*PRODUCTION AND ELIMINATION OF DISRUPTIVE
CLASSROOM BEHAVIOR BY SYSTEMATICALLY
VARYING TEACHER'S BEHAVIOR¹*

DON R. THOMAS, WESLEY C. BECKER, AND MARIANNE ARMSTRONG

UNIVERSITY OF ILLINOIS AND THOMAS PAINE SCHOOL, URBANA, ILLINOIS

The effects of teacher behaviors on the classroom behaviors of children were investigated by systematically varying approving (praise, smiles, contacts, *etc.*) and disapproving (verbal reprimands, physical restraint, *etc.*) classes of teacher behavior. Measures were taken on both teacher and child behaviors. Each day a sample of 10 children was observed. The subject pool was a class of 28 well-behaved children in a middle-primary public school class. The results demonstrated that approving teacher responses served a positive reinforcing function in maintaining appropriate classroom behaviors. Disruptive behaviors increased each time approving teacher behavior was withdrawn. When the teacher's disapproving behaviors were tripled, increases appeared most markedly in the gross motor and noise-making categories of disruptive behavior. The findings emphasize again the important role of the teacher in producing, maintaining, and eliminating disruptive as well as pro-social classroom behavior.

Teachers are sometimes unaware of the effects of their actions on the behavior of their students. Many teachers assume that if a child performs disruptive acts in the classroom then the child must have a problem at home, or at the very least, must not have reached a stage of sufficient maturity to function adequately in the school situation. However, an increasing body of evidence indicates that many of the behaviors which teachers find disruptive are actually within their control. A teacher can modify and control the behavior of her students by controlling her own responses.

Contingent use of social reinforcement has been shown to control such motor behaviors as walking, standing, and running (Bijou and Baer, 1963), talking and crying (Kerr, Meyerson, and Michael, 1965; Hart, Allen, Buell, Harris, and Wolf, 1964), and classroom conduct (Becker, Madsen, Arnold, and Thomas, 1967; Zimmerman and Zimmerman, 1962).

¹The authors wish to thank Urbana School District #116 and the principal of Thomas Paine School, Mr. Richard Sturgeon, for their cooperation. The observers (Loretta Nielson, Barbara Goldberg, Marilyn Goldberg, and Darlene Zientarski) deserve thanks for their conscientious work. This research was supported, in part, by National Institute of Child Health and Human Development Grant HD-00881-05. Reprints may be obtained from Wesley C. Becker, Bureau of Educational Research, University of Illinois, Urbana, Illinois 61801.

Becker *et al.* (1967) worked in public schools with teachers who had problem children in their classes. Behaviors exhibited by the students were observed and the frequency of these behaviors was estimated for each child. Each teacher was taught to use praise, smiles, *etc.* to reinforce good behavior. The rate of appropriate classroom behaviors increased in most cases as soon as teacher approval and recognition were made contingent on such behavior.

The present study evolved from prior research showing the importance of social reinforcement, and Becker's work, which suggests that specific procedures, or definable classes of teacher behaviors can be used by the teacher to increase appropriate classroom behaviors. In order to provide more convincing data on the role of different teacher behaviors, the present study was designed to produce and remove problem behavior in students by systematically varying teacher behaviors in an initially well-behaved class.

METHOD

Subjects

Students. A class of 28 elementary students at the middle-primary level was selected. According to the teacher her class was "a good class, with an above-average distribution of ability and no 'bad' kids." Most of the children were from upper-middle- and middle-

income-range families. Ages at the beginning of the study ranged from 6 yr, 11 months to 7 yr, 11 months; I.Q. range (group test) was from 99 to 134.

Teacher. The teacher, age 23, obtained her student teaching experience with a class of "maladjusted" children. In addition, she had 1-yr experience with a class of "slow learners". Preliminary observations indicated that she rarely attended in an approving manner to children who behaved inappropriately, and rarely reprimanded children who were performing their assigned tasks. She volunteered to participate in the study because of its potential contribution to teacher training in the future.

