This study investigated the effects of training on reciprocity judgments in 43 white, middle class children, ages 38 to 68 months. Hypotheses were (1) children trained to make mature responses to interpersonal conflicts would make more mature judgments in illustrated story conflict situations than those receiving no training, (2) children exposed to discussion and role playing during training would make more mature reciprocity judgments on a story posttest than those exposed to operant discrimination training, and (3) there would be no sex differences on reciprocity judgments. Treatment A children acted out, discussed, and suggested solutions to a conflict situation, and then re-enacted the situation using the trainer's mature response to resolve the conflict. Treatment B consisted of operant discrimination training without discussion. Treatment C children listened to a story and answered questions. Children verbally solved 10 illustrated conflict situations on a posttest. Hypotheses (1) and (3) were supported; however, operant discrimination training was more effective than role playing and discussion training. (DO)
THE EFFECTS OF TRAINING ON RECIPROCITY JUDGMENTS IN PREKINDERGARTEN CHILDREN

Barbara Vance           Linda Cropper
Brigham Young University Brigham Young University

Larry C. Jensen
Brigham Young University

Research on moral development during the past 15 years has focused primarily on the cognitive rather than the affective aspect. This cognitive aspect is called moral judgment or moral reasoning. Studies in moral judgment attempt to infer the way the child thinks rather than the way he might act in a given situation. The implicit assumptions of such studies are that moral development is sequential and that cognitive development (judgment) is antecedent to affective development (action). "The man who understands justice is more likely to practice it (Kohlberg, 1970, p. 115)."

Many moral issues face children in a free-enterprise society such as sharing one's possessions, telling the truth, resisting the temptation to take things belonging to others, and obeying rules of adult authorities (parents, teachers, policemen, clergy, and so on).

This research was supported by the U.S. Office of Education, Bureau of Research, Project No. 0-0264. The authors wish to express appreciation to the head teachers of the prekindergarten groups at Brigham Young University involved in this study, the graduate student in Psychology who developed the story posttest, and the six graduate students in Child Development and Family Relationships who served as trainers. The senior author's address is Instructional Research and Development, Brigham Young University, Provo, Utah 84601.

presented in summary form at the biennial meeting of the Society for Research in Child Development, held in Philadelphia, Pennsylvania, March 29-April 1, 1973
Selman and Kohlberg (1970) claim that all such moral issues have four basic moral elements in common: intention, internality, relativity, and reciprocity. Using an analogy, the moral elements (such as reciprocity and relativity) are seen as the steel superstructure of a building, while the moral issues (such as telling the truth or sharing possessions) are seen as the bricks, doors, and windows. The building can be completed only when the superstructure is properly developed. Intention is an element of this superstructure dealing with movement from judging acts in terms of their physical consequences to judging them in terms of the motives of the doer. Internality is an element concerned with judgment of the "goodness" or "badness" of an act independent of, rather than on the basis of, sanctions such as verbal approval or physical punishment. Relativity is moral judgment that moves from an absolutistic, rigid, single-perspective view of an act to flexibility based on consideration of other points of view. Reciprocity, perhaps the keystone element in the development of the moral superstructure, is based on movement from exact distributive and retributive equality or justice or fairness (such as "an eye for an eye and a tooth for a tooth") to a concept of fairness and justice based on the other person's needs and point of view (such as the Golden Rule). Forgiveness would be characteristic of mature reciprocity.

Piaget ([1932] 1965) views movement from one stage of moral judgment to another as a slowly maturing result of the child's social
interaction with his peers rather than direct teaching. He suggests that cooperation—the "dignity of the individual and respect for general opinion"—is a vital part of moral judgment learning. Maier (1965) indicates that the transition is a result of increased social contact combined with more accurate imitation of models in the environment. Berkowitz (1964) states that imitation and modeling are the primary methods through which children gain values and moral judgments.

