Order Number 9325113

Self-reinforcement and persons with developmental disabilities

Kaplan, Howard Abraham, Ph.D.

City University of New York, 1993

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

SELF-REINFORCEMENT AND PERSONS WITH DEVELOPMENTAL DISABILITIES

by

Howard Kaplan

A dissertation submitted to the Graduate Faculty in Psychology in partial fulfillment of the requirements for the degree of Doctor of Psychology, The City University of New York

1993

Approval Page

This manuscript has been read and accepted by the Graduate Faculty in Psychology in satisfaction of the dissertation requirement for the degree of Doctor of Psychology.

4/19/93
Date

Nancy L. Kenmes
Chair of Examining Committee

4/27/93
Date

Herbert O. Saltzstein
Executive Officer

Dr. Claire Poulson

Dr. Thom Verhave

Dr. John O'Neil

Dr. David Coleman

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

ABSTRACT

Self-Reinforcement and Persons with Developmental Disabilities

by

Howard Kaplan

Adviser: Professor Nancy Hemmes

The study attempted to dissociate the effects of self-reinforcement from those of self-monitoring on the level of on-task behavior of employees in supported work settings. The subjects were developmentally disabled men employed as cafeteria workers in two supported work enclaves. Prior to treatment, all workers required a high level of prompting to remain on task, defined as performing one of the tasks depicted in a set of photographs of each worker performing all aspects of his job. In a multiple baseline design, three workers were trained to reliably self-monitor and self-reinforce their on-task behavior every 30 minutes at the sound of a timer beep. In a second treatment phase, they self-monitored but did not self-reinforce. Response contingent praise and a reinforcement contingency for accurately self-monitoring were programmed during both treatment phases. Self-monitoring with self-reinforcement produced an increase in level of on-task behavior which decreased when self-reinforcement was discontinued. Order

of treatment phases was reversed for four additional workers. For three of these workers, level of on-task behavior was higher when they self-reinforced. These results indicate that self-reinforcement exerts behavioral control beyond that attributable to self-monitoring and concomitant social reinforcement.

KEY WORDS: Self-Reinforcement, Supported Employment, Developmental Disabilities, Self-Management

TABLE OF CONTENTS

LIST OF FIGURES/TABLES

INTRODUCTION

The purpose of the current study was to dissociate behavioral control by self-reinforcement from that attributable to self-monitoring and conditioned social reinforcement. Self-reinforcement is defined after Mahoney and Bandura (1972) as the self-delivery of reinforcement contingent upon emission of a response meeting a pre-established criterion. Reinforcers must be continuously available to the individual, but they must be delivered only when a criterion response has been emitted.

A considerable body of literature indicates that self-monitoring coupled with self-reinforcement does, in fact, exert a potent effect on establishing, maintaining or increasing the frequency of a target response (Agran, Fodor-Davis, Moore, and Deer, 1989; Albion and Salzberg, 1982; Bellamy, 1975; Bolstad and Johnson, 1972; Burgio, Whitman and Johnson, 1980; Drabman, Spitalnick and O'Leary, 1980; Glynn, 1970; Hughes and Peterson, 1989; McNally, Kompik and Sherman, 1984; Ninnes, Fuerst, Rutherford, and Glenn, 1991; Sowers, Verdi, Bourbeau, and Sheehan, 1985). Nevertheless, there is debate on the mechanism responsible for these observed effects. Whereas social learning theorists and others (Bandura, 1977a&b, 1978, 1981; Mahoney and Bandura, 1972; Thoresen and Wilbur, 1976; Karoly, 1982a) suggest that self-monitoring and contingent self-delivery of reinforcers are the agents of observed changes in behavior, Catania

(1975) and Goldiamond (1976) suggest that these effects are artifacts of the process of social standard setting with its accompanying opportunities for social reinforcement. Catania (1975) argued that circumstances generating the behavior of self-reinforcement (i.e. the reinforcement of one's own response) make it impossible to distinguish behavioral control by self-reinforcement from control attributable to other aspects of the process, such as identification and setting of goals; establishment of the target behavior as a valuable event (conditioned social reinforcer); and learning to discriminate among aspects of one's own behavior.

The implications of Catania's criticism are far reaching. In "The Myth of Self-Reinforcement" (1975) he argued that the factors contributing to behavioral control by self-reinforcement are difficult, if not logically impossible, to distinguish. Review of the empirical literature seems to support this conclusion. Problems include the use of experimenter-identified (rather than subject-identified) target responses (Lovitt and Curtiss, 1969; Glynn, 1970); the presence of an observer who may exert antecedent control over the subject's behavior (Ninnes et al. 1989); the failure to establish self-delivered reinforcers as functional reinforcers, along with the confounding of self-reinforcement with external (social) reinforcement (Nelson, Hayes, Spong, Jarrett and McKnight,

1983). There is debate on the basic question of whether personal evaluative standards can be independent of the behavior being evaluated (Goldiamond, 1976; Thoresen and Wilbur, 1976).

Hayes, Rosenfarb, Wulfert, Munt, Korn, and Zettle (1985) attempted to avoid these methodological difficulties in a study designed to separate the effects of self-delivery of reinforcers from those of public standard setting. The first variable they manipulated was whether treatment took place in a public or private context. The second variable was whether there was goal setting with or without the self-delivery of reinforcement. No performance differences between groups were attributable to contingent self-reinforcement; however, there was an effect of public as opposed to private goal setting. Hayes et al. (1985) concluded that "self-reinforcement procedures work by setting a socially available standard against which performance can be evaluated, which then functions as a discriminative stimulus for stringent or lenient social contingencies" (p. 214). This conclusion must be tempered, owing to methodological and statistical considerations. No proof is offered that the public/private ruse used by the authors actually worked, and no attempt was made to ensure the reliability of the independent variable (i.e., that reinforcers were consumed contiguously and in proportion with target behavior). The conclusion that there was no

effect of self-reinforcement per se is subject to threats to statistical conclusion validity because there may not have been enough power in the design to discern a difference where one might really exist.

The present study attempted to avoid the methodological problem of the reactive nature of public and private goal setting by holding this variable constant, while manipulating the self-delivery of reinforcers. For this purpose, a single subject as opposed to between-group design was used. By conducting this research with individuals already working for money, (i.e., using an established potent reinforcer), we hoped to avoid questions regarding the establishment of money as a functional reinforcer, while studying self-reinforcement in a "clinically relevant" environment (Hayes et al. 1985). This study utilized target responses that were selected by the subjects themselves, and the observers who gathered data were a natural part of the work environment and were therefore as unobtrusive as possible.

METHOD

Subjects And Setting

The subjects were seven developmentally disabled males between the ages of 21 and 30 working at two Supported Work Enclaves operated by the Association for Children with Retarded Mental Development (ACRMD). A Supported Work enclave is defined by Public Law 98-527 (the Developmental Disabilities Act of 1984) as "a location where disabled individuals work for at least twenty hours per week, ... are compensated in accordance with the Fair Labor Standards Act, ...the opportunity to interact with non-handicapped co-workers exists and ...the individual is constantly supervised as part of a work crew no larger than eight individuals."

Site 1 was an outdoor cafeteria at a large public zoo. Site 2 was an indoor college cafeteria. All seven of these individuals were identified by their on-site supervisors as exhibiting a level of on-task behavior below that necessary to independently perform their assigned tasks. They required frequent prompts to attend to task and were therefore in danger of loosing their jobs due to their intense need for supervision.

The seven subjects were all male, between the ages of 21 and 30. All seven subjects were able to travel to and from their work sites independently after a period of travel training ranging from one to five days, and had been

employed at these sites for at least several weeks prior to the beginning of the experiment. All had attended Sheltered Workshops operated by ACRMD for various lengths of time, ranging from a few months to several years. In an agreement with the employers, ACRMD provided on-site supervisors at each location who were directly responsible for eight trainees at each site. All seven were highly motivated to succeed on their jobs, as evidenced by their excellent attendance and punctuality records. All had sought help with their on task behavior deficits from their supervisors. Informed consent was obtained for all participants in the study.

Materials

A countdown pillbox timer (Micronta, Catalogue # 63-881) was used to time 30 minute intervals. It was equipped with a compartment large enough to hold the eight reinforcers a subject could earn on any given day. This timer was attached to a clip that subjects wore on their shirt collars or on their belts, and resembled a telephone paging beeper in size, appearance and sound. The subjects readily wore their timers possibly because they were similar in appearance and operation to the telephone paging beepers worn by their counselors.

Design

A multiple baseline design across subjects was used at two sites. After a baseline period during which the operant

levels of on-task behavior and productivity were determined for each subject, three of the four subjects at the first site were taught to a self-monitor and self-reinforce. During the next phase of the experiment, the same three subjects self-monitored but did not self-reinforce on-task behavior. This sequence of treatment phases was reversed for the fourth subject at this first site: He was initially required to self-monitor his level of on-task behavior and then, in the next phase, to both self-monitor and self-reinforce on-task behavior. A fifth non-disabled control subject did not self-monitor or self-reinforce, although data on his level of on-task behavior and productivity were collected in the same manner by the on-site trainer. At the second site, all three subjects first self-monitored their level of on-task behavior, then, in a subsequent phase, self-monitored and self-reinforced for on-task behavior. During a final follow-up phase at the second site, on-task behavior and productivity were measured after the treatment procedure had ended. During this phase, subjects continued to carry their timers, but placed their own selection of reinforcers inside. The contingency for accuracy of self-report was not in operation at this time.

Although an attempt was made to ensure that all phase changes at both sites occurred when responding achieved stability, as measured by a visual inspection of graphed data, the time limited nature of the subjects' stay at each

of the job sites prevented baseline measures from being gathered for a sufficient length of time to ensure that stability had been achieved. Ideally, baselines for Mike, Stu and Walt at the first site and for all three subjects at the second site should have been extended for several additional daily sessions in order to ensure that this was the case. Additionally, the second phase change for Jim at the second site should also have been extended to ensure that the last data point in that phase (self-monitoring alone) was not just a temporary aberration but rather a real drop in the level of on-task behavior. Due to the need to complete all experimental phases before employment for subjects ended at these sites, some compromises were made. The need for improvement in subjects' level of on-task behavior in order to prevent them from losing their jobs was also a consideration.

Dependent Measures

A task analysis that was uniform across the two sites was developed for all seven subjects. It consisted of seven basic tasks: sweeping floors; mopping floors; clearing tables of debris; wiping tables and seats with a damp rag; wiping trays; changing garbage bags and disposing of full garbage bags. In a manner similar to Sowers et al. (1984), photographs were taken of each subject performing each of these seven tasks prior to the collection of baseline data. The resulting seven pictures were assembled into a small

billfold that was kept by subjects in their back pockets. These pictures constituted an instruction manual for the subjects, listing each of the diverse tasks they had to perform in a manner that did not require any reading ability. This performance aid was carried by the subjects throughout the study and currently continues to be used by them for reference purposes. Whenever they asked their supervisor what they should have been doing next, they were prompted to refer to their pictures and select a task that had not been performed recently.

On-Task Behavior. The primary dependent measure was the percentage of observations during which a subject was observed to be on task. For the purpose of this study, on-task behavior was operationally defined as the subject performing one of the tasks depicted in the picture prompts during any portion of a five second observation.

Productivity. The secondary dependent measure was the supervisor's rating of each subject's productivity, using a five point Likert type scale of polar adjectives: 1 = poor; 2 = slightly below acceptable; 3 = acceptable; 4 = slightly better than acceptable; 5 = very good. These data were collected on a daily basis by the on-site supervisor for each subject, as well as on a weekly basis by the security guard at the second site (a college student cafeteria), who was blind to the purpose and procedures of the study.

Data Collection Procedure

A momentary time sampling strategy was used to collect data, with the optimal time sampling interval being empirically determined in the manner described by Sanson-Fisher, Poole and Dunn (1980). Estimates of level of responding were taken using intervals of 15 seconds, 30 seconds, 1, 5, 10, 20, 30 and 60 minutes. These estimates were correlated with the same measures taken during continuous recording of on-task behavior. Pearson product-moment correlations with the continuous measure were .97, .93, .91, .90, .88, .88, .86 and .68 respectively. The rapid decrease after .86 at 30 minutes and the small increment to be gained by cutting the size of the interval in half (.86 to .88 at 15 minutes) suggested that the 30 minute interval be selected as a good compromise between efficiency and accuracy.

Data were collected by the on-site supervisor at each site, who was responsible for supervising eight trainees. Their daily workload required them to move about almost constantly, providing a combination of prompts and instruction, as well as actually performing some of the required tasks themselves. Both sites were laid out so that all trainees would be within visual range from any vantage point, so that no more than a brief glance was required to determine if trainees were working on their assigned tasks. At the end of each 30 minute interval, supervisors recorded, with a coded mark, whether the subject was on or off task at

that time, or whether the subject was unavailable for observation.

Measurement of Procedural Reliability

A measure was taken on the extent to which each subject actually carried out the self-monitoring and self-reinforcement procedures as required by the phase of the experiment. All observers noted whether the subject appropriately responded to each timer signal. This is operationally defined as reaching for the timing device after it sounded, opening the compartment, taking out a coin/washer if he was on-task or not removing one if he was off-task and then resetting the timer and replacing it on his belt or shirt collar. Appropriate responding was scored only if the entire process was correctly performed.

Experimental Procedures

Baseline. After picture prompts were issued to each subject, on-task and productivity data were collected for each subject by his supervisor. No other experimental procedures were in effect but subjects did receive feedback and contingent social reinforcement (praise) from supervisors at this time.

Training in Self-Monitoring. After baseline data were collected, each subject was trained to recognize whether he was on or off task using the operational definition described above. This was accomplished by having the supervisor and subject simultaneously rate the subject, with

feedback being provided, until a criterion level of interobserver agreement of .85 or better (Cohen's kappa) was achieved.

At this point, subjects were trained to use the timer and washer data collection system. The subject was given his timer filled with steel washers. The timer was set to a 30 minute fixed interval. When the timer sounded, the subject was to determine whether he had been on-task immediately prior to the signal. If he had been on-task, he was to take a washer from the timer and place it in his pocket. No washer was to be taken if he was off-task. The subject then reset the timer by pushing the right hand button three times, replaced the timer in its spot, and returned to work. If he had been off-task, the subject was instructed to refer to his set of picture prompts to determine what task he should be performing. Training continued until agreement between the trainer and subject reached .85 or better (Cohen's kappa).

Self-Monitoring Only. During the self-monitoring only phase, subjects were given their timers filled with 8 steel washers each morning. Timers for all subjects were synchronized to go off at the same time, while that of the supervisor was set to go off 30 seconds prior to that of the subjects. The 30 second figure was established during pilot testing of the procedure to be of sufficient length to ensure that synchronization continued throughout all eight

intervals, while allowing for some delay in subjects resetting of their timers. Subjects were required to self-monitor as described in the previous paragraph. The washers were returned to the supervisors at the end of each day, who would provide feedback and social reinforcement to each subject based on his performance.

Self-Monitoring With Self-Reinforcement. During this phase, each subject had nickels rather than steel washers in his timer. He was permitted to keep any nickels earned during this phase. Otherwise, this condition was identical to self-monitoring alone.

A contingency was added to the experimental manipulation in order to increase and maintain the accuracy of each subject's self-ratings. At the end of each day, supervisor's records of on-task behavior were compared with each subject's self-ratings (as measured by the number of nickels/steel washers they self-administered). If the two records differed by no more than one on-task observation, the subject received a supplementary reward of fifty cents. Under this contingency, subjects could earn more by accurately reporting their level of on-task behavior (50 cents plus 5 cents for any interval during which they were on task) than they could by just taking all of the available reinforcers (40 cents, or 8 intervals times 5 cents per interval). This contingency was in effect for both the Self-Monitoring with Self-Reinforcement phase as well as the

Self-Monitoring only phase, but not for the maintenance phase at Site 2.

Maintenance. At the end of the final experimental phase, a counseling session was held with each of the subjects. They were given their timing devices and encouraged to continue using them. A discussion was held with each subject about what they would like to continue putting in their timers as rewards. Three weeks later at only the college cafeteria (site 2), data were collected in the same manner as above regarding the maintenance of high levels of on-task behavior and productivity. The contingency for accuracy was not in effect at this time.

Interobserver Agreement

On-Task Measure. Interobserver agreement was measured during 10% of all trials during each phase of the experiment. The procedure consisted of this experimenter observing each subject's performance simultaneously along with the on-site supervisor and recording these data on a separate data sheet without providing any feedback to the on-site supervisor. Because only three response categories existed, the possibility of chance agreement was quite high, tending to inflate the resulting measure of interobserver agreement. Cohen's kappa was used to adjust the calculated level of interobserver agreement in these data. Kappa is computed by the following formula: $k = (P1-P2)/(1-P2)$ where P1 = the proportion of agreements between observers on

occurrences and nonoccurrences (or agreements on occurrences and nonoccurrences divided by the total number of agreements and disagreements) and P2 = the proportion of expected agreements on the basis of chance. P2 is computed by multiplying the number of occurrences for observer number 1 times the number of occurrences for observer number 2 plus the number of nonoccurrences for observer 1 times the number of nonoccurrences for observer number 2. The sum of these is divided by the total number of intervals squared (Kazdin, 1982, p. 66).

Observer Training. Before collecting baseline data, the first author and each of the on-site supervisors simultaneously observed and rated each subject's level of on-task behavior using the above described operational definitions until a criterion level of .85 agreement or better (Cohen's kappa) was achieved.