Observation Procedures

The basic data for the study consisted of the relative frequency of occurrence of classes of child behaviors in relation to classes of teacher behaviors utilizing rating schedules to be described. One to three observers were placed in the classroom each morning from approximately 9:15 to 10:00 a.m. while the students were completing reading and reading workbook assignments. To insure obtaining a daily sample of both child and teacher behaviors during this 45-min work period, a 20-min observation time was decided on for both child and teacher observations. Thus, even if only one observer was present, the relevant information could be obtained. This time restriction limited the number of children who could be observed each day. Ten children were selected for observation each morning by drawing numbers from a hat. During Baseline₁ and the first No Praise condition a no-replacement procedure was used so that all children had to be observed before a child's number could be drawn a second time. At the start of Baseline₂ this restriction was removed. Through the use of a numbered seating chart, the observers recorded the behaviors of selected children in the order in which they were chosen. Five extra numbers were drawn each day to provide observation targets in case one or more of the first 10 subjects drawn were not available for observation. Target children were observed for 2 min each. Each minute was divided into six 10-sec intervals. Observers were trained to record classes of behavior which occurred in a given interval. Recordings were made during the

first five intervals of each minute. During the sixth 10-sec interval the observers made notes, checked for synchronization, and/or prepared to switch to a new child. Thus, the daily child observation sample consisted of ten 10-sec observation intervals on each of 10 children.

Teacher behaviors were recorded on a similar schedule, the only difference being that for teacher behaviors each occurrence of a response in a specified class was recorded (frequency measure), whereas for child behaviors a given class of behavior could be rated only once in a 10-sec interval. This difference in procedure was necessitated by the greater difficulty in separating child behaviors into discrete response units. Observers used a clipboard, stopwatch, and a recording sheet which had spaces for 100 observation intervals, guides for computing reliability, and a place for comments.

Undergraduate university students were hired and trained to collect the data. Each observer memorized the definitions of classes of child and teacher behaviors. Pre-baseline training in recording of behavior was carried out in the experimental classroom to allow the children to become accustomed to the presence of the observers. The children were already well adapted to the classroom before observer training was started. Observers were instructed to avoid all interactions with the students and teacher while in the class or on the school grounds. At the scheduled time they would enter the class, walk directly to chairs provided for them, sit down, and begin the observations. A hand signal was used to insure synchronization of observation times. Initially two observers were scheduled to observe on Monday, Wednesday, and Friday, and two on Tuesday and Thursday. When a systematic difference developed between the two sets of observers, one of the Tuesday-Thursday observers was placed on a three-day-a-week schedule to tie the two sets of observations together with reliability checks. Thus, on some days there were as many as three observers in the classroom. The number of observers in the classroom varied from one to three. Due to illness or the need to obtain observations in other classroom, there were times when only one observer was available. Observers were not informed of changes in experimental conditions.

Classes of Teacher Behaviors: The Independent Variable

The behaviors emitted by the teacher were defined as belonging to three general classes: (1) Disapproving Behavior, (2) Approving Behavior, and (3) Instructional Behavior. Disapproving and Approving Behaviors were rated only when they immediately followed discriminable child behaviors falling into inappropriate or appropriate classes (see below).² Listings were made of the teacher behaviors that could occur within each class.

The general class of Disapproving Behaviors included Physical Contact, Verbal, and Facial subclasses. The subclasses of Physical behaviors included forcibly holding a child, grabbing, hitting, spanking, shaking, slapping, or pushing a child into position. The Verbal subclass of Disapproving Behaviors included yelling, scolding, raising voice, belittling, or making fun of the child, and threats. Threats included "if-then" statements of loss of privilege or punishment at some future time. For example, the teacher might say to the class, "If you don't remain quiet, you will have to stay in from recess." The Facial subclass of Disapproving Behaviors included frowning, grimacing, side-to-side head shaking, gesturing, *etc.*

The general class of Approving Behaviors also included Physical Contact, Verbal, and Facial subclasses. Approving Physical Contacts included embracing, kissing, patting, holding hand or arm of child, or holding the child in the teacher's lap. Approving Verbal comments included statements of affection, approval, or praise. Approving Facial response was rated whenever the teacher smiled, winked, or nodded at one or more of the children.

The general class of Instructional Behavior included any response from teacher to children which involved giving instructions, information, or indicating correct responses.