Several studies have examined the possibility of accelerating movement from one stage of moral judgment to another. Some of these studies have dealt specifically with stages suggested by Piaget (e.g., Bandura and McDonald, 1963; Crowley, 1968; Durkin 1959a, 1959b, 1961; Larm, 1969). Others have examined stages suggested by Kohlberg (e.g., Turiel, 1966).

The moral element of reciprocity, or the concept of fairness and justice, has received little experimental attention in the literature. Reciprocity is defined as a fairness and equality of interaction wherein the participants act upon the basis of the Golden Rule ("Do unto others as you would have others do unto you"). Selman and Kohlberg (1970) see the child going through stages in his understanding of reciprocity. The child first has no conception of equality, sharing, or taking turns. Next, he shares in obedience to instructions of an authority figure. Thirdly, he perceives reciprocity on the basis of an "eye for an eye." Finally, he pro-
gresses to the stage where he is considerate of others as he would like them to be considerate of him.

Piaget ([1932] 1965) suggests that young children tend to ignore concepts of reciprocity and seek an authority figure in conflict situations. Older children respond in ways that reflect the acceptance of reciprocity. In a series of studies with children divided among grades 2, 5, 8, and 11 and using the definition of reciprocity as a return of identical behavior, Durkin (1959a, 1959b, 1961) found that reciprocity judgments change with age but not as predicted by Piaget. She found that older and younger children tend to seek justice through an authority figure, rather than the older children consistently turning to mature reasoning and mature behavior. Children also had a tendency to favor returning aggressive acts rather than reasoning out a conflict. The aggression returned was to be no more severe than the aggression received.

In an experimental study of the effects of certain modeling conditions on sharing behavior in fourth and fifth graders, Harris (1970) found that children who observed a model share with them tended to share with the model. Those who observed a model share with charity shared in turn with charity. The amount shared, however, was not reciprocal. Those children who observed no sharing showed little or no evidence of sharing when they had the opportunity. Though this study investigates the element of reciprocity in children, it studies reciprocity behavior rather than thinking. However, it
supports the claim that modeling affects reciprocity behavior.

The present investigation studied the effects of training on reciprocity judgments in prekindergarten children. There is little experimental research related to this important element of moral judgment, particularly in prekindergarten children. Prekindergarten children appear to be at an early stage of moral judgment development, therefore making changes in judgment relatively easy to detect.

Studies indicating that changes in moral judgment occurred after short training periods (Bandura & McDonald, 1963; Crowley, 1968; Larm, 1969; Turiel, 1966) provided the rationale for the hypothesis in this study that prekindergarten children trained to make mature responses to interpersonal conflict situations would make more mature judgments in illustrated story conflict situations than those receiving no training.

Because the evidence indicates that modeling affects moral judgment responses in children (Bandura and McDonald, 1963; Harris, 1970), and discussion and role playing tend to be more effective in changing moral judgments than operant discrimination training (Crowley, 1968; Larm, 1969; Turiel, 1966), the present study predicted that prekindergarten children exposed to discussion and role playing during training would make more mature reciprocity judgments on a story posttest than those exposed to operant discrimination training.
Though inconsistent, the moral judgment literature reveals few significant sex differences. Therefore, this investigation also hypothesized there would be no sex differences on reciprocity judgments of prekindergarten children in a story posttest.

METHOD

Sample

The original sample consisted of 60 children, 31 boys and 29 girls, ranging in age from 38 to 68 months. These children were enrolled in the Child Development Laboratories at Brigham Young University during the second session of summer school, 1971. This was the total enrollment in the four laboratory groups with the exception of two children who were non-verbal. The subjects were from white, middle-class homes. The 60 children in the sample were randomly assigned to three training groups by sex and laboratory group. Twenty-one Ss did not complete the study resulting in a final sample of 43. The mean age for the children who completed the study was 56.2 months. The mean age of the children in group 1 (45.8 months) was considerably less than that of the other three laboratory groups. The mean age of laboratory group 4 (54.7 months) was nine months older than the mean of laboratory group 1. The mean ages of groups 2 and 3 were 62.4 and 61.7 respectively. Age, therefore, was used as a covariate in the statistical analysis of the data in this study.
Treatments

The treatment consisted of three fifteen-minute training sessions on three consecutive days for each of the experimental and control groups and a fifteen-minute posttest for each individual child four to six days following the last training session. Four trainers (two male and two female graduate students) were instructed in each treatment and randomly assigned to the experimental or control groups so that each trainer conducted at least one training session for each kind of treatment in the study. Each training group received a different trainer for each training session, thus randomizing the effect that any individual trainer had on the children. The trainers were not informed of the study hypotheses until the study was completed.

