variance in GPI ($p <$ condition is also VAE with r-squared = .42 ($p<.005$). The best four variable model in the teacher condition again contains a positive weighting on VAE along with negative weightings on the diagnostic categories of neurologically impaired and psychotic and on another categorical variable of medium range on the PPVT-R SSE. Therefore, although there is a moderate positive correlation with VAE and GPI, membership in the categories of psychotic, neurologically impaired, or medium range of PPVT-R standard scores predicts poorer performance in the teacher condition. This model accounts for 63% of the variance in GPI ($p<.025$). The best four variable model in the computer condition contains positive weightings for VAE and the diagnostic category of emotionally disturbed. This model also contains negative weightings for the diagnostic category of neurologically impaired and for the Vineland Social Quotient (VSQ). Therefore, in addition to the moderate correlation with VAE, membership in the diagnostic category of emotionally disturbed predicts better task performance with the computer as measured by GPI. Having a diagnosis of neurologically impaired predicts poorer performance. In addition, the negative weighting on the VSQ suggests a moderation of the relationship between GPI and VAE. This model accounts for 67% of the variance in GPI in the computer condition ($p<.01$).

The GPI difference scores were calculated by subtracting the computer GPI from the teacher GPI for each subject (i.e., Teacher GPI - Computer GPI = GPI difference score). Therefore, the valence of the weightings on the predictor variables in the regression models for the difference scores indicate in which condition the variables predict better relative performance. In other words, a negative weighting on a predictor suggests better relative performance with the computer while a positive weighting suggests better relative performance with the teacher. The best one variable model on the GPI difference scores contains a positive weighting on the observational pretest measure of compliance to task directions (CT) with r-squared = .22 ($p < .06$). Although this relationship accounts for a modest proportion of total variance, it suggests that children who have low CT scores perform relatively better with the computer and children who have high CT scores perform relatively better with the teacher. The best four variable model for GPI difference scores contains negative weightings for the categorical variables of emotionally disturbed and medium range PPVT-R standardized scores. It also contains a negative weighting on teacher ratings of noncompliance to task directions. The fourth variable in this model is teacher ratings on noncompliance to behavior directions and has a positive weighting. This model suggests that children diagnosed as emotionally disturbed, children scoring in the medium range of standardized scores on the PPVT-R, and children who are rated as high on noncompliance to teacher delivered task directions are all likely to perform relatively better with the computer. However, children who are rated high on noncompliance to behavior directions are likely to perform relatively better with the teacher. The interpretation of this interaction will be addressed in the forthcoming Discussion Section.

Table 15 displays the results of the regression analyses on TDS in both conditions and the difference scores. The best one variable model for predicting TDS with the teacher employs VAE and accounts for 52% of the variance ($p<.005$). The best three variable model for TDS in the teacher condition contains negative weightings for chronological age (CA), VAE, and membership in the high standard score range on the PPVT-R. Therefore, older children, higher VAE, and high PPVT-R standardized scores predict less deviant behavior in the teacher condition. This model accounts for almost 76% of the variance in TDS ($p < .0005$). The best one variable model for TDS with the computer employs an inverse relationship with teacher ratings on noncompliance to behavior directions (NRB) and accounts for 33% of the variance in TDS ($p<.01$).

The best three variable model for predicting TDS in the computer condition contains positive weightings on NRB and on medium range standard scores on the PPVT-R while also providing a negative weighting for VAE. This model accounts for 68% of the variance in TDS in the computer condition ($p<.005$). The best four variable model for TDS with the computer is also reported in that a number of differences from the three variable model are present. This model assigns negative weights to emotionally disturbed and psychotic and positive weights to VSQ and teacher ratings on noncompliance to task directions (NRT). Therefore, emotionally disturbed and psychotic are predictive of lower TDS with the computer while NRT plays a strong predictive role for TDS. Since VSQ alone has a zero level correlation with TDS and a high correlation with emotionally disturbed ($r=.77$, $p<.01$), VSQ may be serving as a suppressor variable for the variance it shares with emotionally disturbed (cf. Wiggins, 1973). Overall, this four variable model accounts for 70% of the variance in TDS in the computer condition ($p<.005$).

In addition to the consideration of the calculation of the TDS difference scores (Teacher TDS - Computer TDS = TDS difference scores), interpretations of the following regression analyses also require consideration of the previously demonstrated higher magnitude of TDS in the teacher condition in comparison to the computer condition. Therefore, the intercept terms in the regressions on the TDS difference scores assume special importance in interpretation along with the valence and magnitude of the regression weights and the raw score values by which they are multiplied. With a high positive value for the intercept term, it may require a combination of several substantially negative weighted factors to predict higher TDS in the computer than the teacher. The best one variable model for the TDS difference scores contains VAE with a negative weight, but only accounts for 20% of the variance ($p < .01$). The best four variable model for the TDS difference scores has an intercept value of 124 and contains negative weightings for high range standard scores on the PPVT-R, CA, and the observational measures on compliance to task directions. This model also contains a negative weighting on the diagnosis of neurologically impaired. This model accounts for 66% of the variance ($p<.01$) on these difference scores and predicts that older children with high standard scores on the PPVT-R and high scores on compliance to task directions are likely to exhibit more deviant behavior in the computer condition than in the teacher condition. Although the amount of predictive power added to the model by this variable is modest, the diagnostic category of neurologically impaired has a positive weighting and is suggestive of higher TDS with the teacher.

DISCUSSION

The results of this investigation demonstrate that for a diverse group of children with severe learning and behavior problems, there is no overall difference in task performance between computer-delivered and teacher-delivered instruction on a progressively more difficult discrimination task. This is contrary to previously reported results which suggest that automated instruction is less effective than human delivered instruction with similar populations (Richmond, in press, and Russo et al., 1978). This current finding may be attributed to the improvement in the similarity of the delivery of the instructional task in this comparison. A major aspect of this improved similarity is the use of a high quality speech synthesizer programmed to deliver verbal instruction and verbal feedback in the computer condition. This element has been missing from previous comparisons and these results suggest that it may be crucial for equating the delivery of the instructional task between conditions with this population. Buzzers, flashing lights, and dispensers alone may not be an equitable substitute for "Ready....Please point....No, try again...That is right." On the other hand, this study also employed a variety of constraints on the teachers' delivery of the instructional task. Task presentation was rigidly paced; no additional prompting of correct responses was allowed; physical contact was not allowed as a form of reinforcement, and redirection of the children's behavior could not take place unless the child's behavior was potentially dangerous or interfered with the teacher's ability to present the instructional task. However, beyond the notion of experimental control, there are sound empirically-based educational rationales for employing some of these constraints. In addition, as experimental controls, the high degree of similarity allowed for a better analysis of the stimulus functions of the presence of the teacher in the instructional situation. In this respect, the variability in individual performances across children comes into play. Despite the fact that there were no overall group differences, some children clearly performed better with the teacher, while other children clearly performed better with the computer. This raises the question of whether these differences are merely error variance or whether certain child characteristics can systematically account for these differences. For the children who perform better with the teacher, a reasonable hypothesis from an operant learning paradigm is that the teacher possesses relatively strong stimulus control for compliance to task directions. For children who perform better with the computer, a parallel explanation is that the teacher not only possesses weak stimulus control for compliance to task directions, but that the presence of the teacher elicits avoidance responses.

The results of the regression analyses on the General Performance Index difference scores do provide some support for these hypotheses. The best one variable model for predicting the difference scores contains the pretest observational measure of compliance to task directions. This measure was designed to assess the degree of stimulus control that the teacher possesses for compliance to task responses in the day to day instructional setting. Out of the 15 measures of prediction tested, this compliance measure produced the highest correlation and in the predicted direction, although it only accounts for a marginal proportion of the variance in the GPI difference scores.

The best four variable model for the prediction of the GPI difference scores provides further evidence that compliance and avoidance issues are related to differences in performance between the conditions. In this model, teacher ratings on noncompliance to task directions are correlated with the GPI difference scores in a manner that is consistent with the observational compliance measures in the previous model. Higher ratings on noncompliance to task directions are predictive of better performance with the computer. However, it is more difficult to interpret the meaning of the opposite relationship for teacher ratings on noncompliance to behavior management directions. This difficulty lies in the high correlation between the two teacher rating measures ($r = .90$). If these two measures are tapping distinct types of noncompliance and these types of noncompliance happen to be highly correlated, then this model suggests that these two types of noncompliance have opposite influence on predicting these GPI diffferences. That is, children who are rated high on noncompliance to behavior directives are more likely to perform better with the teacher, while children who are rated high on noncompliance to task directions are more likely to perform better with the computer. However, given the high correlation between these two noncompliance measures, near zero correlation between ratings of noncompliance to task and the GPI difference scores, and the slightly higher although modest correlation between ratings of noncompliance to behavior directives and the GPI difference scores, a reasonable statistical interpretation is that the noncompliance to task measure is serving as a suppressor variable that is removing the variance it shares with noncompliance to behavior directions in the prediction of the GPI difference scores (cf. Wiggins, 1973). This is further supported by the larger absolute b value (regression coefficient) assigned to the teacher ratings on noncompliance to behavior directions (b = 25.23) than the absolute value of b assigned to the teacher ratings on noncompliance to task directions (b = -22.42). However, given the moderate level of reliability on these measures of noncompliance, the generalizability of such subtle interaction effects is not highly probable.

The remaining variables in this model are the categorical measures of medium range standard scores on the PPVT-R and having a diagnosis of emotionally disturbed i.e., membership in either of these categories suggests better performance with the computer. Therefore, this model predicts that emotionally disturbed children who have significant delays in receptive language (PPVT-R standard score of greater than 40 but less than 70) are likely to perform much better with the computer. Again, it must be emphasized that multiple regression analyses conducted on a relatively small sample and testing a relatively large number of possible predictor variables are merely exploratory in nature. Given the heterogeneity of the population, these analyses were conducted to tease out combinations of variables that may be important in predicting variations and differences in performance and behavior with respect to these instructional conditions. Knowing the limitations of these statistical procedures, these findings should be viewed as preliminary in that they may provide direction for further investigation of these relationships.

