Conflicting evidence as to the presence or absence of generalization in classroom behavior modification programs prompted this study of the conditions of generalization. During the experiment, behaviors operationally defined as competitive or cooperative were reinforced in certain game situations. Then the generalization of this training over variables of task, type of response, and time periods was measured. It was predicted that generalization of the reinforced response would occur most strongly in the testing situation most like the original one. Experimental results confirmed these expectations. Results showed in addition that there was no significant difference in competitiveness between boys and girls. (RB)
GENERALIZATION OF REINFORCED BEHAVIORS IN A GAME SITUATION

Gail Hay Marchase
The Johns Hopkins University

In the large literature dealing with behavior modification in the classroom, reinforcement contingencies are reported to have affected such diverse target behaviors as cooperative play (Hart, Reynolds, Baer, Brawley, and Harris, 1968), reading skills (Haring and Houck, 1968), instruction-following (Zimmerman, Zimerman, and Russell, 1969), study behavior (Bushell, Wrobel, and Michaelis, 1968; Hall, Lund, and Jackson, 1968), and adaptive classroom behaviors (Madsen, Beckor, and Thomas, 1968; Wolf, Hanley, King, Lachowicz, and Giles, 1970). The range of behaviors and situations the behavior modifier may be affecting when he reinforces a target behavior is an important issue, yet the determinants of generalization in the classroom are unclear. Few systematic studies of generalization in the classroom setting have been reported; often informal observation is the only measure of generalization given. Generalization may be studied across three parameters: time, stimulus situation, and type of response. Studies relevant to each of these factors will be considered in turn.

Generalization across time, with the task and response required constant, implies that extinction of the response has not occurred, either because few unreinforced responses have been emitted or because other reinforcers are operating to maintain the behavior. Hall, Lund, and...
Jackson (1968) reinforced the study behavior of first and third grade pupils with teacher attention and obtained increased study rates. They reported that these higher study rates were maintained after the formal reinforcement program ended. O'Leary and Becker (1967) reinforced in-seat behavior of eight third-graders while ignoring disruptive behaviors. The authors give anecdotal evidence that children's improved behavior carried over to school sessions in which reinforcement was not administered.

In contrast to the above studies, Meichenbaum, Bowers, and Ross (1968) report no generalization of improved behavior from an afternoon to morning session with teen-aged delinquent girls when money was the reinforcer. Similarly, O'Leary, Becker, Evans, and Saudargas (1969) observed that there was no generalization of good behavior to the a.m. when praise for appropriate behavior was administered in the p.m. It is unclear what differences in technique, environment, or observation account for differential generalization over time in the above studies.

The probability of a reinforced response generalizing to a new stimulus has elsewhere been found to be a function of the similarity of the new stimulus to the originally reinforced stimulus (Hilgard and Marquis, 1961). In a study of generalization, Horton (1970) manipulated aggressive responses by alternately making reinforcement contingent upon hitting or not hitting in a game played by adolescent delinquent boys. In a test of generalization, another game, higher rates of hitting occurred following a hitting-reinforced treatment condition than following a treatment condition in which non-aggressive behavior was rein-
forced. Horton’s study provides evidence that an experimentally altered social response is subject to generalization to new stimulus situations. Wahler (1969) failed to obtain generalization across stimulus situations when deviant behavior in the home setting was modified through reinforcement contingent upon good behavior. Children's deviant behavior in the school situation was unaffected by the home program. These conflicting results might be explained on the basis of stimulus similarity. Horton was testing for generalization in a situation that had some similarity to the original reinforcement condition, i.e., both were games. Wahler, however, tested for generalization from home to an entirely different setting. Stimulus similarity, then, was greater in the Horton study, and the probability of generalization greater.

The probability that a new generalized response will be given is partially a function of its similarity to the originally reinforced response (Hilgard and Marquis, 1961). Haring and Houck (1969) obtained higher responding rates and an improvement in reading when they instituted a token reinforcement system in programmed reading sessions. Back in the regular classroom, behavior and performance in other academic areas were reported to have been improved. Although the specific responses required in arithmetic, for instance, are different than those in reading, they have a commonality in that they are all responses to academic tasks.

The conflicting evidence regarding the presence or absence of generalization in classroom behavior modification programs points to the
need for further study of the conditions of generalization. In the present study, behaviors operationally defined as competitive or cooperative were reinforced in a game situation and their generalization over task, type of response, and a short time were measured. It was predicted that generalization of the reinforced responses would occur most strongly in the generalization situation most like the original one.

METHOD

Subjects. Subjects were sixteen male and twelve female Caucasian kindergarten children, ranging in age from 5-4 to 6-3, and reported by their teachers to be of normal intelligence. Two like-sex subjects, each from a different kindergarten class, were employed in a single session. In each case, members of the pair reported being unacquainted with each other prior to the experimental session.