Treatment A consisted of a series of discussions wherein the children acted out a conflict situation, discussed their feelings about the situation, suggested possible solutions to the conflict, and then re-enacted the situation using a mature response supplied by the trainer to resolve the conflict. Each child in this treatment had an opportunity to act out a role and verbalize the mature response of discussing the problem with the offending child to resolve the conflict. A different conflict situation was used for each of the three training sessions. The first was a child taking a book away from another child; the second, a child pushing another child off the slide;
and the third, a child verbally threatening to take away a ball from another child.

Treatment B consisted of operant discrimination training without discussion. Two wood figures, looking vaguely like human figures, were used during these training sessions. The figures had the name of a boy and a girl during the first training session, the names of two boys during the second training session, and the names of two girls during the third training session. The trainer explained during each training session that these children sometimes had problems getting along together. The trainer then presented to the children in each group in this treatment a conflict situation and a verbal model of a mature way to resolve the conflict. In subsequent conflict situations, each S was asked to indicate what he thought should be done to solve the problem. If he answered with a solution that included trying to resolve the problem with the other children without physical or verbal aggression, he received a gold star by his name on a special chart. If he suggested trying to enlist the aid of an adult to help solve the problem, he received a silver star. If the child did not give an answer related to either of these, the trainer went to the next child, ignoring the inappropriate response.

Children in Treatment C, the control group, listened to a popular children's story during each of the three training sessions. These stories had no relationship to the reciprocity training in
Treatment A and B. At the conclusion of each story, the trainer asked selected recall questions.

Posttest

The posttest consisted of ten illustrated conflict situations. Five stories involved a child being verbally or physically attacked by another child and five stories involved a child having his property taken away or destroyed by another child. The stories include conflict situations familiar to the average prekindergarten child. An example of physical attack would be: "Ann/Roger is playing in the sand when another girl/boy dumps sand on Ann's/Roger's head. What do you think Ann/Roger will do about it? Why?"

An example of the property destruction would be: "A girl/boy tears the pictures out of Margaret's/Bobby's story book. What do you think Margaret/Bobby will do about it? Why?"

Parents of prekindergarten children not in the study rated a list of possible ways of solving the conflict situation in each of the ten stories from the least mature to the most mature. An item analysis was made of these parent ratings to determine the rating scale to be used on child responses in the story posttest. Subject responses to each story were coded according to the results of the parent rating scale, with 12 coded the most mature response and 1 coded the least mature response.

During the posttesting a female experimenter selected the children at random from their laboratory groups and brought them
to the experimental room. This room is 10' by 15 1/2' with two one-way mirrors. The examiner, an undergraduate student in Child Development, was seated at a small table. She was introduced to the child by the experimenter. The child sat down at the table with the examiner. The experimenter left, indicating that she would return when the child was through. An observer, also an undergraduate student in Child Development, observed through the one-way mirror and recorded the nonverbal responses of the child during the fifteen-minute testing session. To avoid examiner and observer bias, neither the examiner nor the observers were informed of the hypotheses of the study.

Each child was presented the ten illustrated conflict situations one by one in a random order by the examiner. After the presentation of each situation, the child was asked what the "wronged" child would do and why. Answers were recorded verbatim by the examiner on a response form. Female names were used in each conflict situation when S was female; the names of boys were used when S was male. At the end of the posttest, the child was thanked for his help and given a balloon.

