JOURNAL OF APPLIED BEHAVIOR ANALYSIS 1972, 5, 443-454 NUMBER 4 (WINTER 1972)

SELF-REGULATION IN THE MODIFICATION OF DISRUPTIVE CLASSROOM BEHAVIOR[1]

ORIN D. BOLSTAD AND STEPHEN M. JOHNSON

UNIVERSITY OF OREGON

This study compared self-regulation and external regulation procedures in the treatment of children's disruptive classroom behavior. After baseline data were collected, three of the four most disruptive children in each of 10 first- and second-grade classrooms received reinforcement for achieving low rates of disruptive behavior. The fourth child served as a control subject throughout the experiment. Two of the three experimental subjects were then taught to self-observe their own disruptive behavior. In the final reinforcement period, these subjects were given control over dispensing reinforcers to themselves, based on their self-collected behavioral data while subjects in the other experimental group continued with the externally managed reinforcement. In extinction, reinforcement was discontinued for all subjects, but one of the self-regulation subjects in each classroom continued overtly to self-observe. Results indicated that both reinforcement programs reduced disruptive behavior. The self-regulation procedures were slightly more effective in reducing disruptiveness than was the external regulation procedure, and this advantage persisted into extinction. These results suggest that self-regulation procedures provide a practical, inexpensive, and powerful alternative in dealing with disruptive behavior in children.

Disruptive classroom behavior has often been the target of behavior modification technology. Many studies have demonstrated that rates of disruptive behavior can be substantially reduced by the systematic application of externally managed contingencies (*e.g.*, Allen, Hart, Buell, Harris, and Wolf, 1964; Patterson, 1965; Homme, DeBaca, Devine, Steinhorst, and Rickert, 1963; Schmidt and Ulrich, 1969; Wasik, Senn, Welch, and Cooper, 1969; O'Leary, Becker, Evans, and Saudargas, 1969; Thomas, Becker, and Armstrong, 1968). However, few attempts have been made to explore the utility of self-managed contingencies in affecting desirable behavior changes in the classroom setting.

Lovitt and Curtiss (1969) demonstrated the potential of self-regulation for increasing a student's academic response rate. They found that higher academic rates occurred when the pupil arranged the contingency requirements than when the teacher specified them. In another classroom study, Glynn (1970) found that self-determined reinforcement was as effective as experimenter-determined reinforcement in increasing academic response rate and that differential token reinforcement experience influenced subsequent rates of self-determined reinforcement. Several other studies conducted in laboratory settings have further suggested the potential utility of self-monitoring and self-reinforcement in the modification of behavior. The results of studies on self-administered reinforcement have consistently demonstrated that behavior may be modified and maintained as well with a self-administered token reinforcement system as with an externally managed reinforcement system. Marston and Kanfer (1963) found that self-reinforcement procedures were effective

[1]The authors wish to express their gratitude to District 4J Schools in Eugene, Oregon, and especially to Willagillespie and Awbrey Park grade schools, for their cooperation and helpful suggestions. We also appreciate the dedicated work of our observers: Mike Dillon, Jeanne Bisenius, Don Christiansen, Joyce Christiansen, Diane Hillis, Laurie Harris, Jim Neely, Connie Sakamoto, Kathy Persinger, Jeanne Staples, and Rick Grant. Thanks also to Dr. G. R. Patterson for his advice and encouragement. Reprints may be obtained from the authors, Psychology Clinic, 1679 Agate, University of Oregon, Eugene, Oregon 97403.

in maintaining previously learned verbal discriminations. Also, Bandura and Perloff (1967) demonstrated that self-managed reinforcement was as effective as externally managed reinforcement in maintaining effortful motor behavior with children. Furthermore, there is evidence that suggests that behaviors maintained by self-reinforcement may be more resistant to extinction than those maintained by external reinforcement (Kanfer and Duerfeldt, 1967; Johnson, 1970; Johnson and Martin, 1972). Johnson and Martin (1971) have suggested that these results may be a result of the conditioning of self-evaluative responses as secondary reinforcers. These authors proposed that the secondary reinforcing properties of positive self-evaluation served to maintain children's attention to task in the absence of token reinforcers.

The present study attempted to apply self-regulation procedures to reduce disruptive behavior in the classroom. Within this context, the study was designed to test the relative effectiveness of self- and externally managed reinforcement systems during reinforcement and extinction.

Baseline observations on the frequency of three disruptive behaviors (talking out, aggression, and out-of-seat behaviors) were collected on the four most disruptive children in each of 10 classrooms. Following baseline, three of the four most disruptive children in each class received reinforcement for achieving low rates of disruptive behavior (external regulation). The fourth child served as a no-regulation (NR) control subject. Two of the three experimental subjects were then taught to self-observe their own disruptive behavior accurately. In the final reinforcement period, these subjects were given complete control over dispensing reinforcers to themselves (self-regulation, SR), based on their self-collected data. Subjects in the other experimental group (ER) continued with the externally managed reinforcement system. In extinction, reinforcement was discontinued for all subjects, but one of the self-regulation subjects in each of the classrooms continued overtly to self-observe.

