The Good Behavior Game: A Token Reinforcement System for Preschoolers

Naomi B. Swiezy
Johnny L. Matson
Peggy Box

ABSTRACT. The Good Behavior Game was implemented with four children attending a church-affiliated preschool in efforts to increase concurrent compliance and cooperation. Sessions were conducted individually with each child pair while in free play. The children were given various instructions by "Buddy Bear" (a puppet). Children were praised individually for compliant behavior while noncompliant and inappropriate behaviors were ignored. Further instances in which children complied (i.e., cooperation) were rewarded with tokens made of felt, shaped like happy faces or dinosaurs and placed on a large felt board. If the children earned their criterion levels of tokens, they received animal snacks. Improvements in compliance were noted, with generalization occurring across therapists but not settings. Factors influencing the results and implications of the findings are discussed.

In treating behavior problems of young children, both positive and token reinforcement programs have been effective (e.g., Allen, 1972; Bailey, Wolf, & Phillips, 1970; Brooks, 1975; Gould & Jansma, 1986; Hawkins, 1982; Robinson, Newby, & Ganzell, 1981; Wagener, 1977; Wolfe, Boyd, & Wolfe, 1983; Yoshimuta, Takayama, & Sonoda, 1975). Further, in recent years there has been clear recognition of the fact that social behaviors of children require considerable attention.
and remediation to head off more serious problems (Matson, Esveltd-Dawson, Andrasik, Ollendick, Petti & Hersen, 1980; Matson & Ollendick, 1988).

The token reinforcement program known as the Good Behavior Game has been applied to children in late childhood and adolescence. It employs cost procedures rather than more positive components (e.g., Salend, Reynolds, & Coyle, 1989; Warner, Miller, & Cohen, 1977). As emphasized, utilizing such a game may create a positive atmosphere and structure for appropriate classroom behavior (Martin, 1981). In the game, the class is divided into teams. Rules are then developed regarding undesirable behaviors. Teams are generally informed that the entire team will lose points (tokens) for undesirable behaviors performed by any team member. The game is usually designed to allow one or both teams to win by reaching a criterion number of points. The winning team or teams then receives a group reinforcer (Barrish, Saunders & Wolf, 1969; Martin, 1981).

Few studies have been conducted utilizing this behavior management technique despite its promise. Harris and Sherman (1973) used the technique in a fifth-and sixth-grade classroom in an attempt to reduce talking and out-of-seat behavior. Marks were recorded for instances of disruptive behavior and the winning team(s) earned early dismissal from school.

Similarly, Warner et al. (1977) compared the relative effectiveness of the Good Behavior Game and basic teacher attention in modifying disruptive classroom behavior of fourth-and fifth-grade children. Though both methods were effective, the Good Behavior Game was more so. In addition, because the game required less effort, teachers preferred it over teacher attention.

Lastly, an individualized Good Behavior Game was devised by Salend et al. (1989) for emotionally disturbed adolescents in special education classrooms. Effectiveness of the procedure was attributed to the ease of implementation of the system (i.e., inappropriate behavior was represented by slashes on the blackboard and behavior was changed by decreasing the groups' criterion levels through successive approximations), immediate feedback (through the slashes on the blackboard), and positive peer pressure.

Although the Good Behavior Game has been considered most appropriate for preschool and elementary school children (Martin, 1981), the technique has only been investigated with children in elementary school and adolescence (Barrish et al., 1969; Harris & Sherman, 1973; Medland & Stachnik, 1972; Warner et al., 1977; Salend et al., 1989). The present study, the first to evaluate the efficacy of the game with preschoolers, focuses on rewards rather than cost procedures. Children are rewarded for good behavior and ignored for inappropriate behavior. Finally, this study utilizes pairs of children as teams rather than large groups of classmates, thus allowing for treatment that deals with particularly difficult to manage children. Both members of the team must exhibit good behavior concurrently (i.e., they must coordinate their behavior or cooperate) for the team to earn a point. The existence of a common goal for each team member may facilitate behavior change by promoting positive peer pressure (Slavin, Madden, & Leavey, 1984).

**METHODS**

**Children, Therapists, and Setting**

The students participating in this study were four preschoolers (two boys and two girls) attending a local church-affiliated preschool. They were white middle class children between the ages of four and five years. These children were observed to have distinct interpersonal and behavior problems which could influence their ability to cooperate. The children of each treatment pair were also observed by the researcher (and confirmed by the teacher) to be particularly disruptive and problematic when paired together. The individuals serving as therapists in the study were two female graduate students in psychology, one first year and one second year.

Observations and treatment sessions were conducted in a resource room containing a table, chairs, and shelves of school supplies and the school kitchen, depending on availability. Probe sessions for generalization were conducted on the school playground.

**Behaviors Recorded and Recording Methods**

Ten practice observations were conducted before the data was formally collected. This allowed the children to adapt to the presence
of observers and the observers to become familiar with the identification and recording of target behaviors.

Observations and intervention were conducted for 15-minute periods three times a week for each dyad studied. Observers were seated in one corner of the room (or playground) and expected to be as unobtrusive as possible.