RESULTS

On-Task Behavior

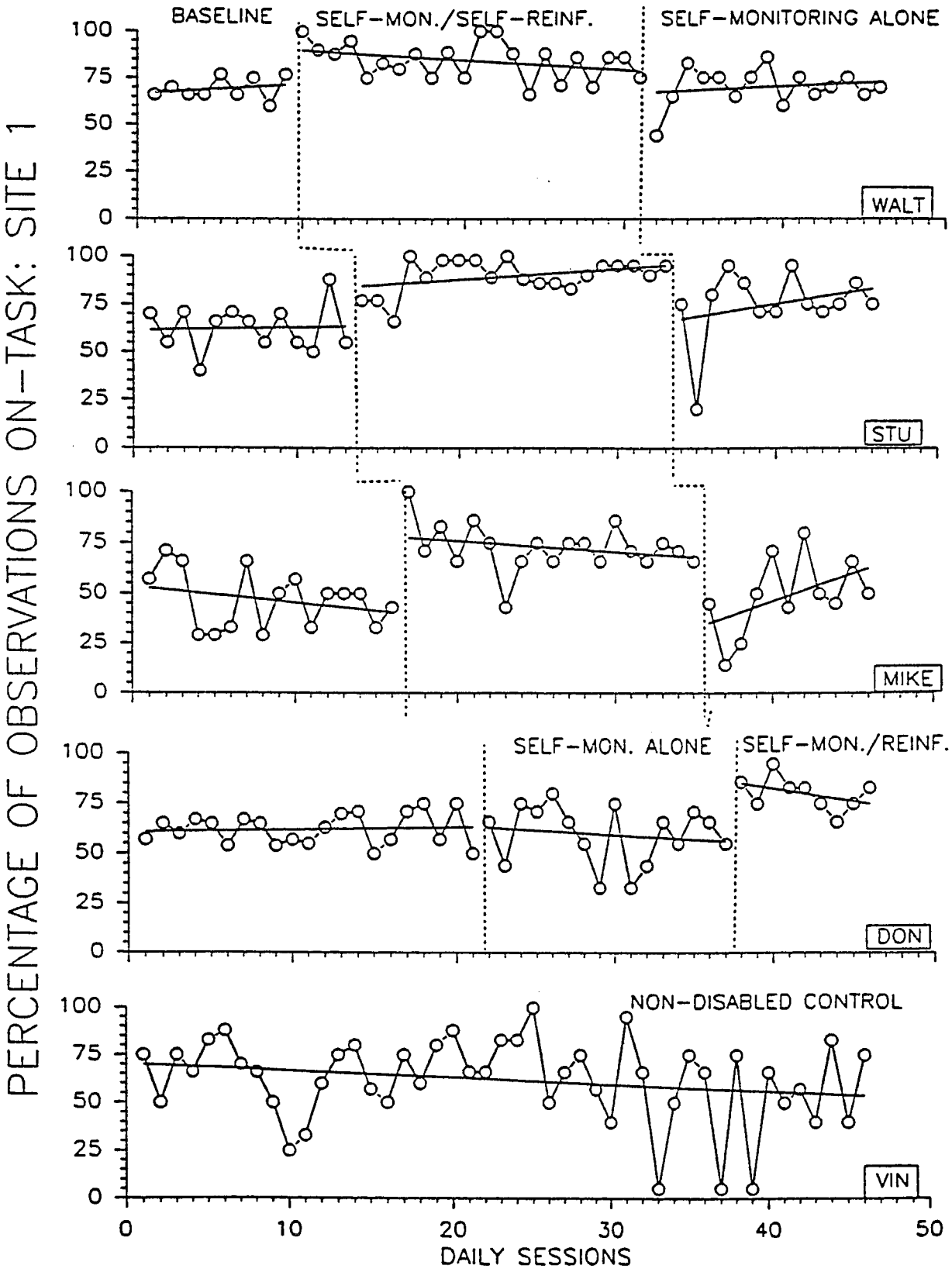
Figures 1 and 2 show the percentage of time on-task at sites 1 and 2 respectively.

Visual inspection of Figure 1 shows that the percentage of observations scored as on-task for Walt, Stu and Mike improved with the introduction of treatment from baseline levels below 75% to up to 100% during the self-monitoring/self-reinforcement phase. The validity of this apparent effect would have been strengthened if subjects' initial baselines were extended for several more daily sessions. However, as previously noted, this was not possible due to the limited amount of time available to carry out all planned phase changes, as well as the need to take immediate action to preserve their jobs. Subsequently (with the same qualifications as above), there was a decrease in the level on-task performance between self-monitoring with self-reinforcement and self-monitoring alone for all three subjects. Don served as a partial control for possible order effects between self-monitoring/self-reinforcement and self-monitoring alone. The percentage of observations in which on-task behavior was observed for Don did not increase above baseline levels with self-monitoring alone, but rather increased only following the introduction of the self-monitoring/ self-reinforcement procedure. Trend lines have been added to the figures to assist with

FIGURE CAPTION FOR FIGURE 1

Figure 1. Daily percentage of momentary time sampling intervals during which subjects' on-task behavior was observed at site 1. Trend lines are computed using the ordinary least squares solution.

FIGURE 1


BASELINE
SELF-MON./SELF-REINF.
SELF-MONITORING ALONE
WALT
STU
MIKE
SELF-MON. ALONE
SELF-MON./REINF.
DON
NON-DISABLED CONTROL
VIN
PERCENTAGE OF OBSERVATIONS ON-TASK: SITE 1
DAILY SESSIONS
0
10
20
30
40
50
25
75
100

between-phase comparisons. The equations for these lines have been calculated using the ordinary least squares solution (Edwards, 1976). This formula is provided at the bottom of Table 1.

Table 1 displays several parameters the for trend lines drawn in Figure 1. These parameters are: the range of data within each phase, the slope of the regression line, its adjusted Y intercept (see figure 2 for explanation of adjustment), the mean for each phase and the standard deviation for each phase.

Figures for the standard deviation for each phase are given as a measure of the variability of the data, but no test for significance was performed on the data due to the observation by Cook and Campbell (1979, p. 234-235) that although the regression lines calculated by the ordinary least squares solution are not biased when data is serially correlated, the estimates of standard deviations, and therefore of significance, are biased in time series experiments. This is because the ordinary least squares solution assumes independence of error terms, an assumption which is violated in time series experiments. As a result, standard deviations are given for rough comparative purposes only. Each phase does not contain a sufficient number of data points to estimate this degree of serial correlation and remove it statistically. Usually, 50 data points or more per phase are required for this purpose.

FIGURE CAPTION FOR TABLE 1

Table 1. Trend line parameters in Figure 1.

TABLE 1

VALUES FOR EQUATIONS OF LINES PLOTTED THROUGH GRAPHED DATA IN FIGURE 1

SUBJ	PHASE	RANGE	SLOPE	Y INT	MEAN	SD
WALT	BASELINE	60 TO 77	.53	66.6	69.2	5.9
	SELF-MON/REINF	70 TO 100	-.54	90.5	84.3	9.9
	SELF-MON ALONE	44 TO 86	.38	67.1	70.0	10.0
STU	BASELINE	40 TO 88	.15	61.4	62.5	12.3
	SELF-MON/REINF	66 TO 100	.56	83.9	89.8	8.9
	SELF-MON ALONE	20 TO 95	1.30	65.8	75.0	18.6
MIKE	BASELINE	29 TO 66	-.81	53.5	46.6	14.4
	SELF-MON/REINF	43 TO 100	-.54	78.1	72.7	11.5
	SELF-MON ALONE	14 TO 80	2.70	32.5	49.0	18.9
DON	BASELINE	50 TO 75	.12	60.8	62.1	7.8
*	SELF-MON ALONE	44 TO 80	-.44	63.4	59.7	14.8
	SELF-MON/REINF	66 TO 95	-1.30	86.6	80.1	8.4
VIN	NON-DISABLED	5 TO 100	-.36	70.3	61.8	22.2

* Note reversed sequence of treatment conditions

All Y Intercepts are adjusted so that the initial X value for their coordinates = 1 and all subsequent X values are decreased by the initial value in the series. This is done to minimize distortion to the value for the Y Intercept resulting from the curve being displaced a distance from the origin.

The formula for the ordinary least squares solution (Edwards, 1976) for the trend lines is

$$\sqrt{\sum (Y-\hat{Y})^2}$$

where Y=the observed Y values for each data pair within the phase and

$$\hat{Y}=$$

the values predicted for Y from the equation of the trend line for each data pair, (ie., the square root of the sum of the squares of the deviations of the predicted Y values from the actual Y values for each data pair within each phase.

For Walt, all parameters change across the first phase change. The range increased, the slope became negative, the Y intercept jumped by 24 units, the mean increased by 15 units and the standard deviation increased by 4 units. For the second phase change, all parameters reverted to approximately their initial levels except for the slope, which again became positive.

Stu's data followed a similar pattern, but the slope increased across both the first and second phase changes and the standard deviation increased substantially across the second phase change, indicating an increase in the variability of the data during the self-monitoring alone phase.

For Mike, the slope increased markedly during the phase change from self-monitoring/reinforcement to self-monitoring alone while the standard deviation again jumped during the self-monitoring alone phase.

For Don, who received the reversed sequence of treatment conditions, the slope was again much more negative during the self-monitoring/reinforcement phase than it was during the self-monitoring alone phase and the standard deviation was also elevated during the self-monitoring alone phase. The other parameters behaved in a manner consistent with the observation that effect on the dependent measure, (i.e., on-task behavior) increased when self-monitoring was combined with self-reinforcement more than when self-

monitoring alone was performed.

Vin's data provide a partial control for maturation effects in addition to those provided by the multiple baseline design. They also provide comparison data for non-disabled control subjects.

A more complete control for order effects between self-monitoring /self-reinforcement and self-monitoring alone was provided by changing the order of treatments across experimental sites.

Figure 2 displays the measure of on-task behavior for Jan, Chet and Jim. For Jan, the percentage of observations in which he was observed to be on-task during baseline fell generally below 80% with one initial outlier. The negative slope for this phase was repeated during the next phase in which he self-monitored only, and the range remained essentially the same. After the introduction of self-monitoring/self-reinforcement, there was a positive change in both level and slope, which continued into the maintenance phase.

Table 2 shows that the parameters did not vary substantially between baseline and self-monitoring alone for Jan. For the next phase change, the slope changed from a substantial negative one to a level slope and the variability, as measured by the standard deviation changes from 16.2 to 9.6. The mean also increased from 71.4 to 82.9.

FIGURE CAPTION FOR FIGURE 2

Figure 2. Daily percentage of momentary time sampling intervals during which subjects' on-task behavior was observed at site 2. Trend lines are computed using the ordinary least squares solution.

FIGURE 2

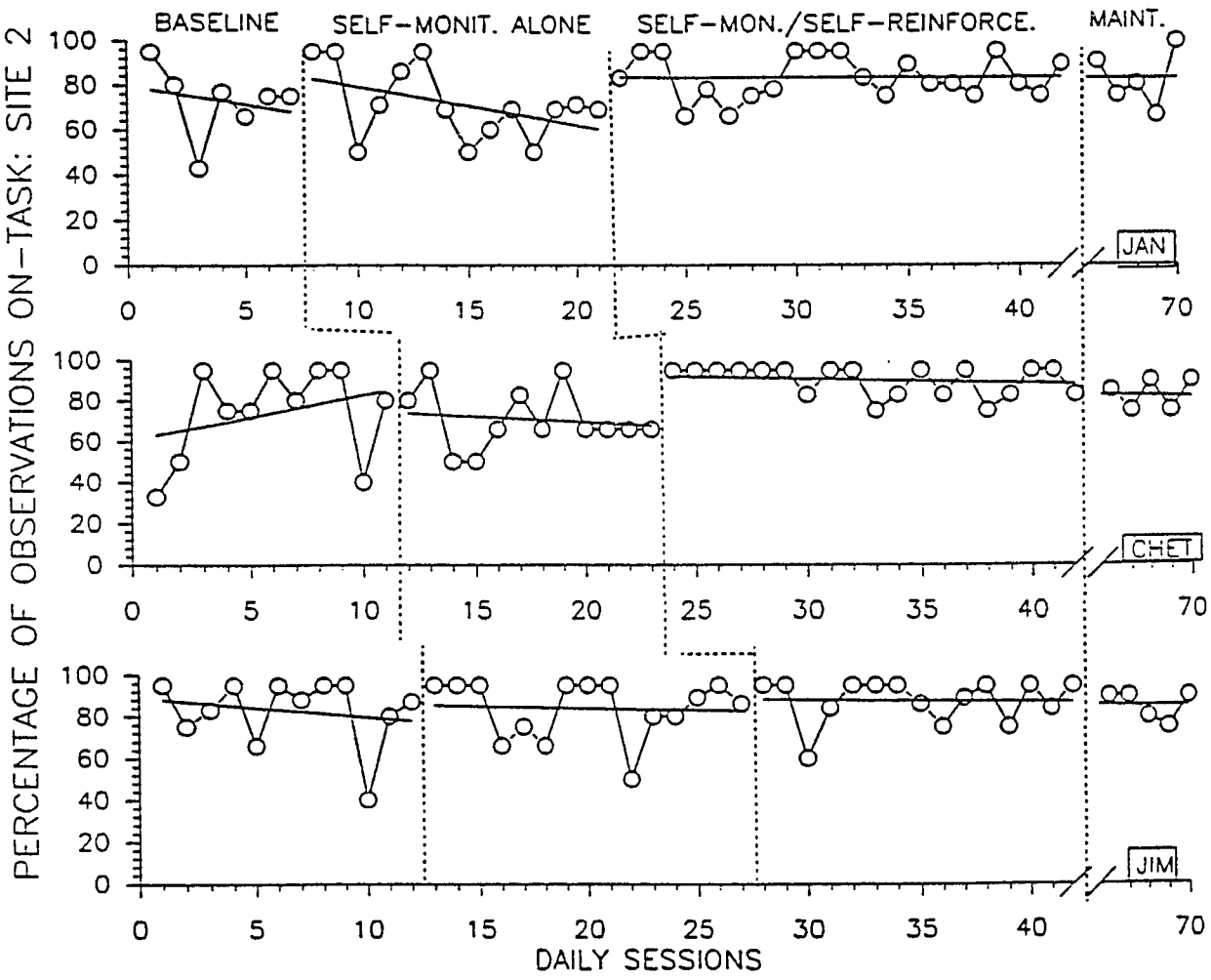

BASELINE
SELF-MONIT. ALONE
SELF-MON./SELF-REINFORCE.
MAINT.
PERCENTAGE OF OBSERVATIONS ON-TASK: SITE 2
JAN
CHET
JIM
DAILY SESSIONS

FIGURE CAPTION FOR TABLE 2

Table 2. Trend line parameters in Figure 3.

TABLE 2

VALUES FOR EQUATIONS OF LINES PLOTED THROUGH GRAPHED DATA IN FIGURE 2

SUBJ	PHASE	RANGE	SLOPE	Y INT	MEAN	SD
JAN	BASELINE	43 TO 95	-1.68	79.7	73.0	15.8
	SELF-MON ALONE	50 TO 95	-1.73	84.3	71.4	16.2
	SELF-MON/REINF	66 TO 95	.02	82.3	82.9	9.6
	MAINTENANCE	66 TO 99	.09	79.3	82.0	12.9
CHET	BASELINE	33 TO 95	2.18	60.8	73.9	22.9
	SELF-MON ALONE	50 TO 95	-.54	74.2	70.8	14.8
	SELF-MON/REINF	75 TO 95	-.51	94.9	89.7	7.4
	MAINTENANCE	75 TO 90	1.0	80.0	83.0	7.6
JIM	BASELINE	40 TO 95	-.92	88.8	82.8	16.4
	SELF-MON ALONE	50 TO 95	-.19	85.3	85.8	14.1
	SELF-MON/REINF	60 TO 95	-.15	86.4	87.5	10.5
	MAINTENANCE	75 TO 90	-1.5	89.5	85.0	7.1

All Y Intercepts other than those for baseline are adjusted so that the initial X value for their coordinates = 1 and all subsequent X values are decreased by the initial value in the series. This is done to minimize distortion to the value for the Y Intercept resulting from the curve being displaced a distance from the origin

For Chet, the slope changed from 2.18 to -.54 between baseline and self-monitoring alone phases, along with a substantial decrease in the standard deviation (14.8 to 7.4). Although there was little change in the percentage of sessions in which Chet was observed to be on-task across baseline to self-monitoring alone phase, the introduction of the self-monitoring/self-reinforcement procedure was accompanied by an increase to over 80% with a burst of 6 consecutive days over 90%. This increase for these two subjects after the introduction of the self-monitoring/self-reinforcement procedure was similar to that of the previous four subjects at the first site.

Jim, the third subject at the second site did not demonstrate any increase across phases of the experiment but rather remained above 80 %, although there was some decrease in variability with the introduction of the self-monitoring/self-reinforcement procedure. (The standard deviation decreased from 16.4 during baseline to 7.1 during maintenance.)

Productivity

Figures 3 and 4 present the secondary dependent variable, subjects' productivity, as measured by supervisor's performance ratings on a five point Likert scale of polar adjectives (where 1 = poor, 2 = slightly below acceptable, 3 = acceptable, 4 = slightly above acceptable, and 5 = very good), followed the same pattern as

FIGURE CAPTION FOR FIGURE 3

Figure 3. Daily supervisory performance ratings of subjects' productivity at site 1. (1=poor, 2=below acceptable, 3=acceptable, 4=slightly above acceptable, 5=very good). Trend lines are computed using the ordinary least squares solution.