In addition to recording the above classes of teacher behavior, note was taken of those

times when the teacher terminated social interaction by turning out lights and saying nothing, turning her back on the class and waiting for silence, or stopping talking and waiting for quiet.

As noted earlier, the observers recorded every teacher response falling in a given class. Thus, the measures of teacher behaviors are frequency counts.

Child Behaviors: The Dependent Variable

The classes of child behaviors were developed by categorization of behaviors occurring with some frequency in the repertoire of problem children (Becker *et al.*, 1967). It was assumed that certain behaviors, because of their common topography, could be grouped together. Five classes of Disruptive Behavior (Gross Motor, Noise Making, Orienting, Verbalizations, and Aggression) and one class of Appropriate Behavior (Relevant) were defined. Behaviors not specifically defined were rated in a separate category (Other Task). Disruptive Behaviors were essentially behaviors apparently incompatible with good classroom learning conditions.

Included in the category of behaviors labeled as Gross Motor activities were: getting out of seat, standing up, walking around, running, hopping, skipping, jumping, rocking chair, moving chair, sitting with chair in aisle, kneeling in chair, arm flailing, and rocking body without moving chair.

The category of Noise Making was rated with the stipulation that the observers must hear the noise as well as see the noise-making action, and included tapping feet, clapping, rattling papers, tearing papers, throwing books or other objects onto desks, slamming desk top, tapping objects on desk, kicking desk or chair, and scooting desk or chair.

The Verbalization category was rated only when the observer could hear the response. Lip movements alone were not rated. Carrying on conversations with other children, calling out teacher's name to get her attention, crying, screaming, singing, whistling, laughing, and coughing were included in the category.

The Orienting class of behaviors required that the child be seated. Turning of head or head and body toward another person, showing objects to another child, and looking at another child were rated. Looking behaviors

²As it turned out, approval following inappropriate behavior occurred only three times and disapproval following appropriate behavior did not occur. Also, this teacher did not make non-response-contingent approval or disapproval comments. Thus, we were dealing essentially with two response-contingent classes of teacher behavior.

of less than 4-sec duration were not rated except for any turn of more than 90 degrees from the desk. When an Orienting response overlapped two rating intervals, and could not be rated in the first interval, because it began too late in the interval to meet the 4-sec criterion, it was rated in the second interval.

Aggression was defined as hitting, pushing, shoving, pinching, slapping, striking with objects, poking with objects, grabbing objects or work belonging to another, knocking neighbor's property off desk, destroying another's property, throwing objects. No judgments of intent were made.

Appropriate behaviors were labeled Relevant and were made more easily identifiable by restricting the observations to a period in the morning when all of the children were preparing reading and workbook assignments. Specific Relevant Behaviors were: looking at the teacher when she was speaking to the entire class or to the child being observed, answering questions of the teacher, raising hand and waiting for teacher to respond, writing answers to workbook questions, looking at pages of text in which reading was assigned. It was required that the entire 10-sec interval be filled with on-task behavior before the Relevant rating was made.

When a child being observed performed a response not defined by one of the categories of Disruptive Behaviors or by Relevant Behavior, a rating of Other Task was made. The Other Task rating was incompatible with Relevant, but could be recorded in the same interval as any or all of the categories of Disruptive Behavior.

When rating the children, the observers were instructed to record each class of behaviors which appeared in an interval regardless of how many other classes had already been recorded in that interval. All five categories of Disruptive Behaviors and the Other Task category were compatible with each other. Relevant Behavior was incompatible with the other categories. No category of behavior was rated more than once in an interval. If a child was conversing with his neighbor, and he made two verbal responses in one interval, this class of behaviors was recorded only once. Thus, each child-behavior measure was a record of intervals in which the response occurred, rather than a count of the number

of discrete responses as in the recording of teacher's behavior.

The overall level of Disruptive Behaviors was defined as the percentage of intervals in which one or more Disruptive Behaviors occurred.

