This study of the cognitive-developmental approach to primary grade value education emphasizes the involvement of the reasoning process in solving moral dilemmas. This approach is in contrast to those which focus on clarifying values and developing an awareness of other cultures' values. Educational intervention in the cognitive-developmental approach is based on the principles of conflict and near-matching, and presented through the media of sound filmstrips, debate, and open discussion. Both "expert teachers" (those trained in cognitive developmental theory and experienced with primary grade children), and lay teachers conducted the study's biweekly sessions. Units presented covered the moral topics of Truth Telling, Sharing and Taking Turns, Promise Keeping, Property Rights, and Rules. Results show that, while change to a higher stage in reasoning may be greater with a teacher who is familiar with the stage theory and is interested in the program than with one who has just read the training manual, the two groups did not differ significantly. (CS)
An Evaluation of a Cognitive-Developmental Values Curriculum for Primary Grade Children

Marcus Lieberman and Robert L. Selman
Harvard University

Chicago, Illinois; April, 1974
An Evaluation of a Cognitive-Developmental Values
Curriculum for Primary Grade Children

Marcus Lieberman
Robert L. Selman

A succinct review of cognitive-developmental theory and research in the moral domain (Kohlberg & Turiel, 1971) indicates the following:
1) Children may move through stages of moral reasoning at varying speeds and may be found exhibiting characteristics of an adjacent stage as well as a particular stage. 2) Higher stages cannot be directly taught to children at much lower stages, i.e., children exposed to reasoning more than one stage above their own, merely translate them into ideas at their own level. 3) Changes in stages of thinking are not immediate, i.e., the term stage should not be taken to imply that the changes in thinking that a child goes through are instantaneous, but that stages are a series of qualitatively more adequate ways of looking at moral problems. 4) Movement to a higher stage requires the experience of conflict or difficulty in the child's attempt to apply his current level of thought to moral problems. 5) Movement to a higher stage also requires an exposure to the next level of thought. It involves a sense of active participation in the social problem-solving process and the opportunity to take the role of others, to see their point of view when it differs from his own.

Kohlberg has argued that research on stages of moral development
implies a new approach to value-education which is nonindoctrinative.

"Stimulation of development (of reasoning) as an aim avoids the critical objection to value education that teachers have no right to indoctrinate children with their particular values, which may be different than those of the child and his family. The existence of moral stages indicates that there is progression to greater moral awareness which teachers and researchers can define independent of their particular culture and religious affiliation." (Kohlberg and Mayer, 1972).

If it is true as research indicates, that moral thought passes through this natural sequence of stages, a moral education program of any type (filmstrips, readings, classroom discussion) should focus on helping the child reach the next step of development rather than directly teaching him fixed rules and values of the adult world. This was the aim of the curriculum devised for primary grade children by the second author (Selman, 1973). According to cognitive-developmental theory, two basic mechanisms are necessary for the development of the next level of moral reasoning. First, the child must experience the social situation as a moral dilemma, he must feel some conflict or indecision over what is the right or moral action. Second, exposure to moral reasoning slightly more adequate than his own (one stage above) may facilitate development to the next stage.

It has been shown in various studies that higher stages of moral reasoning relate positively to moral behavior, and that educational intervention using the developmental principles of conflict and near matching (presenting reasoning one stage above) can lead to significant upward stage movement in a variety of settings, e.g. junior high and high schools and correctional institutions (Kohlberg and Turiel, 1971).

However this approach has never been formally tested in the primary grades. While, in theory, the developmental principles of change should work regardless of the age or stage of the population under study, different age groups present different problems to both research and curriculum development. There have been two unsolved issues which, until now, have left primary grade value-education relatively barren of educational curriculum based on developmental theory.

First, the dilemma approach used by Kohlberg and his associates to instill conflict in junior high and high school students needed modification;
dilemmas relevant for older children were neither interesting nor relevant to primary graders. More appropriate dilemmas needed to be developed. Second, the oral or verbal presentation of hypothetical dilemmas was not appropriate for younger children who had difficulty grasping the details and social facts crucial to the construction and understanding of the dilemma. A more explicit, visual presentation of the dilemmas was needed to hold the attention of young children.

Using the method of sound filmstrips, Selman (1972) constructed a set of dilemmas suitable for stimulating development in primary grade children in the following ways:

1. They present dramatic stories which are enjoyable and involving to watch for children of this age.
2. They present a conflict between two or more values understood by children of this age.
3. They are open - children of this age disagree about what is right and have difficulty making up their minds.
4. Without giving "right answers", they present reasons above the level of most of the children in the class which may help stimulate the child to make his own level of reasoning more adequate.