With this caveat in mind, a comparison of the four-variable models for the prediction of teacher condition GPI and computer condition GPI provides further evidence that different factors contribute to the prediction of performance in each condition while also presenting the common factors related to performance in the two conditions. The single best predictor of performance in both conditions is the VAE, which alone accounts for more than 40% of the variance in GPI. In addition to the concept of social maturity, the items on the Vineland can also be conceptualized as assessing general developmental functioning. Therefore, it is not surprising that the level of developmental functioning would weigh heavily in the prediction of efficiency of task performance. The other three predictors in the four-variable model for task performance in the teacher condition are all categorical variables with negative b value weights suggesting that membership in these categories will result in poorer performance. The model indicates that a diagnosis of neurologically impaired has a particularly strong negative relationship with performance. This model also suggests that a diagnosis of psychotic predicts poorer performance, but this effect is not as clear or as strong as the relationship with neurologically impaired. Only two subjects in this sample have a diagnosis of psychotic and the inclusion of this variable does not significantly contribute to the variance accounted for by regression from the three variable model without this factor. A PPVT-R standard score of greater tr 40 and less than 70 is indicative of substantially impaired receptive language development. Five subjects in this sample fall into this category. This model predicts that membership in this category is also predictive of poor performance with the teacher although the level of significance for the inclusion of this variable is marginal, suggesting considerable variability within this group. None of the variables in this model are clearly suggestive of nonzero intercorrelations. Therefore, it is interesting that it is the medium range of PPVT-R standard scores that contributes to the prediction of poor performance with the teacher in this model rather than the low range (standard score < 40), which is indicative of extremely severe impairment in receptive language abilities. To further investigate this phenomenon, another regression analysis was conducted on this model after substituting low range of PPVT-R standard scores for medium range. This model was very similar to the previous four variable model. Both models account for an equivalent amount of variance at similar levels of significance. The three variables that the two models have in common play similar predictive roles across models. However, the low range PPVT-R scores in the new model are weighted in the opposite direction from the medium range PPVT-R scores of the old model. These results do indicate that the medium range PPVT-R standard scores are predictive of poorer performance rather than the low range PPVT-R scores after accounting for the variance that the low range scores share with the other predictors in this model. This effect may also be related to the higher correlation between deviant behavior and medium range PPVT-R scores than between deviant behavior and low range scores in light of the inverse relationship between deviant behavior and performance. In other words, it appears that the children with substantial, but not profound, receptive language impairment may exhibit more deviant behavior that interferes with task performance with the teacher.

Similar to the four variable models predicting performance in the teacher condition, the four variable model in the computer condition includes a positive relationship with VAE and a negative weighting on the diagnosis of neurologically impaired in the prediction of GPI. A difference in this model is the inclusion of the diagnosis of emotionally disturbed which is positively weighted, i.e., predictive of relatively better performance. Another major difference in this model is that the Vineland Social Quotient (VSQ) is negatively weighted despite its moderately high correlation with VAE (r= .52). VSQ is also highly correlated with emotionally disturbed (r= .77) in this sample. It appears that VSQ is moderating the effects of VAE and emotionally disturbed in this model.

In summarizing the results of the models for predicting performance in the computer and teacher conditions, the more substantive implications are the following:

1. Performance in both conditions is positively related to overall level of social development relatively independent of age.
2. Children diagnosed as neurologically impaired tend to perform poorly in both conditions.
3. Children with substantial, but not profound, levels of receptive language impairment tend to perform better with the computer than with the teacher.
4. Emotionally disturbed children tend to perform better with the computer than with the teacher.

In addition, there is some evidence to suggest that children who are generally more noncompliant to task directions tend to perform better with the computer than with the teacher.

Differences in Deviant Behavior

The results of the measures on collateral behavior provide evidence that the children in this sample exhibited more deviant behavior with the teacher than with the computer. However, this difference does not emerge on the individual scoring categories of out of seat or self-stimulatory behavior and is only marginally apparent on the category of disruptive behavior. This difference only clearly emerges in the Total Deviance Score (TDS) which is the sum of the proportion of occurrence of the individually defined scoring categories. One reason for this might be the surprisingly low rates of occurrence of these behaviors in both conditions for many of the children in this study. Given the nature of this population and familiarity with the specific children, it was anticipated that higher rates of deviant behavic. would occur. In speculating about these lower than expected rates of deviant behavior, it seems reasonable to suggest that the teaching task employed may be an important factor. The errorless learning paradigm allowed children to progress at their own rates while usually maintaining a relatively high level of success at each stage. This level of successful performance along with the corresponding reinforcement for performance may have effectively competed with cff task and counterproductive maladaptive responding for some of these children who were expected to exhibit more deviant behavior. Another interrelated factor for the low rates of deviant behavior may be attributable to the relatively brief total duration of session time for those children who reached the final performance criteria very quickly in both conditions. For these children, the novelty of leaving the classroom and receiving special attention for their participation may have promoted good behavior.

Different task parameters producing longer durations of participation may have modified this effect.

For these reasons, it seems reasonable to propose that the lower than expected base rates for the individual categories of collateral behavior make it more difficult for group differences to emerge as a function of the studied effect. Therefore, the Total Deviance Score does seem to provide a good alternative for assessing differences between conditions. Although it would be desirable to speak more definitively about differences between groups on specific categories of behavior, correlational analyses on the various measures of performance and behavior do provide further information about the differences between conditions. This topic will be discussed at greater length in a subsequent section.

Predictors of Total Deviance Scores

The best single predictor of TDS in the teacher condition is the VAE which was also found to be the best single predictor for performance in both conditions. This variable alone accounts for more than 50% of the variance in TDS with the teacher. Again, it seems logically consistent and intuitively reasonable to expect a strong inverse relationship between the level of overall social development and the amount of deviant behavior exhibited. This relationship continues to operate in the three variable model which also includes inverse relationships between the criterion measure and both chronological age and the categorical variable of high range PPVT-R standard scores. This three variable model accounts for more than 75% of the variance in TDS in this sample. This model predicts that older children with a high level of social maturity and functioning in the normal range of receptive language abilities will exhibit very low levels of deviant behavior with the teacher, i.e., normal children will be well-behaved with the teacher. It is surprising that this predictive model was derived from a population of children with severe learning and behavior problems!

Contrary to the model for the teacher condition, the best single predictor for TDS in the computer condition is the teacher ratings on noncompliance to behavior management directions. This finding is somewhat surprising and counterintuitive in that this teacher rating measure plays a more important role in predicting deviant behavior with the computer than it does with the teacher. To investigate the generality of this effect, a subsequent analysis was conducted. The alternate set of independent teacher ratings on noncompliance to behavior were analyzed with respect to TDS. These independent ratings produced a similiar level of correlation with the computer TDS ($r=.53$) as compared to the initially tested set of ratings ($r=.58$). Teacher ratings on noncompliance to behavior directions continue to play a similiar role in the three variable model for TDS with the computer. This model also predicts an inverse relationship with VAE and TDS while giving a negative weighting to the categorical variable of medium range PPVT-R standard scores. This model which accounts for almost 70% of the variance in TDS predicts that children with substantial deficits in receptive language, who are rated high on noncompliance to behavior directions and who have lower levels of social maturity, will engage in high rates of deviant behavior with the computer. The four variable model for TDS with the computer is also reported, not that it significantly increases R-squared, but because it implicates two diagnostic categories in the prediction of TDS. In this model, emotionally disturbed and psychotic are predictive of lower rates of deviant behavior with the computer.

The best one variable model for predicting the TDS difference scores contains an inverse relationship with VAE. However, given this model only accounts for 20% of the variance, the intercept constant is 31.33, and TDS is significantly higher in the teacher condition, this model should not be interpreted as higher levels of social development predict more deviance in the computer condition. This model is probably accounting for those subjects whose overall rates of deviant behavior were low in both conditions in that VAE is inversely related to TDS in each of the conditions separately.

The four variable model for predicting TDS difference scores suggests that neurologically impaired children are likely to exhibit more deviant behavior with the teacher than the computer. The other three variables in this model have negative weightings and must be viewed in reference to the rather large intercept term. Therefore, some combination of high range scores on the PPVT-R, high scores on compliance to task directions, and older chronological age may predict more deviant behavior with the computer according to this model. An analysis of the observed versus predicted scores for this model was conducted with the following results: 1) four of the predictions of more deviant behavior with the computer were accurate, 2) one prediction for more deviant behavior with the computer was inaccurate, and 3) the predicted TDS difference scores for two subjects who were more deviant with the computer were missed. These findings are generally consistent with the value of R-squared for this model (.66). Therefore, this model does appear to be doing more than just accounting for the range of TDS in the teacher condition.

In reviewing the results of these three sets of regression models, the following conclusions can be made:

1. Variables pertaining to higher levels of social competence and developmental functioning appear to play a stronger role with the teacher than with the computer in the prediction of lower rates of deviant behavior. In the teacher condition, the Vineland Age Equivalent alone has a moderately strong inverse relationship with deviant behavior.
2. Children with high teacher ratings on noncompliance are likely to engage in high rates of deviant behavior with the computer.
3. Children with substantial deficits in receptive language skills are more likely to exhibit high rates of deviant behavior with the computer.
4. Children diagnosed as emotionally disturbed or psychotic will engage in lower rates of deviant behavior with the computer than children diagnosed as neurologically impaired or autistic.
5. Neurologically impaired children may be likely to engage in more deviant behavior with the teacher than the computer.

Another more tentative interpretation of these results suggests that although higher functioning children may tend to exhibit low rates of deviant behavior in both conditions, they may tend to engage in relatively more deviance with the computer.

Discussion of Compliance Measures

The compliance measures were constructed to assess dimensions of avoidance behavior. It was hypothesized that instructions, directions and demands specifically related to performing a task response may have a different degree of stimulus control over child behaviors than directions and demands on a child's behavior that are collateral to task performance. The observational data on compliance to task and compliance to behavior suggest differences do exist between these categories of compliance (r=.31). However, the teacher rating scales on noncompliance to these two categories of demands produced very different information with r=.90, suggesting that perceptions of a child's noncompliance are highly consistent across these demand types. In addition, no significant correlations occurred on each type demand across measurement procedures. There are a number of possible explanations for these findings. The observational measures are derived from a highly operationalized quantitative approach whereas the teacher ratings are more likely to be influenced by qualitative factors such as the topography and intensity of the behaviors exhibited when a child is being noncompliant. In addition, the teacher's global ratings are derived from experience with the child across a wide range of situational circumstances that may vary on many dimensions, e.g., degree of structure, clarity of expectations, etc. The observational compliance measures were sampled from a narrower range of circumstances, i.e. a limited number of highly structured teaching sessions. Clearly, further empirical investigation is needed to understand the relationship between these measures as well as the characteristics of the measures.

