Research has shown that prosocial behavior can be encouraged in an educational setting through social reinforcement and appropriate role modeling. To test the combined effect of media presentations and teacher behavior to promote cooperative behavior in preschool students, 34 children were divided into four groups, each of which was given one of four different treatments. Each group viewed three videotapes which encouraged cooperation, and three of the groups were subsequently subjected to varying degrees of teacher sanctions which reinforced the content of the films. Each group was pre- and posttested for cooperative behavior. The unexpectedly small final sample prohibited any confident conclusions. Tabular data and suggestions for improved research design are included. (EMH)
Interim Report

Effects of Co-Viewing Teachers' Sanctions
Upon the Modeling of Televised Cooperation:
A Pilot Study

Edward E. Hill
Houston, Texas

An Abridged Supplement to
a Presentation at the
July 30, 1975 Meeting
of the Research Staff of
the Children's Television Workshop
New York, New York
Abstract

Using measures developed by Paulson, McDonald and Blittemore (1973) for their formative evaluation of *Sesame Street* cooperation inserts, a pilot study was performed to determine the possible impact of evaluative comments by co-viewing day-nursery teachers upon the modeling and recognition of the cooperative strategies displayed in selected CTW social goal videotapes. The role of positive or no teacher sanctions and subsequent co-viewing teacher presence or absence during testing was examined in relation to pretest and posttest scores of S's in viewing groups. A non-viewing control group was also pre- and posttested. Given an unexpectedly small final sample N and vastly discrepant pretest cooperative levels for each of the five classes participating, treatment conditions were irremediably confounded with pretest dispositions by classroom. However, a review of trends in mean scores by classroom reveals that a ceiling effect may have operated on cooperative performance at posttest since gain scores are inversely related to pretest cooperative levels. Previous television impact research which confirms either a ceiling effect (for cognitive program material) or a "socialization void" influence (for social or moral development programming) is cited to support the present results. Improvements in design for follow-up research on teacher comment intervention are discussed along with observations concerning the use of this technique for in-school television viewing activities.
N.B., For more extensive background information on this and related commenter-intervention methods, readers are directed to a recent thesis by the author (Will, 1975) from which the present study was adapted.
There now exist a considerable number of empirical studies which demonstrate that the behavior of young children can be controlled by peers and adults using only social reinforcement. Behavior modification techniques with children have been used to decrease problem behavior, to increase prosocial behavior, or to do both at the same time (Gelfand and Hartmann, 1968). Most studies of this kind have used adult social reinforcement, usually administered by a teacher or parent, to modify various forms of social behavior in the classroom or the home. Allen et al. (1964), for example, demonstrated that contingent teacher attention increases a shy preschool girl's interaction with her peers. Similarly, Hart et al. (1968), working with a five-year-old girl, made teacher approval contingent upon the child's increased cooperation with her preschool classmates.

It has been unequivocally shown that modeling is also an effective means for teaching various forms of social behavior via both live and symbolic displays (cf. Bandura, 1969). Contingent social rewards or punishments delivered by one live or filmed model to another have heightened postviewing levels of imitation for the behaviors presented, be they prosocial (Bandura, Grusec & Menlove, 1967) or antisocial in nature (Bandura, Ross & Ross, 1963). An equally important issue is whether or not vicarious social rewards or punishments in the form of approving or dis-
approving evaluative comments by co-viewing teachers, parents or peers would affect postviewing replication or retention of those behaviors by other viewers. At issue in the latter case is neither the nature of the social interaction viewed, be it live, filmed, or televised, nor its consequences, but rather the type of comment made by co-viewers concerning the behaviors displayed and the impact such commentary might have upon another observer. Preliminary evidence from studies of modeled film aggression shows that for antisocial behaviors, comments from co-observing adults can influence children's subsequent performance levels in situations designed to test for imitative aggression.