Based upon the research cited above, it was hypothesized that self-regulation procedures would be as effective as externally managed procedures in maintaining low rates of disruptive behavior. It was also predicted, in light of Johnson (1970) and Johnson and Martin (1972), that the reduction in disruptive behaviors achieved through self-regulation procedures would be more resistant to extinction than that achieved through external regulation. The act of recording disruptive behaviors in the self-regulation procedure might be expected to have acquired conditioned aversive properties because rates above criteria levels have been previously consequated by response cost punishment (see Weiner, 1962), or the loss of a token reinforcer. On the other hand, recording rates below criteria levels might set the occasion for positive self-evaluation by virtue of its association with the receipt of token reinforcers. In extinction, the conditioned aversive or reinforcing properties of self-monitoring and self-evaluation were presumed to continue in the absence of primary reinforcement. These conditioned properties (aversive or reinforcing) were assumed to provide the mechanism by which the self-regulation procedures would retard extinction.

In order to increase the likelihood that self-monitoring and self-evaluation would occur during extinction, one of the two SR subjects in each class was asked overtly to self-observe during extinction. It was predicted that this group of subjects (Group SR_1) would demonstrate even greater resistance to extinction than the other SR subjects (Group SR_2) who were not required overtly to self-monitor in this phase. Without previous training in self-observation, it was expected that the ER group would show the least resistance to extinction of the three experimental groups. The predicted direction, then, of resistance to extinction was $SR_1 > SR_2 > ER$. All three groups were expected to have lower rates of disruptive behavior than the NR control group.

METHOD

Subjects

This experiment required the initial selection of disruptive students. Ten teachers of combination first- and second-grade classes in two schools were asked to pre-select six to eight of their students who typically emitted high rates of disruptive behavior. This group was observed for six days of pre-baseline and, on the basis of the resulting data, the four most disruptive students in each class were selected. The four subjects in each classroom were assigned to one of four groups. Restricted randomization procedures were used in assigning subjects to conditions in order to maximize the similarity of baseline averages. The experiment thus began with 40 subjects, 10 in each group. Any subject who did not emit 0.4 disruptive behaviors per minute or more during the subsequent baseline phase was dropped from the study. Also, subjects who were absent four or more successive days or for more than a total of six days during the study were not included. As a result of these *a priori* decision rules, seven subjects were dropped from the experiment, leaving nine subjects in the ER group and eight subjects in each of the SR and NR groups. Five subjects were later added in a new NR control group. The total number of subjects in the experiment was 38.

Experimental Setting

This experiment was conducted in five classrooms in each of two elementary schools of similar middle class socio-economic representation. A single observer was situated in each classroom in a position from which he could clearly see and hear each subject. An additional observer rotated among the five classes at each school to obtain observer-agreement data. All observers were uninformed regarding the hypotheses of the study and the assignment of subjects to conditions.

The experimental setting was defined by the situation in which the student was sitting in his desk and working independently on his assignment. Periods during which subjects were receiving individual attention from the teacher, or were involved in assigned activities that deviated from the experimental setting as defined, were not included in the sample. The experiment was conducted for one 30-min session per day for a period of eight weeks.

Procedure

An initial pre-baseline period of six sessions served both as a subject-selecting device and as a period of adaptation to the observers' presence. Immediately before baseline data were collected, teachers informed their classes that four named students would be involved in a university experiment and that the remainder of the class would not be included or involved. Phase I of the experiment constituted the measure of a baseline of disruptive behavior for subjects in all four groups. Before the five baseline sessions, all subjects were informed that observers would be counting the frequency of certain of their behaviors, which were then specified to them. Disruptive behaviors included:

1. Talking-out or making inappropriate noise without the permission of the teacher.
2. Hitting or physically annoying other students.
3. Leaving desk to do unassigned or inappropriate activities.

Pre-baseline analyses revealed that observers were capable of both simultaneously and continuously observing each of the four subjects in their classrooms on these salient categories of behavior with respectable observer agreement. Observers continued to count the frequency of disruptive behaviors for all subjects throughout the experiment. Observers recorded all occurrences of each disruptive category in 5-min time blocks. Given the above restrictions on the experimental setting, the average number of minutes each child was observed per day was 22 min. The total frequency of disruptive behaviors per day was divided by the total number of minutes observed on that day for each subject. The de-

Table 1
Experiment design: treatment by groups and phases.

Groups	*I*	*II*	*III*	*IV*	*V*
ER	Baseline	ER	ER	ER	EXT
SR_1	Baseline	ER	SR*	SR	EXT**
SR_2	Baseline	ER	SR*	SR	EXT
NR	Baseline	NR	NR	NR	EXT

*Monitored
**Self-observation continued

pendent variable was thus measured in terms of disruptive behaviors per minute.

The present experiment involved four phases beyond the baseline period. Some part of the treatment program was systematically changed in each phase for one or more groups. An outline of procedures for each phase is presented in Table 1.