Interrelationships between Dependent Measures

The results of the correlations between dependent measures provide support for the validity of GPI and TDS as representative measures of task performance and maladaptive behavior. In addition, these results suggest that the relationship between maladaptive behavior and performance is somewhat different in the two conditions. Correlations between the measures of GPI and TDS suggest that maladaptive behavior is more counterproductive to performance with the teacher than with the computer. Part of the contrast in these correlations may be accounted for by the generally lower rates of deviant behavior with the computer than with the teacher while performance remains variable, but similar, across subjects and between conditions. Nonetheless, these results are consistent with the hypothesis that the teacher elicits more avoidance behavior than the computer. This is further supported by the relationship between the specific category of disruptive behavior and the various performance measures. With the computer, disruptive behavior is more strongly related to response accuracy while with the teacher, disruptive behavior is more clearly related to no response trials. These comparisons are based on the conceptual inference that no response trials are more clearly indicative of avoidance responding than are incorrect responses. However, when a relationship exists between a category of maladaptive behavior and response accuracy in one condition, but appears to be unrelated to performance in the other, the nature of the interference is brought into question. Such is the case for self-stimulation which appears to interfere with response accuracy with the teacher but not with the computer. This suggests that self-stimulatory behavior serves a different function in the presence of the teacher than with the computer, but that this difference may not be clearly explained in the framework of avoidance responding.

Interpretations based on these correlational analyses are somewhat tenuous. With the exception of GPI and TDS, the measures analyzed are proportions of occurrence across sessions for each subject. High correlations between behaviors lead to inferences about interresponse relationships when temporal disparity may exist between behaviors that are highly correlated. These correlational representations of interresponse relationships may be the result of third variables that mediate these relationships. A more fine-grained analysis employing brief time intervals and analyzed by conditional probabilities would provide more definitive information on these relationships. However, the existing analyses are discussed in the context of hypothesized differences based on a conceptualization of avoidance behavior, rather than mere post hoc speculation. As such these interpretations do provide meaningful support for these hypotheses as well as an impetus for further investigation of these relationships.

Relationship and Function of Redirection

In both conditions, redirection has a clear inverse relationship to performance. Children who received more redirection performed more poorly. Relationships between redirection and measures of deviant behavior are less clear and do merit some discussion. A significant correlation between redirection and TDS does exist in the teacher condition while the correlation with the component behavioral measures of TDS are in a consistent direction but of lower, nonsignificant magnitudes. In the computer condition, the strongest relationships exists between redirection and disruptive behavior while correlations with the other deviant behavior measures are merely suggestive. This difference can be explained by intrinsic differences in the nature of the teacher delivered instruction versus the computer delivered instruction. The major elements of the criteria to redirect a child's behavior in either condition were potential physical harm to the child and interference with the presentation of the task. Despite constructing a task presentation vehicle that would enhance the teacher's control of task materials and minimize the child's access to these materials, some children persisted at grabbing, blocking, or otherwise attempting to manipulate the task materials. For this they were redirected in accordance with the defined criteria. However, to be redirected in these instances, they did not necessarily meet the criteria for disruptive behavior. Conservatism was employed to maintain the comparability of disruptive behavior between conditions. Since there is no clear analogue for this type of behavior in the computer condition, it was not included in the definition of disruptive behavior. With this decision in mind, it is highly consistent that redirection in the teacher condition is substantially correlated with TDS while having a more modest correlation with disruptive behavior. In the absence of analogous opportunities to interfere with task presentation, the more substantial correlation between redirection and disruptive behavior in the computer condition probably reflects the proportion of disruptive behavior which had a topography or intensity that met the criteria for redirection.

From this discussion, redirection can be construed as another indirect measure of disruptive behavior, which provides further support for the finding that children are more disruptive with the teacher than with the computer. However, as a measure of teacher behavior, the role of redirection becomes more complex. Given the correlation between redirection and TDS in the teacher condition, one could hypothesize that the redirective interactions served as functional reinforcement for deviant behavior and therefore served to escalate the rates of deviant behavior with the teacher. The essence of the comparison between conditions is maintained if one views the intrinsically distinct properties of the teacher delivered instruction as providing discriminative stimuli for engaging in deviant behavior. Therefore, the role of unplanned reinforcement effects requires further investigation to provide a more complete understanding of these behavioral differences between conditions.

Other Methodological Considerations

One possible problem with the design of this experiment is the potential for the attenuation of performance differences between conditions as a function of carry-over effects. However, both Russo et al. (1978) and Richmond (in press) employed variations of within subject designs which did produce clear differences in performance in favor of the teacher-delivered instruction. Also, difficulties with generalization across stimulus situations and overdependency on irrelevant cues in discrimination learning are characteristic of psychotic, autistic, and severely disturbed child populations (cf. Lovaas and Newsom, 1975). In addition, some children performed notably better with the teacher while some children performed better with the computer. Although the possible influences of carry-over effects cannot be entirely discounted, the combination of the above points suggest that carry-over effects were not a major factor in the results of this investigation.

Another methodological issue is the use of multiple teachers for each subject in comparison to the unchanging source of instruction in the computer condition. It may be hypothesized that the use of multiple teachers may detract from child performance in the teacher condition. On the other hand, given the highly controlled precision of the instructional procedures, the use of multiple teachers enhances the generality of these findings by distributing any idiosyncratic nonspecific effects across subjects. In addition, the children who participated in this investigation routinely work with a variety of professional and paraprofessional staff in their educational placement.

Therefore, multiple teachers delivering the same instructional program is the status quo for these children rather than a deviation from their routine. Although these points do not completely resolve this issue, alternative strategies for assigning children to teachers were not pragmatically feasible and would have been at the expense of decreased generality.

CONCLUSIONS AND IMPLICATIONS

This study investigated the relative efficacy of teacher-delivered instruction versus computer-delivered instruction with a sample population of severely disturbed children. The instructional task was a progressively more difficult discrimination problem presented in an errorless learning format employing intrastimulus fading techniques. The analyses of performance and collateral behavior produce two major findings:

1. As a group, the children's task performance with the computer was equivalent to their performance with the teacher.
2. As a group, the children exhibited more deviant behavior with the teacher than with the computer.

In addition there is some evidence to support the conceptualization that children who are more avoidant of task directions perform better with the computer. Also, there is some evidence to suggest that disruptive behavior with the teacher more clearly serves an avoidance function than disruptive behavior with the computer. Furthermore, avoidance behavior with the teacher appears to partially be a function of inadvertent reinforcement via teacher attention.

The implications of these findings are clear. If computerized delivery of instruction effectively produces acquisition of skills while reducing the occurrence of maladaptive behavior that interferes with skill acquisition, computer-delivered instruction could become a valuable component of the educational and treatment plan. This is not to say that the computer can ultimately replace teachers and other service providing adults. Clearly, the development of adaptive prosocial skills is a high priority for treatment. Improving compliance and cooperation to task directions ultimately requires the context of a teacher delivering instructions. However, the performance of skills that are effectively acquired with the computer would need to be generalized to individualized and group instruction with the teacher. The child who has acquired additional skills in a positive manner will have a larger repertoire of desirable skills to employ in the context of responding to teacher instruction. A well-synchronized program of computer-delivered instruction and generalization of task performance to teacher instruction may not only facilitate the acquisition of academic skills, but might also provide a more positive context for promoting compliance and cooperation. In this manner, computer-assisted instruction may provide a valuable positive alternative to the treatment of maladaptive behavior for a subset of children with severe learning and behavior problems.

In providing educational services to children with severe learning problems, a technology of systematic instruction has evolved. This technology of systematic instruction generally incorporates the following: 1) a discrete trial format, 2) graduated presentation of stimuli defining progressively more difficult performance steps, 3) advancement to a more difficult step dependent on performing to criterion on the current step, 4) consistent response contingent consequences to motivate performance, and 5) objective measures of performance (Brown, Crowner, Williams, and York, 1975; Lovaas, 1977; Romanczyk and Lockshin, 1981; and Schreibman and Koegel, 1981). This investigation demonstrates the positive effects of using a microcomputer system to incorporate all of these instructional components. In all likelihood, the computer is also capable of executing these components more efficiently and with greater precision than a person. For methodological considerations, this investigation does not exploit the full potential of computer instruction, e.g., more elaborate audio-visual displays. Clearly, the potential exists to program from fundamental conceptual/perceptual skills through traditional academic subjects.

Although this investigation is an important step in exploring the potential utility of computer-delivered instruction with severely disturbed children, further research is necessary to substantiate these findings. It remains an empirical question whether these results obtained on an analogue discrimination task will generalize to functional skill acquisition tasks. It is also necessary to investigate the generalization of skills acquired by computer-delivered instruction to performance in the classroom with the teacher. The generalization of differences in collateral behavior is another topic for further investigation. Although this investigation produced some very preliminary findings on the identification of children that may profit from computer-assisted instruction, further investigation of these child characteristics is also desirable. In conclusion, the results of this current investigation should provide a new optimistic impetus for investigating the potential utility of computerized instruction in the education and treatment of children with learning and behavior problems. Furthermore, comparisons between teacher and computer delivered instruction establish a new paradigm for investigating stimulus conditions and response relationships with severely disturbed children.

REFERENCES

Azrin, N.H. and Holz, W.C. Punishment. In W. Honig (Ed.), Operant behavior: Areas of research and application. Appleton-Century-Crofts, 1966.

Bijou, S.W. and Baer, D.M. Child development I - A systematic and empirical theory. New York: Appleton-Century-Crofts, 1961.

Bijou, S.W. A functional analysis of retarded development. In N.R. Ellis (Ed.), International review of research in mental retardation, Vol. 1. New York: Academic Press, 1966.

Brown, L., Crowner, T., Williams, W., and York, R. (Eds.). Madison's alternative for zero exclusion: A book of readings. Madison, Wisconsin: Madison Public Schools, 1975.

Bucher, B. and Lovaas, O.I. Use of aversive stimulation in behavior modification. In M. Jones (Ed.), Miami symposium on the prediction of behavior, 1967: Aversive stimulation. Coral Gables, Fla.: University of Miami Press, 1968.

Carr, E.G., Newsom, C.D., and Binkoff, J.A. Escape as a factor in the aggressive behavior of two retarded children. Journal of Applied Behavior Analysis, 1980, 13, 101-118.

Carr, E.G., Newsom, C.D., and Binkoff, J.A. Stimulus control of self-destructive behavior in a psychotic child. Journal of Abnormal Child Psychology, 1976, 4, 139-153.

Doll, E.A. Vineland Social Maturity Scale. Circle Pines, Minn.: American Guidance Services, Inc., 1965.

Dunn, L.M. and Dunn, L.M. Peabody Picture Vocabulary Test-Revised. Circle Pines, Minn.: American Guidance Services, Inc., 1981.