FIGURE 3

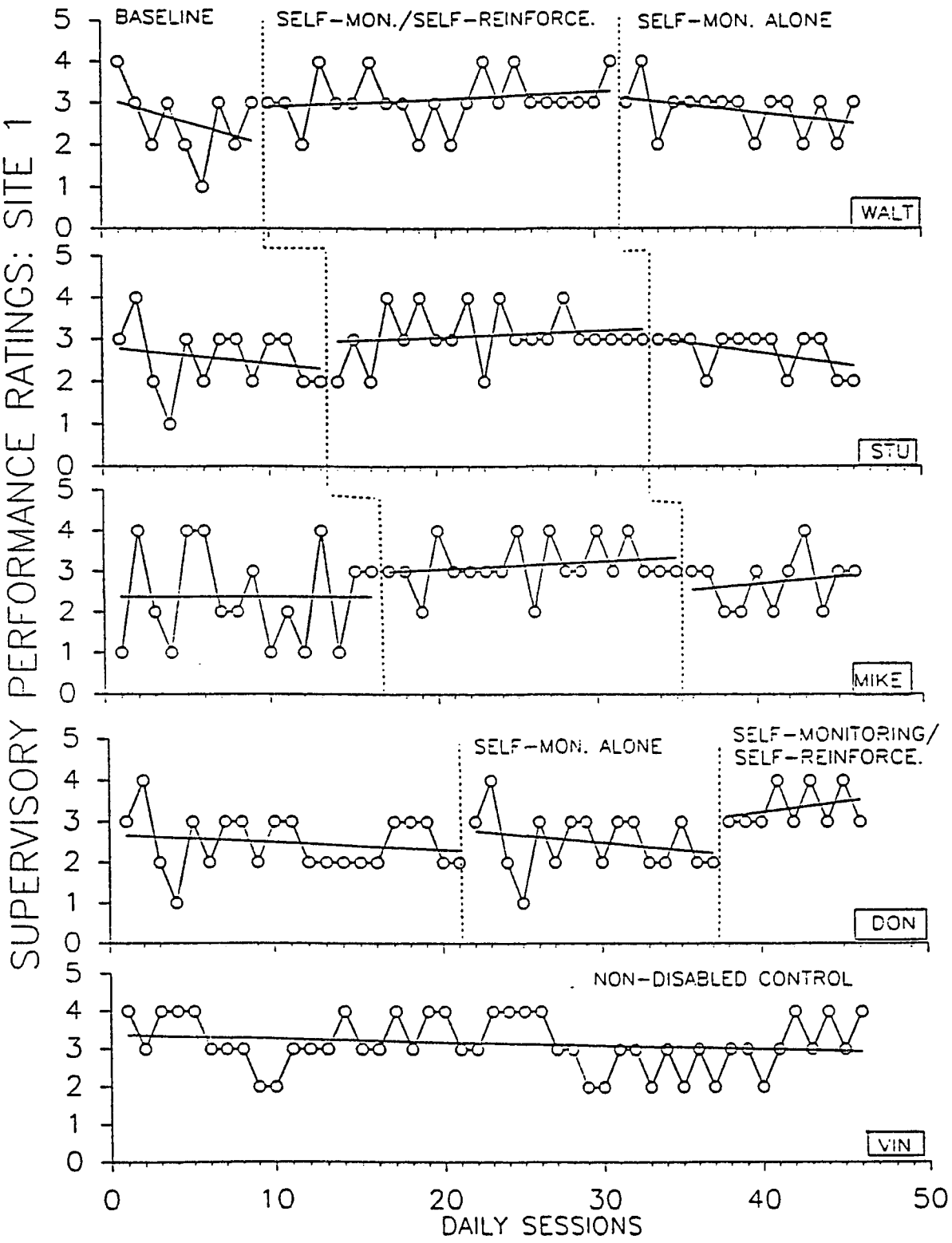
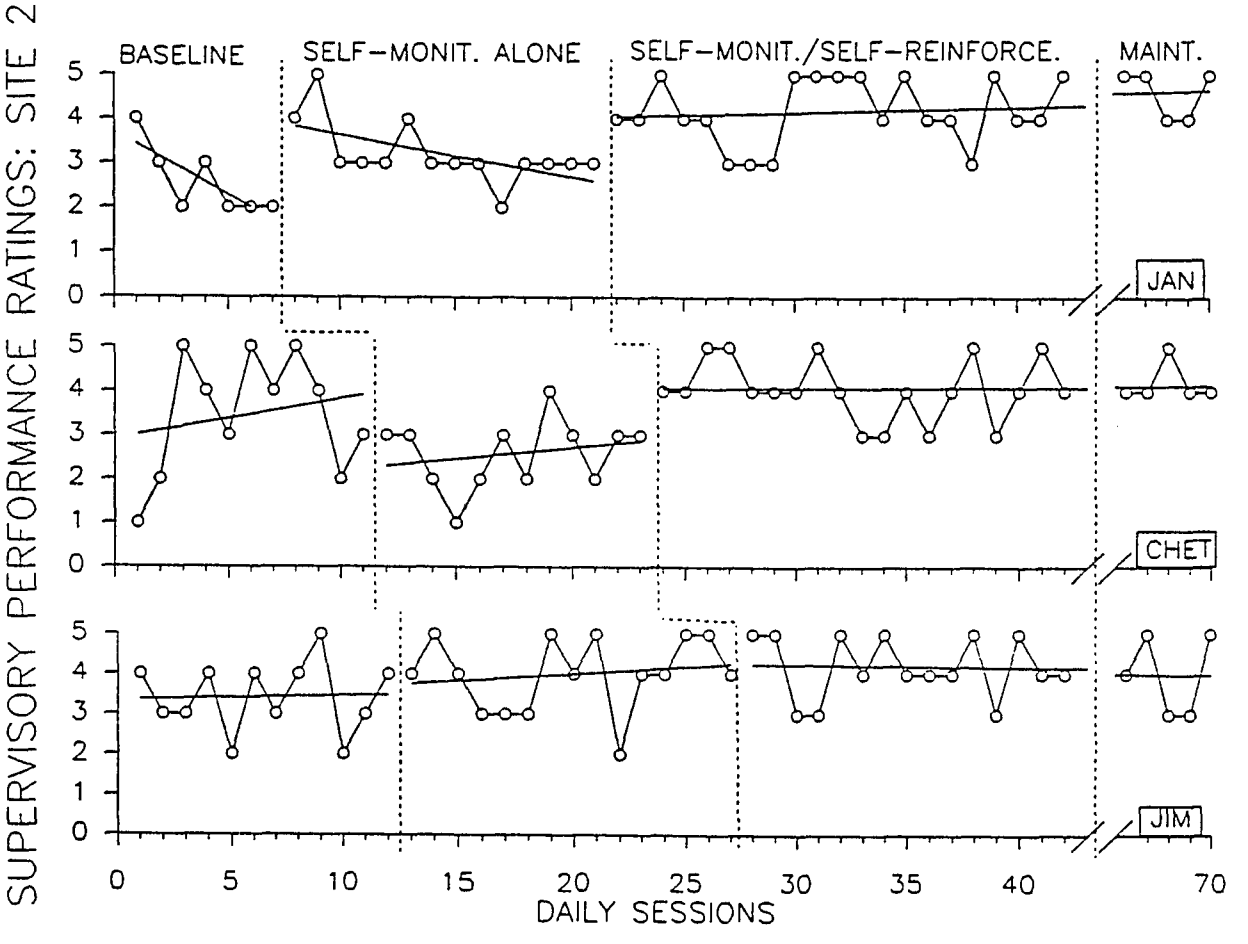

BASELINE
SELF-MON./SELF-REINFORCE.
SELF-MON. ALONE
WALT
STU
MIKE
SELF-MON. ALONE
SELF-MONITORING/
SELF-REINFORCE.
DON
NON-DISABLED CONTROL
VIN
SUPERVISORY PERFORMANCE RATINGS: SITE 1
DAILY SESSIONS
0
10
20
30
40
50

FIGURE CAPTION FOR FIGURE 4

Figure 4. Daily supervisory performance ratings of subjects' productivity at site 2. (1=poor, 2=slightly below acceptable, 3=acceptable, 4=slightly above acceptable, 5=very good). Trend lines are added using the ordinary least squares solution.

FIGURE 4


BASELINE
SELF-MONIT. ALONE
SELF-MONIT./SELF-REINFORCE.
MAINT.
JAN
CHET
JIM
SUPERVISORY PERFORMANCE RATINGS: SITE 2
DAILY SESSIONS
0
5
10
15
20
25
30
35
40
70

the on-task behavior data presented in Figures 1 and 2, although the effect was less pronounced.

Interobserver Agreement

The level of interobserver agreement between the first author and the site supervisors on the on-task variable was measured during ten percent of all daily sessions at randomly selected times during the day during each phase of the experiment. At the first site the level of agreement was .86 during baseline, .88 during self-monitoring/self-reinforcement and .88 during self-monitoring alone. At the second site, the correlations were .90 during baseline, .94 during self-monitoring/self-reinforcement, .95 during self-monitoring alone and .93 during maintenance. Interobserver agreement was calculated using Cohen's kappa to adjust for chance levels of agreement.

The mean level of interobserver agreement between subjects and their supervisors on the on-task measure exceeded .85 (Cohen's kappa) during all phases at both sites, and no retraining was required after initial training.

Reliability on the productivity variable was measured at the second site by comparing the supervisor's ratings of subjects' work performance on a five point Likert type scale with that of a Security Guard at that site. The correlation obtained was .92 (Pearson Product-Moment, $p \leq .01$). This level of reliability was calculated across phases for all

three subjects at this Site.

Procedural Reliability

Procedural reliability of the independent variable, as measured by the percentage of trials in which subjects either self-monitored and self-reinforced (Self-Monitoring with Self-Reinforcement phase) or self-monitored only (Self-Monitoring phase), exceeded .80 for all subjects during all phases. Figures 5 and 6 show that the first two or three data points after baseline for each subject fall at about the .85 level and therefore represent the maximum discrepancy obtained between supervisor and subject on the dependent variable. After the initial phase change, interobserver agreement increased to close to 1.00 for all subjects other than Mike at the first site. For Mike, the level of interobserver agreement, although remaining above .85, fluctuated between .85 and .90 throughout the remaining daily sessions.

Consumer Satisfaction

At the conclusion of the study at each site, all subjects indicated (via their verbal responses to questions posed by their job coaches) that the combination of picture prompts and auditory cues improved their on-task as well as general performance on the job, and that they preferred the Self-Monitoring with Self-Reinforcement phase to Self-Monitoring alone. During the maintenance phase at site 2, which followed the end of the self-monitoring with self-

reinforcement phase by three weeks, all three subjects were using their timers and filling them with their own selection of reinforcers 100% of the time. (See Figures 5 and 6).

FIGURE CAPTION FOR FIGURE 5

Figure 5. Site 1: Percentage of trainee and trainer agreement and percentage of sessions self-monitoring actually performed averaged across blocks of five sessions.

FIGURE 5

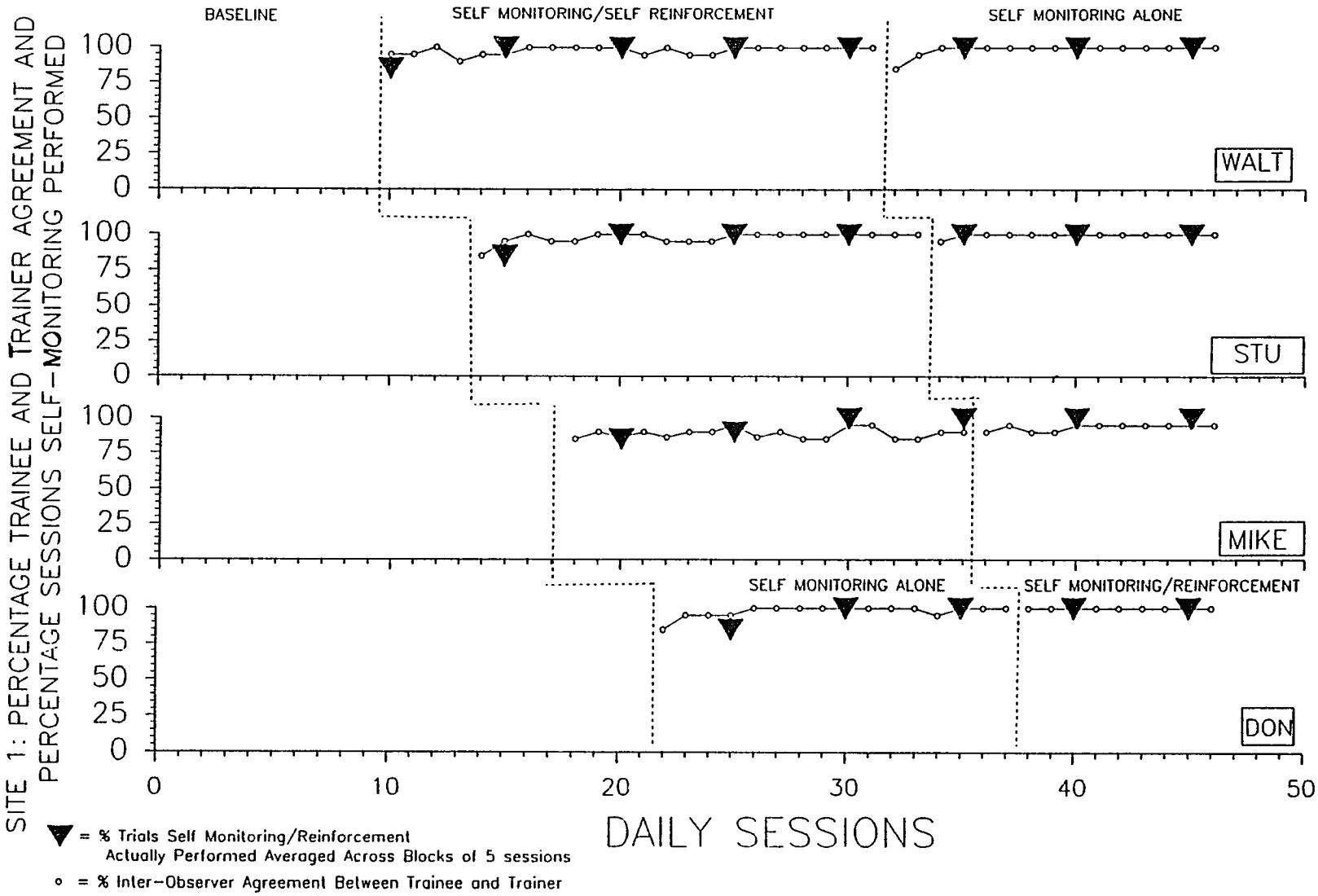

BASELINE
SELF MONITORING/SELF REINFORCEMENT
SELF MONITORING ALONE
SELF MONITORING ALONE
SELF MONITORING/REINFORCEMENT
WALT
STU
MIKE
DON
100
75
50
25
0
0
10
20
30
40
50
DAILY SESSIONS
SITE 1: PERCENTAGE TRAINEE AND TRAINER AGREEMENT AND PERCENTAGE SESSIONS SELF-MONITORING PERFORMED
▼ = % Trials Self Monitoring/Reinforcement Actually Performed Averaged Across Blocks of 5 sessions
○ = % Inter-Observer Agreement Between Trainee and Trainer

FIGURE CAPTION FOR FIGURE 6

Figure 6. Site 2: Percentage of trainee and trainer agreement and percentage of sessions self-monitoring actually performed averaged across blocks of five sessions.

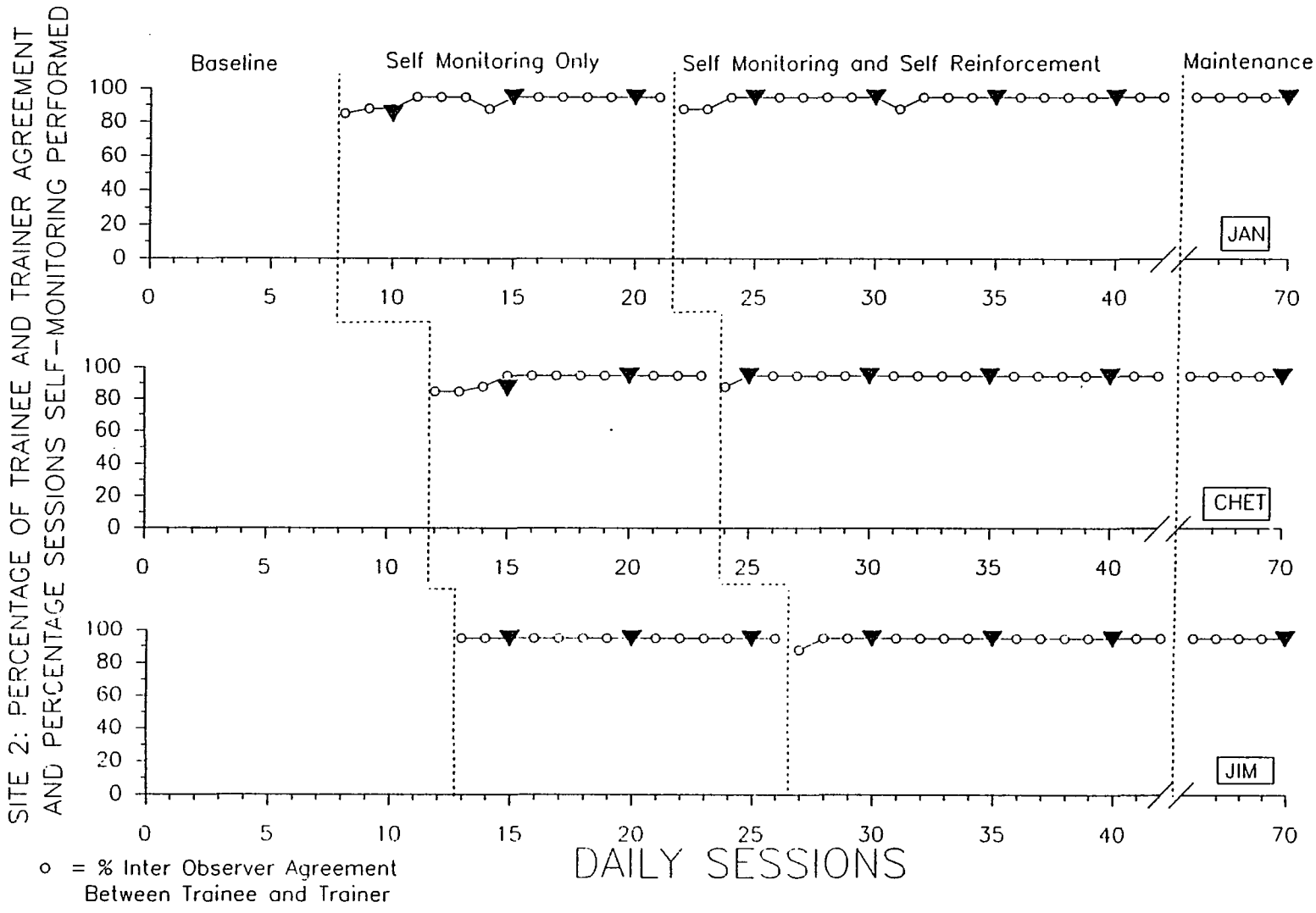


FIGURE 6

SITE 2: PERCENTAGE OF TRAINEE AND TRAINER AGREEMENT AND PERCENTAGE SESSIONS SELF-MONITORING PERFORMED

DISCUSSION

The current study demonstrated that the combination of picture cues and self-management training increased subjects' level of on-task behavior and productivity, and that the observed improvement was socially valid, both from the perspective of the subjects' own perceptions as well as those of their supervisors. Further, the observed changes persisted even after the experimental manipulation ended and responsibility for prompting, providing and administering reinforcers shifted to the subjects themselves, in the absence of the added contingency for accuracy of self-report. Finally, the observed effect was over and above that attributable to any conditioned social reinforcers the subjects may have received.