Pairs of subjects were randomly assigned to a cooperative or a competitive treatment in a schedule which assured counterbalanced order of the two dependent variable measures. The same experimental treatment was administered on two consecutive days (Phase A) with one dependent variable measure being taken each day immediately after treatment (Phase B).

Treatment Conditions (Phase A) and Materials. For both the cooperative and competitive conditions of the experimental treatment, a board game was employed. The "board" was a large cotton sheet on the floor marked off into a winding trail of squares upon which a subject could
stand. Dice were thrown by each subject in turn to determine the number of spaces that could be travelled. Tokens of a unique color for each of the two players served as reward at the termination of each game. These were exchangeable for candy at the rate of three tokens for one piece of candy. The treatment condition lasted for fifteen minutes after the instructions had been given, several games being played in the course of this time.

In the competitive condition (A-1) the goal given to the subject was to win the game by moving ahead of the other subject on the game board. Praise was given to the subject who was ahead on each turn, and at the termination of the game, the subject who was ahead was given tokens, the size of his lead up to three spaces determining the number of tokens he received.

In the cooperative condition (A-2) the goal emphasized was to minimize the number of spaces between players, thus working together to win a joint reward. Subjects moved in either direction on the game board; both subjects were praised for moves that brought them into the proximity of one another. Reward of tokens was given to each player if they were within three spaces of one another at the termination of the game.

The difference in the two treatments, then, was that in A-1 one subject was rewarded for competitive success and in A-2 both subjects were rewarded for cooperative success.

Dependent Variable Measures (Phase B) and Materials. The same generalization measures were administered to all pairs of subjects. For
generalization measure B-1 (to be called "Hunt") the same colored tokens that were used as reinforcers in Phase A were employed. Each subject was given the task of finding fifteen tokens of a designated color, different for each subject, which were exchangeable for a lollipop reward. Twenty tokens of each color were hidden about the room, with six pairs composed of the two colors hidden together, to insure that subjects would see tokens other than their own. In every case, both children found fifteen tokens and were given the lollipops. Two observers, one of whom was the experimenter, each scored instances of "hindering" and "helping" responses given by each subject. "Hindering" responses were broken down into subcategories of
- push partner or snatch token from him
- keep partner's tokens.

"Helping" responses were broken down into subcategories of
- tell or hint where partner's token is
- give partner his token.

These subcategories were pooled into the categories of "hindering" and "helping" responses for the purpose of analysis.

The second measure (3-2) was "Paste and Color." Each subject had six animal scene cut-outs, six outlines on a sheet of paper corresponding to the animal forms, two crayons, and paste before him. Subjects were given the task of pasting the cut-outs on the sheet of paper, with the suggestion that the cut-outs could be fitted into the outline forms, and coloring the resulting picture. Completion of the task was facilitated if
subjects exchanged pictures and crayons, for three of each subject's pictures fit his partner's outlines, and the total number of crayon colors was four for the two subjects, rather than just two.

Two independent observers recorded instances of "helping" responses into subcategories of

- giving cut-out or crayon to partner in response to a request from partner
- giving cut-out or crayon to partner without his request.

"Hindering" responses were scored in subcategories of

- refusing to give a requested crayon or cut-out
- using partner's cut-out in own picture.

This phase lasted for ten minutes after the subjects began work.

An inter-observer reliability coefficient was obtained by correlating all scores for the two measures between observers. When sub-categories of "helping" and "hindering" responses were pooled, this coefficient was .90. Where disagreements occurred, the average of the observers' scores was used in the analysis.

For each dependent variable measure, a composite score for each pair of subjects was calculated. First, a "hindering" score and a "helping" score for each pair was obtained by pooling the sub-categories of these responses and adding the two individual's scores together. Second, a composite score was obtained by assigning each pair a neutral score of +10, then subtracting the "hindering" score and adding the "helping" score.

In formula form
\[ X_{stp} = 10 + "\text{Help}_{stp} - "\text{Hinder}_{stp} \]

where \( s = \text{sex} \)

\( t = \text{treatment} \)

\( p = \text{pair number} \)

for each subject pair.

**PROCEDURE**

**Phase A.** Subjects were brought into the experimental room and introduced to each other. In both treatments, subject pairs were told that they would be playing a game in which they could win prizes. One subject was told that he would win red tokens, the other that he would win green.