Hicks (1968) discovered that a co-viewing adult's approving or disapproving comments, or lack thereof, concerning the aggressive behavior portrayed in a film that preschoolers viewed were always reflected in subsequent modeling levels of those behaviors. Hence, the adult's disapproving comments, for example, resulted in less modeling of filmed aggression (i.e., inhibition) whereas his approving comments increased it (dissinhibition). Negative comments included "that's awful," (general) and "he shouldn't hit Bobo," (behavior-specific). Grusec (1973) replicated Hicks' (1968) results and found, as did Hicks, that the differential effects of the adult's comments manifested themselves only if the adult accompanied the child to the experimental playroom after the viewing session. The children did not, it would appear, "internalize" the sanctions or criticisms of aggressiveness from the co-viewing
adults, but they did act in accordance with the adult's presumed wishes— and their presumed social rewards for pleasing the adult. There is further evidence, from subjects who both modeled the aggressive acts and concurrently repeated the older viewer's disapproving remarks about those behaviors, that the children did not understand the evaluations of content provided by the adult. A number of other recent studies confirm that the young child, especially the preschooler, neither understands nor remembers very much of what live or symbolic models do or say (cf. Hill, 1975). This is not to suggest that the child cannot be influenced by those models or by program comments from parents, teachers, or even peers and siblings.

The present series of investigations is intended to help determine whether or not the modeling of prosocial behaviors such as cooperation are also subject to the mediating influence of evaluative commentary by adults or peers who co-view prosocial filmed or televised programs with youngsters. The study reported here is the first in what should be a series of formative probes treating only the teacher-child-television interaction question. The strength of such commenter influences as well as their situational determinants will be explored as refinements in testing and measurement designs are developed.
Method

Subjects

The subjects were 34 children from the Little Friends Day Nursery in Portland, Oregon. (This center was, coincidentally, one of those chosen by Paulson, McDonald & Whittemore (1973) for their Sesame Street Social Goals Study in 1971-72.) Subjects ranged in age from 42 to 79 months with a mean age of 61 months (5.10 years). The slightly higher than normal mean age is explained by the presence of numerous five-year-olds and some six-year-olds who had not yet entered elementary school, since the state of Oregon has relatively few kindergartens. A breakdown of subjects by age reveals that four- and five-year-olds predominated (N=12, N=13 respectively), with a few three- and six-year-olds completing the sample (N=4, N=5 respectively). A total of 15 boys and 19 girls were both pre- and posttested. Nearly all children in the center are from low income households (i.e., having below $4,000/yr. mean salary according to administrative records), and virtually all are receiving financial assistance from the local Child Services Division to permit them to attend. In addition, a majority are from one parent homes but very few are drawn from racial minorities. Hence subjects were, on the whole, slightly older and less racially diversified than those usually included in CTW studies of Sesame Street impact. The sample is, however, characteristic of the program's "target" audience of preschool children who have not yet begun formal education.
Subjects were drawn from all five classes in the day nursery. Although normal class sizes range from 12 to 14 children with two teachers per class, summer attrition was particularly severe during the June 1975 testing period. Thus, from an expected sample size of 60-70, only 47 children were available for pretesting. Of these, a total of 34 S's were both pre- and posttested.

Procedure

A specific room in the center was selected for the entire procedure. As the children normally view television in this room, it was here that they were pretested and viewed the videotapes prior to posttesting.

For pretesting, the children were taken in dyads by classroom, with the same teacher present for all dyads drawn from a given class. The teacher present was asked not to prompt or otherwise engage the children during testing.

The materials used to determine baseline levels of cooperation were selected from among those developed by Paulson, McDonald & Whittemore (1973). These included two situational tests of cooperation, "Coat Hanger" and "Paint Brush," and one picture recognition test, "Orange Juice," briefly described below:

Coat Hanger . . . Children given two coats and one hanger.
Instruction: Hang up the coats.

Paint Brush . . . One child given paint, another brushes.
Instruction: Paint a picture.

Orange Juice . . . Question: In which picture are they cooperating? (pp. 19-20)
The first two tests were systematically alternated from dyad to dyad, so as to avoid any order effect. The picture recognition test was consistently given after the two situational tests. In addition to a teacher and the experimenter, two observers were present at all times in the testing room. They implemented a scoring procedure adopted from McDonald and Paulson (1971), in which a Likert-style scale from 0-4 was used, with a score of 4 signifying genuine "cooperation," and a score of 0 signifying "obstructive behavior." Although the 0-6 scale later developed by Paulson, McDonald and Whittemore (1973), represents a more theoretically refined measure, the earlier scoring procedure proved more manageable for the current study. Using the 0-4 point scale, trial scoring in free play observations prior to pre-testing established an interobserver agreement level of .79. The 0-4 point scale was of course only utilized for scoring the two situational tests. In the case of the picture recognition test a correct response translated into 1 point, whereas incorrect responses were scored 0.