Both therapist and child behaviors were recorded. However, in that the therapist behaviors were specifically programmed to follow a standard format (i.e., set number of instructions per session, praise to follow each compliance), these behaviors were recorded solely as a treatment integrity measure. Only the child behaviors were included in the final analyses.

The recorded therapist behaviors were as follows:

Instructions (IA). This term referred to any statements from the therapist requesting a response from both children. Ten such statements occurred during each 15-minute observation period.

Positive attention (PA). This behavior included any positive verbal attention directed toward a child or both children when they were exhibiting good behavior (i.e., sitting quietly, complying to an IA). Examples of verbal positive attention include comments which specifically referenced the behaviors being praised. Such comments are found to be especially reinforcing to preschoolers and include statements such as “Good job, you raised your hand to ask a question!” rather than just saying “Good job” (Bernhardt & Forehand, 1975). Comments such as “You have a nice dress” or “I like your stuffed animal” were not included in this category in that they reference belongings rather than behaviors or child characteristics.

Reward (PT). This behavior included the point gain that followed from both children in the team complying to instructions directed towards them.

The child behaviors recorded were:

Compliance (C). This behavior referred to the initiation or completion of the response appropriate to therapist’s instruction by either or both children within five seconds of the instruction.

Noncompliance (NC). Noncompliance was defined as failure to initiate or terminate the response appropriate to the therapist’s instruction by either or both children within five seconds of the instruction.

Compliance Ratio

This figure was derived by dividing the number of compliance behaviors by the number of instructions given and multiplying by 100.

Training Procedures

Prior to direct child involvement, therapists conducted a session in which definitions of target behaviors were reviewed. Notecards were provided as cues to the therapists about the correct format and content of instructions. Feedback from one therapist to another was given during and after sessions. Once all necessary problem solving and adaptation to the setting occurred, the formal baseline sessions were initiated.

Reliability

Data were collected by three graduate psychology students. The two primary raters were also the therapists in the study. They served as raters when they were not functioning as therapists. The third student was not a therapist and served as the reliability rater. Percent agreement was calculated by dividing the total number of agreements by the total number of agreements plus disagreements and multiplying by 100. The observers achieved at least 80% agreement before formal data collection. Reliability data were collected for 29% to 40% of probe and overall observation sessions.

Experimental Design

This study used a multiple baseline design-within subject pairs and across therapists. Three to 15 baseline and six to 17 treatment sessions were conducted. Effects of the game procedure on concurrent compliance (i.e., cooperation) of the children in each pair was evaluated.

Experimental Phases

Baseline. In this condition, normal rates of responding were recorded while children engaged in free play and responded to instructions given by the therapist.
Game Procedure Plus Praise and Ignoring. Children engaged in the same activities as in baseline. The therapist announced that the pair was going to play a game, emphasizing that the game was only to be played during the designated time periods each week. The therapist then announced that “Buddy Bear,” a puppet, would be asking the pair to do things for him. She explained that each time both members cooperated by exhibiting the appropriate behavior concurrently, the team would earn a brightly colored smiley face or dinosaur which would then be displayed on a large black felt board. Finally, the therapist explained that the pair could win the game if they met the criterion number of points. This criterion was determined individually for each pair of children according to the mean level of performance during baseline. The criterion was increased for a pair once their original criterion was met and maintained during at least two sessions. If the pair won, they received a prize (e.g., three bear-shaped graham crackers or one animal cookie). To maintain incentive for appropriate behavior, the therapist explained that pairs who surpassed their criterion by achieving 150% of their criterion or complying to all 10 instructions (if their criterion was greater than six) were to be given bonus points and extra rewards at the session’s end. Children were allowed to ask questions and received answers once explanation of the procedures was given.

A standard number of instructions (e.g., “X and Y, shake each others’ hand,” “X and Y, go to the bookcase and bring back a book,” “X and Y, tell Buddy Bear that you like him”) were given to the children. In that the study was not conducted in the natural classroom setting, a standard number of instructions was given to ensure the ability to make adequate comparisons between sessions and to have ample opportunities for the subjects to respond. Compliance by either child was rewarded with praise to that child. Points were awarded along with praise when both children complied concurrently (i.e., cooperated). Noncompliance was ignored.

Probe Sessions. These sessions (four to five in number) were conducted during both baseline and treatment phases of the study. During these sessions, only the setting was altered. That is, rather than observing the children’s compliance while they played in the resource room or kitchen, compliance was observed while the children engaged in free play in the school yard.

RESULTS

Interrater Reliability

Total mean interrater reliabilities irrespective of raters (i.e., rater 1 or 2) and types of sessions (i.e., regular or probe) were 86% and 84% for pair A (boys) and pair B (girls), respectively. Mean interrater reliabilities for separate recorded behaviors are presented for both pairs in Table 1.