Phase II of the experiment, lasting six sessions, immediately followed the baseline period and involved external regulation procedures for all three experimental groups. External regulation was defined as an external source evaluating behavior and dispensing reinforcement contingent upon that evaluation. In this ER condition, three arbitrary levels of disruptive behavior were designated:

1. Fewer than five disruptive behaviors.
2. Fewer than 10 disruptive behaviors.
3. More than 10 disruptive behaviors.

Subjects were instructed that if their behavior were at Level 1 during the session, they would receive eight points; if their behavior were at Level 2, they would receive four points; if their behavior were at Level 3, they would receive no points. Points were redeemable for reinforcements dispensed by the experimenter. Reinforcers were of a school-related nature and included pencils, erasers, notepads, *etc.* These reinforcers were placed in three boxes. With four points, a subject was allowed to choose out of a box labelled "4", which contained the least expensive reinforcers (less than 7¢). Eight points earned the choice of a prize from a box labelled "8", with slightly more expensive prizes (7¢ to 15¢). Subjects were also allowed to save points to earn children's readers (25¢) from a "12"-point box. Experimental subjects received points and chose prizes immediately after each session in the counselor's office to avoid jealousy of control subjects and classroom peers. All prizes were picked up by subjects at the end of the school day.

In Phase III, lasting seven sessions, two of the three experimental groups began training in self-regulation procedures. Self-regulation is here defined as the case in which the individual evaluates his own behavior and dispenses his own reinforcers contingent with previously learned criteria. The two SR groups were given self-observation cards and instructed in recording their own behavior within the three disruptive behavior categories. At the end of each session in Phase III, the SR group's subjects' observation cards were matched against the observers' data to check on the subjects' accuracy. If a subject's self-observation rating was within a range of three disruptive behaviors above or below the observer's score, the subject received the equivalent number of points as dispensed in Phase II under the ER condition. If, however, a subject rated himself beyond the range of three disruptive behaviors above or below the observer's rating, the subject received two points less than he would have received for the coder's rating. These measures were added to improve the accuracy of self-observation. The ER and NR groups continued in Phase III with the same treatment as in Phase II.

In Phase IV, lasting seven sessions, both SR groups self-regulated independently, without checking accuracy with the observers. In this phase, the self-observation data constituted the sole determinant of the number of points SR subjects received, irrespective of the observers' ratings. The criteria for the dispensing of points were the same as in Phase II. Both ER and NR groups continued as before.

In Phase V, lasting seven sessions, subjects in all groups underwent extinction. Subjects in the experimental groups were informed by the exper-

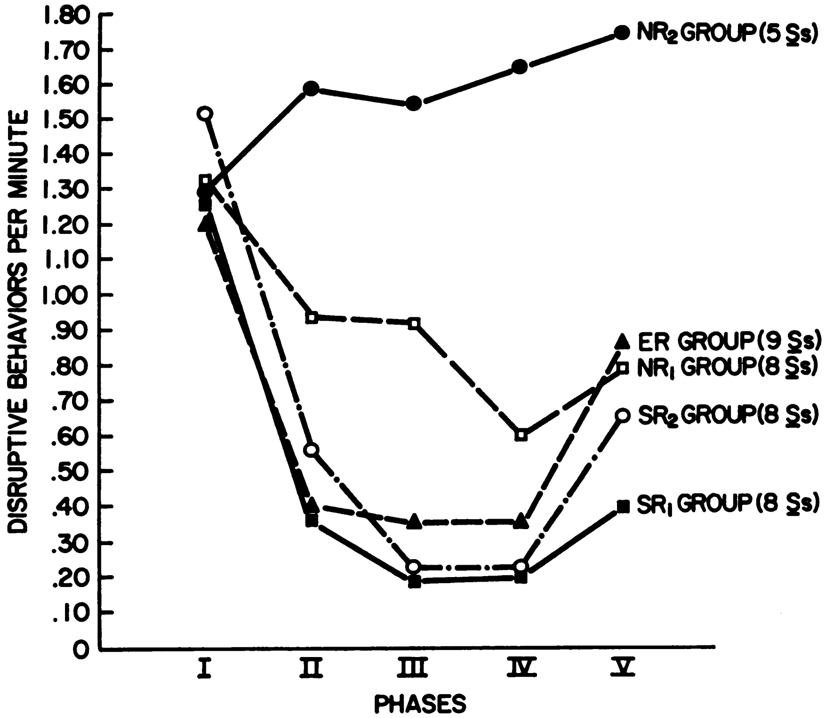

Fig. 1. Average disruptive behavior per minute of groups. During Phase I (baseline), disruptive behaviors were observed in class. During Phase II (external regulation), children in all experimental groups (ER, SR_1 and SR_2) were awarded points after class for fewer disruptions. During Phase III, ER group remained on external regulation while the self-regulation groups (SR_1 and SR_2) were trained to record and report their own behavior and were given points for accurate reports of fewer disruptions. Phase IV was the same as Phase III, except that points were given for children reporting few disruptions regardless of the accuracy of their reports. In Phase V (extinction), no points were given and only the SR_1 group was still required to record and report on their disruptive behavior.

imenter that prizes were no longer obtainable. One SR group (SR_1) was asked to continue to self-observe their frequency of disruptive behaviors on the self-observation cards.