Reliability

Two types of reliability were calculated. Reliability I reflects simply the degree to which two observers obtained the same score for each category of behavior during a 20-min observation period. The smaller score is divided by the larger. Reliability I most appropriately applies to the data as reported in Fig. 1, since these are averages for an observation period. Random errors tend to cancel each other out when a score is based on a series of observations and a reliability measure should reflect the gain in accuracy obtained by averaging. For training purposes, and for greater confidence in the accuracy of the observation procedure, a second type of reliability was also calculated (Reliability II). Reliability II required that the same behavior category be recorded in the same interval by each observer to define an agreement. Reliability II was calculated by dividing the number of agreements by the number of agreements plus disagreements.

During the pre-baseline observer training, reliability checks were required for every observation. Before baseline observations were started, consistent reliabilities (Type II) above 80% were required. Reliability I data based on a weighted average of the reliabilities of the child-behavior codes are reported in Fig. 1, as are the average reliabilities by conditions for teacher behaviors. Comparable Reliability II data averaged 82.6% for child behaviors and 83.2% for teacher behaviors. Reliabilities for individual categories are well represented by these averages.

Sequence of Conditions

The first phase of the study (Baseline₁) consisted of measuring both teacher and child behaviors. No attempt was made to manipulate teacher behavior.

The second phase (No Approval₁) was defined by the absence of Approval Behaviors. The teacher discontinued the use of praise statements and used only contingent Disapproving Behaviors to control the children.

These phases were then repeated (Baseline₂, No Approval₂). At the beginning of No Approval₂ and throughout the rest of the study, the teacher carried a small "supermarket" adding machine with her to count the frequency of Disapproval Behaviors so that she could better monitor her behavior.

The fifth phase of the study, Frequent Disapproval, involved increasing the level of Disapproving Behaviors to approximately three times that given during Baseline₁ while continuing to withhold Approving Behaviors.

Phase 6 returned to the lower level of Disapproval (No Approval₃) and Phase 7 again returned to the baseline conditions (Baseline₃).

The teacher was instructed to maintain experimental conditions throughout the day, not just during the observation period. During the periods when praise was withheld beginning with No Approval₂, checks of the daily counts of Disapproving Responses obtained by the teacher with her counter corresponded closely to those which would have

been predicted by extrapolation from the observation periods.

RESULTS

The relationships of greatest interest are the effects of presence and absence of Approval Behaviors on Relevant Behaviors and the effects of levels of Disapproval Behaviors on Disruptive Behaviors. Because of a systematic rater bias which entered into the data for Other Task Behavior (discussed later), and therefore also affected Relevant Behaviors incompatible with Other Task, greater emphasis is given to the analysis of Disruptive Behaviors in presenting the results.

Average Level of Disruptive Behavior

In Baseline₁ Disruptive Behaviors occurred in an average of 8.7% of the intervals observed. When Approving Behaviors were discontinued (No Approval₁), Disruptive Behavior increased to an average of 25.5% (Fig. 1). Approving Behaviors were again provided

Fig. 1. Disruptive classroom behaviors as a function of nature of teacher behavior. Data points represent 2-min samples on 10 children each day. Dotted lines cross observations where the regular teacher was absent due to a recurrent illness, including a 10-day hospitalization between Days 39 and 41. The dotted line connecting days 44 and 45 represents the Easter vacation break. The data for Day 26 were taken with the teacher out of the room.

(Baseline₂) and Disruptive Behavior dropped to an average of 12.9%. In order to show more conclusively that the changes in Disruptive Behavior were related to the changes in teacher behavior, Approving Behaviors were again discontinued (No Approval₂) and the level of Disruptive Behaviors stabilized near the same level as in No Approval₁ condition (average 19.4%). When the Disapproving Behaviors (critical comments) were tripled (Frequent Disapproval), while Approving Behaviors were still withheld, Disruptive Behavior increased to an average of 31.2% with high points far above any observed before. The behavior stabilized, however, near the level at which the two previous No Approval phases had stabilized. When the rate of disapproval was lowered (No Approval₃), no great reduction in Disruptive Behavior occurred. The average level of Disruptive Behaviors over No Approval₂, Frequent Disapproval, and No Approval₃ was 25.9%. At the end of No Approval₃, Approval was again added to the low level of Disapproving Behaviors, and Disruptive Behavior dropped to an average of 13.2%, with the trend indicating a level far below this average.³