Though a fixed time sampling technique was not used, the children were scored for the highest behavior obtained during a two minute period of interaction. In the event of a conflict between the two independently assessed scores, the lower score prevailed. Since the children were tested in dyads, it was determined that a perfect score of 4 could only be obtained by both members of a given pair. Lower scores within a dyad could, however, be at variance
with one another.

For posttesting, each class of children was divided into pairs that would be exposed to differing treatment conditions. This precautionary measure was taken to reduce classroom-related predisposition from confounding results by treatment. Assignment to groups was done randomly by classroom with no systematic pairing by age, sex, or pretest disposition. Children could not be pre- and posttested with the same partner since the sample size would have been further reduced had any pretest partners been unavailable at posttest. Control subjects were drawn from all five of the participating classes. They were pre- and posttested, as with the experimental subjects, but did not view any "cooperation" videotapes. As with pretests, the materials utilized for posttesting were selected from among those developed by Paulson, McDonald and Whittemore (193). Subjects in the viewing groups watched three brief videotapes, of which two were directly related to the situational tests of cooperation that followed. These were "Truck and Blocks" and "Draw-A-House" (p. 19). A third segment, "Neighborhood Library," was also used since in it the word "cooperation" is repeated several times. At posttest another picture recognition test was administered ("Drinking Fountain"; p. 20), which, unlike the situational tests, did not directly parallel the specific content of the third videotape segment shown. This segment was thus included to establish the correlation of posttest performance levels in the situational tests with knowledge of the word "cooperation."
tion," i.e., as a measure of "symbolization."

During the presentation of the videotapes, the co-viewing teachers provided either positive or no evaluative statements. Comments were presented during the time the models were performing the cooperative act. The experimenter instructed teachers to make the following general as opposed to behavior-specific comments: "I like that; cooperation is good; I like cooperation."

Teachers assigned to the no-comment viewing groups either remained in the viewing room with the S's during immediate post-viewing testing, or were replaced by the other non-viewing teacher from the same class. During the testing of comment viewing groups, co-viewing teachers were sometimes likewise replaced by their non-viewing counterparts, according to treatment condition. Due to an insufficient sample N in the two no-comment viewing conditions (i.e., with teacher presence or absence during posttesting), these two groups were combined for the purpose of analysis.

The same scoring system was used for pre- and posttests, with the 0-4 point scale again employed for the two situational tests for cooperation, and the score of 0 or 1 given for the picture recognition test. The two observers were at all times blind to the experimental treatment group that they were rating.

A summary of treatment conditions along with a breakdown of experimental and control groups by component classrooms follows.
A Summary of Treatment Conditions

Group 1 (N=8) Teacher co-views three cooperation inserts (about six minutes total), comments favorably about them, and remains in the viewing room during posttesting.

Group 2 (N=8) Teacher co-views, comments favorably, then leaves the room during posttesting to be replaced by his/her colleague, who has neither viewed the segments with the children, nor commented about their content.

Group 3 (N=8) Teacher co-views, but makes no comment concerning the segments. The co-viewing teacher-present vs. teacher-absent (during posttesting) groups were combined to yield a minimally large sample size of eight children.

Group 4 (N=10) Control Group. Children were pre- and posttested with the same teacher who was present at testing for the viewing group from their particular class, but they neither viewed the videotapes, nor did they receive any prompting concerning "cooperation."

Children from each class were assigned to a minimum of two different treatment groups:

Group 1 (Teacher comments and remains for testing); Total N=8
Class A, N=4
Class B, N=4

Group 2 (Teacher comments and leaves for testing); Total N=8
Class C, N=4
Class B, N=4

Group 3 (Teacher makes no comment); Total N=8
Class D, N=4
Class A, N=2
Class E, N=2

Group 4 (Controls, no-viewing); Total N=10
Class A, N=2
Class B, N=2
Class C, N=2
Class D, N=2
Class E, N=2
The above outlined procedure required eight school days to execute.

Results

The unexpectedly small final sample size precludes the reporting of assertive results. In an effort to reduce the number of independent variables while strengthening the power of the available data, the two no-comment groups were combined as were scores on each of the two pairs of situational tests of cooperation used at pre- and posttest. The latter technique changed individual scores of cooperation originally recorded from 0 through 4 to paired scores reported from 0 through 8. A random selection of tests of significance (at $p < .05$) were then performed on the data. However, these tests failed to disclose affirmative results at acceptable levels of error. Although a complete analysis of available data was not undertaken, any positive results which might be thereby uncovered would remain tentative, given the unfortunately small sample size from which they would be drawn. Accordingly, the review of trends that follows is, of course, highly speculative in nature.