The conclusion that six of seven subjects (Walt, Stu, Mike and Don at site 1 depicted in Figures 1 and 3 as well as Jan and Chet at site 2 depicted in Figures 2 and 4) improved their level of on-task behavior and productivity above baseline using the intervention package was to be expected because the efficacy of the procedure has been well established in the previously reviewed experimental literature. The interesting finding was that in six of the seven subjects, these two measures were higher when self-monitoring alone was supplemented with self-reinforcement. Contrary to the position of Hayes et al. (1985), the present finding indicates that social reinforcement resulting from

subjects' meeting of publicly announced goals does not, by itself, account for the behavioral control exerted by the self-reinforcement procedure. This is because social reinforcement for publicly announced goals was continually available throughout all three phases of the study, (i.e., baseline, self-monitoring alone and self-monitoring/self-reinforcement), although increases in both the productivity as well as the on-task behavior measure occurred for Walt, Stu, Mike, Don, Jan and Chet only following the introduction of the self-monitoring/self-reinforcement procedure.

In terms of Mahoney and Bandura's (1972) three possible explanations of why self-reinforcement procedures can increase target responses, the current experiment supports the notion that self-reinforcement operates in a manner analogous to operant reinforcement of a target response by an outside agent. Both Goldiamond (1976) and Catania (1975) state that what defines a self-reinforcement procedure is not who delivers the reinforcer, but who determines when the response contingency is met. The subjects in the current study were responsible for determining when the response contingency was met, as well as for actually delivering the reinforcer, so that both Goldiamond's (1976) and Catania's (1975) criteria were satisfied. The present study also addressed Hayes' et al. (1985) concern that findings be extended to different populations, and that other methods should be employed to ensure that the public/private

distinction is clearly made.

Several factors converge to account for the effectiveness of the Self-Reinforcement procedure under study. Virtually all of the authors of studies cited reported that the social reinforcement available when an individual meets a publicly announced goal accounts for at least a proportion of the observed effect. Self-Monitoring may also make the employer's evaluative criteria more salient to individual subjects. The current study demonstrates that, at least in the context of individuals with a developmental disability, the self-reinforcement procedure itself accounts for some proportion of the observed effect over and above the effect accounted for by contingent social reinforcement.

In considering potential threats to the validity of the study's findings, several issues need to be considered. Only six of seven subjects demonstrated the expected results (i.e., change with the introduction of the procedure), although a reduction in the variability of Jim's data, as indicated by the change in standard deviations across phases suggests another possible beneficial effect of the procedure. All subjects other than Walt at Site 1 seem to demonstrate this same pattern of decreased variability during self-monitoring with self-reinforcement as opposed to self-monitoring alone.

A problem with drawing conclusions from the data is the

apparent lack an effect during the self-monitoring alone phase for experimental subjects at the second site only (Figures 2 and 4). Much of the experimental literature suggests that self-monitoring alone should have had some effect on the target behavior over that attained during baseline (Glynn, 1970; Sowers et al., 1985), although the decrease in variability of the dependent measure may reflect this effect. Of course, this effect might have been demonstrated had the self-monitoring alone phase been continued for an additional number of daily sessions.

Supervisors were not blind to the purpose of the study, and their biases may have influenced the data. This argument is mitigated by the high level of agreement between the supervisor's performance ratings with those of the security guard at the second site, who was blind to the design of the study.

Another possible explanation of the lack of and effect during the self-monitoring alone phase is that the subjects in the study were all formerly Sheltered Workshop employees. While at the workshop, their actual income did not vary much from week to week. This is because their Supplemental Security Income (S.S.I.) or Supplemental Security Disability Insurance (S.S.D.I.) income maintenance check would be decreased when they earned more and increased when they earned less for productivity. As participants in the supported employment program, their income from these income

maintenance sources ceased, with their earnings from work now accounting for all their income. As a result, thy may have been sensitized to the need to maximize their earnings, making the nickel reinforcers they earned from being on-task an especially potent reinforcer in this context.

The possibility of an order of treatment effect across the self-monitoring alone phase and self-monitoring/self-reinforcement phase was addressed through the use of a counterbalanced order of experimental conditions within the first site (ie, Walt, Stu and Mike receiving one order and Don receiving the opposite order, Figures 1 and 3) as well as being counterbalanced across the two sites, (with Jan, Chet and Jim receiving the same treatment order as Don, Figures 2 and 4). This procedure detected a subtle order effect (see Figures 1-4). The magnitude of change above baseline in the dependent variable (on-task behavior) was greater for those subjects for whom the order of treatment was self-monitoring alone followed by self-monitoring with self-reinforcement (Don at Site 1 and Jan and Chet at site 2) than it was for those subjects receiving the reversed order of treatment (Walt and Stu at Site 1). Mike, (at Site 1, Figure 1), was an exception to this observation, in that his level of on-task behavior improved more sharply and more immediately than his peers at the same site who underwent the same order of treatment.

Caution must also be exercised in generalizing these

findings to other populations because Zeaman and House (1963) have demonstrated qualitative as well as quantitative differences in the learning strategies of developmentally and non-developmentally disabled individuals.

Questions relating to the stability of baseline data would be best addressed by increasing the length of this phase of the experiment in future replications. In fact, increasing the length of the experiment by extending all phases to include at least 50 observations per phase would facilitate the statistical analysis of many of the parameters of the trend lines drawn through the data. This would supplement visual inspection of the data, allowing the serial correlation or autocorrelation of the error terms to be statistically modelled and then removed. If this could be accomplished, standard statistical tests could then be used to more objectively analyze resulting patterns in the data.

The current findings need to be extended in future studies. Replications with larger samples, different populations, using different target behavior and measures that avoid the ceiling effect observed above would serve to strengthen the current findings. The ceiling effect could be better addressed by beginning to collect on-task behavior data before the picture prompts are introduced in a pre-baseline phase, and then eliminating subjects whose resulting level of on-task suggested the possibility of a

ceiling effect. This strategy would also better identify the performance increment yielded by each of the respective components of the current intervention package.

Reinforcers of greater magnitude than the nickels used in the current experiment would perhaps have a less ambiguous effect on the dependent measures.

Other ways that self-reinforcement is analogous to operant reinforcement by an outside agent also need to be explored. For example, Herrnstein's Law (Herrnstein, 1970) and its subsequent formulations predict that reinforcers have a differential effect depending on their magnitude and the delay involved in their delivery. If it could be demonstrated that self-reinforcement follows similar patterns (as suggested with the density of reinforcement observation), then the current finding that it operates in a manner analogous to operant reinforcement would find more diverse support.

Catania (1975) and Goldiamond (1976) make the point that because Social Reinforcement may be the most potent component of the process referred to as Self-Reinforcement, perhaps the procedure should be described in that language. They both discount the relative contribution of self-delivery of reinforcers, but, based on results of the current study, this conclusion may be premature, because the self-delivery of reinforcers cannot be completely discounted as contributing to the efficacy of the procedure.

Additional research is required to quantify the relative contribution of the various component factors.

APPENDIX A

Literature Review of Self-Reinforcement

Self-reinforcement has been variously referred to in the experimental literature as a special case of the larger class of operant reinforcement, or as a behavioral relation that is completely distinct from operant reinforcement and whose effects on behavior are better described in the language of verbal behavior.

In Science and Human Behavior, Skinner (1953) examined the concept of self-control, and stated that:

> The place of self-reinforcement in self-control is not clear. In one sense, all reinforcements are self-administered, since a response may be regarded as "producing" its reinforcement. (p. 237)

Aristotle treats the distinction between intrinsic and extrinsic good in much the same way. In his "Ethics", he defines an intrinsic good as one that is inherently good without the need for reference to any external objects or concepts. He further states that extrinsic good, by definition, must eventually be reducible to intrinsic good in order for it to be manifest.

This is a point that comes up again and again in the literature. In a sense, the case of a rat pressing a bar for food can also be considered to be an example of self-reinforcement, since the behavior of bar pressing brings about the opportunity to eat. Skinner (1953) then goes on to qualify that something else is also involved in the process of reinforcing one's own behavior. The qualifying

factor that distinguishes self-reinforcement from operant reinforcement is that the organism has it in its power to provide reinforcement at any time, but chooses not to do so until some criterion behavior has been emitted. Here, the continuous availability of the reinforcement and the determination of when that criterion has been satisfied are both under the control of the organism being conditioned.

Conditioning is defined by Skinner (1953, p. 65) as the strengthening of behavior that results from reinforcement. He also states that we can judge whether an event is reinforcing solely in terms of the effects of stimuli on behavior, and that we can't identify any other factor that allows us to judge whether an event is reinforcing and still retain objectivity. Skinner then deals with an example of self-reinforcement:

> Self-reinforcement of operant behavior presupposes that the individual has it in his power to obtain reinforcement but does not do so until a particular response has been emitted. This might be the case if a man denied himself all social contacts until he has finished a particular job. Something of this sort unquestionably happens, but is it operant reinforcement? It is certainly roughly parallel to the process of conditioning the behavior of another person. But it must be remembered that the individual may, at any time, drop the work at hand and obtain reinforcement. We **have to account for his not doing so.** It may be that such indulgent behavior has been punished, say, with disapproval, except when a piece of work has been completed. The indulgent behavior will therefore generate strong aversive stimulation except at such a time. The individual finishes the work in order to indulge himself free of guilt. The ultimate question is whether the consequence has any strengthening effect upon the behavior which precedes it. Is the individual more likely to do a similar piece of work in the future? It would not be

> surprising if he were not, although we must agree that he has arranged a sequence of events in which certain behavior has been followed by a reinforcing event. (Skinner, 1953 p. 237).

Here, we see that the situation is much more complicated than simple operant reinforcement. The reinforcer is freely available throughout the process. The organism itself determines when the criterion behavior is present. Verbal behavior mediates the entire process, and guilt is brought into play as a potential aversive stimulus. The distinction between a private event and a public event is blurred. Inclusion of this verbally mediated behavior is critical, for it has far reaching implications in terms of whether social reinforcement, either implied or overt, comes into play.

Mahoney and Bandura (1972) operationalized self-reinforcement for animals in much the same way as it had been previously done for human subjects. The criteria were that the reinforcers were made freely available prior to each trial so that the animal had full control over them. Further, that the organism refrained from using the reinforcer until the criterion behavior had been emitted. Initially, pigeons were trained to peck discs for grain reinforcement. There was a training phase in which the birds had to peck the response key in order to produce the grain filled hopper. Gradually, the hopper was presented earlier in the pecking sequence until it appeared before the birds began to emit any pecking responses. If the bird

entered the hopper before performing any pecking responses, the hopper was removed and the bird was unable to feed until the next trial. If the pigeon pecked the disc before entering the hopper, it was allowed 3.5 seconds of feeding time, after which the hopper was again removed and the next trial began. Training continued until the criterion behavior of 100 consecutive trials, in which the birds pecked before entering the hopper to feed, was met. Test conditions, under which there was no punishment for the birds' eating without pecking were then instituted and the persistence of the pecking prior to eating was examined.

The authors present three possible explanations for why the animals continued to work for freely accessible reinforcers. The first possibility was that the pecking behavior was inherently reinforcing. They cite an experiment by Neuringer (1969) that failed to show any intrinsic reinforcement value in pecking behavior. The second possibility was that there was a punishment-induced effect, also alluded to by Skinner (1953) above. They state that "the remarkable durability of the self-reinforcement pattern bears some resemblance to the refractory quality of avoidance behavior" (Mahoney and Bandura, 1972 p. 302). The third possibility is that the behavior persists because it was self-reinforced, and because self-reinforcement acts in a manner analogous to operant reinforcement.

This animal analog of self-reinforcement in humans has

been questioned by a number of investigators, primarily on the grounds that the animal does not independently select an important behavior for self-reinforcement. It is only by the intervention of the experimenter, i.e., experimenter punishment of movements towards the food hopper, that the connection is established. These types of examples are routinely dealt with under the heading of self-control (Rachlin and Green, 1974), while examples with human subjects are more commonly referred to as self-reinforcement or self-management.

Self-Reinforcement And Behavior Change

Returning to Mahoney and Bandura's (1972) third possible explanation for the effects of self-reinforcement, a large number of studies have demonstrated that behavior can in fact be sustained through self-administered rewards as well as through reinforcers delivered by an external agent. The authors cite two studies of human subjects, Bandura and Perloff (1967) and Glynn (1970) that demonstrate the maintenance of behavior through self-administered rewards.

Bandura and Perloff (1967) set up 4 groups of children performing a motor task. The first group self-selected criteria and self-administered rewards for adequate performance. In a second yoked control group, subjects were reinforced by an observer for achieving performance levels selected as criteria by the first group. The other groups

were a noncontingent reinforcement and no incentive control group. The authors found no difference between the first two groups that both performed at a higher level than the two control groups.

Glynn (1970) examined the relative effectiveness of self-determined criteria and self-administered reinforcement on the test scores of a group of ninth grade girls. These two groups were compared to a no token control group. The performance of the first two groups was equally improved, and both improved at a much faster rate than the third control group.

A study by Lovitt and Curtiss (1969) assessed the relative effectiveness of self- or teacher-imposed contingencies on the academic response rate of 12-year-old students at a school for children with behavioral disorders. They performed two separate experiments. The first explored the relative effectiveness of self- as opposed to teacher-imposed contingencies for increased academic response rate, which was evaluated via a single subject ABAB design (Kazdin, 1982). When it became clear that the self-imposed contingencies were more effective, the authors varied the rate of reinforcement, and found that the rate was not as potent an agent in increasing the response rate as was the selection of the reinforcing agent. These results were replicated with several students. The selection of target behaviors as well as the rate of

reinforcement were made by the teacher and not by the student. It was only the provision of the reinforcement itself that was accomplished by the subject of the experiment. A qualitative difference between self-reinforcement and operant reinforcement is suggested by the observation that rate of reinforcement was not as potent a variable as was selection of a reinforcing agent.

Felixbrod and O'Leary (1973) examined the difference between self- and externally-reinforced behaviors. They found that the various reinforcement arrangements yielded no difference in the second graders' accuracy or rate of performing simple math tasks, but that self-imposed criteria and self-delivered reinforcers improved their persistence on the tasks.

A study by Mahoney, Bandura, Dirks and Wright (1974) investigated whether monkeys had a relative preference for self-monitored or externally imposed sources of reinforcement. Of the two subjects examined, one monkey clearly preferred the self-reward system, while the second had little preference initially but shifted towards acceptance of the self-delivery system at the end of the experiment. The important observation here was that both subjects would perform extra work in order to have their rewards self-delivered rather than externally delivered.

Ninnes, Fuerst, Rutherford, and Glenn (1991) investigated the effect of self-management and self-

reinforcement on improvements in conduct in the absence of supervision. The subjects were three emotionally disturbed adolescents, who, when covertly filmed in their classrooms in the absence of supervision, engaged in extremely high levels of off- task behavior. They were trained to self-assess, self-monitor and self-reinforce for the presence of on-task behavior in the absence of direct supervision. These adolescents were easily able to learn the three-step process and apply it in the classroom environment where it was trained, but it did not generalize to the between class activities of the three until it was explicitly trained in that environment. No effort was made to separate the differential effects of the self-reinforcement from social reinforcement that they also received contingently upon improvement in target behaviors. An important finding, though, was the maintenance of the improvements in target behaviors in the absence of outside interventions after the experiment had been completed.

The list of similar studies is quite extensive, and they all seem to suggest that self-imposed or self-administered contingencies can be effective in changing behavior.

Self-Reinforcement And Self-Management

A number of theorists have offered models, variously called self-control or self-management, that describe how the Self-Reinforcement procedure promotes the process of

behavior change.