Procedure

Subjects were sixty-eight second graders, half from an integrated blue collar and half from a middle-class school district in the Cambridge, Massachusetts Public School System. Subjects were divided equally by sex.

Six classrooms participated in the experiment. While entire classes were exposed to the program, random samples of children within each class were selected to be interviewed as subjects.

Three schools were represented to increase the generalizability of the results. In School 1, one class was randomly assigned to each of three treatments: expert-led, informed-lay teacher led, and control. In School 2, one class each was assigned to either an expert-led or an informed-lay led group. To estimate and control for carry-over effects form treatment classroom children who communicate casually with control

1. The classroom teachers participated in the expert-led groups but were not required to prepare for the sessions.
group children a second control group was drawn from a third school geographically but not demographically separate from the second.

The design is summarized as follows:

<table>
<thead>
<tr>
<th>Expert-led group</th>
<th>Informed-lay led group</th>
<th>Control group</th>
</tr>
</thead>
<tbody>
<tr>
<td>School 1</td>
<td>School 1</td>
<td>School 1</td>
</tr>
<tr>
<td>6 female</td>
<td>6 female</td>
<td>5 female</td>
</tr>
<tr>
<td>6 male</td>
<td>6 male</td>
<td>5 male</td>
</tr>
<tr>
<td>School 2</td>
<td>School 2</td>
<td>School 3</td>
</tr>
<tr>
<td>6 female</td>
<td>6 female</td>
<td>5 female</td>
</tr>
<tr>
<td>6 male</td>
<td>6 male</td>
<td>5 male</td>
</tr>
</tbody>
</table>

Interviews were conducted three times: once in the fall before the instructional program, once in the winter immediately after the program, and at the end of the school year (to determine long-term effects).

Procedure

The filmstrip presentation to the group was only a small part of total instruction. The remaining portions involved asking the children what should be done to resolve the dilemma, reasons for each choice, and debate about whether some reasons are better than others. Children discussed the dilemmas in small groups, and teachers acted as guides, keeping the discussion on the topic.

The experimental groups began after the pretesting of all subjects. They met twice a week, each session running for approximately 30 to 40 minutes. Each week both dilemmas from each of five units were presented to the class, one at each session. The units were the following: Truth Telling, Sharing and Taking Turns, Promise Keeping, Property Rights, and Rules.

The "expert teachers" had training in cognitive-developmental theory as well as experience with primary grade children. The lay teachers simply
used the teacher's training guide. The control group teachers were told to conduct the classrooms in their usual manner.

**Instrumentation**

Three types of moral dilemmas were used to evaluate the intervention program: one based directly on situations in the filmstrip series, one analogous to those of the filmstrip situations but not identical with them, and one, standard dilemmas used in previous research on moral development.

An estimate of the test-retest reliability of the moral dilemma interviews was found by correlating control groups' pre- and post-testings with a five-week interval between administrations. Filmstrip dilemmas correlated .67 and standard dilemmas correlated .62. These results were judged adequate given the small number of subjects (20) and the narrow range of possible scores.

To determine inter judge reliability, 18 protocols were picked from each testing session, three protocols randomly picked from each class (six classes). Correlations between two trained scorers on moral judgment scores were as follows: Pretest, $r = .844$; Post-test, $r = .900$; Post-post-test, $r = .816$.

Since the filmstrip dilemmas were not the standard ones used by Kohlberg in his research, it was necessary to determine the degree to which both types of dilemmas reflected the stage of an individual subject. One estimate of this relation is a correlation between the filmstrip dilemmas and the standard dilemmas. For the control group alone the correlation was .500 on the pre-test and .575 on the post-test. For all groups the correlation was .473 on the pre-test and .593 on the post-test.

The two hypothesis entertained in this study were:

1. Intervention in the form of a filmstrip dilemma presentation and discussion will cause significant change in moral reasoning level compared with a comparison group which does not receive the treatment.
2. Although the amount of change may be greater with a teacher familiar with the stage theory (expert-led) using the material than when used by a lay teacher who has only read the manual and
instructions, it was hypothesized that both the "expert-led" and the "teacher-led" groups would not differ significantly from each other but their mean amount of change would differ from that of the control group.

Results
Stage score analyses can be either qualitative or quantitative in nature. For this study, the variable employed was the Moral Reasoning score, (MMR) where 100 represents pure stage one reasoning, 200 represents pure stage two reasoning, etc.

