The Effect of Different Teaching Strategies on the Moral Development of Physical Education Students

Sandra L. Gibbons  
University of Victoria

Vicki Ebbeck  
Oregon State University

This study examined the effectiveness of social learning (SL) or structural-developmental (SD) teaching strategies on the moral development of elementary-age students. Participants were 204 physical education students in Grades 4, 5, and 6; three classrooms in each grade were randomly assigned to control, SL, or SD groups. Self-report measures assessed moral judgment, reason, and intention; teachers rated prosocial behavior. By mid- and postintervention class-level analyses, the SL and SD groups scored significantly higher than the control on moral judgment and/or intention; by postintervention, the SD group was significantly higher on moral reason. Mid- and postintervention student-level analyses showed that the SL and SD groups scored significantly higher on moral judgment, intention, and behavior; the SD group was significantly higher on moral reason. These results provide support for the effectiveness of both social learning and structural-developmental teaching strategies on the moral development of children in physical education.

Concepts associated with moral development are consistently identified as learning outcomes in school physical education programs. Physical education teachers often use such terms as sportspersonship and fair play to describe the type of behavior and attitude expected of participants in sport. In addition, researchers identify the potential for physical education to develop these particular learning outcomes. Shields and Bredemeier (1995) describe physical education as “probably the most significant physical activity context for developing moral character” (p.199). Two major challenges, however, confront attempts to develop the potential of physical education to achieve this outcome. The first challenge involves the attempts of researchers and practitioners alike to identify the types of teaching strategies that may best promote moral development in the physical education setting. Secondly, both groups struggle with the challenge of conceptualizing and measuring moral development. It was our intent to address both of the preceding challenges in the design of the present study. Thus, the general purpose of this study was to compare the effectiveness of teaching strategies from two different theoretical perspectives on several indicators of moral development.

Sandra L. Gibbons is with the School of Physical Education at the University of Victoria, P.O. Box 3015, Victoria, B.C., Canada V8W 3P1. Vicki Ebbeck is with the Department of Exercise and Sport Science at Oregon State University, Langton Hall 214, Corvallis, OR 97331-3302.
To date, moral development teaching strategies utilized in sport and physical education have fallen into either of two theoretical perspectives: social learning or structural developmental. Social learning theorists (Bandura, 1977) define moral development as the demonstration of behavior that is consistent with societal norms. The processes by which individuals develop morally include those that focus primarily on the observation of and reward for socially acceptable behavior. Examples of teaching strategies associated with social learning theory include modeling, verbal instruction, and direct reinforcement. It follows that the major task of a physical education teacher within this context is to show students by example and reward that which is considered socially acceptable behavior during game play.

Structural-developmental theorists (e.g. Haan, 1991; Haan, Aerts, & Cooper, 1985; Kohlberg, 1976) define moral development as the change in reasoning patterns associated with an individual's cognitive structure. Development of these cognitive structures is dependent on active involvement with the environment. Accordingly, the learning processes associated with structural-developmental theory require an environment that will facilitate productive interaction between the self and others. Examples of moral development teaching strategies consistent with structural-developmental theory include dialogue and problem-solving tasks associated with the resolution of both naturally occurring and created moral conflicts and dilemmas.

As mentioned earlier, a frustrating element involved in research on moral development has been the difficulty in defining the concept of morality. This has presented researchers with the additional difficulty of assessing moral development. Work by James Rest has provided a considerable contribution to the preceding challenges. Rest's (1984) notion of a multidimensional model of morality has provided a productive framework from which to examine various aspects of moral development. Rest describes the following four components in his model of morality: (a) the identification of a moral dilemma; (b) making a judgment and reasoning about what ought to be done; (c) an individual's intention to act; and (d) the actual behavior. Rest describes these components as interacting in a multifaceted pattern, with each component exerting influence on and being influenced by the other components in a given situation. This model, then, provides us with a productive framework for examining a variety of different aspects of moral development.

A small group of researchers in the physical activity setting have examined the effectiveness of a variety of social learning and/or structural-developmental teaching strategies on several aspects of moral development. Only one study compared these two different types of teaching strategies. Bredeemeier, Weiss, Shields, and Shewchuk (1986) compared the effectiveness of social learning and structural-developmental teaching strategies on the moral development of children (N = 84), aged 5 to 7, attending a 6-week summer sports program. The authors noted significant within-group pre-to-post gain scores for the experimental groups on the measures of moral intentionality and distributive justice; however, no significant between-group differences for either experimental condition were found in comparison to the control group. Although this study was stated as a comparison of social learning and structural developmental teaching strategies, there were no hypotheses made regarding the efficacy of either experimental teaching strategy relative to the other.

Giebink and McKenzie (1985) examined the effects of three social learning teaching strategies (modeling, praise, and contingent reward) on the sportsmanship
behaviors of a small group of preadolescent boys \((N = 4)\) during physical education and recreational activities. Results indicated that the reward system was the most effective of the three teaching strategies in effecting positive change in sportsmanlike behavior and decreasing incidents of unsportsmanlike behavior. The authors did not include a possible explanation for the effectiveness of the reward system compared to the other two teaching strategies. In addition, any interpretation or generalization of these results must be tempered with the small sample size of this study.