Forehand, R. and King, H.E. Noncompliant children-efforts of parent training on behavior and attitude change. Behavior Modification, 1977, 1, 93-108.

Gelfand, D.M. and Hartmann, D.P. Child behavior analysis and therapy. New York: Pergamon Press, Inc., 1975.

Holland, J.G. Teaching machines: An application of principles from the laboratory. Journal of the Experimental Analysis of Behavior, 1960, 3, 275-287.

Keller, F.S. "Good-bye, Teacher....". Journal of Applied Behavior Analysis, 1968, 1, 79-89.

Koegel, R.L. and Rincover, A. Some detrimental effects of using extra stimuli to guide learning in normal and autistic children. Journal of Abnormal Child Psychology, 1976, 4, 59-71.

Leitenberg, H. Handbook of behavior modification and behavior therapy. New Jersey: Prentice-Hall, Inc., 1976.

Lovaas, O.I. The autistic child-language development through behavior modification. New York: Irvington Publishers, Inc., 1977.

Lovaas, O.I., Koegel, R.L., and Schreibman, L. Stimulus overselectivity in autism: A review of research. Psychological Bulletin, 1979, 86, 1236-1254.

Lovaas, O.I. and Newsom, C.D. Behavior modification with psychotic children. In H. Leitenberg (Ed.), Handbook of behavior modification and behavior therapy. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1976.

Lovaas, O.I., Schreibman, L., Koegel, R.L. and Rehm, R. Selective responding by autistic children to multiple sensory input. Journal of Abnormal Psychology, 1971, 77, 211-222.

Marholin II, D. Child behavior therapy. New York: Gardner Press, Inc., 1978.

Peterson, L., Homer, A.L., and Wonderlich, S.A. The integrity of independent variables in behavior analysis. Journal of Applied Behavior Analysis, 1982, 15, 477-492.

Plummer, S., Baer, D.M., and Le Blanc, J.M. Functional considerations in the use of procedural timeout and an effective alternative. Journal of Applied Behavior Analysis, 1977, 10, 689-705.

Richmond, G. Comparison of automated and human instruction for developmentally retarded preschool children. Journal of the Association for the Severely Handicapped, in press.

Rincover, A. Variables affecting stimulus fading and discriminative responding in autistic children. Journal of Abnormal Psychology, 1978, 87, 541-553.

Romanczyk, R.G. and Lockshin, S. How to create a curriculum for autistic and other handicapped children. Lawrence, Kansas: H & H Enterprises, Inc., 1981.

Ross, A.O. Child behavior therapy: Principles, procedures and empirical basis. New York: John Wiley and Sons, 1981.

Russo, D.C., Cataldo, Michael F., and Cushing, P.J. Compliance training and behavioral covariation in the treatment of multiple behavior problems. Journal of Applied Behavior Analysis, 1981, 14, 209-222.

Russo, D.C., Koegel, R.L., and Lovaas, O.I. A comparison of human and automated instruction of autistic children. Journal of Abnormal Child Psychology, 1978, 6, 189-201.

SAS user's guide - 1979 edition. Cary, North Carolina: SAS Institute, Inc., 1979.

Schreibman, L. and Koegel, R.L. A guideline for planning behavior modification programs for autistic children. In S.M. Turner, K.S. Calhoun, and H.E. Adams (Ed.), Handbook of clinical behavior therapy. New York: John Wiley and Sons, 1981.

Schreibman, L. Effects of within-stimulus and extra-stimulus prompting on discrimination learning in autistic children. Journal of Applied Behavior Analysis, 1975, 8, 91-112.

Sidman, M. and Stoddard, L.T. Programming perception and learning for retarded children. In N.R. Ellis (Ed.), International review of research in mental retardation, Vol. 2. New York: Academic Press, 1966.

Skinner, B.F. The technology of teaching. New Jersey: Prentice-Hall, Inc., 1968.

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

Tatsuoka, M.M. Validation studies: The use of multiple regression equations. Champaign, Illinois: The Institute for Personality and Ability Testing, 1969.

Terrace, H. Stimulus control. In W.K. Honig (Ed.), Operant behavior: Areas of research and application. New York: Appleton-Century-Crofts, 1966.

Thorkildsen, R. Educating handicapped students via microcomputer/videodisc technology. Education and Training of the Mentally Retarded Child, 1981, 16, 264-26[illegible].

Touchette, P.E. The effects of graduated stimulus change on the acquisition of a simple discrimination in severly retarded boys. Journal of the Experimental Analysis of Behavior, 1968, 11, 39-48.

Wahler, R.G. Oppositional children: A quest for parental reinforcement control. *Journal of Applied Behavior Analysis*, 1969, 2, 159-170.

Wiggins, J.S. *Personality and prediction: Principles of personality assessment*. Reading, Massachusetts: Addison-Wesley Publishing Co., 1973.

Table 1

Age, Diagnosis, and Developmental Scores on the

PPVT-R and Vineland Social Maturity Scale

			PPVT-R		Vineland	
Subject Number	Diagnosis	CA[a]	SSE	AE[a]	SQ	AE[a]
1	Emotionally Disturbed	52	High	48	93	50
2	Neurologically Impaired	97	Low	38	52	51
3	Neurologically Impaired	171	Low	27	36	62
4	Emotionally Disturbed	74	Medium	41	102	76
5	Autistic	70	Medium	41	77	54
6	Psychotic	78	High	66	86	78
7	Emotionally Disturbed	83	High	79	95	90
8	Emotionally Disturbed	61	High	67	116	87
9	Autistic	119	Medium	61	84	100
10	Neurologically Impaired	101	Medium	71	71	72
11	Neurologically Impaired	176	Low	31	28	49
12	Autistic	55	Low	23	58	32
13	Emotionally Disturbed	65	Medium	31	91	56
14	Autistic	143	Low	49	57	84
15	Psychotic	162	Low	44	58	92
16	Autistic	65	Low	23	47	30
17	Autistic	77	Low	28	40	31

[a]All age values expressed in months.

Table 2

Observational Compliance Measures and

Noncompliance Ratings by Subject

SUBJECTS	PERCENT COMPLY		NONCOMPLY RATINGS	
	TASK	BEHAVIOR	TASK	BEHAVIOR
1	63	42	3	3
2	96	86	3	3
3	100	78	3	4
4	98	63	5	5
5	92	90	3	4
6	96	88	6	6
7	77	83	3	4
8	100	67	3	4
9	98	92	1	2
10	94	58	4	4
11	100	79	3	2
12	100	40	3	3
13	71	63	3	3
14	100	100	2	2
15	86	80	4	3
16	99	75	5	5
17	97	56	7	7

Table 3

Interrater and Intercategory Correlations for

Teacher Ratings on Noncompliance

		Task		Behavior	
		Primary	Rel	Primary	Rel
Task	Primary	1.000			
	Reliability	.780*	1.000		
Behavior	Primary	.902*	.649*	1.000	
	Reliability	.674*	.844*	.680*	1.000

Note. D.F. = 15

* $p < .01$

TABLE 4

Programmed Changes in the Schedule
of Primary Reinforcement as a Function of
Consecutive Correct Responses and Stage of Presentation

STAGE	SCHEDULE	CHANGE CRITERIA
1	CRF	5 consecutive correct
1-2	FR2	advance to Stage 3
3	FR3	advance to Stage 4
4-5	FR4	advance to Stage 6
6	FR3	back to Stage 5 or advance to Stage 7
7	FR2	back to Stage 6 or advance to Stage 8
8	CRF	back to Stage 7

Table 5

Summary of Procedural Reliability Measures

in the Teacher Condition

I. Adherence to stage change criteria

   A. All 119 sessions reviewed

   B. Stage advancement errors

      1. One within session stage advancement before subject reached criteria.

      2. Two sessions began on next stage before subject reached criterion on previous stage.

      3. One extra trial presented before advancing to next stage occurred three times.

      4. Total of six errors out of 108 stage advancements, i.e. 102/108 = 94.4%.

   C. Stage dropback errors

      1. One error consisting of stimulus card not changed.

      2. Total of one error out of 26 stage dropbacks, i.e. 25/26 = 96.2%.

   D. Total stage change errors

      1. Seven errors out of 134 stage changes, i.e. 127/134 = 94.8%.

II. Accuracy of scoring trial responses

   A. Eight sessions sampled, one randomly selected from each teacher.

   B. One disagreement out of 320 trials, i.e., 319/320 = 99.7%.

Table 6

Description of Performance Measures

1. Percent correct - the proportion of correct responses for all trials presented.
2. Attempts correct - the proportion of correct responses for all trials on which the child made a response, i.e., total number of trials minus the number of trials in which the child failed to make a response.
3. Difficulty level - the number of trials presented at each stage multiplied by the stage number and then divided by the total number of trials.
4. Percent incorrect - the proportion of incorrect responses for all trials presented.
5. No response - the proportion of the total number of trials in which the child failed to make a task response.

Table 7

Operational Definitions for Measures of Collateral Behavior

1. Disruptive Behavior - any of the following:

   a. Out of area - child completely removes self from the immediate task vicinity, the area of which is defined by the front of the task presentation structure and extending along perpendicular lines from the edges of the task presentation structure to the wall. All parts of child must be out of this defined space to be scored as disruptive behavior with the exception of kneeling, lying, or sitting on the floor which are scored as disruptive.

   b. Displacement of task presentation structure - In the teacher condition, this includes tipping or vigorous shaking of the desk/lecturn. In the computer condition, this includes visible movement of the plywood casing which encloses the teaching apparatus.

   c. Removal of attached parts of fixtures - includes complete disengagement with no contact remaining between the previously attached part and the previously adjoined surface. Examples include removal of the liner in the M & M dispenser receptacle in the computer condition and removal of the attached plastic M & M receptacle from the lecturn in the teacher condition.

   d. Throwing objects - includes either the detection of projected object in flight or child's arm engaging in definitive throwing or dropping motion followed by hand-release movement.

2. Out of seat - child's buttocks break contact with seat of chair. When child is out of area (see disruptive behavior definition), do not score out of seat.

3. Self-stimulation - three continuous cycles of rhythmic motion of head, torso, arm, hand or finger. The directional plane or intensity of the response may vary across cycles.

4. Redirection - Any occurrence of the teacher or monitoring adult providing a verbal direction (other than the specified task instructions) or making any physical contact with the child.