Responses of the children to the story posttest were rated independently from least mature to most mature by two female graduate students in Child Development and Early Childhood Education using the 12-point coding scale. For instance, the response "He/she should wipe it off," to the story about having sand dumped on the
child was coded 11, meaning that the child ignored the behavior of the offending child. A response calling for help from a teacher in the conflict situation was coded 10, indicating the child would ask an adult to intervene. These graduate students were unaware of the hypotheses of the study and the training sessions which preceded the posttesting. Interrater reliability was computed using the following formula:

\[
\text{Percent of agreement} = \frac{A}{A + D}
\]

\[A = \text{number of agreements}\]
\[D = \text{number of disagreements}\]

Interrater reliability was .98.

Statistical Analysis

The Modified Abbreviated Doolittle (MAD) analysis of variance and covariance for unequal cells and unequal numbers of observations was performed by treatment, group, story, and sex to test the hypotheses of this study.

RESULTS

Table 1 is an analysis of variance summary by treatment, group, and sex of the scores of the 43 children in the final sample on the story posttest.

---

Insert Table 1 here
---
There was a significant treatment effect and also treatment by group interaction. The mean scores are presented in Table 2.

---

Insert Table 2 here
---

Groups 1, 2, 3, and 4 performed relatively alike on the posttest for the control groups, but groups 2 and 3 in Treatment B scored more than Groups 1 and 4. Apparently Treatment B was more effective for groups 1 and 4. Treatment B had a mean score of 9.4, but Treatment A had a mean of 7.3, and the control treatment a mean of 6.9, indicating that the main effect for treatment was due to Treatment B.

The analysis of sex differences revealed no differences between the sexes on the story posttest.

**DISCUSSION**

This study supported the hypothesis that a short training period related to appropriate behavior in child conflict situations can increase the level of cognitive development inferred from moral reasoning in prekindergarten children when responding to illustrated story conflict situations. However, it did not support the hypothesis that role playing and discussion training would be more effective than operant discrimination training; instead operant discrimination training seemed to account for the response differences in this study. This difference might be explained by a careful analysis of the task required of each child on the posttest and the relationships of the two types of training.
to this posttest task.

The posttest task required the child to provide a verbal solution to each of several conflict situations not previously presented during training. The solution was a particular rule or principle such as, "Try to resolve the difference with the other child," or "Walk away from or ignore the other child" (stated in general adult terms). For instance, if a child gives this solution to the story conflict of one child having his tricycle taken from him by another, "He'll tell Don he's angry and to wait his turn," this is a specific example of the rule, "Try to resolve the difference with the other child." If he says, "He'll walk away and get another trike," he is giving an instance of the rule, "Walk away from or ignore the other child." If rule using is the task to be performed, instruction (training) should be designed to help the child use the appropriate rule or rules with appropriate instances. The role playing and discussion training (Treatment A) in this study allowed the children to role play one specific rule (that judged most mature) related to three specific situations only. In contrast, the operant discrimination training (Treatment B) allowed the children to respond to 21 different story conflict situations using the same rule. If the child is presented similar but not identical situations in a posttest, he is being required to transfer a rule to a new but similar situation. It is reasonable to expect children to make this transfer more effectively if they have experienced more instances related to that task during
training. This might explain the difference between the role-playing
discussion and the operant discrimination training groups in this
study.

In this study, coding was done using response sheets on which
collectors and observers had written the exact verbal responses of each
child to the posttest stories. The investigators who heard the children
respond during testing agreed that hearing the child's voice and reading
the printed word can produce very different perceptions of the same
thing. For instance, in a story about a child being blocked by another
while riding his tricycle, the subject is asked, "What will (the "wronged"
child) do?" and replies, "He'll run over him." On the written response
sheet such a response would be classified in Category 5 (Revenge himself/herself against the other child). However, the tone of voice used
by the child sounded very matter of fact and certainly not filled with
revenge. The response to a story of a child being pushed down a slide
by another might be, "Fall down and cut his head open." In this case,
the written response was also coded 5, but the sound of the child's
voice leads one to wonder who S was talking about--the child being pushed
or the child doing the pushing. Even the written response could lead to the
same query. It thus becomes clear that voice inflection may be an
important element in determining a child's level of judgment, and
that verbal responses from prekindergarten children are often very
difficult to interpret because the words of such young children tend to be
far richer (general) in meaning than those of adults.