Analysis of Classes of Behavior

Discontinuation of approving behaviors. In reviewing the changes in the individual categories of behavior through the first two with-

drawals of Approving Behavior, the majority of the increase in Disruptive Behaviors could be attributed to changes in Verbalization and Orienting categories (Table 1). The mean of Verbalization in No Approval₁ was 22.6% due to one extremely high observation on the second day of the condition; however, these behaviors stabilized between 9% and 17% (Fig. 2). Orienting showed a slight decrease across No Approval₁ (Fig. 2). The second time Approval was discontinued, Orienting increased across the condition while Verbalization remained relatively stable except for two high observations. Gross Motor behaviors followed the same pattern as Orienting and Verbalization through No Approval (1 and 2), increasing each time Approving Behavior was discontinued and decreasing when Approving Behaviors were present (Fig. 2).

Noise Making and Aggression followed a pattern through No Approval₁ and ₂ which was distinctly different from the other categories of disruptive behavior. Both of these categories of behavior were already occurring at a low frequency in the Baseline condition (Table 1), but they occurred even less often when only Disapproving Behavior was given.

Increase of disapproving behaviors. In the Frequent Disapproval condition, Noise Making, Gross Motor, and Orienting all increased (Table 1). Verbalizations showed a decline

³A conservative statistical analysis was performed (*F* test) to compare those three conditions where approval responses were available with those two conditions where approval responses were withdrawn. For this test the Frequent Disapproval and No Approval_{2,3} conditions were collapsed into one condition. In order

to insure independence of observations, the average values within each condition were used, thus providing four degrees of freedom. Significant differences were found for Relevant Behavior (*p* < 0.01), Noise Making (*p* < 0.05), Gross Motor (*p* < 0.025), and for the overall level of Disruptive Behavior (*p* < 0.01).

Table 1
Average Percentages for Specific Behavior Classes for Each Experimental Phase

Behavior Classes	Baseline ₁	No Approval ₁	Baseline ₂	No Approval ₂	Frequent Disapproval	No Approval ₃	Baseline ₄
Disruptive Behaviors*	8.7	25.5	12.9	19.4	31.2	26.8	13.2
Gross Motor	2.7	6.7	2.0	4.8	12.3	10.4	2.4
Noise	0.9	0.1	0.7	0.09	4.1	4.4	0.9
Verbalization	4.6	22.6	7.7	9.6	7.9	6.0	3.9
Orienting	1.4	6.5	4.1	7.1	11.5	10.2	7.6
Aggression	0.25	0.01	0.2	0.01	0.04	0.04	0.1
Other Task	7.0	10.4	5.9	10.7	5.9	4.2	1.2
Relevant	84.1	65.3	83.9	72.1	64.3	69.4	85.6

*The addition of percentages for the five classes of Disruptive Behaviors will usually lead to a sum higher than that reported as percentage of Disruptive Behaviors, since the latter does not reflect the occurrence of more than one subclass of Disruptive Behaviors in a given 10-sec interval.

Fig. 2. Analysis of specific behavior classes by condition. Data points represent 2-min samples on 10 children each day. See notes under Fig. 1.

over this condition and continued to decline through the rest of the study.

Changing from a high level of Disapproving Behaviors to a lower level did not markedly change the frequency of the various categories of Disruptive Behaviors relative to their

terminal level under the Frequent Disapproval condition.

When Approving Behaviors were again used by the teacher (Baseline₃), the frequency of Gross Motor, Noise Making, and Orienting behaviors decreased noticeably (Fig. 2). Ver-

balization continued to show the steady decrease in frequency which had started in the Frequent Disapproval condition. In Baseline₃, Aggression again occurred, but rarely. All Disruptive Behaviors except Orienting dropped to the level of the initial Baseline (or below) during the final Baseline.

Relevant behavior. Appropriate behaviors were initially high in the classroom (Fig. 2). Behaviors such as getting out of seat to move to a reading group or to check a completed workbook assignment were rated in the Gross Motor category. The requirements for such behaviors, however, remained constant through all conditions so changes in the level of Relevant Behaviors cannot be attributed to changes in classroom requirements. Relevant Behavior decreased each time Approving Behavior was discontinued and increased each time the Approval was reinstated. Relevant Behavior was at a slightly higher level during the final Baseline than during the initial Baseline.