Goetz and Etzel (1978) defined Self-Control as "responses made by an individual that, in turn, modify the individual's own behavior." (p. 58)

Karoly and Kanfer (1982) present a model they call Self-Management, that consists of the three processes of self-monitoring, self-evaluation and self-reinforcement. These are described in more detail by Karoly (1982) as self-observation, the setting of self-evaluative standards and finally self-consequation. The three stage model (Kanfer, 1971; Kanfer and Phillips, 1970; Karoly and Kanfer, 1972) is described by the author as:

> The experimenter who would shape the behavior of a pigeon in an operant chamber must observe the bird's actions in regards to the response level, decide on some criterion for the delivery of a food reward, (i.e., establish a schedule), and arrange for the machine to provide the rewards contingent upon the requisite tempo or frequency of responding. When the purveyor and recipient of shaping efforts are the same person (as in human self-control), then the subjectively administered program is built upon the process of self-observation, self-evaluation (criterion or standard setting) and self-consequation (self-reward or self-punishment). (Karoly, 1982, p.17)

Karoly (1982) further states that the matching law (Herrnstein, 1970) offers some explanation of the efficacy of self-management procedures. A variant of the original matching law (Rachlin and Green, 1972) indicates that the value of a reinforcer is related to the ratio of its magnitude to its delay. Self-Management procedures may diminish the delay involved in the delivery of reinforcers.

In an applied setting, for example, a target subject would be able to recognize a target response immediately after it is emitted, and contingently reward it, while delivery of the same reward by an independent observer may involve a longer delay interval. Therefore, since the delay is reduced in self-reinforcement, the reinforcement value is enhanced.

The Mechanism Of Behavior Change Through Self-Reinforcement

Does the self-delivery of reinforcers bring about change in behavior in a manner analogous to that under operant reinforcement, or does some other factor, such as mediation by verbal behavior or social reinforcement, account for the observed change? None of the previously-cited studies examined the difference between publicly-and-privately set goals. Skinner (1953) alludes to the importance of this distinction.

> The contradiction that private events depend on public or societal contingencies has a possible explanation in that the terms are established as part of a repertoire when the individual is behaving publicly....Verbal behavior always involves social reinforcement and derives its characteristic properties from this fact (Skinner, 1953 p. 261-299).

Goldiamond (1976) raises some interesting problems with the definition of the term self-reinforcement. He reiterates Skinner's observation that in a sense, all operant behavior is self-reinforced, since by responding, the consequence is brought about. Goldiamond suggests that the entire process is more complicated than just who supplies the

reinforcement. He claims that it is not as much who supplies the reinforcement, but who evaluates whether the response requirement has been met. He claims that operant contingency is defined as the independent definition of a response as a requirement for the delivery of a consequence (i.e., the required response is defined by an external observer rather than by the organism emitting the response). Many self-reinforcement procedures do not have the contingency independently defined, (i.e., by an outside agent). Therefore, he claims that it is misleading to talk of self-reinforcement because it confounds self-monitoring with the self-delivery of contingencies.

Thoresen and Wilbur (1976) debated the extent to which self-evaluative standards are independent of the behavior to be evaluated. They see Goldiamond's objection as implying that personal performance standards cannot be independent, and therefore are not bona-fide criteria for administering positive contingencies contingently. The authors then cite evidence from three previous studies that suggests this independence may be possible. Bolstad and Johnson (1974); Drabman, Spitalnick and O'Leary (1973); Turklewitz, O'Leary and Ironsmith (1975), suggest that with the proper training, in which the responsibility for the self-administration of consequences and the self-determination of standards is gradually introduced, subjects can contingently self-reinforce appropriate behavior.

Catania (1975) questioned the logical foundation of the process of self-reinforcement. He argued that the circumstances that generate the behavior of self-reinforcement (i.e., the reinforcement of one's own response) make it impossible to tell whether increases in self-reinforced responses can be attributed to their having been reinforced rather than to other salient aspects of the process, such as the identification and setting of goals, the verbal behavior of recognizing the target behavior as a desirable consequence, etc. In his examination, he claims that the essence of self-reinforcement, in the Mahoney and Bandura (1972) pigeon paradigm, is not that the opportunity to eat maintains the pecking behavior of the pigeons, but rather that pecking provides a discriminative stimulus for eating. This operation is better described in terms of self-discrimination or self-awareness, rather than self-reinforcement. He says that the relevant variable here is that the subject develops the capability of discriminating its own behavior and that reinforcement is only incidental to maintenance of pecking. He uses the example of a student who reinforces his studying with a trip to the movies. The student in the example would only reinforce the completion of an assignment if its completion were already sufficiently important that the completion alone would be reinforcing anyway.

Catania summarizes by saying that evidence suggesting

that self-reinforcement is just as effective as external reinforcement may not be wrong, but rather, that attention to the more relevant factors involved, (i.e., goal setting and learning to discriminate aspects of one's own behavior) might be more productive fields of investigation than the reinforcement procedure itself.

The implications of Catania's criticism of the concept of self-reinforcement are far reaching, and, as he states in an article titled "the Myth of Self-Reinforcement" (Catania, 1975), the contributing factors are difficult if not logically impossible to tease out. Let us assume for the moment that one means of testing this would be to examine whether goal setting alone or in combination with self-reinforcement is capable of increasing the frequency of some target behavior. A further distinction is whether private as opposed to public goal setting is more effective in accomplishing this end. How does one examine whether goals, that are set privately without the knowledge of the experimenter, are being contingently reinforced. This is a reactive measurement problem, for if the experimenter gathers information about the reinforcement of privately set goals, then they are no longer private. We could try to delude the subject into thinking that the process is a private on when in fact it is not, but then we could never be sure that our ruse was successful.

In order to explore the factors contributing to the

efficacy of Self-Reinforcement, a series of experiments was performed, initially by Nelson, Hayes, Spong, Jarrett and McKnight (1983) and by Hayes, Rosenfarb, Wulfert, Munt, Korn and Zettle (1985).

This reviewer found the first in the series of experiments to be beset by several methodological flaws. The experiments performed by Nelson et al. (1983) used undergraduate females who were participating in the study to meet a course requirement. The paradigm was a 2x2 repeated measures design, where the variable manipulated was whether reinforcement for exercising was administered prior to or following the activity, and whether the goal was publicly or privately established. These researchers found no improvement on the dependent variables above that of a control group where no contingent reinforcement was administered. The experiment included no measures to ensure that the reinforcers were actually delivered contingently (i.e., the integrity of the independent variable was not verified). Further, the use of a very small sample size calls into question whether the statistical test had sufficient power to tease out real differences between the groups in the sample. The finding of no significant difference between the groups in the study was cited as supporting the view that self-reinforcement per se had no effect on the dependent measure. This could be a problem, because the subjects in the study had no real desire to

improve in the target area (i.e., freshmen psychology students volunteering to participate in the study as partial fulfillment of their course requirements), and therefore, the effect of social standard setting may have been compromised.

A second study by Hayes et al. (1985) avoided many of the methodological flaws of the first study. It used subjects that had some real desire to increase the target behavior, the measures were well developed, and perhaps most importantly, the method used to fool the subjects into thinking that their goal setting and self-reinforcement were truly private events clearly was more credible. Due to the fact that this second study is frequently cited as providing clear evidence that self-reinforcement alone cannot account for behavioral change observed when the procedure is used in applied settings, the study will be described in more detail.

The study attempted to separate the effects of self-delivery of reinforcers from those of public versus private standard setting. Hayes et al. (1985) proposed that all previous research in this area has confounded these two processes. The authors used a 2 X 2 X 2 factorial between groups design, with repeated measures on the final factor. A feedback-only control group was also run. The first variable to be manipulated was whether treatment took place in a public or a private context. The second variable was

whether there was goal setting with or without the self-delivery of reinforcement that was contingent upon a criterion response. The final variable was whether the dependent variable was measured during the baseline or experimental phase of the study.

Each subject in the study was given a series of reading passages modeled after those presented in the reading comprehension section of the Graduate Record Examination. After each passage, subjects were given six multiple choice questions, each containing 5 possible answers. The dependent measure was the number of correct answers following each passage.

Subjects were recruited by placing ads in local newspapers and with local radio and television stations calling for people who wanted help preparing for the GRE examination. Fifty subjects volunteered. A pre-test was administered, and those subjects scoring more than 3 correct answers per passage were eliminated. Hayes et al. (1985) claims that this was necessary to create a clinically relevant sample and to prevent a ceiling effect. Twenty-six subjects remained.

During baseline, subjects were given their choice of food reinforcers (bags of M&M's, peanuts or raisins). There were three trials in baseline, with each trial being defined as answering the 6 questions following one reading passage. Passages and questions were presented on a teaching machine.

The sequence used was uniform across all subjects, as were the actual passages selected.

During the experimental phase, subjects were assigned to one of five possible groups as follows: 1. Control. (N = 5) Subjects were told to continue reading passages and to eat reinforcers as they wished. 2. Private goal setting with no self-consequation. (N = 4) These subjects set a goal for the number of correct answers to each passage and were led to believe that no one could know their goal. They wrote down their goal and put it in a box that contained other subjects goals. They were told that the reason for this procedure was to enable another yoked control group to choose their goal without knowing whose it was. In actuality, all of the other papers in the box were blank so that the goal was known by the experimenter. They were also told to eat the reinforcers whenever they liked. 3. Public goal setting with no self-consequation. (N=6) These subjects were given the same instructions as the previous group, but when they wrote down their goal, the experimenter read the goal out loud to the rest of the group. 4. Public goal setting with self-consequation. (N=6) This group set their goals exactly as the previous group. In addition, the experimenter described the process of self-reinforcement and subjects were instructed to use the food as a reinforcer to be delivered according to a fixed schedule when the goal set for each passage had been met. 5. Private goal setting

with self-consequation. (N=5) These subjects set their goal as did the other private goal group and were given the same guidance in self-reinforcement.

The authors performed an ANOVA on the mean scores of the five groups during the pre-experimental condition and found no significant difference between them. As a result of this finding, they subtracted the pre-from the post-scores to yield a difference score. It was this difference score that was subjected to a detailed statistical analysis. The authors' assumption that the groups could be treated as being equivalent because the ANOVA yielded no significant difference, and therefore difference scores could be used in their data analysis, is questionable, since, in several of the sub-scores, the finding of a ratio of less than one but more than zero calls into question whether the extent of variability among subjects within groups may have been in excess of that between groups. Since power to tease out between group differences needs to be maximized here, the authors might have assigned subjects to experimental conditions by a randomized block design, using the conveniently available pre-test scores as the blocking factor. This would have increased the power of the subsequent ANOVA to tease out significant differences between the small group sizes and therefore made the assumption of inter-group equivalence more plausible. Another alternative would have been to increase the size of

N in each group, a tactic that could have been used in conjunction with blocking to further increase power. The authors also switched back and forth between use of one or two tailed tests of significance, based not on the nature of data being examined, but rather on the liberal rejection region afforded by the change. Keppel's Handbook of Experimental Design (1982) treats this as a questionable practice that should be avoided.

Diffusion of treatment effects may have taken place here. It is actually possible that the treatment condition described as private goal setting may have been a public one without the knowledge of the experimenters. This would have been the case if these subjects went home and discussed their privately set goals with their friends or families.

Differential treatment and the lack of reliability of the independent variable are other potential sources of confounding. The amount of food eaten by each subject was not controlled. They could have taken the food home with them rather than immediately reinforcing their own behavior. There was therefore no assurance that self-reinforcement was reliably administered in the conditions where this was part of the experimental manipulation (i.e., whether reinforcement was actually consumed at times appropriate and in proportion to the number of modules completed to criterion levels).

There was no proof offered that the public/private ruse

of the authors actually worked, and that its effects on the dependent measure were not confounded with other relevant variables.

The authors finding that there was no difference between the performance of groups based on contingent self-reinforcement, but that there was a difference between groups when goals were set in a public as opposed to a private context is still quite interesting. The threats to Statistical Conclusion Validity reviewed above call into question whether there was enough power to discern a difference where one might really exist. The fact that a difference was found in spite of the lack of power tends to support their analysis on the public versus private dimension, but not their findings on the question of whether self-delivery of reinforcement was effective in improving performance on the dependent measure. The authors' use of the difference score, based on a possibly faulty assumption of between group equivalence during the pre-experimental condition, makes even this conclusion questionable. The obtained data could have been subjected to appropriate statistical tests of significance without the use of the difference score (that functionally cut in half the number of analyses to be performed), and the results would have been more credible, even given the reservations described above. They discovered that the groups that set goals publicly achieved a much greater improvement than did those

with privately set goals, and that the contingent self-delivery of reinforcement added nothing to the effectiveness of the treatment. In the words of the authors:

> The two experiments make more plausible the view that self-reinforcement procedures work by setting a socially available standard against which performance can be evaluated. The procedure itself functions as a discriminative stimulus for stringent or lenient social contingencies (Hayes et al. 1985 p. 214).

The authors suggest that areas for future research should include replications of the Hayes et al. (1985) study with populations other than college students, and the use of alternate designs.

Perhaps, as described above, it really is impossible to make the public versus private goal setting distinction experimentally because of the reactive nature of the variable. An alternative strategy would be to keep the public/private variable constant and to manipulate the self-delivery of reinforcers. The current experiment used this strategy to accomplish this with a novel population in a "clinically relevant" environment, to use the same terminology as Hayes et al. (1985). A single subject experimental design is used to address the question of statistical conclusion validity with small samples usually found in clinical settings. Results arising from this design would suggest that something other than social standard setting and the social reinforcement resulting from meeting publicly set goals accounts for the perceived efficacy of the self-reinforcement process.

The question is not purely theoretical in nature but rather has very practical implications. Many procedures currently used to change behaviors include a self-reinforcement component. It would therefore be of value to determine whether this procedure is superfluous. If public standard setting is identified as the potent agent of behavior change, then it can be made a more explicit component of the package.

To summarize, social learning theorists have proposed that humans can learn to set performance standards, monitor their behavior and self-deliver reinforcement contingently, all of which tends to increase the frequency or strength of a target response (Bandura, 1976, 1978, 1981; Mahoney and Bandura, 1976; Thoresen and Wilbur, 1976). This view has been criticized by other theorists who are loath to place the ultimate cause of behavior in more behavior of the same organism (Catania, 1975, 1976; Goldiamond, 1976; Rachlin, 1974). A considerable body of literature indicates that self-monitoring coupled with self-reinforcement does in fact have a potent effect on establishing, maintaining or increasing the frequency or strength of a targeted response. The debate really hinges on the mechanisms responsible for these observed effects. Whereas the Social Learning theorists suggest that self-monitoring and contingent self-delivery of reinforcers are the potent agents of observed changes in behavior, the school of thought most eloquently

represented by Catania suggests that these effects are artifacts of the process of social standard setting and its accompanying opportunities for social reinforcement.

Qualitative Differences In Cognitive Strategies

One final consideration involves the use of these procedures with developmentally disabled adults. In an influential series of studies, Zeaman and House (1963) indicated that there are qualitative as well as quantitative differences between the way developmentally disabled and non-disabled individuals learn. The authors found that when learning a visual discrimination task, developmentally disabled individuals were better able to learn reversals of previously trained discriminations within a stimulus dimension, while non-disabled subjects were better able to learn changes in which a new stimulus dimension came to govern reinforcement. Reversals within a stimulus dimension were easier to handle for the developmentally disabled, while changes in the discriminative stimulus to a different stimulus dimension were easier to learn for non disabled subjects. The study indicated that there may be major differences between the way developmentally disabled individuals and their non disabled age peers learn. This could have important implications for the development of training procedures for the population, and suggests caution when applying procedures developed for the "normal" population to developmentally disabled individuals.

In the context of the current study, this line of research suggests that information about the extent of behavioral control exerted by self-reinforcement may not be easily generalized to the non-disabled population.

In a literature review of self-instructional procedures by Whitman, (1987) the author stated that because of the nature of disabilities found in the developmentally disabled, self-instructional training procedures are particularly useful with populations that have a limited knowledge base and limited linguistic skills. The mentally retarded as well as young children fall into this category. This further suggests that there may be a unique relationship between self-reinforcement procedures and the developmentally disabled.

APPENDIX B

Literature Review of Supported Employment and Individuals with Developmental Disabilities

Supported work is an alternative to traditional service delivery models for individuals with developmental disabilities emphasizing independence and self-sufficiency. Supported Employment provides opportunities for disabled individuals to become integrated into regular mainstream work environments. In doing this, it provides access for the disabled into one of the major social institutions in the life of the community, thereby empowering them with all of the rights and privileges available in this important cultural setting (Wehman, 1988, Kiernan, McGaughey and Schalock, 1988). Supported Employment offers occupational choices to members of this population who have traditionally been considered unemployable in the competitive labor market (Rusch and Hughes, 1989).

The Developmental Disabilities Act of 1984 defines Supported Employment as:

> Paid employment which is for persons with developmental disabilities for whom competitive employment at or above the minimum wage is unlikely and who, because of their disabilities, need ongoing support to perform in a work setting. It is conducted in a variety of settings, particularly work sites in which persons without disabilities are employed; and is supported by any activity needed to sustain paid work by persons with disabilities, including supervision, transportation and training. (Federal Register, 1984)

It emerged as a program model as a result of the simultaneous interaction of six major factors:

First, the civil rights movement of the 60's and 70's, with its emphasis on the responsibility of society to provide opportunities for all of its members.

Second, the process of deinstitutionalization, coming after several landmark lawsuits brought by advocates for the disabled (i.e., Wyatt vs. Stickney, Rockefeller vs. N.Y.) that resulted in the closing of many of the larger institutions where the developmentally disabled had been housed, and their subsequent release back into their local community.