In the competitive treatment (Phase A-1) subjects were given motivating instructions which emphasized winning and obtaining a prize. In turn, subjects advanced the number of spaces indicated by rolling the dice. The experimenter praised the subject who was ahead on each turn. After 2, 4, 6, or 8 turns (randomly determined), a buzzer signal indicated the termination of a game and token reinforcers were administered to the child who was ahead. Three tokens were awarded for a lead of three or more spaces, two tokens for a lead of two spaces, and one token for a lead of one space. No tokens were awarded if both subjects were on the same space. Thus, winning by a large margin was reinforced more strongly than winning by a small margin. The game was played several times in the course of fifteen minutes, the range being 4-7 games and the
mean 5.86. After the first game, the experimenter intervened only to
give praise and tokens as described above.

In the cooperative treatment (Phase A-2), motivating instructions
emphasized working together to win identical prizes by staying close to-
together on the game board. Partners rolled the dice for each other, and
the subject whose move it was could move in either direction on the game
board. Both subjects were praised for moves which brought them into
the proximity of one another. At the termination of each game, each sub-
ject was given one token if they were three spaces apart, two if they
were two spaces apart, and three if they were one space apart or on the
same space. No tokens were awarded if subjects were more than three
spaces apart. Again, doing "very well" resulted in a larger reward than
doing "well." The range of the number of games played in fifteen minutes
was 4-7, with a mean of 5.64 games per session.

Subjects in both treatments were given the opportunity to exchange
tokens for candy at the rate of three tokens per piece of candy when the
series of games was concluded each day.

Phase B-1 ("Hunt"). The first generalization task was to find fifteen
tokens of a designated color to obtain a lollipop reward. Subjects were
given the following instructions:

Have you ever gone on an egg hunt? We are going
to have a hunt today. Hidden around the room are
more tokens. [S's name], here is a green token.
If you bring me fifteen more, you may have this
lollipop (show lollipop). [S's name], here is a
red token. If you bring me fifteen more, you may
have this lollipop (show another lollipop). You
see, I have a lollipop for each of you. You have seven minutes to find fifteen tokens. Do you have any questions? O.K. Go ahead.

The observers scored dependent variable behaviors during this seven-minute period.

Phase B-2 ("Paste and Color"). The task for the second generalization phase was to paste and color animal scene cut-outs on sheets of paper. As the subjects sat at the table with the materials spread out before them, the following instructions were given:

Now you are going to have a chance to make some pictures. (Show example, different from both subjects' pictures). Here I have a cut-out that looks like a chick. Can you see an outline on this paper where the chick fits? (That's right) it goes right here, and I'm going to paste it on. I have another cut-out here. There is no outline on my sheet where it fits, is there? Well, I could fit it someplace else on my sheet, like here, or I could give it to my partner, who does have an outline for it. When the cut-outs are pasted on, I will color the picture. Let's look at yours. See, [S1's name], here is a red crayon and here is a brown crayon, and these cut-outs are (list cut-outs and hold them up). [S2's name], here is an orange crayon and a blue crayon, and your cut-outs are (list cut-outs as before). Go ahead and work on your pictures now.

The experimenter withdrew to another part of the room. The two observers each recorded dependent variable behaviors as the subjects worked on the task.

RESULTS

The composite Cooperation Scale score for each pair (as described under the Dependent Variable heading) was treated as one observation in
a 2x2 (sex by treatment) unweighted means analysis of variance for each of the two dependent variable measures. The results are summarized in Table 1.

On the "Hunt" measure, the mean score for competitive treatment subjects was 10.63, while cooperative treatment subjects obtained a mean score of 19.67. This difference was statistically significant (F = 10.64, p < .01). It is interesting to note that the cooperative treatment increased the mean score above the neutral value of 10, but the competitive treatment did not decrease the score. Analysis of variance on "helping" response scores alone revealed that the difference between the treatment means of 2.00 and 10.13 was significant (F = 8.41, p < .05). "Hindering" responses alone were not significantly different for the two groups.

On the "Paste and Color" measure the obtained mean Cooperation Scale score for cooperative treatment subjects was 15.63; for competitive treatment subjects this score was 11.42. This was, however, a non-significant difference (F = 4.00, .05 < p < .10).

There were no significant sex differences on either Cooperation Scale measure.

DISCUSSION

Generalization of a reinforced response across task and type of response have been observed. Although the reinforced and generalized responses are both described conceptually as competitive or cooperative,
<table>
<thead>
<tr>
<th></th>
<th>&quot;Hunt&quot;</th>
<th>&quot;Paste and Color&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Competitive</td>
<td>Cooperative</td>
</tr>
<tr>
<td></td>
<td>$\bar{x}_{st}$</td>
<td>$\bar{x}_{st}$</td>
</tr>
<tr>
<td>boys</td>
<td>11.25</td>
<td>20.00</td>
</tr>
<tr>
<td>girls</td>
<td>10.00</td>
<td>19.67</td>
</tr>
<tr>
<td>$\Sigma \bar{x}_{st}/2$</td>
<td>10.63</td>
<td>19.84</td>
</tr>
</tbody>
</table>