Other task: Behavior not specifically defined. As indicated earlier, a systematic rater difference was encountered early in the study in rating Other Task behaviors. In Fig. 2 this bias can be seen by contrasting data collected on Days 2, 4, 6, 8, and 10 from one set of observers with that collected on Days 1, 3, 5, 7, and 9 by another set of observers. While an attempt was made to correct this bias by interlocking reliability checks, it is apparent that the bias continued to some extent throughout the study. Since Other Task is by definition incompatible with Relevant Behavior, Relevant Behavior shows the same bias. By looking at Disruptive Behavior, defined so as to exclude Other Task behaviors, the systematic bias was largely eliminated from the data presented in Fig. 1.

Teacher Behaviors

The behavior of the teacher remained under good control throughout the study. Averages by conditions for Approving and Disapproving Behaviors are given in the upper part of Fig. 1. As the conditions were changed, little difficulty was found in withholding behaviors in the Approving category. Some difficulty was reported by the teacher in regulating the frequency of Disapproving Behaviors while withholding Approving Behaviors, but a partial solution to this problem was

found. The teacher found that by carrying a small hand-counter (mentioned earlier) she could more accurately judge the frequency of her critical comments. In the Frequent Disapproval Phase there were days when the children were not emitting enough Disruptive Behaviors for critical comments to be appropriate at the programmed frequency. Rather than make inappropriate comments, the rate of Disapproving comments was adjusted to the frequency of the Disruptive Behaviors. When enough Disruptive Behaviors were available, Disapproving Behaviors were dispensed at a maximum rate of one per minute throughout the day; thus, many of the responses of the children were reprimanded very quickly.

General frequency of instructional comments did not change appreciably across conditions. However, the teacher did increase the frequency with which she would say in a neutral tone whether responses were correct or incorrect in the phases where Approval was not given.

The behaviors characterized as Terminating Social Interaction occurred only twice during the study and were, therefore, not subject to further analysis.

Substitute teachers. Observations taken on the days when a substitute teacher was in charge of the classroom appear in four conditions of the study. The frequency of Disruptive Behaviors increased in the presence of a temporary teacher as long as the regular teacher was in either Baseline or No Approving Behavior phases. When the Disapproving Behavior was being dispensed at a high rate, however, the level of Disruptive Behaviors decreased in the presence of a temporary teacher (Fig. 1).

Day 26. The data for this day were taken while the teacher was out of the room. Since the experimental conditions were not operative, this point should have been omitted altogether.

DISCUSSION

The results indicate that some aspects of the behaviors included in the category of Approving Behaviors were reinforcing for task-appropriate behaviors. The frequency of Relevant Behaviors was high whenever Approving Behaviors followed Relevant child

Behavior, and decreased whenever Approving Behaviors were discontinued.

In each change of conditions that involved discontinuation of Approving Behaviors, there appeared a reliable transition effect (observation Days 11 and 27). This effect may be an example of the typical increase in rate found when a positive reinforcer is removed. In support of this explanation, the teacher reported, "When I stop praising the children, and make only negative comments, they behave very nicely for three or four hours. However, by the middle of the afternoon the whole classroom is chaotic." Since observations were taken during a study period in the morning, the periods of good behavior show up in the data each time a condition was changed. A similar low deviant behavior point occurred at the transition to Frequent Disapproving Behaviors (day 37), but it is not clearly explained. "The children seemed stunned."

Reviewing the individual classes of Disruptive Behaviors brings out certain similarities and differences among the classes. During the first alternations of Baseline with discontinuation of Approving Behaviors, Gross Motor, Orienting, and Verbalization Behaviors increased with discontinuation of Approval, while Noise Making and Aggressive Behaviors remained at their already low frequency. The increases are interpreted as suggesting that some responses in the disruptive classes may be reinforced by peer attention or other environmental circumstances when control through approving teacher responses to incompatible behaviors is withdrawn. For example, Orienting behaviors, such as looking around the room or out the window may be reinforced by seeing other children playing, by observing a custodian cleaning up the schoolyard, or by seeing any of numerous events which have no relationship to the classroom. Observational evidence for this inference was clearest in the Frequent Disapproval phase (below). It is also possible to attribute the increases in Disruptive Behaviors during No Approval₁ to the increase in use of Disapproval. However, the data for No Approval₂, where Disapproval was held to the Baseline level, would argue that the effect was primarily related to the withdrawal of approval.