Table 8

Three Forms of Reliability for Out of Seat and Redirection by Condition

	Direct Interval		Adjacent Interval		Session Totals			
	A/(A + D)[a]	Reliability	A/(A + D)[a]	Reliability	T_1/T_2	Reliability	Sessions	Intervals
Out of seat teacher	1,543/1,737	.89	1,600/1,656	.97	1,612/1,668	.97	40	6,400
Out of seat computer	1,204/1,464	.82	1,275/1,366	.93	1,274/1,390	.92	41	6,560
Redirection teacher	151/199	.76	163/186	.87	172/188	.91	40	6,400
Redirection computer	48/62	.77	52/55	.95	52/58	.90	41	6,560

[a] A = agreements; D = disagreements

Table 9

Three Forms of Reliability for Disruptive Behavior and Self-stimulation by Condition

	Direct Interval		Adjacent Interval		Session Totals			
	A/(A + D)[a]	Reliability	A/(A + D)[a]	Reliability	T_1/T_2	Reliability	Sessions	Intervals
Disruptive Behavior teacher	471/609	.77	513/573	.90	540/565	.96	50	7,945
Disruptive Behavior computer	285/358	.80	310/350	.89	320/347	.92	51	8,066
Self-stimulation teacher	830/1,099	.76	907/1,096	.83	866/1,038	.93	49	7,785
Self-stimulation computer	711/1,085	.66	822/971	.85	823/977	.84	48	7,586

[a] A = agreements; D = disagreements

Table 10

Differences Between Means on All

Dependent Measures by Condition

DEPENDENT MEASURES	TEACHER		COMPUTER		DIFFERENCE SCORES			
	MEAN	SD	MEAN	SD	MEAN	SD	t(16)	P
PERFORMANCE								
General Performance Index	80.59	58.68	77.29	63.33	3.30	23.70	0.574	n.s.
Percent Correct	75.92	14.52	78.65	12.96	-2.73	14.52	-0.775	n.s.
Attempts Correct	83.61	18.37	85.56	9.02	-1.95	15.05	-0.535	n.s.
Difficulty Level	5.13	1.50	4.55	1.76	0.58	1.69	1.404	$p < .10$
Percent Incorrect	14.4	16.67	12.82	7.93	1.58	14.47	.449	n.s.
Percent No Response	9.79	12.35	8.58	7.86	1.21	10.97	0.456	n.s.
COLLATERAL BEHAVIOR								
Total Deviance Scores	35.29	27.06	27.34	20.08	7.95	18.21	1.800	$p < .05$
Out of Seat	19.59	20.98	16.59	17.54	3.00	13.39	0.924	n.s.
Disruptive Behavior	5.71	11.40	2.88	1.04	2.82	8.29	1.405	$p < .10$
Self-Stimulation	10.00	8.93	9.18	7.5[illegible]	0.82	6.47	0.525	n.s.
Redirection	2.55	2.49	0.44	0.6[illegible]	2.11	2.19	3.976	$p < .005$

Table 11

Correlations on Session Performance and

Subject Assignment to Teachers

	Teacher	Subject	Percent Correct	Difficulty Level
Teacher	1.000			
Subject	- .140	1.000		
Percent Correct	- .057	- .046	1.000	
Difficulty Level	.039	- .144	.258*	1.000

*$p < .05$, d.f. = 117

Table 12

Correlation Matrices for all Dependent Measures by Condition

Teacher

	GPI	PC	AC	DL	PI	NR	TDS	OS	DB	SS
General Performance Index	1.000									
Percent Correct (all trials)	.733***	1.000								
Attempts Correct (%)	.621***	.876***	1.000							
Difficulty Level (mean)	.399	.683***	.725***	1.000						
Percent Incorect	-.577**	-.814***	-.993***	-.708***	1.000					
Percent No Response	-.540**	-.617***	-.162	-.216	.046	1.000				
Total Deviance Score	580**	-.579**	-.290	-.405	.226	.681***	1.000			
Out of Seat	371	-.293	-.079	-.065	.041	.439*	.864***	1.000		
Disruptive Behavior	.425*	-.454*	-.024	-.285	-.075	.882***	.666***	.354	1.000	
Self-stimulation	-.343	-.485**	-.665***	-.711***	.683***	-.093	.152	-.183	-.087	1.000
Redirection	-.591**	-.676***	-.553**	-.451*	.506**	.470*	.562**	.374	.361	.363

Computer

	GPI	PC	AC	DL	PI	NR	TDS	OS	DB	SS
General Performance Index	1.000									
Percent Correct (all trials)	.951***	1.000								
Attempts Correct (%)	.921***	.887***	1.000							
Difficulty Level (mean)	.569**	.417*	.509**	1.000						
Percent Incorrect	-.886***	-.817***	-.988***	-.523**	1.000					
Percent No Response	-.662***	-.805***	-.444**	-.178	.320	1.000				
Total Deviance Score	-.396	-.337	-.[illegible]86	-.181	.164	.428	1.000			
Out of Seat	-.340	-.307	-.141	-.131	.123	.429*	.948***	1.000		
Disruptive Behavior	-.503**	-.505**	-.411	-.341	.374	.467*	.681***	.644***	1.000	
Self-stimulation	-.177	-.146	-.197	-.307	.170	.059	.003	-.197	-.250	1.000
Redirection	-.637***	-.629***	-.672***	-.720***	.646***	.389	.302	.309	.626***	.242

Note. D.F. = 15

* $p < .10$

** $p < .05$

*** $p < .01$

Table 13

Correlations Between Predictor Variables and Criterion Measures Used in Regression Analyses

	Emot. Dist.	Neur. Imp.	Aut.	Psychotic	LOW	MED	HIGH	CA	PPVT-R AE	V-SQ
Low range PPVT-R SSE	-.609***	.311	.290	.022	1.000					
Medium range PPVT-R SSE	.150	-.054	.064	-.236	-.60[illegible]***	1.000				
High range PPVT-R SSE	.555**	-.308	-.410	.228	-.523**	-.358	1.000			
CA	-.470*	.558**	-.173	.186	.497**	-.187	-.385	1.000		
PPVT-R AE	.299	-.110	-.328	.207	-.670***	.143	.635***	-.109	1.000	
Vineland SQ	.767***	-.524**	-.286	.029	-.880***	.391	.616***	-.599**	.638***	1.000
Vineland AE	.220	-.149	-.311	.345	-.452*	.214	.302	.287	.767***	.519**
Noncomply ratings-task (NRT)	-.086	-.132	-.046	.364	.108	-.177	.063	-.271	-.167	-.136
Noncomply ratings-behavior (NRB)	.017	-.211	.037	.199	-.097	.079	.199	-.415*	-.013	.055
% Comply-behavior	-.359	.076	.113	.241	.074	.010	-.097	.487**	.256	-.126
% Comply-task	-.611***	.269	.370	-.039	.436*	-.093	-.414[illegible]	.316	-.180	-.411
Total Deviance Scores with teacher	-.051	-.033	.306	-.338	-.024	.357	[illegible]	-.516**	-.468*	-.109
Total Deviance Scores with computer	-.104	-.088	.378	-.297	-.108	.260	-.152	-.446*	-.259	-.073
Total Deviance difference scores	.038	.049	.038	-.175	.085	.243	-.359	-.276	-.410	-.081
General Performance Index with teacher	.312	-.461*	-.041	.227	-.251	-.014	[illegible]	.062	.483**	.406
General Performance Index with Computer	.453*	-.461*	-.168	.215	-.344	.107	[illegible]	-.204	.484**	.493**
General Performance Index difference scores	-.438*	.091	.346	-.013	.272	-.322	[illegible]	.19[illegible]	-.098	-.313

	V-AE	NRT	NRB	%CB	%CT	TDS-T	[illegible]	TDS DIF	GPI-T	GPI-C
Vineland AE	1.000									
Noncomply ratings-task (NRT)	-.308	1.000								
Noncomply ratings-behavior (NRB)	-.268	.902***	1.000							
% Comply-behavior	.503**	-.276	-.205	1.000						
% Comply-task	-.031	.122	.146	[illegible]	1.000					
Total Deviance Scores with teacher	-.723***	.301	.267	-.424*	[illegible]	1.000				
Total Deviance Scores with computer	-.568**	.536[illegible]	.576[illegible]	-.337	[illegible]	.740***	1.000			
Total Deviance difference scores	-.449*	-.145	-.240	-.257	-.399	.671***	-.003	1.000		
General Performance Index with teacher	.674[illegible]	-.109	.031	.287	-.133	-.580**	-.322	-.506**	1.000	
General Performance Index with computer	.651***	-.137	-.079	.151	-.298	-.486**	-.396	-.2[illegible]6	.927***	1.000
General Performance Index difference scores	-.071	.095	.289	.306	.467*	-.137	.259	-.489**	[illegible]	-.376

	DFGPI
General Performance Index difference scores	1.000

Note. D.F. = 15

* $p < .10$

** $p < .05$

*** $p < .01$

Table 14

Regression Models Predicting General Performance Index

Variables	Regression Coefficient	Intercept Value	DF	T	P	R^{2a}	F	DF	P
				One variable model-teacher					
Vineland AE	1.76	-32.35	15	3.53	.005	.454			
				Four variable model-teacher					
Neurologically impaired	-53.37		12	-2.21	.025				
Psychotic	-27.16		12	-.76	.25				
Medium range PPVT-R SSE	-28.09		12	-1.18	.15				
Vineland AE	1.86		12	3.63	.005				
		-15.40				.634	5.19	4,12	.025
				One variable model-computer					
Vineland AE	1.83	-40.4[illegible]	15	3.32	.005	.423			
				Four variable model-computer					
Emotionally disturbed	74.21		12	1.99	.05				
Neurologically impaired	-63.60		12	-2.20	.025				
Vineland AE	2.10		12	3.56	.005				
Vineland SQ	-1.36		12	-1.56	.10				
		30.35				.668	[illegible]	4.12	.01
				One variable model-difference scores					
% Comply Task	[illegible]	-86.97	15	2.05	.05	.218			
				Four variable model-difference scores					
Emotionally disturbed	-26.55		12	-2.99	.01				
Medium Range PPVT-R SSE	-18.54		12	-2.07	.05				
Noncomply ratings-task	-22.32		12	-3.12	.005				
Noncomply ratings-behavior	25.66		12	3.69	.005				
		.036				.652	5.63	4,12	.01