An additional control group (Group 5) was added to the study when the experiment was conducted in the second school. This group was distinguished from the initial control group by having all its subjects in a single classroom. That is, there were no experimental subjects present in the Group 5 classroom.

RESULTS

The dependent variable in this study was the frequency of disruptive behaviors per minute. Observer agreement on this variable was measured by the correlation between the daily recorded frequencies of the regular observers and the observers who alternated between classes. The average Pearson Product Moment correlation over all five phases was 0.93. Figure 1 shows the mean number of disruptive behaviors per minute for each group in each of the five phases.[2] As will be clear from the ensuing presentation of these results, the group data reflect rather well the direction of change for the individual case.

Results were analyzed separately for each phase. In all phases beyond baseline, one-way analyses of variance were performed on the raw scores for that phase. Orthogonal comparisons were made between the three experimental and two control groups and between the two self-regulation groups and the one external regula-

[2]Individual subject data can be obtained from the authors on request.

tion group, where appropriate. Summaries of individual subject data are also presented that provide a perspective on the magnitude and breadth of change.

Phase I: Baseline

A one-way analysis of variance on the mean rates of disruptive behaviors per minute across all five days of baseline revealed no significant differences between groups ($F < 1$, $df = 4,33$). Thus, all groups were essentially the same in their display of disruptive behavior before the experimental procedures were introduced.

Phase II: External Regulation

Analysis of raw score data in this phase showed a significant groups effect ($F = 6.35$, $df = 4,33$, $p < 0.01$). The orthogonal comparison between the three experimental and two control groups revealed a significant difference in favor of the experimental groups ($t = 4.75$, $df = 33$, $p < 0.01$). In this phase, 96% of the experimental subjects reduced their rate of disruptive behavior and 76% reduced their rate to less than one-half of their baseline level.

There was a noticeable reduction from the baseline in the number of disruptive behaviors for Control Group 4 in this and all subsequent treatment phases (see Figure 1). In Phases II, III, and IV, the frequency of disruptive behavior for Group 4 was reduced from baseline by 28%, 38%, and 54% respectively. This finding might partially be accounted for by the fact that 23% of the disruptive behaviors of these control subjects during baseline involved interactions with the experimental subjects in the same classroom. In the subsequent phases, however, only 7% of the disruptive behaviors in Group 4 involved interaction with subjects in the experimental groups. The subjects in Group 5 were all in a separate classroom in which no experimental subjects were present. This group did not show the same reduction from baseline level as did subjects in Control Group 4. Multiple comparisons using Duncan's New Multiple Range Test showed that Group 4 emitted a significantly lower rate of disruptive behaviors than Group 5 ($p < 0.05$) in this phase. There were no significant differences between these two groups in the previous baseline phase.

Phase III: Self-Regulation and External Regulation

One-way analysis of variance revealed a significant groups effect in Phase III ($F = 12.49$, $df = 4,33$, $p < 0.01$). Again, an orthogonal contrast demonstrated that the three experimental groups had significantly lower rates of disruptive behavior than the two control groups ($t = 6.88$, $df = 33$, $p < 0.001$). In this phase, 96% of the experimental subjects decreased their rate of disruptive behavior to less than one-half of their baseline level. And, 84% reduced their rates to less than one-third of baseline level.

An orthogonal comparison of the self-regulation and external regulation groups yielded no significant difference ($t = 0.88$, $df = 33$). That is, in Phase III, subjects who continued to receive ER treatment did not differ significantly in the raw score analysis from subjects who, in this phase, received training in SR procedures. Nonetheless, it was found that only 33% of the subjects in the ER condition decreased their disruptive rate from their Phase II level, whereas 88% of the subjects in the SR condition decreased their rate from the previous phase. Also, the SR groups averaged 42% less disruptive behavior in this phase than the ER group.

The subjects who received SR training were checked as to the accuracy of recording their frequency of disruptive behaviors. In Phase III, it was found that 75% of the subjects' self-observation ratings fell within the permissible range of plus or minus three disruptive behaviors as recorded by the experimenter. The median discrepancy between the subjects' ratings and the experimenter's was 2.07 disruptive behaviors. Forty-four per cent of the discrepancies were in the direction of underestimation by the subjects; 28% of the discrepancies represented overestimation; and 27% of the accuracy checks showed the subjects and the experimenter to be in perfect agreement.

Subjects in the SR condition in Phase III were fined two points for inaccuracy or ratings that did not fall within the permissible range of plus or minus three disruptive behaviors as recorded by the observer. An average of 6.1 points per day was obtained by SR subjects in this phase, which is contrasted by an average figure of 6.4 points that could have been earned had fines not been imposed. As such, the subjects in the SR groups received only 5% less points than they would have without the two-point fine for inaccuracy. The ER group in Phase III earned an average of 5.3 points per day. There were no significant differences between the ER and SR groups in terms of number of points received in Phase III ($t = 1.04$, $df = 23$), in spite of slightly lower rates of disruptive behavior and fines for inaccuracy for the SR groups.