$\bar{x}_{st} = \text{mean score within a cell} = \frac{\sum_{p=1}^{n} \bar{x}_{stp}}{n} = \frac{\sum_{p=1}^{n} (10 + \text{Help}_{stp} - \text{Hinder}_{stp})}{n}$

where $s = \text{sex}$
$t = \text{treatment}$
$p = \text{pair within cell}$
$n = \text{number of pairs in a given cell}$
Help = number of helping responses for a given pair
Hinder = number of hindering responses for a given pair
their motor aspect differs in the treatment and "Hunt" situations. Similarly, the treatment and "Hunt" situations are both games, but the specific stimuli are different.

The failure of the second generalization measure to reach significance indicates that it is not as sensitive to generalization as the first measure. This situation is apparently further removed along a generalization gradient than the Hunt measure is. To clarify the rationale for such a generalization gradient interpretation the similarities and differences between the treatment and generalization situations will be examined.

"Hunt" and "Paste and Color" have in common that facilitation in reaching one's own goal can be achieved by giving and receiving aid (as in the cooperative Phase A treatment) or an individual may try to do better than his partner at reaching that goal (as in the competitive Phase A treatment). Thus, tokens may be exchanged or withheld in the "Hunt" and cut-outs and crayons may be exchanged or withheld in "Paste and Color."

Differences in the three stimulus settings exist in the degree to which a goal is explicitly defined, the presence or absence of definitive criteria that the goal has been reached, and the presence or absence of reward for achieving a goal. The responses or procedures required also vary for the three tasks.

The goal in each treatment (Phase A) condition was clearly defined; subjects were to win rewards by achieving particular positions on the
game board relative to one another, i.e., they were to maximize or minimize the number of spaces between them. Similarly, the "Hunt" game had an explicitly defined goal, and attainment of the goal was clearly evaluated by counting the number of tokens a subject had. Material reward was presented. This made it easy for a subject so inclined to construe the game as having a "winner"--the first one to collect fifteen tokens.

While the treatment conditions and the "Hunt" measure were similar in that they both had a clearly defined goal, evaluation, and reward for reaching that goal, the "Paste and Color" situation had none of these elements specifically delineated. The instructions emphasized that a variety of responses would be appropriate, and no mention was made of evaluation or reward.

Differences in the responses required for the three situations are even more striking. The treatment response was moving on a game board; the "Hunt" and the "Paste and Color" responses are described by their names. The most obvious similarity between the responses for the treatment and "Hunt" situations is that they both involved the acquisition of tokens, even though the motor behaviors were different. No such similarity in response existed in the "Paste and Color" situation.

The conclusion that "Paste and Color" differs in more ways from the treatment situation than the "Hunt" does imply that it is further removed along a generalization gradient. The experimental results bear out the expectation that generalization is less likely to spread to dissimilar than to similar situations.
It is commonly asserted that boys are more competitive than girls are, perhaps because they are more often reinforced for competitive behavior. The lack of significant sex differences on the dependent variable measures indicates that boys and girls alike are subject to the influence of reinforcement for these behaviors. Perhaps the strength of the treatment condition overrode original tendencies. Further investigation should include a measure of base rates of cooperative and competitive behaviors before treatment to clarify this question.

That the behaviors of interest in this study were so easily influenced is noteworthy. In the classroom, children are rewarded for succeeding in competition, while at the same time they are rewarded for being helpful and cooperative. Perhaps children gradually learn to switch easily from one behavior to another as they discriminate the desired behavior for a situation. It is to be noted that the experimenter in the treatment phase became one of the observers in the dependent variable phase. Even though the experimenter moved to a far corner of the room and said nothing during the dependent variable phase, the possible carry-over effect of giving the response most likely to be approved by the experimenter is not to be discounted.

The results of this study demonstrate the possible generalization of reinforced behaviors over time, type of situation, and type of response. This points to the necessity that planners of behavior modification programs consider the possible range of concomitantly affected behaviors and situations when target behaviors are reinforced.
The author wishes to thank Mr. Timothy Sheehan for his assistance in collecting the data, Miss Joan Pinucci for suggesting the general form of one of the dependent variable measures, and Dr. John T. Guthrie for providing many helpful comments throughout the course of this research.
REFERENCES


Haring, N. G., and Houck, M. A. Improved learning conditions in the establishment of reading skills with disturbed readers. Exceptional Children, 1969, 35, 341-351.


Madsen, C., Becker, W., and Thomas D. Rules, praise, and ignoring:


elements of elementary classroom control. Journal of Applied Behavior Analysis, 1968, 1, 139-151.