[a] r^2 for one variable models

BEST COPY AVAILABLE

Table 15

Regression Models Predicting Total Deviance Scores

Variables	Regression Coefficient	Intercept Value	DF	T	P	R^{2a}	F	DF	P
				One variable model-teacher					
Vineland AE	-.87	91.25	15	-4.06	.005	.523			
				Three variable model-teacher					
High Range PPVT-R SSE	-27.38		13	-2.64	.025				
CA	-.35		13	-3.39	.005				
Vineland AE	-.51		13	-2.65	.01				
		108.78				.758	13.54	3,13	.0005
				One variable model-computer					
Noncomply ratings-behavior	8.33	-4.00	[illegible]	2.74	.01	.578			
				Three variable model-computer					
Medium Range PPVT-R SSE	17.56		13	2.55	.025				
Vineland AE	-.47		13	-3.18	.005				
Noncomply ratings-behavior	6.74		13	2.87	.01				
		27.23				.678	9.15	3,13	.005
				Four variable model-computer					
Emotionally disturbed	-27.05		12	-2.30	.025				
Psychotic	-46.29		12	-4.03	.005				
Vineland SQ	.43		12	2.03	.05				
Noncomply ratings-task	11.48		12	4.76	.0005				
						[illegible]	7.04	4,12	.005
				One variable model-difference scores					
Vineland AE	-.36	31.33	15	-1.95	.05	.202			
				Three variable model-difference scores					
High Range PPVT-R SSE	-31.18		13	-3.75	.005				
CA	-.16		13	-1.97	.05				
% Comply Task	-.95		13	-3.04	.005				
		118.13				[illegible]	6.73	3,13	.01
				Four variable model-difference scores					
Neurologically impaired	11.61		12	-1.36	.10				
High Range PPVT-R SSE	-30.22		12	-3.74	.005				
CA	-.21		12	-2.44	.025				
% Comply Task	-.98		12	-3.24	.005				
		124.00				.661	5.84	4,12	.01

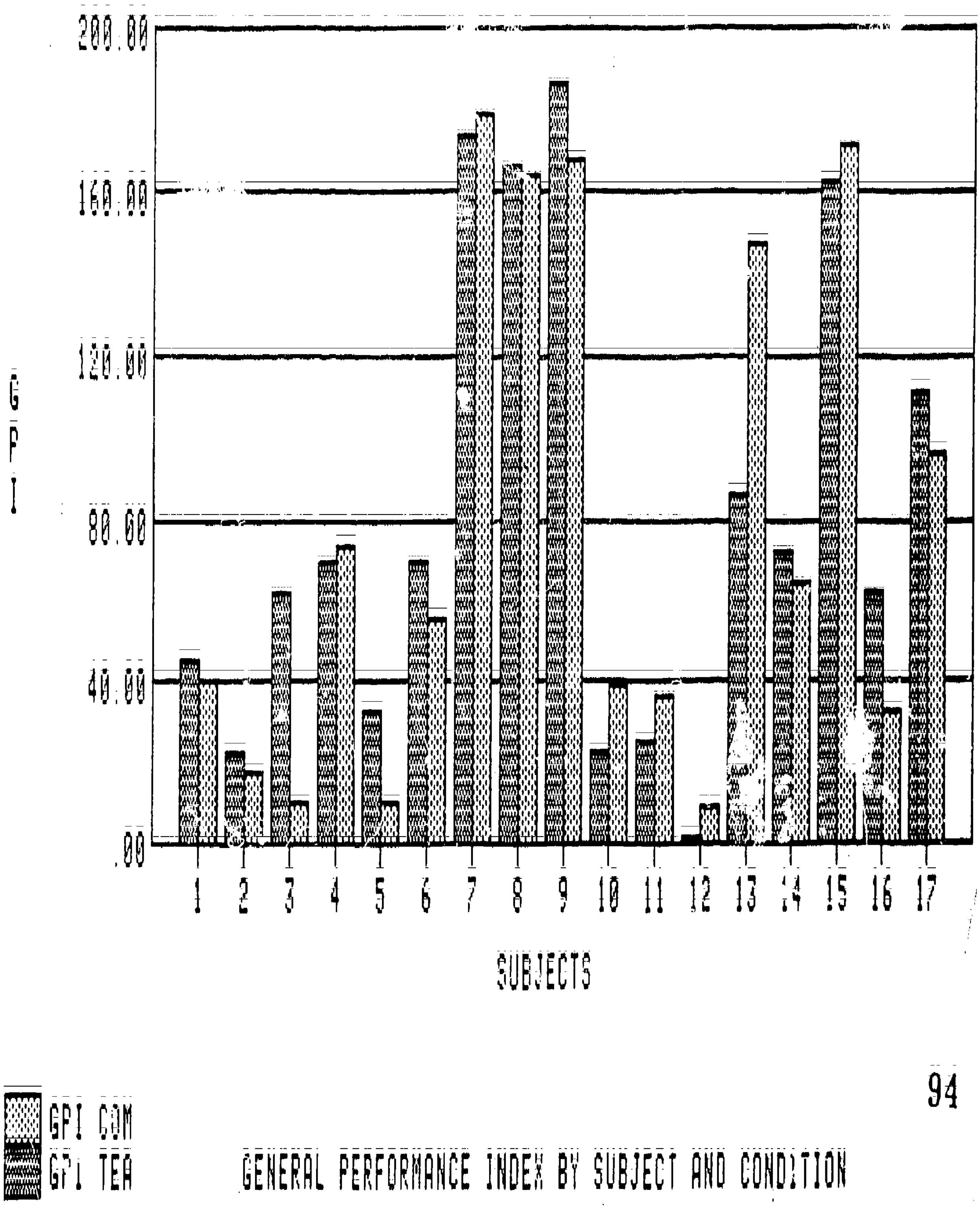
[a] r^2 for one variable models

FIGURE 1

$$\frac{\sum_{1 \text{ to } 8}^{J} (\text{J X PRECENT CORRECT AT STAGE J X NUMBER OF TRIALS AT STAGE J}) \;/\; \text{TOTAL TRIALS}}{\text{TOTAL NUMBER OF SESSIONS}}$$

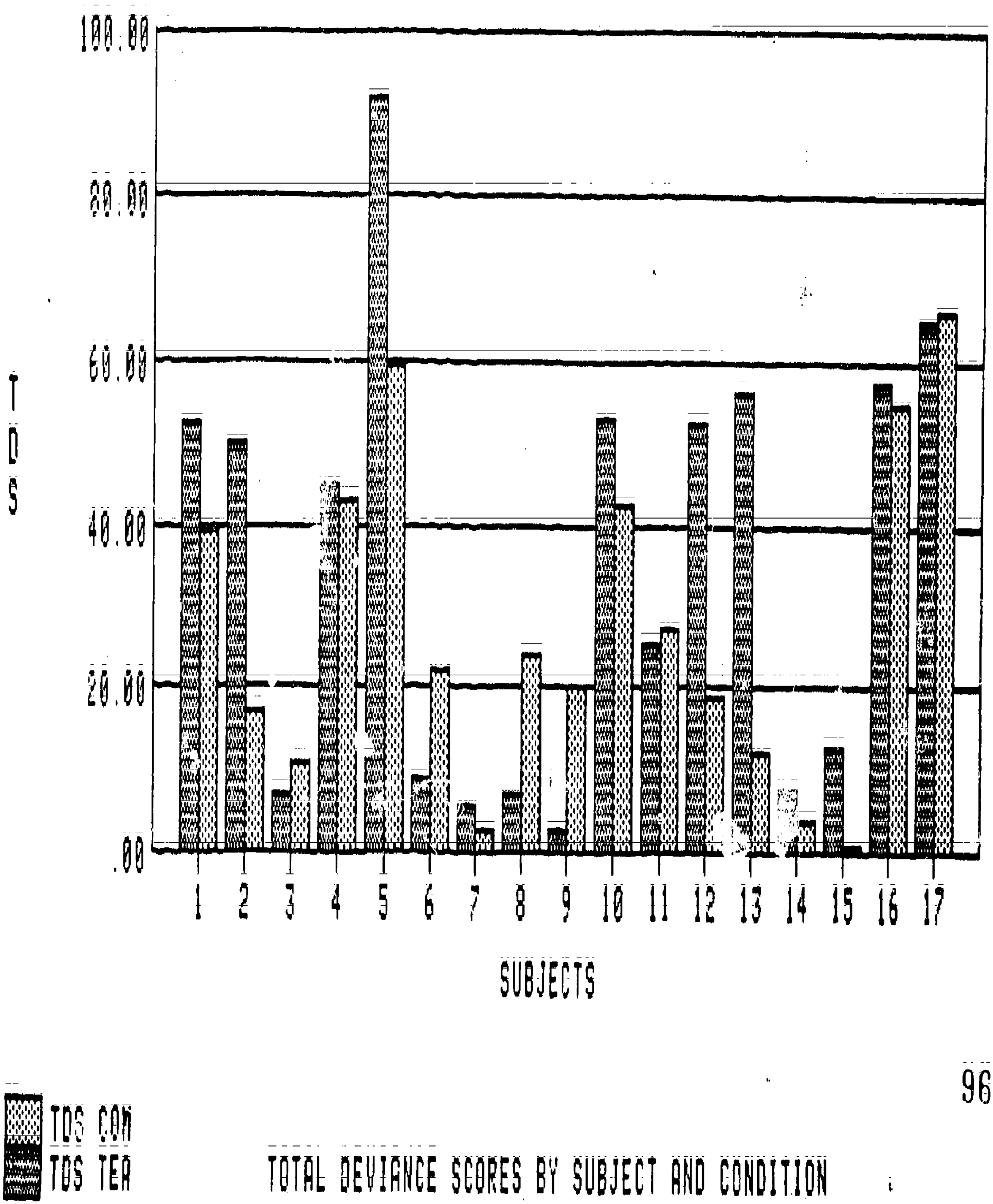
Formula for the calculation of General Performance Index (GPI)