Group 4 continued to emit a significantly lower rate of disruptive behaviors than Group 5 in this phase ($p < 0.01$), as revealed by Duncan's New Multiple Range Test.

Phase IV: Self-Regulation and External Regulation

Analysis of variance for means in Phase IV revealed a significant groups effect ($F = 17.73$, $df = 4,33$, $p < .01$). As in the previous two phases, an orthogonal comparison indicated that the three experimental groups differed significantly from the two control groups ($t = 7.22$, $df = 33$, $p < 0.001$). An orthogonal contrast comparing raw score means of self-regulation and external regulation groups revealed no significant difference ($t = 0.98$, $df = 33$).

It can readily be seen from Figure 1 that all three experimental groups maintained the same rates of disruptive behavior in Phase IV as they had in Phase III. Again, 96% of the experimental subjects continued to maintain their rate less than one-half of their baseline rate; and, of disruptive behavior in Phase IV at a level 80% maintained at a level less than one-third of their baseline rate. The breakdown of these experimental groups in Phase IV revealed that 89% of the ER subjects reduced their rate of disruptive behavior to less than one-half of baseline level; 67% reduced their rate to less than one-third of baseline level. In the SR groups, 100% of the subjects reduced their rate to less than one-half of baseline levels and 92% less than one-third baseline levels. In this phase, the SR groups averaged 39% fewer disruptive behaviors than the ER group.

Although the subjects in the two SR groups were neither informed of nor rewarded for self-observation accuracy in this phase, 71% of the subjects' self-observation ratings fell within the range of plus or minus three disruptive behaviors as recorded by the experimenter. The median discrepancy between the subjects' ratings and the experimenter's was 1.75 disruptive behaviors. As in the previous phase, the majority of the discrepancies were in the direction of underestimation of disruptive behavior by the subjects. Forty-five per cent of the discrepancies represented underestimation, 21% represented overestimation, and 34% of the accuracy checks showed perfect agreement between observers and subjects.

In Phase IV, subjects in the ER condition earned and received an average of 5.7 points per daily session. It was found that the SR subjects awarded themselves an average of 7.4 points per session, whereas their corresponding rates of disruptive behavior merited only 6.4 points. That is, SR subjects received an average of 16% more points than they deserved. An examination of the individual subject data indicates that this discrepancy between points awarded and deserved was accounted for by fewer than half of the subjects involved. More specifically, nine of the 16 SR subjects consistently awarded themselves veridically. Two subjects tended to over-reward themselves by approximately one point per day, while five others awarded themselves considerably more points per day than were earned (2.5 points or more). The nine accurate SR subjects obtained significantly lower rates of disruptive behavior than the seven inaccurate subjects ($t = 2.61$, $df = 14$). It should also be pointed out that these nine subjects were not significantly lower in disruptive rates at baseline ($t = 1.14$, $df = 14$) than the seven inaccurate subjects.

Multiple comparisons using Duncan's New Multiple Range Test again revealed that control Group 4 emitted significantly fewer disruptive behaviors than control Group 5 ($p < 0.001$).

Phase V: Extinction

An additional observer was introduced in each school in this phase. This observer also alternated between classes to provide additional observer agreement data. This observer was totally naive as to the previous four phases and thereby had no opportunity to infer hypotheses or establish bias. The average Pearson Product Moment correlation between the daily rates obtained by the naive observers and the regular observers was 0.93 in Phase V. An average correlation between the regular alternating observers and the observers who remained in the same class was 0.98 in this phase.

During Phase V, the predicted direction of resistance to extinction was $SR_1 > SR_2 > ER$ and, it was predicted that all three experimental groups would have lower rates of disruptive behavior than the NR control condition. A trend analysis on this predicted direction obtained significance ($t = 3.17$, $df = 33$, $p < 0.01$).

A significant main effect for groups was obtained in Phase V in the raw score analysis ($F = 4.11$, $df = 4{,}33$, $p < 0.01$). Three orthogonal comparisons were performed on these Phase V raw score results. The first comparison revealed that the three experimental groups continued to be lower in rate of disruptive behavior than the two control groups even in the absence of reinforcement ($t = 2.99$, $df = 33$, $p < 0.01$). The second comparison indicated that the two SR groups were not significantly lower in raw score disruptive rates than the ER group during extinction ($t = 1.31$, $df = 33$). The third comparison showed that Group 2 (SR_1), whose subjects continued to self-observe during extinction, was not significantly lower ($t = 0.50$, $df = 33$) than Group 3 (SR_2).

While the analyses given above test the significance of observed differences in the extinction period, they do not test differences in resistance to extinction. The following analyses were performed to analyze that question.

Two orthogonal comparisons were performed on the difference scores derived by Phase IV subtracted from Phase V. These analyses revealed that there were no significant differences between the ER group and the two SR groups in terms of increased rate of disruptive behaviors from the last phase of treatment to extinction ($t = 0.85$, $df = 33$) and that there were no significant differences between the two SR groups ($t = 0.80$, $df = 33$).