Differences in baseline performance levels in the situational tests seem to have wiped out any treatment effects per se, excluding differences between view and non-view groups. However, some potentially interesting trends are apparent on a classroom by classroom level. Barring the equally convincing possibility of regression
toward the mean, the results of the situational tests of cooperation suggest an inverse relationship between pretest scores for a given classroom and posttest gain scores for that classroom. Children from classes which performed substantially above the overall pretest mean score of 4.65 (i.e., Classes A and C) either declined or remained at essentially the same level at posttest. Conversely, children who began at or below average baseline levels of cooperation improved substantially at posttest (i.e., Classes B, D and E). These trends are particularly apparent when control S's are deleted and only the mean scores of experimental group S's are reviewed (see Table II). Although the changes from pre- to posttest were nearly always in the predictable direction for the situational tests of cooperation, the corresponding changes in posttest scores in the picture recognition test was not. Differences were marginal here, but it should be pointed out that these scores are necessarily less reliable than those obtained from the situational tests, since they represent only one independent measure at pre- and posttest instead of two.
TABLE I

Mean Scores by Treatment Condition for Situational and Picture Recognition Tests of Cooperation

<table>
<thead>
<tr>
<th>Group</th>
<th>View, Comment, Remain</th>
<th>Pretest Cooperation</th>
<th>Posttest Cooperation</th>
<th>Gain</th>
<th>Pretest Picture Recognition</th>
<th>Posttest Picture Recognition</th>
<th>Gain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>View, Comment, Remain</td>
<td>5.63</td>
<td>5.75</td>
<td>.12</td>
<td>.250</td>
<td>.250</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>N=8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group 2</td>
<td>View, Comment, Leave</td>
<td>4.50</td>
<td>6.50</td>
<td>2.00</td>
<td>.375</td>
<td>.500</td>
<td>.125</td>
</tr>
<tr>
<td></td>
<td>N=8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group 3</td>
<td>View, No Comment</td>
<td>4.00</td>
<td>5.38</td>
<td>1.38</td>
<td>.375</td>
<td>.625</td>
<td>.250</td>
</tr>
<tr>
<td></td>
<td>N=8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group 4</td>
<td>Non-View</td>
<td>4.50</td>
<td>5.10</td>
<td>.60</td>
<td>.500</td>
<td>.400</td>
<td>.100</td>
</tr>
<tr>
<td></td>
<td>N=10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall Means</td>
<td></td>
<td>4.65</td>
<td>5.65</td>
<td>1.00</td>
<td>.382</td>
<td>.444</td>
<td>.052</td>
</tr>
<tr>
<td>N=34</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### TABLE II
Mean Scores by Classroom for Situational Tests of Cooperation

<table>
<thead>
<tr>
<th></th>
<th>Experimental S's</th>
<th>Control S's</th>
<th>Exp. &amp; Control S's</th>
<th>Deviation Score (of Overall Mean at Pretest-Class Mean at Pretest)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre</td>
<td>Post</td>
<td>Gain</td>
<td>Pre</td>
</tr>
<tr>
<td>CLASS A (N=6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>7.50</td>
<td>5.50</td>
<td>-2.00</td>
<td>5.50</td>
</tr>
<tr>
<td>CLASS B (N=10)</td>
<td>4.00</td>
<td>6.87</td>
<td>2.87</td>
<td>3.00</td>
</tr>
<tr>
<td>CLASS C (N=6)</td>
<td>5.00</td>
<td>5.50</td>
<td>.50</td>
<td>7.00</td>
</tr>
<tr>
<td>CLASS D (N=6)</td>
<td>2.00</td>
<td>3.75</td>
<td>1.75</td>
<td>4.00</td>
</tr>
<tr>
<td>CLASS E (N=4)</td>
<td>4.00</td>
<td>8.00</td>
<td>4.00</td>
<td>3.00</td>
</tr>
<tr>
<td>OVERALL MEANS (N=34)</td>
<td>4.70</td>
<td>5.88</td>
<td>1.18</td>
<td>4.50</td>
</tr>
</tbody>
</table>