During the posttest each S was asked not only what the wronged child would do in each case, but also the question "Why?" Most replies to the why question were, "Because." Therefore, the "Why" answers could not be used in the analysis of data in this study. Such a question is probably a cause-and-effect concept that is beyond the experience level of the children in this study. If they do understand the concept they don't have the necessary language skill to appropriately express their understanding. This problem of interpreting what the prekindergarten child means when he verbalizes may be one reason why so few studies of moral reasoning have been done with prekindergarten children. Perhaps this is one reason why Selman and Kohlberg (1970) have suggested a Stage 0 or amoral orientation during these earliest years of life. Valid evaluation of a child's moral reasoning ability is highly dependent on his verbal ability.

Gagné (1970) and Merrill (1971) have suggested a general sequence of cumulative learning that holds promise for the study of moral reasoning in young children. They, like Kohlberg and Piaget, have postulated invariant sequences. However, they have defined each stage and the conditions of each in observable terms making it possible for the researcher to more clearly define his dependent and independent variables. A refinement of this sequence in the cognitive domain is suggested by Merrill and Boutwell (1972).
Inasmuch as investigations of moral reasoning deal with the cognitive domain, it might be advisable for those interested in such research to carefully examine the general sequences of cumulative learning. Merrill and Boutwell (1972) suggest that all cognitive behavior can be classified in four categories: discriminated recall, classification, rule using, and higher rule using. **Discriminated recall** means the individual, when presented with any given object, symbol, or event he has experienced before can provide the appropriate symbol for it or indicate the associated object, symbol, or event. For example, the child can point to his dog and say, "Spot" (object-to-symbol) or pick out the letters of the alphabet as they are named by his mother (symbol-to-object). This is a memory process dealing with one-to-one relationships. However, complex cognitive behavior, consisting of the next three categories in the cumulative learning sequence, is dependent on the child's ability to transfer what he has learned to new instances not experienced before. **Classification** means the child can indicate the class membership of something he has not previously encountered because of common attributes related to past experience. For example, a child looks at a dog he has never seen before and says, "Dog," because the dog has characteristics similar to other dogs he has seen before. After hearing a pair of incidents he has never heard before about children, he is asked to pick out the incident about the child who was "naughty" or the child who did the "right" thing. If
he has had adequate past experience with incidents related to the concepts "naughty" and "good" the child can transfer his learning to a new instance. Rule using occurs when the child can demonstrate he understands the relationship between two or more concepts when presented a situation not encountered before. In this study children were presented illustrated stories during posttesting showing conflict situations between children. These stories had not been encountered in previous training. If children could transfer from their previous training or past experience the general rule, "Try to resolve the difference with the other child," to these new situations they were demonstrating the desired rule-using behavior. But rule using is not possible until the child can first demonstrate understanding of the concepts that comprise the rule. With appropriate rule-using behavior the child can then move to higher rule using where two or more rules are used to solve a previously unencountered problem. For example, if a child has demonstrated ability to use alternative rules in a conflict situation, he might come up with a strategy that says, "First, try to resolve the difference with the other child. If this doesn't work, then ask an adult to intervene." Of course, children would not say the rules in these words, but their cognitive behavior would indicate whether or not they were operating in terms of this sequence of rules to solve a problem. Gagne (1968) suggests such a sequence for learning conservation tasks, though he labels the categories differently.
If moral reasoning is a function of this four-step sequence of cumulative learning, it becomes important to identify the task we wish the child to perform. This task can be classified as one of the four cognitive categories. Once the desired task is identified and classified, instructional activities can be planned to help the child achieve the task. These activities become the independent variables and the child's responses to the task become the dependent variables. Each task category probably requires a different set of instructional activities. That is, learning activities to help children develop classification behaviors probably should be different from those to teach rule-using behavior. Vance (1973) uses this idea in the preparation of learning activities for prekindergarten children.