To arrive at MMR, each subject was given a qualitative score on each dilemma which represented a predominant and a minor level of reasoning on that dilemma (e.g. a score of 1(2) indicates that stage 1 reasoning was predominant but some stage 2 reasoning was evident). Qualitative scores were transformed into quantitative scores (e.g. in the above case 1(2) transforms to 133, a score of 2(1) would transform to 166) and the MMR was the average of scores across dilemmas for each S.

MMR scores for social class and treatment groups on the three administrations of the interview are presented in Table 1.

Table 1
Mean Moral Reasoning Score

<table>
<thead>
<tr>
<th></th>
<th>Controls</th>
<th>Expert-Led</th>
<th>Lay Teacher-Led</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mixed Middle</td>
<td>103.6</td>
<td>109.1</td>
<td>116.7</td>
</tr>
<tr>
<td></td>
<td>110.7</td>
<td>125.0</td>
<td>125.0</td>
</tr>
<tr>
<td>and Working</td>
<td>117.9</td>
<td>163.6</td>
<td>147.2</td>
</tr>
<tr>
<td>Class</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle Class</td>
<td>116.7</td>
<td>112.5</td>
<td>122.9</td>
</tr>
<tr>
<td></td>
<td>122.2</td>
<td>137.5</td>
<td>135.4</td>
</tr>
<tr>
<td></td>
<td>130.6</td>
<td>156.2</td>
<td>189.6</td>
</tr>
<tr>
<td></td>
<td>Pretest</td>
<td>Post test</td>
<td>Post-post test</td>
</tr>
<tr>
<td></td>
<td>Pretest</td>
<td>Post test</td>
<td>Post-post test</td>
</tr>
</tbody>
</table>
Since post test and post-post test scores were highly correlated, with pretest scores ($r = .67$ and $.46$ respectively), and regression lines were not significantly non-parallel, an analysis of covariance was performed on the data.

Orthogonal contrasts among treatment groups showed significant differences between the control group and the mean of the two intervention groups on both the post and the post-post testings ($p = .012$ and $p < .0001$ respectively) and no difference between the expert-led and lay teacher-led groups ($p = .155$ and $p = .494$). One interaction, that of treatment by social class, was significant on the adjusted post-post testing ($p = .024$) mostly due to a "lay teacher" continuing to use the classroom discussion method during the remainder of the school year and producing the highest score of any group.

Analyses of the change in scores from pretest to post test and pretest to post-post tests adjusted for pretest scores produced identical results.

The most interesting result was the differences in level of reasoning between control and both experimental groups at the time of the post-post testing, i.e., at the end of the school year. Taken across both types of treatment, expert and lay-teacher, and across both social classes the mean amount of change was about one half a stage of reasoning for the experimental groups over the control group.

In effect these results reflect the fact that the four teachers whose classes defined the experimental group continued to use the methods of small group discussion to resolve interpersonal and moral conflicts which arose in the classroom throughout the school year.

It is obvious that in our experimental design it is impossible to sort out "teacher effects" from "intervention techniques." In other words, we cannot tell how much of the difference in level of reasoning at the post-post test is attributable to the competency of the teachers involved, how much is a function of the curriculum technique, and how much a function of some interaction of the two. For this reason alone, this can only be regarded as a pilot study. It is very possible that such post-post test gains would not have been obtained without a certain level of teacher
competence or cooperation.¹

In fact, the biggest gains occurred, not in an expert-led class, but in the class of the lay teacher who showed the greatest interest in the program, its techniques, and the underlying theory. The treatment X social class effect on the post-post test is largely a function of the change in this one classroom.

A major limitation of this study was the use of only one evaluation procedure, the assessment of social (moral) cognitions. There was no evaluation of whether the children who participated in the intervention gained in other abilities. For example, were children better able to communicate with one another after the group discussion experience, were they more aware of what moral conflicts are, or were they able to articulate their own reasons to others both in hypothetical and real-life situations? Informal reports of the teachers involved in both types of treatment conditions indicated this to be the case. In fact, in many ways, development was more "social" than "moral". By this, it is meant that at the lower primary grades, moral reasoning is not clearly distinct from social reasoning. The acquisition of the conception of intentionality, to take one example, is pertinent to social relations in general, not to just moral judgments. For that reason, we believe that the value intervention techniques at the early stages make children more social than moral, i.e., more socially aware of others' ideas and better able to integrate others' thinking and valuing with their own.

Given the amount of stage change which occurred over the October-May period, it seems worthwhile to design and carry through research which takes teacher factors into consideration and which looks systematically at other social behavioral objectives which might be influenced by this type of program.

¹ Teachers and their classrooms were selected for the study on the basis of their interest and an evaluation of their competence by the Assistant Superintendent in charge of Social Studies.
References