Third, the philosophy of Normalization, (Nirje, 1976, Wolfensberger, 1968) advocated that disabled individuals should be encouraged to live as normal a lifestyle as their abilities permit, and that they should participate as fully as is possible in all aspects of the life of the community.

Fourth, the local labor market. The local economy changed from predominately skilled and unskilled manufacturing industries to emphasize the growing service sector, with its demand for many unskilled workers in new industries like fast foods, cleaning, porter maintenance, messengers, chambermaids, nursing home attendants, etc. (Stark, 1988)

Fifth, demographics. Closely related to the shifting demands in the local labor market, there has been a gradual

aging of the American population (Turnbull and Turnbull, 1988). The median age of the population has progressively increased, and fewer and fewer teenagers are introduced into the population to assume these entry level jobs just as the demand for them is increasing. Until the recent recession, many suburban communities were unable to fill available jobs in fast food and retail establishments even when they offered wages well above traditional minimum levels, and the employees they were able to hire were plagued with poor attendance, punctuality and performance levels, resulting in a high level of employee turnover

Sixth, the refinement of Applied Behavioral technology. Numerous early studies (Rusch and Hughes, 1989, p. 352; Fergusen and Fergusen, 1986; Wehman, 1988) demonstrated that the technology was available to enable severely disabled individuals to perform complex tasks at levels comparable to those of regular employment settings. Much of this research grew out of the Mark Gold, Lew Brown, Seattle Special Employment Project work done in the late 60's and early 70's (Bellamy, Inman and Horner, 1979). These researchers used applied behavior analysis techniques to train the reliable and accurate performance of complex assembly tasks by severely and profoundly retarded adults. Once this capability was plausibly demonstrated (Mark Gold, 1973), it was a short step to asking what other skill-related obstacles prevented these people from working on jobs in the

community, since, as outlined above, the legal and ethical mandate and emerging technology conspired to meet a real labor market demand. If these obstacles could be identified and a task analysis developed, then these individuals could be trained and ultimately employed in the community. The technology to ensure maintenance of these behaviors and generalization to novel situations existed (Stokes and Baer, 1977). These researchers then asked why these individuals couldn't move out of their sequestered or sheltered settings into mainstream employment. So, a flurry of activity began to develop task analyses of:

1. Independent travel Skills
2. Janitorial Tasks
3. Maintenance Tasks
4. Food Service tasks
5. Social and Interpersonal skills required in job settings.

At the time, competitive employment into industry had been occurring in this population at a rate of approximately 10% of the Sheltered Workshop population (Gold, 1973). This was based on a Train and Place model, where an individual would attend a training program that would initially evaluate his/her skills and then develop a service plan to remediate any identified deficit areas. Then, theoretically, he or she would be ready to succeed in whatever job opportunity arose. Staff would then assist the individual in finding a

job, interviewing for and obtaining a job, requesting accommodations at the job site that made job performance possible given his/her limitations (i.e., counting jigs, reassigning job responsibilities among other employees to minimize prerequisite reading or math skills, etc.). The counselor would stay on the job with the new trainee for the first few days and then gradually fade out his/her intervention so that after one month, the trainee would either sink or swim on his/her own. Follow-up visits then occurred on a regular basis, but too infrequently to maintain an adequate level of prompting. Approximately 10% of these trainees remained employed (Gold, 1973). Several programs, such as Altro Workshops, received a grant to provide more intensive long term follow up and found that this percentage could be significantly increased. The Fountain House model was also well known for its emphasis on intensive follow-up services (Sands and Radin, 1978, Postgraduate Center for Mental Health).

Problems With The Supported Work Model

Several difficulties arose. Many interventions were structured so that control of relevant behaviors was not successfully transferred to stimuli in the work environment (Rusch, Schutz and Agran, 1982; Salzberg, Agran and Lignugaris-Kraft, 1986). Many of the behaviors required on the job were not in the realm of specific job-related skills that needed to be learned, but rather were more in the realm

of social and interpersonal skills (Rusch et al. 1982). Integration into the work environment and its social culture did not occur as simply as planned. As the Civil Rights movement of the 60's had indicated, simply providing increased opportunities for interaction did not, by itself, ensure that real integration would take place. Both Rusch et al. (1982) and Salzberg et al. (1986) observed that these new employees were not being absorbed into the social system of the work places into which they were introduced.

Initial studies in the area of supported employment were somewhat vague as to how to transfer control of target behaviors from the trainer to stimuli indigenous to the work place, how to perform a task analysis of some of these grey areas such as social and interpersonal skills that might occur spontaneously and with low frequencies, how to accomplish travel training and supervision of recreational opportunities, deal with inappropriate role models on the part of co-workers, supervisors, employers or customers, or prevent the development of maladaptive behaviors such as the use of alcohol and drugs, unwarranted emotional displays, distinguish between sexual victimization and legitimate sexual expression, and finally to ensure generalization of skills learned to cope with a specific problem to other similar situations that had not been specifically trained. In the last fifteen years, a flurry of research activity has attempted to address many of these problems, and as a

result, the program model of supported employment has emerged as a new discipline with a substantial body of literature.

An important early exploration of strategies that could be used to transfer responsibility for prompting target behaviors from the trainer to the supported employee was a study by Connis (1979). A combination of antecedent cue regulation and self-monitoring was used with four mentally retarded adults working in a food service setting. The target behavior was the trainees' dependence on their trainer to cue them when to change tasks, as well as which task to perform next. The tasks consisted of dish washing, dish drying, table cleaning and tending garbage receptacles. Visual cues were provided that were pictures of the trainees performing each of the four tasks. These were arranged in the sequence they were to be performed during their work tour. Each picture had a check off sheet, so that the trainees could record when they had completed that task. The combination of picture cues and self-monitoring was effective in prompting independent job changes in the four subjects studied by the author. The components of the intervention were not systematically withdrawn at the end of the project, since the author felt that the training package could be left in place to continue providing necessary prompts to the trainees.

In an editorial preface to an entire issue of the

Journal of Applied Behavior Analysis devoted to Supported Employment, Brandon Greene (1989) discussed the need for or advisability of completely withdrawing supports that enable supported employees to function more independently at their work sites. He states that although the aim of a well designed intervention by a behavior analyst is to gradually fade the intervention after training has been accomplished, it is clear that the substantial deficit in cognitive functioning demonstrated by a developmentally disabled adult will not be removed as a result of even a well designed intervention. Supports that assist individuals in coping with their limitation may have to remain in place, creating a prosthetic environment, suggested in an article by Lindsley (1966). The disabled individual is then left able to perform a task independently that previously required the intervention of a trainer. Ingenuity comes into play in ensuring that the tools provided are economical, unobtrusive, and capable of being used in situations different from those initially trained, ensuring long term maintenance and generalization. Sowers, Verdi, Bourbeau and Sheehan (1985) extended Connis' (1979) findings by describing the self-monitoring procedure used in more detail. They used a multiple baseline across subjects design to assess the effectiveness of a combination of picture cues and self-monitoring in increasing the job independence and flexibility of mentally retarded adults

working in a cafeteria setting. Each of four subjects was assigned seven (of a possible pool of fifteen) different tasks each day. Initially, each of the subjects needed frequent verbal prompts from an on-site trainer to initiate each of the different tasks. Pictures were taken of the subjects performing the fifteen tasks and these were place into a loose-leaf folder for reference use by the trainees. Sowers et al. (1985) described four discrete steps in the self-management process: 1. Refer to pictures; 2. mark off each picture as the task is completed; 3. touch the next picture; 4. begin the next task in picture sequence. Generalization was assessed by introducing novel pictures into the folder and observing whether the trainee responded in an appropriate manner. The dependent measure was whether the subject began the designated task in its appropriate order without the need for an additional supervisory prompt. Results indicated that the procedure was successful for all four subjects, and that the procedure generalized to all untrained pictures as well.

Wacker and Berg (1983) examined the effectiveness of picture prompts on facilitating the acquisition of complex vocational tasks by Mentally Retarded adolescents. Here, the authors used a series of picture prompts to prompt their trainees' acquisition of a complex task. A task analysis was done and each step was represented by a photograph of the task being performed. This yielded an instruction

manual for completion of the task that could be used by individuals who lacked reading skills. These manuals were retained by the subjects for later reference, and resulted in accurate task performance without the need for any additional prompting. The procedure generalized to different tasks that had not been explicitly trained. The authors hypothesized that these externally mediated cues worked by providing stable discriminitive stimuli that the student used to guide his/her work performance.

A variant of the issue of job flexibility, (i.e., the ability to change tasks without supervisor's prompts, Sowers et al. 1985) is initiative, defined by McCuller, Salzberg and Lignugaris/Kraft (1987) as recognizing that something needs to be done and attending to it even though it was not an explicitly assigned responsibility. The authors state that attempting to explicitly train all possible exemplars of initiative would be impossible. Therefore, they developed a three step procedure to address the issue. These were: developing a list of possible exemplars; examining the variables that produced generalization; and finally, examining whether new responses could be taught with less effort than was required to teach the initial steps.

The authors interviewed local employers to produce a list of thirteen job initiative situations and then operationalized them for the employment setting of the

subjects. The next step was to explicitly train a subset of these situations to the subjects, using discrimination training, role playing and self-monitoring. Finally, after the subset had been acquired, new exemplars were examined. The authors found that the new exemplars were learned with only discrimination training. In summarizing, they stated that it was not clear whether a stimulus class was established with each exemplar. Further, follow up or maintenance data were unavailable due to a layoff at the employment site. Generalization across settings was evident, but generalization across slight variants of the stimulus materials was not readily apparent.

Wacker and Berg (1986) attempted to apply the work on generalization done by Stokes and Baer (1977) to the environment of the work place. The authors state that too often we assume that generalization has not taken place in a given context because students do not perform errorlessly. As the McCuller et al. (1987) article indicated above, generalization can be assumed to have occurred across different training phases if the training effort required to elicit a behavior has been reduced below pre-training levels. Wacker and Berg (1986) categorize these procedures as falling into two major groups; antecedent and consequent procedures.

Antecedent Procedures

Common Discriminative Stimuli. In a work environment,

multiple antecedents are often present, any one of which may come to guide a particular behavior. The object, then, is to teach the trainee to respond appropriately to any of a range of discriminative stimuli that all may require a specific response. This is particularly important in a simulated training context.

Sufficient Exemplars. Antecedents in a real work environment frequently vary. Rather than hoping that generalization occurs, or attempting to train under all possible conditions, Stokes and Baer (1977) suggest training sufficient exemplars. Training should be conducted with multiple examples of antecedents, desired target responses, settings and supervisors. This includes varying materials used, directions, settings and supervisors.

General Case Instruction. Sprague and Horner (1984) and Horner and McDonald (1982) developed this procedure. It consists of defining the range of stimulus properties that prompt a particular response and then sequentially training them. The two articles by Horner et al. involved training students to spray tables in a cafeteria with a spray bottle before wiping them clean. Both the spray bottles and the types of tables were varied systematically, so that at the end of the experiments, students were generalizing spraying behavior to a wide range of both trained and novel tables and types of spray bottles.

Antecedent Cue Regulation. These types of prompts

precede a response by a subject to a training stimulus. They promote self-control, defined by Goetz and Etzel (1978) as responses made by an individual that, in turn, modify the individual's behavior. They can be self-mediated or externally mediated. Self-instructions are the most commonly used form of self-generated mediators. They have been used with the developmentally disabled to improve social behavior (Matson and Adkins, 1980), the acquisition of math skills (Albion and Salzberg, 1982) and speed of task completion in vocational contexts (Crouch, Rusch and Karlan, 1984). Initially, individuals with a developmental disability were trained to perform a task to a criterion level. They were then taught to self-instruct correct performance and follow this with correct task performance. The above-cited articles indicate that training individuals to produce self-instructional verbalizations prior to task performance enhances the level of task performance. A second variant of self-generated mediators is self-labels. Wacker and Berg (1983) improved the performance level of seven moderately mentally retarded adolescents by teaching them to label relevant dimensions of the objects they were sorting.

Externally Mediated Generators. These are antecedent stimuli that are provided by outside agents. These cues correct responding and eventually come to facilitate the exercise of independent control over the target behaviors by

subjects. Connis (1979); Sowers et al. (1985) and Wacker and Berg (1983) have used this procedure to train developmentally disabled adults to improve their performance of vocational and self-care tasks, as reviewed above. Here, pictures were used to prompt correct responding and outside supervision was faded or eliminated as the individual became more competent at self-prompting correct responding. The pictures need not be faded but serve as a prosthesis to be used by the individual whenever the skill must be used (Greene, 1989).

Consequent Procedures

Natural Maintaining Contingencies. Stokes and Baer (1977) urged that only naturally occurring consequences should be used during training, if at all possible. Trainees who move into work settings should only be rewarded with reinforcers readily available in the context of the work environment, such as money, social reinforcement, praise from bosses and peers, etc. Rusch, Schutz and Agran (1982) observed that employers usually stipulate entry level requirements for potential employees. Supported Work placement counselors should first survey employers for their expectations and then train their potential workers to perform those skills. This ensures that, as workers, they will be engaging in behaviors that normally receive reinforcement.

Naturally-occurring consequences may not control

responding initially for an individual. Natural consequences can be initially paired with reinforcers that have previously demonstrated effectiveness in controlling responding. When a particular target behavior comes under the control of these reinforcers, the extra reinforcer can be faded or gradually withdrawn so that only the naturally occurring reinforcer controls responding (Martin, Pallotta-Cornick; Johnstone and Goyos, 1980; McNally, Kompick and Sherman, 1984). Martin et al. (1980) used verbal praise paired with money as a reward. Initially, the money did not serve as a reinforcer for the 16 individuals in the study. Gradually, as praise was faded after being initially paired with money, the money itself began to control responding.

A final step is to train the supported employees to appropriately solicit readily available reinforcers from the work environment. This involves training workers in the types of reinforcers they can readily expect to receive in the work environment as well as how to request them in an appropriate manner and at appropriate times.

Intermittent Reinforcement Schedules. Rusch, Connors and Sowers (1978) indicated that not only are intermittent reinforcement schedules more resistent to extinction than are regular schedules, but such schedules, such as weekly paychecks, are naturally available in work contexts.

Self-Regulation of Consequences. A large number of studies in both the animal and human experimental literature

have noted that self-reinforcement may be used instead of external reinforcement for target behaviors (Bandura and Perloff, 1967; Bolstad and Johnson, 1972; Drabman, Spitalnick and O'Leary, 1973; Felixbrod and O'Leary, 1973; Glynn, 1970; Mahoney, Bandura, Dirks and Wright, 1974; Mahoney and Bandura, 1972; Turkewitz, Ironsmith and O'Leary, 1974; Sowers et al. 1985; McCuller, Salzberg and Lignugaris/Kraft, 1987; Agran, Fodor-Davis, Moore and Deer, 1989; McNally, Kompik and Sherman, 1984; Haywood and Switzky, 1985). A number of these articles indicate that subjects may work harder for the opportunity to provide reinforcement themselves rather than receiving it through an outside agent. All of the above studies acknowledge the effectiveness of self-reinforcement in changing the frequency of target behaviors. Helland, Palluck and Klein (1976) compared self-reinforcement to supervisor delivered reinforcement in a sheltered workshop for the mentally retarded. They found that both groups effectively increased productivity, but the self-delivered group worked independently of their supervisors, an extremely desirable quality in Supported Work settings as well as in the sheltered workshop. As discussed previously in the review of self-reinforcement literature, several theorists ascribe the effectiveness of the self-reinforcement process to the implicit social praise inherent in the self-reinforcement process, while others see it as operating in a manner

analogous to reinforcement by an outside agent. In order to maximize the effectiveness of self-reinforcement in the absence of contingent social praise from the trainer, as is the case in a supported work setting, it would seem important to separate the two effects. This point needs to be emphasized more strongly. The current study is not attempting to investigate self-reinforcement as an academic exercise. In this case, the social reinforcement associated with the trainees meeting a goal will not necessarily occur in the absence of the trainer. It is critical to ascertain whether the self-delivery of reinforcers operates in a manner analogous to operant reinforcement in order to determine if the intervention package has any value above and beyond occasioning the prompts associated with probes. If not, then self-reinforcement can be deleted from the intervention package and the natural contingencies in the environment that could maintain target behaviors need to be made more salient to the trainee. If it does, then the great pains associated with development of self-control and ensuring a high level of inter-observer agreement are warranted. Since virtually all self-management packages discussed in the literature utilize this as an adjunct, it seems surprising that this question has not been explored up to this point and the need for such exploration is well justified.

The ability of individuals to self-monitor their level

of on- task behavior has been extensively explored. Thomas (1976) explored the ability of non-disabled elementary school children to accurately assess their own level of on task behavior. Average level of accuracy for the 31 children in the study was 78%, and the author demonstrated that subjects were also able to self-reinforce for improved levels. The observed changes were not only significant, but were also durable over time. Hughes and Peterson (1989) used self-instructions and self-reinforcement to increase the level of on task behavior in four moderately retarded adults in a sheltered workshop and found that their ability to report their level of on task behavior accurately was excellent, and that self-reinforcement for goal attainment increased their levels above that of baseline. Morrow, Burke and Buell (1985) performed a similar investigation with two mentally retarded adults and found the level of interobserver agreement between subjects and independent observers to be over 90%. Further, in the context of an academic task, on task behavior increased from a baseline of 40% to close to 80%, a level maintained upon 30 day follow up probes. Burgio, Whitman and Johnson (1980) found a similarly high level of inter-observer agreement on self-assessment of on task behavior in 4 mentally retarded adolescents after training, and was also able to significantly increase their level of on task behavior above baseline levels using a self-reinforcement procedure.