FIGURE 2



GENERAL PERFORMANCE INDEX BY SUBJECT AND CONDITION

FIGURE 3



TOTAL DEVIANCE SCORES BY SUBJECT AND CONDITION

APPENDIX

Table A

Out of Seat and Redirection Reliabilities as a Function

of Pairings with Observer A

	Observers	Direct Intervals		Adjacent Intervals		Session Totals		Sessions	Intervals
		A/(A+D)[a]	Reliability	A/(A+D)[a]	Reliability	T_1/T_2	Reliability		
Seat r)	B	1079/1231	.88	1127/1169	.96	1134/1176	.96	32	5,120
	C	464/506	.92	473/487	.97	476/496	.96	8	1,280
Seat er)	B	706/845	.84	739/785	.94	741/810	.91	25	4,000
	C	494/619	.80	536/581	.92	532/581	.92	16	2,560
tion r)	B	75/111	.68	84/97	.87	87/99	.88	32	5,120
	C	80/104	.77	82/94	.87	88/96	.92	8	1,280
tion er)	B	37/49	.76	41/43	.95	41/45	.91	25	4,000
	C	11/13	.85	11/12	.92	11/13	.85	16	2,560

reements; D = disagreements

Table B

Self-Stimulation and Disruptive Behavior Reliabilities for Each Observer

as a Function of Rotated Pairings

	Observers	Direct Intervals A/(A+D)[a]	Direct Intervals Reliability	Adjacent Intervals A/(A+D)[a]	Adjacent Intervals Reliability	Session Totals T_1/T_2	Session Totals Reliability	Sessions	Intervals
Self-Stimulation (Teacher)	D	679/890	.76	737/849	.87	755/820	.92	34	5,223
	E	599/834	.72	656/794	.83	686/752	.91	34	5,097
	F	217/314	.69	232/284	.82	242/288	.84	10	1,600
	G	99/170	.58	113/151	.75	127/142	.89	11	1,760
	H	296/390	.76	318/360	.88	324/362	.90	13	2,208
Self-Stimulation (Computer)	D	418/658	.64	477/579	.82	494/586	.84	29	4,712
	E	229/381	.60	270/343	.79	281/333	.84	23	3,624
	F	2/6	.33	2/5	.40	3/5	.60	2	318
	G	365/568	.64	429/500	.86	421/512	.82	21	3,342
	H	293/427	.69	346/392	.88	329/391	.84	19	3,020
Disruptive Behavior (Teacher)	D	300/343	.80	356/399	.89	373/399	.93	35	5,673
	E	305/390	.78	318/362	.88	335/360	.93	37	5,865
	F	174/192	.91	176/187	.94	180/186	.97	8	1,280
	G	194/223	.87	198/215	.92	205/212	.95	13	2,080
	H	114/133	.86	117/129	.91	120/127	.94	24	3,822
Disruptive Behavior (Computer)	D	173/203	.85	195/226	.86	202/225	.90	29	4,700
	E	130/155	.84	152/180	.84	158/178	.89	22	3,452
	F	0/0	-	0/0	-	0/0	-	2	318
	G	113/129	.88	116/125	.93	119/123	.97	24	3,822
	H	154/175	.88	157/169	.93	161/168	.96	25	4,108

[a]A = agreements; D = disagreements

Table C

Three Forms of Reliability for Out of Seat

by Subject and Condition

Teacher

Subjects	Direct Interval A/A+D	Direct Interval Rel.	Adjacent Interval A/A+D	Adjacent Interval Rel.	Session Totals T_1/T_2	Session Totals Rel.	Sessions	Intervals
1	207/228	.91	215/218	.99	217/218	.99	2	320
2	8/14	.57	10/14	.71	9/13	.69	3	480
3	0/0	-	-	-	-	-	1	160
4	99/115	.86	102/106	.96	105/109	.96	2	320
5	188/233	.81	204/212	.96	206/215	.96	4	640
6	4/8	.50	4/8	.50	6/6	100	5	800
7	-	-	-	-	-	-	-	-
8	2/4	.50	3/3	100	3/3	100	1	160
9	-	-	-	-	-	-	-	-
10	25/34	.73	26/33	.78	29/30	.97	1	160
11	0/0	-	0/0	-	0/0	-	4	640
12	121/129	.94	123/123	100	123/127	.97	3	480
13	364/402	.91	368/387	.95	370/396	.93	5	800
14	13/20	.65	14/19	.74	16/17	.94	2	320
15	0/0	-	0/0	-	0/0	-	1	160
16	226/252	.90	240/241	100	237/241	.98	3	480
17	286/298	.96	291/292	100	291/293	.99	3	480

Computer

Subjects	Direct Interval A/A+D	Direct Interval Rel.	Adjacent Interval A/A+D	Adjacent Interval Rel.	Session Totals T_1/T_2	Session Totals Rel.	Sessions	Interval
1	159/196	.81	165/179	.92	168/187	.90	3	480
2	24/41	.61	30/40	.75	28/38	.74	3	480
3	0/0	-	0/0	-	0/0	-	3	480
4	109/136	.70	117/136	.86	119/146	.82	3	480
5	227/293	.95	281/285	.99	284/286	.99	3	490
6	95/110	.86	100/100	100	100/105	.95	3	480
7	-	-	-	-	-	-	-	-
8	-	-	-	-	-	-	-	-
9	0/0	-	-	-	-	-	1	160
10	80/125	.64	98/116	.84	89/116	.77	3	480
11	0/2	0	0/2	0	0/2		2	320
12	141/164	.86	148/156	.95	147/154	.95	4	640
13	5/11	.45	6/9	.66	6/10	.60	1	160
14	3/6	.50	3/5	.60	3/6	.50	5	800
15	0/0	-	0/0	-	0/0	-	2	320
16	157/179	.88	163/168	.97	164/172	.95	3	480
17	153/181	.85	164/170	.96	166/168	.99	2	320

Table D

Three Forms of Reliability for Redirection

by Subject and Condition

Teacher

	Direct	Interval	Adjacent	Interval	Session	Totals		
Subjects	A/A+D	Rel.	A/A+D	Rel.	T_1/T_2	Rel.	Sessions	Intervals
1	0/0	-	0/0	-	0/0	-	2	320
2	0/2	0	1/1	100	1/1	100	3	480
3	0/0	-	-	-	-	-	1	160
4	6/10	.60	8/8	100	8/8	100	2	320
5	39/54	.72	43/47	.91	44/49	.90	4	640
6	0/1	0	0/1	0	0/1	0	5	800
7	-	-	-	-	-	-	-	-
8	1/1	100	1/1	100	1/1	100	1	160
9	-	-	-	-	-	-	-	-
10	30/38	.78	32/34	.94	33/35	.94	1	160
11	8/13	.62	10/12	.83	9/12	.75	4	640
12	30/37	.81	30/34	.88	32/35	.91	3	480
13	17/21	.81	17/20	.85	19/19	100	5	800
14	4/6	.66	4/5	.80	5/5	100	2	320
15	0/0	-	-	-	-	-	1	160
16	6/10	.60	7/9	.78	8/8	100	3	480
17	10/16	.63	10/14	.71	12/14	.86	3	480

Computer

	Direct	Interval	Adjacent	Interval	Session	Totals		
Subjects	A/A+D	Rel.	A/A+D	Rel.	T_1/T_2	Rel.	Sessions	Intervals
1	1/3	.33	2/2	100	2/2	100	3	480
2	4/6	.66	4/4	100	4/6	.66	3	480
3	3/3	100	3/3	100	3/3	100	3	480
4	0/0	-	-	-	-	-	3	480
5	5/6	.83	5/5	100	5/6	.83	3	480
6	0/0	-	-	-	-	-	3	480
7	-	-	-	-	-	-	-	-
8	-	-	-	-	-	-	-	-
9	-	-	-	-	-	-	1	160
10	1/2	.50	1/2	.50	1/2	.50	3	480
11	2/5	.40	3/4	.75	3/4	.75	2	320
12	25/30	.83	22/28	.96	27/28	.96	4	640
13	0/0	-	-	-	-	-	1	160
14	0/0	-	-	-	-	-	5	800
15	0/0	-	-	-	-	-	2	320
16	3/3	100	3/3	100	3/3	100	3	480
17	4/4	100	4/4	100	4/4	100	2	320

Table D

Three Forms of Reliability for Redirection

by Subject and Condition

	Teacher								Computer							
	Direct	Interval	Adjacent	Interval	Session	Totals			Direct	Interval	Adjacent	Interval	Session	Totals		
ubjects	A/A+D	Rel.	A/A+D	Rel.	T_1/T_2	Rel.	Sessions	Intervals	A/A+D	Rel.	A/A+D	Rel.	T_1/T_2	Rel.	Sessions	Intervals
1	0/0	-	0/0	-	0/0	-	2	320	1/3	.33	2/2	100	2/2	100	3	480
2	0/2	0	1/1	100	1/1	100	3	480	4/6	.66	4/4	100	4/6	.66	3	480
3	0/0	-	-	-	-	-	1	160	3/3	100	3/3	100	3/3	100	3	480
4	6/10	.60	8/8	100	8/8	100	2	320	0/0	-	-	-	-	-	3	480
5	39/54	.72	43/47	.91	44/49	.90	4	640	5/6	.83	5/5	100	5/6	.83	3	480
6	0/1	0	0/1	0	0/1	0	5	800	0/0	-	-	-	-	-	3	480
7	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8	1/1	100	1/1	100	1/1	100	1	160	-	-	-	-	-	-	-	-
9	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	160
10	30/38	.78	32/34	.94	33/35	.94	1	160	1/2	.50	1/2	.50	1/2	.50	3	480
11	8/13	.62	10/12	.83	9/12	.75	4	640	2/5	.40	3/4	.75	3/4	.75	2	320
12	30/37	.81	30/34	.88	32/35	.91	3	480	25/30	.83	22/28	.96	27/28	.96	4	640
13	17/21	.81	17/20	.85	19/19	100	5	800	0/0	-	-	-	-	-	1	160
14	4/6	.66	4/5	.80	5/5	100	2	320	0/0	-	-	-	-	-	5	800
15	0/0	-	-	-	-	-	1	160	0/0	-	-	-	-	-	2	320
16	6/10	.60	7/9	.78	8/8	100	3	480	3/3	100	3/3	100	3/3	100	3	480
17	10/16	.63	10/14	.71	12/14	.86	3	480	4/4	100	4/4	100	4/4	100	2	320

Table E

Three Forms of Reliability for Disruptive Behavior

by Subject and Condition

Teacher

Subjects	Direct A/A+D	Interval Rel.	Adjacent A/A+D	Interval Rel.	Session T_1/T_2	Totals Rel.	Sessions	Intervals
1	19/27	.70	20/24	.83	21/25	.84	4	635
2	27/33	.82	22/23	.96	27/33	.82	3	468
3	0/0	-	0/0	-	0/0	-	3	442
4	16/26	.62	18/83	.78	20/22	.91	3	480
5	221/253	.87	227/245	.93	237/237	100	4	640
6	0/0	-	0/0	-	0/0	-	2	320
7	0/0	-	0/0	-	0/0	-	1	160
8	0/0	-	0/0	-	0/0	-	2	320
9	0/0	-	0/0	-	0/0	-	1	160
10	140/168	.83	148/158	.94	151/157	.96	5	800
11	5/9	.55	5/9	.55	6/8	.75	3	480
12	2/8	.25	2/8	.25	4/6	.67	5	800
13	1/1	100	1/1	100	1/1	100	2	320
14	1/3	.33	1/3	.33	2/2	100	4	640
15	0/0	-	0/0	-	0/0	-	2	320
16	59/73	.81	64/72	.89	65/67	.97	4	640
17	5/8	.62	5/7	.71	6/7	.86	2	320