(N=24) | (N=10) | (N=34)
### TABLE III

Mean Scores by Classroom for Picture Recognition Tests of Cooperation

<table>
<thead>
<tr>
<th></th>
<th>Experimental S's</th>
<th>Control S's</th>
<th>Exp. &amp; Control S's</th>
<th>Deviation Score (of Overall Mean at Pretest - Class Mean at Pretest)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre</td>
<td>Post</td>
<td>Gain</td>
<td>Pre</td>
</tr>
<tr>
<td><strong>CLASS A</strong> (N=3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.500</td>
<td>.830</td>
<td>.330</td>
<td>1.00</td>
</tr>
<tr>
<td><strong>CLASS B</strong> (N=10)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.375</td>
<td>.625</td>
<td>.250</td>
<td>1.00</td>
</tr>
<tr>
<td><strong>CLASS C</strong> (N=6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.250</td>
<td>.250</td>
<td>0</td>
<td>.500</td>
</tr>
<tr>
<td><strong>CLASS D</strong> (N=6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>.500</td>
<td>.500</td>
<td>0</td>
</tr>
<tr>
<td><strong>CLASS E</strong> (N=4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.500</td>
<td>.500</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>OVERALL MEANS</strong> (N=34)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.333</td>
<td>.533</td>
<td>.250</td>
<td>.500</td>
</tr>
</tbody>
</table>

*(n=24) (n=10) (n=34)*
Discussion

It has been recently shown that the impact of televised social development material for children may be subject to significant alteration by the particular subject's assumed pretest disposition. Friedrich & Stein (1972) found that preschoolers from low-SES backgrounds would benefit substantially more from prolonged viewing of Misterogers Neighborhood than would their higher-SES peers, who ironically manifest increased prosocial behavior when exposed to high-action aggressive programming, such as Batman or Superman episodes. The theoretical framework of a "socialization void" is cited to account for their results. Within this context, it is the novelty of the programming in question which in large part determines the child's subsequent modeling of the behaviors presented in it. Hence a lower-SES child is presumed to have had less contact with prosocial displays in the home and elsewhere than his or her higher-SES peers. For this reason the lower-SES child responds well to the novelty of Misterogers. The higher-SES peer is already more likely to be a frequent viewer of Misterogers and other education-oriented television programs. He or she is assumed to have been exposed to more frequent examples of prosocial behavior than lower-SES peers. Thus action packed, rarely viewed programs like Batman would elicit more interest and subsequent imitation from this sort of youngster than the already trite Misterogers.
To the extent that pretest baseline levels of cooperation are also sources of information on the probable behavioral dispositions of S's, the results of the current study tend to lend very tentative support to the "socialization void" theory as applied to television impact research.

For cognitive goal material like The Electric Company, ceiling effects have been well established (Ball & Bogatz, 1973) for older, more competent viewers, but the nature and scope of related effects upon social development programming are still unclear.

Recommendations for Further Research

1. Replicate current study with sample N of 80 plus; assure matched pairs of S's at pre- and posttest; include at least three situational tests of a given prosocial behavior at pre- and posttest with a minimum of two projective tests to complement them.

2. Adapt or develop a revised scoring procedure to allow for dichotomous, criterion-based judgements of the critical behavior.

3. Develop a methodology to control for the influence of age, sex and intelligence of S's upon scores.

Observations Concerning the Use of the Teacher-Comment Technique with Prosocial Programming for In-School TV Viewing Activities

Perhaps the single most intriguing and unexplored aspect of the pilot study described above is the extent to which the impact
of teacher commentary is modified by the predispositions of the children. Observations made both prior to and during testing of teachers and children from all classes included in this research revealed vast differences in style and frequency of child-child and teacher-child contacts. These differences were undoubtedly reflected in pretest performance levels. Yet the extent to which they carried over into the structured interactions designed by Paulson, McDonald & Whittemore (1973) represents an essential unanswered question for further analysis. A testing procedure which allows for both a well controlled measure of modeling performance by S's with a systematic observation of teacher-child interaction patterns would clarify many of the gray areas in the teacher-commenter intervention issue. A procedure patterned after that employed by Ball & Bogatz (1973) for their Electric Company evaluation might be helpful in this regard. In place of an analysis of the instructor's methods for teaching reading one might substitute systematic observations of such areas as mutual reinforcement patterns between classmates, teacher nurturance and the frequency of child-teacher contacts.
BIBLIOGRAPHY


Hicks, D.J., Effects of co-observers' sanctions and adult presence on imitative aggression. *Child Development*, 1968, 39, 303-309.