At the present time investigations of moral reasoning in children are difficult to compare because of a lack of cohesive theory to pull it all together. The "stage" theories may be useful descriptions of cognitive processes, but they are not yet stated in terms specific enough for two researchers studying the same moral element, such as reciprocity, to provide comparable results. Stage theories need to be stated in terms of specific observable independent and dependent variables. Then experimental research with children of different ages, sex, and socioeconomic and ethnic backgrounds can begin to build a tested theory of the development of moral judgment or reasoning.
REFERENCES


Durkin, D. Children's concepts of justice: A comparison with the Piaget data. *Child Development, 1959, 30, 59-67. (a)

Durkin, D. Children's concept of justice: A further comparison with the Piaget data. *Journal of Educational Research, 1959, 52 (7), 253-257. (b)


Merrill, M. D., & Boutwell, R. C. Instructional development: Methodology and research. Working Paper No. 33, Instructional Research and Development, Division of Instructional Services, Brigham Young University, 1972.


**TABLE I**

Modified Abbreviated Doolittle Analysis of Scores on Story Posttest

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Mean Score</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatments</td>
<td>2</td>
<td>132.8000</td>
<td>17.460*</td>
</tr>
<tr>
<td>Groups</td>
<td>3</td>
<td>8.2937</td>
<td>1.182</td>
</tr>
<tr>
<td>Group x Treatment</td>
<td>6</td>
<td>50.2140</td>
<td>7.159*</td>
</tr>
<tr>
<td>Sex</td>
<td>1</td>
<td>7.5307</td>
<td>.692</td>
</tr>
<tr>
<td>Stories</td>
<td>4</td>
<td>13.2950</td>
<td>1.900</td>
</tr>
<tr>
<td>Treatment x Sex</td>
<td>2</td>
<td>16.8730</td>
<td>2.406</td>
</tr>
<tr>
<td>Property x Physical Stories</td>
<td>1</td>
<td>15.1200</td>
<td>1.130</td>
</tr>
<tr>
<td>Treatment x Property x Physical</td>
<td>2</td>
<td>2.4200</td>
<td>.318</td>
</tr>
<tr>
<td>Property x Physical x Sex</td>
<td>1</td>
<td>.7051</td>
<td>.065</td>
</tr>
<tr>
<td>Treatment x Stories</td>
<td>16</td>
<td>7.6040</td>
<td>1.084</td>
</tr>
<tr>
<td>Sex x Stories</td>
<td>8</td>
<td>10.8890</td>
<td>1.550</td>
</tr>
<tr>
<td>Error</td>
<td>357</td>
<td>7.0141</td>
<td></td>
</tr>
</tbody>
</table>

* Significant at .005
## TABLE 2

Mean Scores from Parent Rated Scale on Story Posttest by Treatment, Group and Sex

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Group A</th>
<th>Group B</th>
<th>Group C</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>P</td>
<td>Tot.</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>7.5</td>
<td>7.4</td>
<td>7.4</td>
<td>7.4</td>
</tr>
<tr>
<td>2</td>
<td>6.9</td>
<td>6.1</td>
<td>6.6</td>
<td>6.8</td>
</tr>
<tr>
<td>3</td>
<td>7.4</td>
<td>7.4</td>
<td>7.2</td>
<td>7.3</td>
</tr>
<tr>
<td>4</td>
<td>7.0</td>
<td>7.0</td>
<td>6.8</td>
<td>7.0</td>
</tr>
<tr>
<td>Totals</td>
<td>7.7</td>
<td>6.9</td>
<td>7.3</td>
<td>7.8</td>
</tr>
</tbody>
</table>

By Treatment, Group and Sex

Mean Scores from Parent Rated Scale on Story Posttest

TABLE 2