The reliabilities are based on comparisons of ratings between the primary observers and a third observer who did not function as a

| Table 1 |
| Mean Interrater Reliabilities |

<table>
<thead>
<tr>
<th></th>
<th>Pair A</th>
<th></th>
<th>Pair B</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rater 1</td>
<td>Rater 2</td>
<td>Rater 1</td>
<td>Rater 2</td>
</tr>
<tr>
<td>IA</td>
<td>Regular</td>
<td>95%</td>
<td>93%</td>
<td>96%</td>
</tr>
<tr>
<td></td>
<td>Probe</td>
<td>91%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>PA</td>
<td>Regular</td>
<td>84%</td>
<td>78%</td>
<td>70%</td>
</tr>
<tr>
<td></td>
<td>Probe</td>
<td>79%</td>
<td>65%</td>
<td>65%</td>
</tr>
<tr>
<td>PT</td>
<td>Regular</td>
<td>100%</td>
<td>86%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>Probe</td>
<td>100%</td>
<td>75%</td>
<td>75%</td>
</tr>
<tr>
<td>G</td>
<td>Regular</td>
<td>88%</td>
<td>90%</td>
<td>84%</td>
</tr>
<tr>
<td></td>
<td>Probe</td>
<td>78%</td>
<td>87%</td>
<td>61%</td>
</tr>
<tr>
<td>NC</td>
<td>Regular</td>
<td>84%</td>
<td>81%</td>
<td>80%</td>
</tr>
<tr>
<td></td>
<td>Probe</td>
<td>50%</td>
<td>42%</td>
<td>85%</td>
</tr>
<tr>
<td>TOTAL</td>
<td>Regular</td>
<td>88%</td>
<td>86%</td>
<td>83%</td>
</tr>
<tr>
<td></td>
<td>Probe</td>
<td>80%</td>
<td>75%</td>
<td>78%</td>
</tr>
</tbody>
</table>

Note. Regular = regular sessions
      Probe = probe sessions
therapist. Given good reliability, bias by therapists on target behaviors are most likely minimal.

**Treatment Effects**

Increased percentages of compliance were noted for both child pairs (see Figure 1). Compliance increased from an average baseline level of 11.7% to a treatment level of 74.7% in pair A and 27.3% to 76.5% in pair B, a significant result given that these individuals were noted to be particularly disruptive and problematic when paired with their respective partner. In addition, comments made by the children to each other (e.g., “Come on, _____”) indicated that concurrent compliance resulted from peer pressure and cooperative efforts rather than from merely parallel compliance.

Generalizability occurred across therapists. The percent of compliance increased from baseline to treatment independent of the particular therapist implementing treatment. To discriminate any effects of one therapist working with the children first during sessions, Therapist 1 always worked with pair A before Therapist 2 and Therapist 2 always worked with pair B before Therapist 1. Again, the results appear to be independent of which therapist worked with the children first. Another measure of generalizability was achieved with the inclusion of probe sessions. The effects obtained during these sessions did not reflect improvement in compliance as consistently or to the same degree as the treatment effects. Overall, evidence for generalization across settings was considered to be non-existent.

**DISCUSSION**

The Good Behavior Game has served as an effective behavioral treatment (e.g., Salend et al., 1989; Warner et al., 1977). The procedure reflects the way in which behavioral science has advanced from the use of basic techniques (e.g., token reinforcement, response cost) to the use of innovative adaptations (e.g., token reinforcement with pairs of children rather than single individuals) as well as combinations (e.g., token reinforcement and differential reinforcement) of
these techniques (Matson, 1988). The present study adapted the Good Behavior Game to a more positive application. The bear puppet, who served as the instruction-giver, seemed to be important for young children. The children gave him hugs and attention in increasing amounts as the study progressed. A variety of games were used with the children (e.g., coloring, legos, blocks, doctors’ kits) to prevent boredom. Also, the felt pieces which served as tokens were brightly colored, designed to reflect the particular interests of young children (i.e., smiley faces for the girls and dinosaurs for the boys), and displayed conspicuously on a large black felt board.

Given that maladaptive preschool behavior may well affect the child’s long-term development (Michelson, Sugai, Wood, & Kazdin, 1983; Ollendick & Hersen, 1983), it seems appropriate that our study extended the literature to this group. Given our findings, more research of this type seems warranted.

This investigation is also the first to utilize treatment teams consisting of pairs of children rather than entire classes. This allows the investigator to monitor the children’s overall understanding of the method and the amount of cooperation involved more accurately.

Generalization occurred across therapists. However, effects were not as promising regarding generalization across settings.

The present study demonstrates that preschool children are capable of understanding the procedure presented and the concept of cooperation implicit in these training methods. Future research on other behaviors, and long term follow-up to determine if extended beneficial treatment effects occur may be warranted.

REFERENCES


Assessing the Duration of First-Aid Treatments by Elementary-Aged Students with Disabilities

Nancy E. Marchand-Martella
Ronald C. Martella
Ann M. Christensen
Martin Agran
K. Richard Young

ABSTRACT. Nine elementary-aged students with intellectual, learning, or behavioral disabilities received first-aid training in the treatment of three types of injuries. Students listened to a task-analyzed story read by the teacher then practiced taking care of the “injured” puppets who were characters in the story. Following training, posttest generalization probe assessments were conducted to measure the ex-