Repeated measures analysis of variance performed across the seven days of Phase V, with the mean of Phase IV as the starting point, provided a measure of resistance to extinction for the three experimental groups. This analysis revealed a significant trials effect ($F = 2.39$, $df = 7{,}154$, $p < 0.05$), reflecting extinction over trials but no main effects for groups ($F = 1.18$, $df = 2{,}22$) or for the groups by trials interaction ($F = 0.74$, $df = 14{,}154$). The three experimental groups, then, did not extinguish differentially during this period.

During extinction, 56% of the ER subjects maintained their reduced rate of disruptive behavior at less than one-half of their baseline level whereas, for the subjects in the two SR groups, 69% maintained at or below one-half of their baseline level. Only 22% of the subjects in the ER group maintained their reduced rate at less than one-third their baseline level, while 56% of the subjects in the two SR groups maintained at or below one-third of their baseline level. As in the previous phase, the SR groups averaged 39% fewer disruptive behaviors than the ER group, even in the absence of reinforcement. And, the SR_1 group, which continued to self-observe during extinction, averaged 53% fewer disruptive behaviors than the ER group.

The superiority of control Group 4 over Group 5 was maintained at a significant level ($p < 0.01$) in this phase (Duncan's New Multiple Range Test).

An examination of Figure 1 reveals that the two SR groups consistently averaged less disruptive behavior than the ER group in Phases III, IV, and V. Summaries of individual data also indicate that the SR groups reduced their rate of disruptive behavior from baseline more than the ER groups in all three of these phases. Nevertheless, orthogonal comparisons on the raw scores analyses of variance do not reflect any superiority of the SR condition. Because of the large within-group variability in rates of disruptive behavior evidenced in these comparisons, difference score analyses were also conducted. Between-subject (within groups) variability is reduced by difference scores analyses in that difference scores can control for initial baseline differences.

Orthogonal comparisons on the difference score analyses of variance reveals that the SR groups did not reduce their rate of disruptive behavior significantly more than the ER group in Phase II ($t = 1.17$, $df = 33$). This is as expected because both conditions received the same external reinforcement treatment in this phase. But, in the next two phases, representing different treatment conditions, the SR groups were found to have reduced their disruptive behavior rate significantly more from baseline than the ER group (Phase III: $t = 2.50$, $df = 33$, $p < 0.05$; Phase IV: $t = 2.32$, $df = 33$, $p < 0.05$). This greater reduction for SR groups persisted in Phase V, extinction ($t = 2.49$, $df = 33$, $p < 0.05$). These findings should be qualified by acknowledging that there are a number of important problems with difference scores, two of the most critical being the reliability of the differences and possible correlation of differences with pre-treatment scores (Lord, 1967).

DISCUSSION

In all phases after baseline, the experimental groups exhibited significantly lower rates of disruptive behavior than the control groups. Clearly, both the external and self-regulation procedures were effective in establishing and maintaining reductions in disruptive behavior.

It was hypothesized that self-regulation procedures would be as effective as externally managed procedures in maintaining low rates of disruptive behavior. Figure 1 indicates that not only were the self-regulation procedures as effective as externally managed procedures but that they were slightly more effective in producing consistently lower rates from their introduction in Phase III through extinction in Phase V. Throughout each of these three phases, the two self-regulation groups evidenced an average of roughly 40% fewer disruptive behaviors than the externally managed group. In spite of this apparent superiority of the self-regulation procedures, significant differences between this condition and the external management condition were not obtained upon direct comparison of raw score data. Large within-group or between-subject variability characterizes these group comparisons. Difference-score analyses, which control for this variability problem, demonstrate a significant superiority for the self-regulation procedures over the external regulation procedures in terms of greater reductions from baseline levels. As previously indicated, however, difference scores are difficult to interpret in group comparisons, due to a number of inherent problems in their analysis. Thus, while a slight superiority for the self-regulation procedures is indicated by the data, differences between the two conditions were not found to be significant upon direct comparison. Therefore, the hypothesis that self- and externally managed procedures would be equally effective in maintaining low rates of disruptive behavior was essentially supported.

Crucial to the effectiveness of any self-regulation procedure is the accuracy of self-observation. It is an encouraging finding in this study that most of the first- and second-grade children deemed disruptive by their teachers and a screening procedure, were capable of self-observing their frequency of disruptive behavior with respectable accuracy. This is especially noteworthy in that the SR subjects did not receive immediate feedback as to their accuracy in training and no feedback during Phase IV and extinction. This relatively high degree of accuracy in self-observation for young children is congruent with other findings (Johnson, 1970; Johnson and Martin, 1972).