Hill, Wehman, Hill and Goodall (1986); Shafer, Brooke and Wehman (1991); Kiernan, Mcgaughey and Schalock (1988); McDonnell, Nofs, Hardman and Chambliss (1989); Wacker, Fromm-Steege, Berg and Flynn (1989); Brickey, Campbell and Browning (1985); Lignugaris/Kraft, Salzberg, Rule and Stowitschek (1988) all report that the major reasons for job rejection of supported employees lie not in the basic performance of job tasks or their inability to learn complex behavioral chains, but rather in the areas of ability to function without supervision and the development of appropriate social skills. Johnson and Rusch (1990) emphasize the point that programs need to be able to withdraw direct supervision of their trainees at their job sites in order for them to succeed as employees. The author states that this is due to two factors. The first and more obvious one is that rapid development of independence is related to the development of competence. The second factor is that the rapid withdrawal of supervisory personnel facilitates the development of supporting services that are provided by regular employees at the work site, such as coworkers, employer's supervisory personnel, union representatives, employee assistance programs, etc. The use of naturally occurring resources to help support the employment of disabled adults has been explored extensively in the literature in the past few years (Linkins, Salzberg and Stowitschek, 1989; Shafer, 1986; Rusch and Minch, 1990).

This strategy is stressed as an effective strategy to help prevent the development of over dependence on an employment trainer or job coach. Wacker, Fromm-Steege, Berg and Flynn (1989) performed a component analysis of factors that predicted positive supported employment outcomes. These include the existence of a mechanism for regular and long term delivery of prompts as well as the existence of a consistent maintenance plan.

APPENDIX C

Literature Review of Methodological Issues

Time Sampling

In the current study, the investigator used a momentary time sampling procedure on a fixed interval schedule (30 minutes) to monitor the level of on task behavior, that is the main dependent measure under study. Several methodological issues are pertinent to the present study, including: time sampling rate; time sampling length and the representativeness of the resulting data; methods for calculating inter-observer agreement; design and observational methods.

Observation Procedures

Kazdin (1979) argues that since most behavioral assessments entail the direct observation of overt behavior, the result is that most investigations involve obtrusive observations. If the subject is aware that he/she is being observed, then the observed behavior may not accurately reflect behavior that would occur in the absence of the observation process. This is the problem of reactivity. The author goes on to state that if the resulting data are going to be used to represent the usual behavior of the subject when he/she is not being observed, then great pains must be taken to minimize the impact of the observational process on the dependent measure. Several methods are suggested, including the use of archival data, physical

artifacts and the use of indirect observational techniques. One method discussed is the use of observers that are common to the natural environment in which the target behavior(s) are exhibited. The current study utilized observers whose presence and scrutiny of subject's performance in the natural environment is routine. These are the group's supervisor and other non-handicapped employees. This strategy minimizes the reactive nature of the observations, but does not completely eliminate it. A factor that complicates the procedure is the presence of this author for the purpose of assessing interobserver agreement during approximately 10% of trials. The presence of this author could have had the effect of prompting on-task behavior due to the obtrusive nature of his presence during interobserver agreement probes, and therefore may represent a confounding variable.

Kazdin (1979) also cautions that any observation process raises the ethical issue of whether the subject agrees to being observed, either obtrusively or unobtrusively. In the current study, all subjects agreed in advance to the need to increase their levels of on task behavior and were informed in advance of the procedures to be used to collect data and agreed to these procedures by giving their informed consent to the process.

Kazdin (1977), in evaluating sources of artifact and bias in assessing reliability, identifies four possible

sources that can confound the results of experiments. These are reactivity of reliability assessment, observer drift, complexity of recording codes and observer expectancies. The first of these sources of bias (reactivity of reliability assessment) is discussed above and remains a potential confound. Observer expectations and drift are also germane to the current experiment. These two factors are at least partially addressed through the use of a third observer, a non-handicapped employee, to gather data on a second dependent measure. The productivity measure gathered by this observer is used to evaluate the effect of both of these potential sources of bias on observations made by the subjects' supervisor.

Research Design

Kazdin (1982) discusses potential problems with several commonly used experimental designs, such as alternating treatments, multiple baseline and the use of control groups. The alternating treatments design makes several assumptions that may not be warranted, such as the reversibility of effect during reversals. Kazdin (1982) states that where this issue is a problem, a multiple baseline design can be used. This design, however, suffers from the problem that the effect of treatment of one subject may influence other subjects. Here, he suggests the use of a control to determine if this is occurring. The current study utilized a modification of the alternating treatment design along

with a multiple baseline design and the use of a control subject to determine the effect these sources of confound have on the dependent measures. This solution is also suggested in an article by Wacker, Fromm-Steege, Berg and Flynn (1989).

Sampling Rate and Sampling Duration

Johnston and Pennypacker (1980) observe that Time Sampling procedures present some serious problems that may result in unrepresentative data. These include poor representation of long or short duration, low frequency behaviors, the inability to discern patterns of responding over time, the need to oversimplify operational definitions of behaviors to fit into presence or absence categories rather than more descriptive functional categories, as well as several other potentially vexing problems.

Harrop and Daniels (1986) noted that although the above identified problems with momentary time sampling as a data collection system exist, it is nevertheless widely used in applied settings because of its inherent economy. They used an empirical method to compare momentary time sampling to partial interval recording. A computer generated several sets of data, that, in effect, simulated continuous recording of hypothetical behaviors. These data were then broken up into intervals and the performance of the two methods then compared to the continuous recording of the behavior. The authors found that duration rather than rate

should be used to estimate absolute levels of behaviors, and that momentary time sampling was better at representing the absolute level of target behavior when similar interval lengths were used. The authors distinguished between sensitivity of the methods, defined as the ability of the measure to accurately reflect small changes in the rate of behavior, and linearity, defined as a measure of the degree to which the average of the recorded rates would be similar in shape to a graph of the average actual rates. For behaviors with high rates, (as is the case in the current experiment), linearity representativeness is higher in momentary time sampling than with partial interval recording. In conclusion, the authors observe that although momentary time sampling is slightly less sensitive, it did not appear to suffer from systematic error in estimating relative change, that is a critical attribute in the context of the current experiment. Harrop and Daniels (1985) did not assess the effect varying the length of the interval had on these variables.

Mudford, Beale and Singh (1990) investigated the representativeness of observational samples of different durations. The authors found that the representativeness of the samples varied as a function of interval length. The authors also concluded that it was imperative for researchers to routinely empirically assess the adequacy of the data collection system.

Sanson-Fisher, Poole and Dunn (1980) examined the effects of various interval lengths on the representativeness of data obtained using time sampling method of data collection in a psychiatric population. They found that increasing the interval length had little effect on the duration measure, suggesting that the interval recording method is a relatively stable method of data collection. The authors reached several conclusions from their examination of the data. They stated that the current practice of arbitrarily decreasing interval length to increase representativeness of resulting data is not necessarily valid. Further, that determining interval length in a arbitrary manner, or on the basis of what they called convention, should not be used by researchers in applied settings. They strongly emphasize that interval length should be determined empirically by each investigator. Finally, they suggested a method by which this could be accomplished. Their suggestion was that prior to beginning an experiment, researchers should get a continuous recording of the target behavior(s) of the subject in the environment in which the investigation will be conducted. The duration of this continuous record should be as long as possible. The record can then be divided into intervals of varying length. The presence or absence of target behavior(s) at the end of each of the intervals would then be recorded, and compared to the total percentage of

intervals in which the behavior(s) is observed via the continuous recording method. The interval length having the highest correlation with the continuously recorded behavioral sample should then be selected, of course, taking into consideration the constraints of economy and feasibility. This method was utilized in the current experiment in order to determine an acceptable interval length, given feasibility constraints.

Saudargas and Zanolli (1990) used the above methodology to assess how closely a momentary time sampling data collection method approximated the actual percentage of time across a range of naturally occurring behaviors in the field setting of an elementary school. Using a fairly brief time sampling interval (i.e., 15 seconds), they observed that data from momentary time sampling correlated very well with data from the continuous recording method. In their examination of several different behaviors, they further concluded that in order to capture very short duration behaviors, a brief interval should be used, while long duration behaviors were accurately reflected by even very long momentary time intervals. Since the dependent measure in the current study, (i.e., on-task behavior), is a very long duration behavior, this research suggests that an interval of 30 minutes might be representative, especially if it is empirically validated.

Methods of Calculating Interobserver Agreement

Kelly (1977) analyzed the observational, data collection and reliability procedures reported in the Journal of Applied Behavior Analysis between 1968 - 1975. He found that three quarters of the studies reported observational data, consisting of event recording, trial scoring, interval recording or time sampling recording. Virtually all of the studies reviewed reported assessment of interobserver agreement, mostly using the point-by-point percentage agreement method of calculation. In approximately half of the studies reviewed, the level of interobserver agreement obtained exceeded 90 percent. Fewer than twenty-five percent of the studies reviewed stated that interobserver agreement data was collected during each phase of the study. The author stressed that this final observation pointed to the need for increased rigor on the part of researchers in applied behavior analysis. This was done in the current experiment.

The issue of selecting a method for calculating interobserver agreement is extremely controversial. Hartman (1977); Hopkins and Herman (1977) and Kratochwill and Wetzel (1977) all describe the various methods of calculating interobserver agreement data. These include percent agreement, point by point agreement, agreement on occurrence and non-occurrence, Cohen's kappa and the phi coefficient, as well as inferential statistics such as the chi square. These authors all point out that the same data can yield

varying correlations depending on which method is used. They further state that no one method is better than the others in all situations. An article by Baer (1977) titled "Just Because It's Reliable Doesn't Mean That You Can Use It" goes so far as to say that the concept of a true measure of reliability that is estimated using the above listed methods is misleading, and that the variability among methods only shows up the arbitrary nature of all of the methods.

The current study used a coding system for the dependent measure with only three categories. As the number of codes decreases, the possibility of chance agreement increases proportionally. Therefore, the possibility of chance agreement would artificially inflate any of the simpler methods of calculating interobserver agreement. As a result, Cohen's kappa (Kazdin, 1982) has been selected as the method of calculation that yields a conservative measure of agreement that controls for the possibility of chance agreement. As a cautionary note, an article by Deitz (1988) suggests that these debates over interobserver agreement should not be stressed at the expense of accuracy of data. He quotes a statement by the philosopher Wittgenstein (1953) relating to the question of stressing agreement at the expense of accuracy "as if someone were to buy several copies of the morning paper to assure himself that what it said was true."

Other Methodological Issues

Mace (1990) examined a collateral effect of reward predicted by matching theory. Here, the authors caution that the Matching Law (Herrnstein, 1970) predicts decreases in the frequency of other responses that may have some functional or adaptive value when one particular target response is differentially reinforced.

Peterson, Homer and Wonderlich (1982) stress the importance of taking pains to maintain the integrity of the independent as well as the dependent variables in an experimental investigation. The authors state that in virtually all studies, researchers go to great lengths to ensure that the dependent variable is closely monitored and accurately and reliably measured. Most studies, though, do not pay the same level of attention to ensuring that the independent variable is accurately and reliably measured. Here, Peterson et. al. (1982) suggest that data collection routinely include probes of whether the independent variable is being consistently applied. In the current study, this involves the observer noting whether the subject actually responds to the alarm of the timer and whether they actually open the device and self-administer a reward contingently upon the presence of on task behavior. These data for the current experiment are presented in the results section.

REFERENCES FOR APPENDIX A

Bandura, A., (1977). Social learning theory. New Jersey: Prentice-Hall.

Bandura, A. (1978). A self-system in reciprocal determinism. American Psychologist, 33, 344-358.

Bandura, A. (1981). Self-referent thought: A developmental analysis of self-efficacy. In J.H. Flavell & I. Ross, (Eds.), Social cognitive development, Cambridge, England: Cambridge University Press.

Bandura, A., & Perloff, B. (1967). Relative efficacy of self-monitored and externally posed reinforcement systems. Journal of Personality and Social Psychology, 7, 11-16.

Bolstad, O. D., & Johnson, S. M. (1972). Self-regulation in the modification of disruptive classroom behavior. Journal of Applied Behavior Analysis, 5, 443-454.

Catania, A. C. (1975). The myth of self-reinforcement. Behaviorism, 3, 192-199.

Drabman, R. S., Spitalnick, R., & O'Leary, K. D. (1973). Teaching self-control to disruptive children. Journal of Abnormal Psychology, 82, 10-16.

Felixbrod, J. J., & O'Leary, K. D. (1973). Effects of reinforcement on children's academic behavior as a function of self-determined and externally imposed contingencies. Journal of Applied Behavior Analysis, 6, 241-250.

Glynn, E. L. (1970). Classroom applications of self-determined reinforcement. Journal of Applied Behavior Analysis, 3, 122-132.

Goetz, E., & Etzel, B. (1982). A brief review of self-control procedures: Problems, solutions. Behavior Therapy, 1, 58-69.

Goldiamond, I. (1976). Self-reinforcement. Journal of Applied Behavior Analysis, 9, 509-514.

Hayes, S. C., Rosenfarb, I, Wulfert, E, Munt, E. D., Korn, Z., & Zettle, R. D. (1985). Self-reinforcement effects: An artifact of social standard setting? Journal of Applied Behavior Analysis, 18, 201-214.

Herrnstein, R. J. (1970). On the law of effect. Journal of the Experimental Analysis of Behavior, 13, 243-266.

Kanfer, F. H. (1971). The maintenance of behavior by self-generated stimuli and reinforcement. In A. Jacobs and L. B. Sachs, (Eds.), The psychology of private events., New York: Academic Press.

Kanfer, F. H., & Phillips, J. S. (1979). Learning foundations of behavior therapy. New York: Wiley and Sons.

Karoly, P. (1982) Perspectives on self-management and behavior change. In P. Karoly and F. H. Kanfer, (Eds.), Self-management and behavior change. New York: Pergamon Press.

Karoly, P., & Kanfer, F. H. (1982). Self-management and behavior change. New York: Pergamon Press.

Kazdin, A. E. (1982). Single case research designs. New York: Oxford University Press.

Keppel, G. (1982). Handbook of design and analysis. New Jersey: Prentice-Hall.

Lovitt, T. C., & Curtiss, K. A. (1969). Academic response rate as a function of teacher and self-imposed contingencies. Journal of Applied Behavior Analysis, 2, 49-53.

Mahoney, M. J., & Bandura, A. (1972). Self-reinforcement in pigeons. Learning and Motivation, 3, 293-303.

Mahoney, M. J., Bandura, A., Dirks, S. J., & Wright, C. L. (1974). Relative preference for external and self-controlled reinforcement in monkeys. Behavior Research and Therapy, 12, 157-163.

Nelson, R. D., Hayes, S. C., Spong, R. T., Jarrett, R. B., & McKnight, D. L. (1983). Self-reinforcement: Appealing misnomer or appealing mechanism. Behavior Research and Therapy, 21, 557-566.

Neuringer, A. J. (1967). Animals respond for food in the presence of free food. Science, 166, 399-401.

Ninnes, H. A. C., Fuerst, J, Rutherford, R. D., & Glenn, S. S. (1991). Effects of self-management training and reinforcement on the transfer of improved conduct in the absence of supervision. Journal of Applied Behavior Analysis, 24, 499-508.

Rachlin, H., (1974). Self-control. Behaviorism, 12, 94-107.

Rachlin, H., & Green, L. (1972). Commitment, choice and self-control. Journal of the Experimental Analysis of Behavior, 17, 15-22.

Skinner, B. F. (1953). Science and Human Behavior. New York: Collier-McMillan.

Thoresen, C. E., & Wilbur, C. S. (1976). Some encouraging thoughts about self-reinforcement. Journal of Applied Behavior Analysis, 9, 518-520.

Turkewitz, H., O'Leary, K. D., & Ironsmith, R. (1974). Producing generalization and maintaining appropriate behavior through self-control. Journal of Consulting and Clinical Psychology, 18, 577-583.

Whitman, T. L. (1987). Self-instruction, individual differences and mental retardation. American Journal of Mental Deficiency, 92, 213-223.

Zeaman, D., & House, B. (1963). The role of attention in retardate discrimination learning. In N. R. Ellis, (Ed.), Handbook of mental deficiency. New York: McGraw-Hill.

REFERENCES FOR APPENDIX B

Albion, F., & Salzberg, C. (1982). The effect of self-instructions on the rate of correct addition problems with mentally retarded persons. Education and Treatment of Children, 5, 212-131.

Baer, D. M. (1977). Reviewer's comment: Just because it's reliable doesn't mean that you can use it. Journal of Applied Behavior Analysis, 10, 117-119.

Bandura, A., & Perloff, B. (1967). Relative efficacy of self-monitored and externally posed reinforcement systems. Journal of Personality and Social Psychology, 7, 11-16.

Bellamy, G. T., Horner, R., & Inman, D. (1978). Vocational habilitation of severely retarded adults. Baltimore: University Park Press.

Bolstad, O. D., & Johnson, S. M. (1972). Self-regulation in the modification of disruptive classroom behavior. Journal of Applied Behavior Analysis, 5, 443-454.

Brickey, M., Campbell, K., & Browning, L. (1985). Five year follow up of sheltered workshop employees. Mental Retardation, 23, 67-73.

Brown, L. (1973). Instructional programs for trainable-level retarded students. In L. Mann & D. Sabatino, (Eds.), The first review of special education, Vol. 2, Philadelphia, Journal of Special Education Press.

Connis, R. T. (1979). The effects of sequential pictorial cues, self-recording and praise on the job task sequencing of retarded adults. Journal of Applied Behavior Analysis, 12, 355-371.

Crouch, K. P., Rusch, F. R., & Karlan, G. R. (1984). Competitive employment: Utilizing the correspondence training paradigm to enhance productivity. Education and Training of the Mentally Retarded, 19, 268-275.