Increasing Disapproving Behaviors to a high level produced four days where Disruptive

Behaviors were above 40%. Several individual categories of behavior also showed marked changes. The increase in Gross Motor Behaviors was related to an increase in interactions with other students. During the Frequent Disapproval condition, two or three children would make alternate trips to check their workbooks at a table provided for that purpose. Only one child was permitted at the table at a time. During Baseline and No Approval phases, it was rare to see a child make more than one trip to the table; in the Frequent Disapproval phase, some of the children would check their papers several times. Others responded by pushing their papers off of their desks and then getting up to get them. There was a noticeable "pairing off" with two or more children exhibiting the same behaviors.

Another consequence of the Frequent Disapproval phase was a marked increase in the noise level in the room. A majority of the noises during this period were created by children scooting their desks and chairs. One observer reported, "I waited for a few minutes after the regular observation period was over and counted the noises. During one 40-sec period, I counted 17 separate chair scraping noises. They came in bursts of two or three at a time. It looked as though the kids were trying to irritate the teacher." The noises in "bursts of two or three" seemed similar to the "pairing off" of children noted with the Gross Motor behaviors, and strengthens an hypothesis that reinforcement from peers is one of the elements which accounts for the increase in Disruptive Behaviors during this time. Peer attention cannot be the only element affecting the behavior of the children, however, because the Verbalization category of behaviors showed a constant decrease throughout the Frequent Disapproval condition. The inhibition of Verbalization could be due to interfering emotional responses being elicited by the high level of critical comments by the teacher. More probable, however, is that the children simply talked more quietly to avoid being caught by the teacher. Observers' reports indicate that a substantial number of verbalizations would have been recorded during the Frequent Disapproving Behaviors condition if there had been no requirement that the responses be heard by the observers. The children could be seen to turn

their heads, and lip movements could be seen frequently, but the verbalizations could not be heard.

Work by Lövaas, Freitag, Kinder, Rubenstein, Schaeffer, and Simmons (1964) suggests that for some children any adult attention may be reinforcing. Some of the present findings under the Disapproval conditions could also be interpreted as indicating that teacher behavior of the Disapproving variety was positively reinforcing. The level of Disruptive Behaviors during each of the conditions when only Disapproving and Instructional attentions were available does appear to vary with the level of Disapproving Behaviors dispensed by the teacher. Unfortunately, the illness-caused absences and Easter break make the results less clear than hoped. It should be apparent that the effect of Frequent Disapproval on the behavior of the children is not subject to a simple interpretation. Some criticized behaviors decreased, some increased, and several possible controlling stimuli could have been operating with contradictory effects on behavior. It is obviously difficult in a field-experimental study of this complexity to maintain control of all the possibly relevant variables at once.

Another limitation of the present design should be noted. Because of a shortage of observation time under the desired classroom conditions, a sample of 10 children was observed daily. A procedure which included all children each day would have provided a stronger basis for analysis of effects on individuals. A rough analysis of individuals with the present data confirms, however, that an average of 76% of the students made changes in the same direction as the group changes. From Baseline, to No Approval, 81% of the students showed increases in Disruptive Behavior. When Approving Behavior became available, 75% of the students improved within two weeks. Discontinuing Approving Behavior a second time resulted in 78% of the students being more disruptive, while the final addition of Approving Behavior showed an increase in appropriate behavior for 71% of the children. Across condition changes, 5% of the children showed no change on the average, and 19% showed change (usually minor) in an opposite direction. Procedures which permitted specifications of which children were praised or criticized for which be-

haviors would be needed to clarify fully individual effects. It is quite possible that the children who changed opposite to the group trend were being responded to differently. Of course, there are many ways one can speculate here. In an as yet unpublished study we have shown that praising some children but not others leads to changes in the behavior only for the children who are praised. Results of this sort emphasize the importance of looking at individual contingencies.