Computer

Subjects	Direct A/A+D	Interval Rel.	Adjacent A/A+D	Interval Rel.	Session T_1/T_2	Totals Rel.	Sessions	Intervals
1	3/5	.60	3/4	.75	3/5	.60	2	314
2	2/2	100	2/2	100	2/2	100	3	450
3	5/11	.45	5/8	.63	5/11	.45	4	636
4	5/8	.63	5/8	.63	5/8	.63	4	636
5	154/171	.90	157/165	.95	161/164	.98	5	800
6	0/0	-	0/0	-	0/0	-	3	477
7	0/0	-	0/0	-	0/0	-	2	314
8	1/2	.50	1/2	.50	1/2	.50	1	160
9	1/1	100	1/1	100	1/1	100	2	320
10	41/58	.71	56/68	.82	59/64	.92	5	787
11	0/0	-	0/0	-	0/0	-	3	475
12	23/38	.61	26/35	.74	29/32	.91	4	622
13	0/0	-	0/0	-	0/0	-	3	480
14	0/0	-	0/0	-	0/0	-	3	480
15	0/0	-	0/0	-	0/0	-	2	318
16	32/39	.82	34/37	.92	34/37	.71	3	477
17	18/23	.78	20/20	100	20/21	.95	2	320

Table F

Three Forms of Reliability for Self-Stimulation

by Subject and Condition

	Teacher								Computer							
	Direct Interval		Adjacent Interval		Session Totals				Direct Interval		Adjacent Interval		Session Totals			
Subjects	A/A+D	Rel.	A/A+D	Rel.	T_1/T_2	Rel.	Sessions	Intervals	A/A+D	Rel.	A/A+D	Rel.	T_1/T_2	Rel.	Sessions	Intervals
1	12/23	.52	16/22	.72	17/18	.94	4	635	9/13	.69	10/12	.83	11/11	100	2	314
2	138/163	.85	146/159	.92	149/152	.98	3	468	26/45	.58	33/42	.79	32/44	.73	3	450
3	41/62	.66	44/52	.85	47/56	.84	3	442	52/114	.46	70/97	.72	79/87	.91	4	636
4	21/39	.54	26/36	.72	30/30	100	3	480	34/50	.68	35/47	.74	39/45	.87	4	636
5	8/12	.67	8/12	.67	10/11	.91	2	320	4/6	.67	4/6	.67	4/6	.67	2	320
6	23/44	.52	27/39	.69	32/35	.91	2	320	31/52	.60	34/46	.74	40/43	.93	3	477
7	25/28	.89	25/27	.92	25/28	.89	1	160	1/5	.2	1/5	.2	3/3	100	2	314
8	5/12	.42	6/11	.56	8/9	.89	2	320	42/55	.76	48/52	.92	48/49	.98	1	160
9	5/7	.71	6/7	.85	6/6	100	1	160	36/57	.63	40/49	.82	44/49	.90	2	320
10	52/79	.66	62/82	.76	67/69	.97	5	800	35/66	.53	43/58	.74	43/58	.74	5	787
11	119/166	.72	125/141	.89	138/147	.94	4	640	161/218	.74	201/203	.99	169/210	.80	3	475
12	220/294	.75	235/277	.85	250/264	.95	5	800	142/165	.86	151/157	.96	148/159	.93	4	622
13	8/14	.57	11/14	.79	10/12	.83	2	320	48/111	.43	58/79	.73	65/94	.69	3	480
14	18/22	.82	18/22	.82	20/20	100	4	640	12/20	.60	12/19	.63	13/19	.68	3	480
15	38/76	.50	45/71	.63	51/62	.82	2	320	2/6	.33	2/5	.40	3/5	.60	2	318
16	20/32	.63	23/32	.72	23/29	.79	4	640	38/52	.73	38/47	.81	41/49	.84	3	477
17	77/96	.80	84/92	.91	83/90	.92	2	320	38/50	.76	42/47	.89	41/46	.89	2	320

Performance Record Form

SUBJECT:_______________ SUBJ.#:_____ TEACHER:_______________ DATE:_________

SESSION TIME:__________ P.A.T. 1 2 3 4 SESSION 1 2 3 4 5 6 7 8 9 10

STAGE BEGIN:_______ REINF SCHED BEGIN:_______

PDS-3

1.R__/__	6.L__/__	11.R__/__	16.R__/__	21.R__/__	26.L__/__	31.L__/__	36.R__/__
2.R__/__	7.R__/__	12.L__/__	17.R__/__	22.L__/__	27.R__/__	32.R__/__	37.R__/__
3.L__/__	8.L__/__	13.R__/__	18.L__/__	23.R__/__	28.R__/__	33.R__/__	38.R__/__
4.R__/__	9.R__/__	14.L__/__	19.L__/__	24.R__/__	29.L__/__	34.L__/__	39.R__/__
5.R__/__	10.L__/__	15.L__/__	20.L__/__	25.L__/__	30.L__/__	35.L__/__	40.L__/__

STG																						FR
___	1	2	3	4	5	6	7	8	9	10	R	R	R	R	R	R	R	1	2	3	4	___
___	1	2	3								L	L	L	L	L	L	L	1	2	3	4	___
___	1	2	3	4	5	6	7	8	9	10	R	R	R	R	R	R	R	1	2	3	4	___
___	1	2	3								L	L	L	L	L	L	L	1	2	3	4	___
___	1	2	3	4	5	6	7	8	9	10	R	R	R	R	R	R	R	1	2	3	4	___
___	1	2	3								L	L	L	L	L	L	L	1	2	3	4	___
___	1	2	3	4	5	6	7	8	9	10	R	R	R	R	R	R	R	1	2	3	4	___
___	1	2	3								L	L	L	L	L	L	L	1	2	3	4	___
___	1	2	3	4	5	6	7	8	9	10	R	R	R	R	R	R	R	1	2	3	4	___
___	1	2	3								L	L	L	L	L	L	L	1	2	3	4	___
___	1	2	3	4	5	6	7	8	9	10	R	R	R	R	R	R	R	1	2	3	4	___
___	1	2	3								L	L	L	L	L	L	L	1	2	3	4	___
___	1	2	3	4	5	6	7	8	9	10	R	R	R	R	R	R	R	1	2	3	4	___
___	1	2	3								L	L	L	L	L	L	L	1	2	3	4	___
___	1	2	3	4	5	6	7	8	9	10	R	R	R	R	R	R	R	1	2	3	4	___
___	1	2	3								L	L	L	L	L	L	L	1	2	3	4	___
___	1	2	3	4	5	6	7	8	9	10	R	R	R	R	R	R	R	1	2	3	4	___
___	1	2	3								L	L	L	L	L	L	L	1	2	3	4	___
___	1	2	3	4	5	6	7	8	9	10	R	R	R	R	R	R	R	1	2	3	4	___
___	1	2	3								L	L	L	L	L	L	L	1	2	3	4	___
___	1	2	3	4	5	6	7	8	9	10	R	R	R	R	R	R	R	1	2	3	4	___
___	1	2	3								L	L	L	L	L	L	L	1	2	3	4	___
___	1	2	3	4	5	6	7	8	9	10	R	R	R	R	R	R	R	1	2	3	4	___
___	1	2	3								L	L	L	L	L	L	L	1	2	3	4	___
___	1	2	3	4	5	6	7	8	9	10	R	R	R	R	R	R	R	1	2	3	4	___
___	1	2	3								L	L	L	L	L	L	L	1	2	3	4	___
___	1	2	3	4	5	6	7	8	9	10	R	R	R	R	R	R	R	1	2	3	4	___
___	1	2	3								L	L	L	L	L	L	L	1	2	3	4	___
___	1	2	3	4	5	6	7	8	9	10	R	R	R	R	R	R	R	1	2	3	4	___
___	1	2	3								L	L	L	L	L	L	L	1	2	3	4	___
___	1	2	3	4	5	6	7	8	9	10	R	R	R	R	R	R	R	1	2	3	4	___
___	1	2	3								L	L	L	L	L	L	L	1	2	3	4	___
___	1	2	3	4	5	6	7	8	9	10	R	R	R	R	R	R	R	1	2	3	4	___
___	1	2	3								L	L	L	L	L	L	L	1	2	3	4	___
___	1	2	3	4	5	6	7	8	9	10	R	R	R	R	R	R	R	1	2	3	4	___
___	1	2	3								L	L	L	L	L	L	L	1	2	3	4	___
___	1	2	3	4	5	6	7	8	9	10	R	R	R	R	R	R	R	1	2	3	4	___
___	1	2	3								L	L	L	L	L	L	L	1	2	3	4	___
___	1	2	3	4	5	6	7	8	9	10	R	R	R	R	R	R	R	1	2	3	4	___
___	1	2	3								L	L	L	L	L	L	L	1	2	3	4	___
___	1	2	3	4	5	6	7	8	9	10	R	R	R	R	R	R	R	1	2	3	4	___
___	1	2	3								L	L	L	L	L	L	L	1	2	3	4	___
___	1	2	3	4	5	6	7	8	9	10	R	R	R	R	R	R	R	1	2	3	4	___
___	1	2	3								L	L	L	L	L	L	L	1	2	3	4	___
___	1	2	3	4	5	6	7	8	9	10	R	R	R	R	R	R	R	1	2	3	4	___
___	1	2	3								L	L	L	L	L	L	L	1	2	3	4	___
___	1	2	3	4	5	6	7	8	9	10	R	R	R	R	R	R	R	1	2	3	4	___
___	1	2	3								L	L	L	L	L	L	L	1	2	3	4	___
___	1	2	3	4	5	6	7	8	9	10	R	R	R	R	R	R	R	1	2	3	4	___
___	1	2	3								L	L	L	L	L	L	L	1	2	3	4	___

											TOTAL	SES %
STAGE												
TRIALS												
CORRECT												
INCOR												
NR												

THE CHILDREN'S UNITS

Dr. Raymond G. Romanczyk
Department of Psychology

CHILD: ____________________

STAFF: ____________________

DATE: ____________________

On the following scales, please indicate ratings that are generally representative of this child's behavior on these characteristics. Please make your ratings independently and do not discuss them with other staff.

Is this child noncompliant to task instructions and task demands?

1	2	3	4	5	6	7
Almost never			About 50% of the time			Almost Always

Is this child noncompliant to direction and demands on behavior?

1	2	3	4	5	6	7
Almost Never			About 50% of the time			Almost Always