It is also crucial to the effectiveness of any self-regulation procedure that subjects reward themselves appropriately. In Phase IV of this study, subjects in the SR condition exercised control over dispensing points to themselves, independent of the observers' data. These subjects had the opportunity to receive more points than they deserved, simply by not recording all occurrences of their disruptive behavior. The previous literature on self-reinforcement suggests that when subjects take over the task of dispensing their own reinforcements, only minimal average increases in reinforcement delivery occur (Kanfer and Duerfeldt, 1967; Johnson, 1970; Johnson and Martin, 1972). The present study revealed that the SR subjects received an average of 7.4 points per session, whereas their rates of disruptive behavior merited only 6.4 points, representing an average daily discrepancy of one point. The results indicate that this discrepancy was accounted for primarily by seven SR subjects who reported inaccurate self-observation data. The remaining nine subjects recorded accurately in this phase and consistently awarded themselves exactly the number of points they deserved. It is obvious that discrepancies in points received *vs* deserved are not consistent across subjects. It is also obvious that this average one-point per day error in reinforcement delivery is inconsequential for purposes of application.

The fact that SR subjects in Phase IV had the opportunity to receive more points than their actual behavior merited has important implications for considering differences in the SR condition *vs* the ER condition. As long as subjects in the SR condition accurately observe and thereby reinforce their behavior contingently, consistent with the same criteria as the ER group, there is no differential advantage for the SR condition. However, when an SR subject awards himself more points than he deserves, a magnitude of reinforcement confound exists in that this subject receives more than his counterpart in the ER condition. This potential confound is problematic for the ER *vs* SR comparison in Phase IV, particularly for the seven inaccurate SR subjects. An important consideration for the application of self-regulation procedures is whether or not this potential discrepancy in points received *vs* merited results in increments in disruptive behavior rate. It is interesting that the SR subjects did not evidence higher rates of disruptive behavior in Phase IV, relative to Phase II or relative to the ER group, when they had the opportunity to receive points independent of their actual behavior.

The second hypothesis predicted that the reduction in disruptive behaviors achieved through self-regulation procedures would be more resistant to extinction than that achieved through external regulation. Even though the groups were aligned in the predicted direction in the extinction period with $SR_1 < SR_2 < ER$, greater resistance to extinction was not clearly evident from these data. It appears that the slight superiority of the SR groups in extinction, previously described, can be accounted for almost entirely by their superiority in the two former periods. This interpretation is substantiated by the nonsignificant findings in the repeated measures analysis of the extinction data. It should also be pointed out that the resistance to extinction comparison between the two procedures was confounded by different magnitudes of reinforcement received by SR and ER groups. This confounding could have been averted by yoking or matching the amount of reinforcement delivered to the ER subjects with the amount that SR subjects dispensed to themselves. However, ER subjects, receiving yoked reinforcement, might learn that the amount of reinforcement they received was unrelated to their actual behavior. The confounding influence of this operation was viewed as far more serious than the anticipated minor differences in receipt of token points.

The resistance to extinction predictions were based on assumptions about the conditioned aversive and reinforcing properties of both overt and covert self-observation and self-evaluation during extinction. It was believed that children in the self-regulation groups would continue more overt (SR_1) and covert (SR_2) self-observation and self-evaluation than would children in the ER group and that these behaviors would retard extinction. While there appears to be no support for this line of reasoning in the direct comparison of the ER groups with the combined SR groups, there may be some weak support for it in the comparisons of the two SR groups. The SR children who continued overtly to self-monitor in extinction displayed 23% less deviant behavior in extinction than those who were not asked to continue self-monitoring. This superiority does not appear to be due to any appreciable prior advantage (see Figure 1) nor is it related to different magnitudes of reinforcement. The significance of the directional prediction also lends weak support to the superiority of the SR group, whose members continued to self-monitor during extinction. Another possible explanation for these differences in extinction might be that the continuing presence of the observers served as a cue to remain well-behaved for the SR_1 subjects. The earlier accuracy checking in Phase III may have had some continuing effect on the subjects' behavior. However, this seems unlikely given that SR subjects were not checked for inaccuracy in the intervening seven days of Phase IV. While the SR_1 *vs* SR_2 differences were not of great magnitude nor statistically significant on direct comparison, it appears to be a noteworthy finding consistent with the conditioned reinforcement hypothesis.

An unexpected finding in this study was the marked decline in disruptive behavior for the control subjects who were present in the same classrooms as the experimental subjects. One possible explanation is that when the disruptive behavior of three of the four most disruptive students in a class is substantially reduced, this will have a dampening or a spread effect on the fourth student. Some evidence for this explanation was provided in that control subjects had fewer disruptive interactions with the experimental subjects during the treatment phases than they did during baseline. An alternative explanation is that the control subjects may have discovered that the other subjects in their classrooms were getting reinforcers at the end of the school day for low rates of disruptive behavior. This knowledge may have set the occasion for vicarious reinforcement or fruitless attempts also to earn reinforcement by emitting low rates of disruptive behavior. The interpretation that the reduction in Group 4 had something to do with the presence of treatment in the classrooms is born out by the comparison of Group 4 with Group 5 in which no such reduction was observed. The significantly lower rate of disruptive behavior for Group 4 represents an important finding in suggesting that the modification of disruptive children in the classroom may have suppressive effects on other disruptive children in the same classroom. Indeed, it is possible that the whole social system in the classroom is affected by intervention with selected children.