Drabman, R. S., Spitalnick, R., & O'Leary, K. D. (1973). Teaching self-control to disruptive children. Journal of Abnormal Psychology, 82, 10-16.

Felixbrod, J. J., & O'Leary, K. D. (1973). Effects of reinforcement on children's academic behavior as a function of self-determined and externally imposed contingencies. Journal of Applied Behavior Analysis, 6, 241-250.

Ferguson, D. L., & Ferguson, P. M. (1986). The new victors: A progressive policy analysis of work reform for people with severe handicaps. Mental Retardation, 26, 331-336.

Glynn, E. L. (1970). Classroom applications of self-determined reinforcement. Journal of Applied Behavior Analysis, 3, 122-132.

Goetz, E., & Etzel, B. (1982). A brief review of self-control procedures: Problems, solutions. Behavior Therapy, 1, 58-69.

Gold, M. (1972). Stimulus factors in skill training of the retarded on a complex assembly task: Acquisition, transfer and retention. American Journal of Mental Deficiency, 76, 517-526.

Gold, M. (1973). Factors affecting production by the retarded: Base rates. Mental Retardation, 11, 9-11.

Greene, B. (1989). Editor's page: The integration of individuals with disabilities into the vocational mainstream. Journal of Applied Behavior Analysis, 22, 343-345.

Hawkins, R. P. (1991). Is social validity what we are interested in? Argument for a functional approach. Journal of Applied Behavior Analysis, 24, 205-214.

Haywood, H. C., & Switzky, H. N. (1985). Work response of mildly mentally retarded adults to self- versus external regulation as a function of motivational orien tation. American Journal of Mental Deficiency, 90, 151-159.

Helland, C., Palluck, R., & Klein, M. (1976). A comparison of self- and external reinforcement with the trainable retarded. Mental Retardation, 14, 22-23.

Hill, J. W., Wehman, P., Hill, M., & Goodall, P. (1986). Differential reasons for job separation by previously employed persons with mental retardation. Mental Retardation, 24, 347-352.

Horner, R. H., & McDonald, R. S. (1982). Comparison of single instance and general case instruction in teaching a generalized vocational skill. Journal of the Association for the Severely Handicapped, 7, 7-20.

Hughes, C., & Peterson, D. L., (1989). Utilizing a self-instructional training package to increase on-task

behavior and work performance. Education and Training of the Mentally Retarded, 16, 114-120.

Johnson, J. R., & Rusch, F. R. (1990). Analysis of hours of direct training provided by employment specialists to supported employees. American Journal on Mental Retardation, 94, 674-682.

Kiernan, W. E., McGaughey, M. J., & Schalock, R. L. (1988). Employment environments and outcomes for adults with disabilities. Mental Retardation, 26, 279-287.

Lignugaris/Kraft, B., Rule, S., Salzberg, C. L., & Stowitschek, J. T. (1986). Social-interpersonal skills of handicapped and nonhandicapped adults at work. Journal of Employment Counseling, 23, 20-30.

Lignugaris/Kraft, B., Salzberg, C. L., Rule, S., & Stowitschek, J. T. (1988). Social-vocational skills of developmentally disabled and nonhandicapped workers in two community employment sites. Mental Retardation, 26, 297-305.

Lindsley, O. (1966). Geriatric behavioral prosthetics. In R. Ulrich, T. Stachnik & J. Mabry, (Eds.), Control of human behavior. p. 156-168. New York: Scott Foresman.

Linkins, M., Salzberg, C. L., Stowitschek, J. T, Lignugaris/Kraft, B., & Curl, R. (1989). Co worker implemented job training: The use of coincidental training and quality control checking procedures on the food preparation skills of employees with mental retardation. Journal of Applied Behavior Analysis, 22, 381-394.

Mahoney, M. J., & Bandura, A. (1972). Self-reinforcement in pigeons. Learning and Motivation, 3, 293-303.

Mahoney, M. J., Bandura, A., Dirks, S. J., & Wright, C. L. (1974). Relative preference for external and self-controlled reinforcement in monkeys. Behavior Research and Therapy, 12, 157-163.

Martin, G., Pallotta-Cornick, A., Johnstone, G., & Goyos, C. (1980). A supervisory strategy to improve work performance by lower functioning retarded clients in a sheltered workshop. Journal of Applied Behavior Analysis, 13, 183-190.

Matson, J., & Adkins, J. (1980). A self-instructional social skills training program for mentally retarded persons. Mental Retardation, 18, 245-248.

McCuller, G. L., Salzberg, C. L., & Lignugaris/Kraft, B. (1987). Producing generalized job initiative in severely mentally retarded sheltered workers. Journal of Applied Behavior Analysis, 20, 413-420.

McDonnell, J., Nofs, D., Hardman, M., & Chambless, C. (1989). An analysis of the procedural components of supported employment programs associated with employment outcomes. Journal of Applied Behavior Analysis, 22, 417-428.

Morrow, L. W., Burke, J. G., & Buell, B. J. (1985). Effects of a self-recording procedure on the attending to task behavior and academic productivity of adolescents with multiple handicaps. Mental Retardation, 23, 137-141.

Nirje, B. (1976). The normalization principle and its human management implications. In M. Rosen, G. R. Clark & M. S. Kivitz, (Eds.), The history of mental retardation, Vol. 2. Baltimore: University Park Press.

Rusch, F. R. (1986). Competitive employment issues and strategies. Baltimore: Paul H. Brookes.

Rusch, F. R., Connors, R. T., & Sowers, J. (1978). The modification and maintenance of time spent attending to task using social reinforcement, token reinforcement and response cost in an applied restaurant setting. Journal of Special Education Technology, 2, 18-26.

Rusch, F. R., & Hughes, C. (1988). Supported employment: Promoting employee independence. Mental Retardation, 26, 351-355.

Rusch, F. R., & Hughes, C. (1989). Overview of supported employment. Journal of Applied Behavior Analysis, 22, 351-372.

Rusch, F. R., & Minch, K. E. (1988). Identification of co-workers involvement in supported employment: A review and analysis. Research in Developmental Disabilities, 9, 247-254.

Rusch, F. R., Schutz, R., & Agran, M. (1982). Validating entry level survival skills for service occupations: Implications for curriculum development. Journal of the Association for the Severely Handicapped, 7, 32-41.

Salzberg, C. L., Agran, M., & Lignugaris/Kraft, B. (1986). Behaviors that contribute to entry level employment: A

profile of five jobs. Applied Research in Mental Retardation, 7, 299-314.

Salzberg, C. L., Lignugaris/Kraft, B., & McCuller, G. L. (1988). Reasons for job loss: A review of employment termination studies of mentally retarded workers. Research in Developmental Disabilities, 9, 153-170.

Sands, H., Radin, J. (1978). The mentally disabled rehabilitant: Post-employment services. Washington D. C.: Rehabilitation Services Administration, Department of Health, Education and Welfare.

Shafer, M. S. (1986). Utilizing co-workers as change agents. In F. Rusch (Ed.), Competitive employment issues and strategies. Baltimore: Paul H. Brookes.

Shafer, M. S., Brooke, V., & Wehman, P. (1985). Developing appropriate social-interpersonal skills in mentally retarded workers. Vocational Evaluation and Work Adjustment Bulletin, 18, 76-81.

Sowers, J. A., Verdi, M., Bourbeau, P., & Sheehan, M. (1985). Teaching job independence and flexibility to mentally retarded students through the use of a self-control package. Journal of Applied Behavior Analysis, 18, 81-85.

Sprague, J., & Horner, R. (1984). The effects of single instance, multiple instance and general case training on generalized vending machine usage by moderately and severely handicapped students. Journal of Applied Behavior Analysis, 17, 273-278.

Stark, J., & Goldsbury, T. (1988). Analysis of labor and economics: Needs for the next decade. Mental Retardation, 26, 363-368.

Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. Journal of Applied Behavior Analysis, 10, 349-367.

Thomas, J. D. (1976). Accuracy of self-assessment of on-task behavior by elementary school children. Journal of Applied Behavior Analysis, 9, 209-210.

Turkewitz, H., O'Leary, K. D., & Ironsmith, R. (1974). Producing generalization and maintaining appropriate behavior through self-control. Journal of Consulting and Clinical Psychology, 18, 577-583.

Turnbull, A. P., & Turnbull, R. T. (1988). Toward great expectations for vocational opportunities: Family-profe ssional partnerships. Mental Retardation, 26, 337-342.

Wacker, D. P., Berg, W. K. (1983). Effects of picture prompts on the acquisition of complex vocational tasks by mentally retarded adolescents. Journal of Applied Behavior Analysis, 16, 417-433.

Wacker, D. P., & Berg, W. K. (1986). Generalizing and maintaining work behavior. In F. Rusch, (Ed.), Competitive employment issues and strategies. Baltimore: Paul H. Brookes.

Wacker, D. P., Fromm-Steege, L., Berg, W. L., & Flynn, T. H. (1989). Supported employment as an intervention package: A preliminary analysis of functional variables. Journal of Applied Behavior Analysis, 22, 429-440.

Wehman, P. (1988). Supported employment: Toward equal employment opportunity for persons with severe disabilities. Mental Retardation, 26, 357-362.

Wolfensberger, W. (1968). The origin and nature of our institutional models. Syracuse, New York: Human Policy Press.

Wyatt v Stickney and New York State Association of Retarded Children v Rockefeller, in Repp, A. C. (1983) Teaching the mentally retarded. p. 81. New Jersey, Prentice-Hall.

REFERENCES FOR APPENDIX C

Deitz, S. M. (1988). Another's view of observer agreement and observer accuracy. Journal of Applied Behavior Analysis, 21, 113-114.

Harrop, A, & Daniels, M. (1986). Methods of time sampling: A reappraisal of momentary time sampling and partial interval recording. Journal of Applied Behavior Analysis, 19, 73-77.

Hartman, D. P. (1977). Considerations in the choice of interobserver reliability estimates. Journal of Applied Behavior Analysis, 10, 103-116.

Herrnstein, R. J., (1970). On the law of effect. Journal of the Experimental Analysis of Behavior, 13, 243-266.

Hopkins, B. L., & Herman, J. A. (1977). Evaluating inter-observer reliability of interval data. Journal of Applied Behavior Analysis, 10, 121-126.

Johnston, J. J., & Pennypacker, H. S. (1980). Strategies and tactics of human behavioral research. Hillsdale, New Jersey: Lawrence Erlbaum.

Kazdin, A. E., (1977). Artifact, bias and complexity of assessment: The abc's of reliability. Journal of Applied Behavior Analysis, 10, 141-150.

Kazdin, A. E. (1979). Direct observations as unobtrusive measures in treatment evaluation. New Directions for Methodology of Behavioral Science, 1, 19-31.

Kazdin, A. E. (1982). Single case research designs. New York: Oxford University Press.

Kelly, M. B. (1977). A review of observational data collection and reliability procedures reported in the Journal of applied behavior analysis. Journal of Applied Behavior Analysis, 10, 97-101.

Kratochwill, T. R, & Wetzel, R. J. (1977). Observer agreement, credibility and judgement: Some considerations in presenting observer agreement data. Journal of Applied Behavior Analysis, 10, 133-139.

Mace, F. C. (1990). A collateral effect of reward predicted by matching theory. Journal of Applied Behavior Analysis, 23, 197-206.

Mudford, O. C., Beale, I. L., & Singh, N. N. (1990). The representativeness of observational samples of different duration. Journal of Applied Behavior Analysis, 23, 323-332.

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

Sanson-Fisher, R. W., Poole, A. D., & Dunn, J. (1980). An empirical method for determining an appropriate interval length for recording behavior. Journal of Applied Behavior Analysis, 13, 493-500.

Saudargas, R. A., & Zanolli, K. (1990). Momentary time sampling as an estimate of percentage time: A field validation. Journal of Applied Behavior Analysis, 23, 533-538.

Wittgenstein, L. (1953). Philosophical investigations. From Deitz, S. M., (1988). Another's view of observer agreement and observer accuracy. Journal of Applied Behavior Analysis, 21, 113-114.

BIBLIOGRAPHY

Agran, M., Fodor-Davis, J., Moore, S., & Deer, M. (1989). The application of a self-management program on instruction following skills. Journal of the Association for Persons with Severe Handicaps, 14, 147-154.

Albion, F., & Salzberg, C. (1982). The effect of self-instructions on the rate of correct addition problems with mentally retarded persons. Education and Treatment of Children, 5, 212-131.

Baer, D. M., Wolf, M., & Risley, T. (1968). Some current dimensions of applied behavior analysis. Journal of Applied Behavior Analysis, 1, 91-97.

Bandura, A. (1977a). Self-efficacy: Toward a unifying theory of behavioral change. Psychological Review, 84, 191-215.

Bandura, A. (1977b). Social learning theory. New Jersey: Prentice-Hall.

Bandura, A. (1978). A self-system in reciprocal determinism. American Psychologist, 33, 344-358.

Bandura, A. (1981). Self-referent thought: A developmental analysis of self-efficacy. In J.H. Flavell & I. Ross, (Eds.), Social cognitive development, Cambridge, England: Cambridge University Press.

Bellamy, G. T. (1975). Habilitation of the severely and profoundly retarded: A review of research on work productivity. In T. Bellamy, (Ed.), Habilitation of severely and profoundly retarded adults: Reports from the specialized training program. Eugene, Oregon: University of Oregon Center on Human Development.

Bijou, S. W., Peterson, R. F., & Ault, M. H. (1968). A method to integrate descriptive and experimental field studies at the level of data and empirical concepts. Journal of Applied Behavior Analysis, 1, 175-191.

Bolstad, O. D., & Johnson, S. M. (1972). Self-regulation in the modification of disruptive classroom behavior. Journal of Applied Behavior Analysis, 5, 443-454.

Burgio, L. D., Whitman, T. L., &Johnson, M. L. (1980). A self-instructional package for increasing attending behavior in educable retarded children. Journal of Applied Behavior Analysis, 13, 443-459.

Catania, A. C. (1975). The myth of self-reinforcement. Behaviorism, 3, 192-199.

Cook, T. D., & Campbell, D. T. (1979). Quasi-Experimentation. Boston: Houghton Mifflin.

Drabman, R. S., Spitalnick, R., & O'Leary, K. D. (1973). Teaching self-control to disruptive children. Journal of Abnormal Psychology, 82, 10-16.

Edwards, A. L. (1976). An introduction to linear regression and correlation. New York, W. H. Freeman and Co.

Glynn, E. L. (1970). Classroom applications of self-determined reinforcement. Journal of Applied Behavior Analysis, 3, 122-132.

Goldiamond, I. (1976). Self-reinforcement. Journal of Applied Behavior Analysis, 9, 509-514.

Hawkins, R. P. (1991). Is social validity what we are interested in? Argument for a functional approach. Journal of Applied Behavior Analysis, 24, 205-214.

Hayes, S. C., Rosenfarb, I., Wulfert, E., Munt, E. D., Korn, Z.,& Zettle, R. D. (1985). Self-reinforcement effects: An artifact of social standard setting? Journal of Applied Behavior Analysis, 18, 201-214.

Herrnstein, R. J., (1970). On the law of effect. Journal of the Experimental Analysis of Behavior, 13, 243-266.

Hughes, C., & Peterson, D. L., (1989). Utilizing a self-instructional training package to increase on-task behavior and work performance. Education and Training of the Mentally Retarded, 16, 114-120.

Karoly, P. (1982) Perspectives on self-management and behavior change. In P. Karoly and F. H. Kanfer, (Eds.), Self-management and behavior change. New York: Pergamon Press.

Kazdin, A. E. (1982). Single case research designs. New York: Oxford University Press.

Lovitt, T. C., & Curtiss, K. A. (1969). Academic response rate as a function of teacher and self-imposed contingencies. Journal of Applied Behavior Analysis, 2, 49-53.

Mahoney, M. J., & Bandura, A. (1972). Self-reinforcement in pigeons. Learning and Motivation, 3, 293-303.

McNally, R. J., Kompik, J. J., & Sherman, G. (1984). Increasing the productivity of mentally retarded workers through self-management. Analysis and Intervention in Developmental Disabilities, 4, 129-135.

Nelson, R. D., Hayes, S. C., Spong, R. T., Jarrett, R. B., & McKnight, D. L. (1983). Self-reinforcement: Appealing misnomer or appealing mechanism. Behavior Research and Therapy, 21, 557-566.

Ninnes, H. A. C., Fuerst, J, Rutherford, R.D., & Glenn, S. S. (1991). Effects of self-management training and reinforcement on the transfer of improved conduct in the absence of supervision. Journal of Applied Behavior Analysis, 24, 499-508.

Repp, A. C. (1983). A history of mental retardation. In Repp, A. C., Teaching the mentally retarded. New Jersey: Prentice-Hall.

Rescorla, R. A., & Solomon, R. L. (1967). Two process learning theory: relationships between Pavlovian conditioning and instrumental learning. Psychological Review, 74, 151-182.

Schwartz, I. S., & Baer, D. M. (1991). Social validity assessment: Is current practice state of the art? Journal of Applied Behavior Analysis, 24, 189-204.

Sowers, J. A., Verdi, M., Bourbeau, P., & Sheehan, M. (1985). Teaching job independence and flexibility to mentally retarded students through the use of a self-control package. Journal of Applied Behavior Analysis, 18, 81-85.

Thoresen, C. E., & Wilbur, C. S. (1976). Some encouraging thoughts about self-reinforcement. Journal of Applied Behavior Analysis, 9, 518-520.

Zeaman, D., & House, B. (1963). The role of attention in retardate discrimination learning. In N. R. Ellis, (Ed.), Handbook of mental deficiency. New York: McGraw-Hill.