Brief mention should be made of the possible ethical considerations involved in producing disruptive behaviors. One needs to weigh the potential gains in knowledge against the short-term or long-term deleterious effects on the children or teacher. On the basis of prior research and the return to baseline after the first No Approving Behaviors condition, the teacher and the experimenters were confident that appropriate behaviors could be readily reinstated at any time it was felt necessary. It may also be reassuring to know that this accelerated middle primary class did achieve well academically during the year. The children completed all second and third grade work and were all performing on a fourth grade level by the end of the year.

IMPLICATIONS

This further demonstration of the importance of specific teacher behaviors in influencing classroom behavior has a double implication. First, the teacher who uses her Approving Behaviors as immediate consequences for good behavior should find that the frequency and duration of appropriate behaviors increase in her classroom (at least for most children). On the other hand, the teacher who cuddles the miscreant, tries pleasantly to get a child to stop behaving disruptively, talks with a child so that he "understands" what he was doing wrong, or who pleasantly suggests an alternative activity to a child who has been performing inappropriately, is likely to find an increase in the very behaviors she had hoped to reduce. This view of the functional importance of teacher's behavior in creating, maintaining, or reducing classroom behavior problems contrasts sharply with that generated by psychodynamic models of problem behaviors and what to do about them. Work of this sort also suggests a need

to re-evaluate the popular cliché about the importance of the interaction of the "personality" of the teacher with that of the child in looking at classroom management procedures.

The suggestive evidence that peer reinforcement (among other stimuli) takes over when social reinforcement is not provided by teacher is given support by the recent work of Wahler (1967). Wahler has shown how preschool children can systematically control the behavior of their peers by differential use of social reinforcement. The more general implication for the teacher is this: unless an effort is made to support desirable classroom behaviors with appropriate consequences, the children's behavior will be controlled by others in ways likely to interfere with the teacher's objectives.

Finally, the possibility that critical comments may actually function to increase some behaviors upon which they are contingent cannot be overlooked. A recent study (Madsen, Becker, Thomas, Koser, and Plager, 1967), gives clear evidence that some forms of critical comment do function to strengthen behavior. The more often a teacher told first graders to "sit down", the more often they stood up. Only praising sitting seemed to increase sitting behavior.

REFERENCES

- Becker, W. C., Madsen, C. H., Jr., Arnold, Carole R., and Thomas, D. R. The contingent use of teacher attention and praise in reducing classroom behavior problems. *Journal of Special Education*, 1967, 1, 287-307.
- Bijou, S. W. and Baer, D. M. Some methodological contributions from a functional analysis of child development. In L. P. Lipsitt and C. S. Spiker (Eds.), *Advances in child development and behavior*. New York: Academic Press, 1963. Pp. 197-231.
- Hart, Betty M., Allen, K. Eileen; Buell, Joan S., Harris, Florence R., and Wolf, M. M. Effects of social reinforcement on operant crying. *Journal of Experimental Child Psychology*, 1964, 1, 145-153.
- Kerr, Nancy; Meyerson, L., and Michael, J. A procedure for shaping vocalizations in a mute child. In L. P. Ullman and L. Krasner (Eds.), *Case studies in behavior modification*. New York: Holt, Rinehart, & Winston, Inc., 1965. Pp. 366-370.
- Lövaas, O. I., Freitag, G., Kinder, M. I., Rubenstein, D. B., Schaeffer, B., and Simmons, J. B. Experimental studies in childhood schizophrenia—Establishment of social reinforcers. Paper delivered at Western Psychological Association, Portland, April, 1964.
- Madsen, C. H., Jr., Becker, W. C., Thomas, D. R., Koser, Linda, and Plager, Elaine. An analysis of the reinforcing function of "sit down" commands. In R. K. Parker (Ed.), *Readings in educational psychology*. Boston: Allyn and Bacon, 1968 (in press).
- Wahler, R. G. Child-child interactions in free field settings: Some experimental analyses. *Journal of Experimental Child Psychology*, 1967, 5, 278-293.
- Zimmerman, Elaine H. and Zimmerman, J. The alteration of behavior in a special classroom situation. *Journal of the Experimental Analysis of Behavior*, 1962, 5, 59-60.

Received 6 November 1967.