The findings of this investigation have important implications for the applied use of self-regulation in the classroom. It appears that self-regulation procedures can be as effective, if not slightly more effective, than external regulation procedures for this population. Disruptive children in the first and second grade have been determined to be quite capable of self-observing and recording their own disruptive behavior with impressive accuracy. Perhaps the simple act of self-recording disruptive behaviors accounted for the slight superiority of the self-regulation groups. The act of self-recording would seem to have the advantage of more immediate loss of reinforcement signals with greater consistency than would the recording of an observer or teacher with external regulation procedures. Furthermore, most children in self-regulation appear capable of self-observing their own behavior, and thereby applying designated contingencies, without the monitoring of an external agent.

In terms of teacher time and effort, self-regulation procedures have the advantage of being more practical and less expensive. Although these procedures might initially require as much teacher involvement as externally managed procedures, it is apparent that a good deal of the responsibility of monitoring a child's disruptive behavior can be turned over to the child, without substantial increases in disruptive rate. It may be necessary for the teacher occasionally to check the accuracy of a child's monitoring, but it is apparent that this procedure would, nevertheless, require less time and effort than the continuous monitoring of a child with external procedures.

In general, the results of this and other investigations (Johnson, 1970; Johnson and Martin, 1972; Lovitt and Curtiss, 1969; Glynn, 1970) are clear. Self-regulation procedures appear to be either equally effective or more effective than external regulation procedures in both establishing and maintaining desired changes in behavior. Furthermore, it appears that most children of this age are capable of self-monitoring their own behavior and applying designated contingencies. It is concluded that self-regulation procedures provide a practical, inexpensive, and powerful alternative to external procedures.

REFERENCES

Allen, K. E., Hart, B., Buell, J. S., Harris, R. T., and Wolf, M. M. Effects of social reinforcement on isolate behavior of a nursery school child. *Child Development,* 1964, **35**, 511-518.

Bandura, A. and Perloff, B. Relative efficacy of self-monitored and externally imposed reinforcement systems. *Journal of Personality and Social Psychology,* 1967, **7**, 111-116.

Edwards, A. L. *Experimental design in psychological research.* New York: Holt, Rinehart & Winston, 1964.

Glynn, E. L. Classroom applications of self-determined reinforcement. *Journal of Applied Behavior Analysis,* 1970, **3**, 123-132.

Homme, L. E., DeBaca, P. C., Devine, J. V., Steinhorst, R., and Rickert, E. J. Use of Premack principle in controlling the behavior of nursery school children. *Journal of the Experimental Analysis of Behavior,* 1963, **6**, 544.

Johnson, S. M. Self-reinforcement *vs.* external reinforcement in behavior modification with children. *Developmental Psychology,* 1970, **3**, 148-149.

Johnson, S. M. and Martin, S. Developing self-evaluation as a conditioned reinforcer. In B. Ashem and E. G. Poser (Eds.), *Behavior modification with children.* New York: Pergamon, 1972, (*in press*).

Johnson, S. M. and White, G. Self-observation as an agent of behavioral change. *Behavior Therapy,* 1971, **2**, 488-497.

Kanfer, F. H. and Duerfeldt, P. H. The effects of pre-training on self-evaluation and self-reinforcement. *Perceptual and Motor Skills,* 1967, **25**, 237-246.

Lord, F. M. Elementary models for measuring change. In C. W. Harris (Ed.), *Problems in measuring change.* Madison: The University of Wisconsin Press, 1967. Pp. 21-38.

Lovitt, T. and Curtiss, K. A. Academic response rate as a function of teacher- and self-imposed contingencies. *Journal of Applied Behavior Analysis,* 1969, **2**, 49-54.

Marston, A. R. and Kanfer, F. H. Human reinforcement: Experimenter and subject controlled. *Journal of Experimental Psychology,* 1963, **66**, 91-94.

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

Patterson, G. An application of conditioning techniques to the control of a hyperactive child. In L. P. Ullman and L. Krasner (Eds.), *Case studies in behavior modification.* New York: Holt, Rinehart & Winston, 1965. Pp. 370-375.

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

Schmidt, G. W. and Ulrich, R. E. Effects of group-contingent events upon classroom noise. *Journal of Applied Behavior Analysis,* 1969, **2**, 171-179.

Thomas, D. R., Becker, W. C., and Armstrong, M. Production and elimination of disruptive classroom behavior by systematically varying teacher's behavior. *Journal of Applied Behavior Analysis,* 1968, **1**, 35-45.

Wasik, B. H., Senn, K., Welch, R. H., and Cooper, B. R. Behavior modification with culturally deprived school children: Two case studies. *Journal of Applied Behavior Analysis,* 1969, **2**, 181-194.

Weiner, H. Some effects of response cost punishment upon human operant behavior. *Journal of the Experimental Analysis of Behavior,* 1962 **5**, 201-208.

Received 17 August 1971.
(Revised 5 June 1972.)