Whereas the preceding authors examined several social learning teaching strategies, Romance, Weiss, and Bockoven (1986) investigated the effectiveness of a selection of structural-developmental teaching strategies utilized in physical education lessons on the moral reasoning of a group of fifth-grade children. Two classes of fifth-grade students, one experimental \((n = 16)\) and one control \((n = 16)\), participated in the same physical activities during the eight-week intervention, but the experimental group also participated in activities associated with the structural-developmental teaching strategies. The authors reported significant pre- to posttest increases in the moral reasoning of children in the experimental group compared to the control group. Of particular note in this study was the deliberate design of teaching strategies to be consistent with concepts of a specific structural-developmental theory of moral development. Romance et al. (1985) grounded the design of their teaching strategies in concepts associated with Haan’s (1977) model of interaction morality. According to this model, moral development occurs as individuals work together to achieve a balance concerning rights and responsibilities in a given situation. In a later work, Haan et al. (1985) suggested that teaching strategies that focus on moral development must provide a student with the opportunity to (a) experience real moral dilemmas, (b) practice resolving moral dilemmas and conflicts, (c) practice dialoguing with peers, and (d) identify moral balance. Romance et al. (1986) utilized these specific concepts in the design of their teaching strategies.

In a more recent study, Gibbons, Ebbeck, and Weiss (1995) conducted a field experiment to investigate the effectiveness of teaching strategies selected from Fair Play for Kids (Commission for Fair Play, 1990) on four indicators of moral development (moral judgment, reason, intention, and prosocial behavior). Fair Play for Kids, a resource manual for teachers, includes a series of teaching strategies that focus on the development of attitudes and behaviors that exemplify the ideals of fair play. The manual includes a blend of both social learning and structural-developmental teaching strategies, with an emphasis on the latter.

The sample in the Gibbons et al. (1995) study was comprised of children in Grades 4, 5, and 6 \((N = 452)\). Six intact classrooms at each grade level were randomly assigned to one of the following experimental groups: (a) control, (b) Fair Play for Kids teaching strategies in physical education only, or (c) Fair Play for Kids strategies in all school subjects. The moral development intervention extended for seven months of the school year (mid-October through mid-May). Results of between-class analyses indicated significant pre- to posttest increases for both experimental groups on moral judgment, reason, and intention scores. Within-class analyses revealed significant pre- to posttest increases for both experimental groups on all four dependent variables compared to the control group.

There are several notable aspects of the Gibbons et al. (1995) study. First, the 7-month intervention was considerably longer than the 6- to 9-week intervention
periods of earlier studies. This longer time period allowed for a more realistic scenario for the implementation of the teaching strategies. It is likely that a teacher would have an ongoing focus on aspects of moral development throughout the school year, rather than a concentrated focus for a shorter period of time. In addition, the Gibbons et al. study examined a much larger sample size than all of the earlier studies, allowing for more confidence in the interpretation of the results. Thirdly, from a practical perspective, the *Fair Play for Kids* manual appears to provide promising teaching strategies that are readily available to teachers and worthy of further examination. Finally, the authors utilized a specific conceptual framework to determine which dependent variables should be included in the study. Rest's (1984) four-component model of moral action mentioned earlier in this paper was used to define and frame the dependent variables. Utilization of this model provided the potential for examining the effectiveness of multiple indicators of moral development within a structured framework. Considering these potential strengths, it was decided to incorporate several aspects of the Gibbons et al. (1995) study into the design of a follow-up investigation. Specifically, the present investigation utilized a similar length of intervention, a selection of teaching strategies from *Fair Play for Kids*, as well as the same conceptual framework and dependent measures.

The present investigation was designed to extend the Gibbons et al. (1995) study in two ways. First, although the Gibbons et al. study demonstrated the effectiveness of the *Fair Play for Kids* program, the design of the study prevented any conclusions regarding which teaching strategies were most essential. The effectiveness of the program might have been due to the social learning strategies, the structural-developmental strategies, or the combination of social learning and structural-developmental strategies. Therefore, the present investigation compared separate social learning and structural-developmental intervention strategies selected from *Fair Play for Kids* to a control group. One previous study conducted with summer sport program participants has compared the effectiveness of social learning and structural-developmental teaching strategies on moral reasoning (Bredemeier, Weiss, Shields, & Shewchuk, 1986). What remains unknown, however, is how these strategies compare in the physical education setting, and whether the strategies differentially impact the various indicators of moral development derived from Rest's (1984) model. The distinct feature of Rest's model is that it incorporates both cognitive and behavioral components that may respond differently to social learning and structural-developmental teaching strategies.

Second, data were collected in the middle of the school year as well as at the beginning and end of the year. Gibbons et al. (1995) reported significant improvement in moral development across the school year, but the rate of change needs to be explored. It may be that all indicators of moral development improve at approximately the same rate following the introduction of the intervention, or the indicators may increase at differing rates. An understanding of which indicators change and when they change may better explain why the intervention is effective and provide additional insights regarding the interrelationships among the components in Rest's (1984) model. In addition, it would be valuable from a practical perspective to provide teachers with information about when they might expect to see change in their students.

The purpose of this study, therefore, was to compare the effect of social learning and structural-developmental teaching strategies on the moral development of
physical education students. Moral judgment, reason, intention, and behavior were assessed at three time periods across the school year. Based on earlier empirical findings (Bredemeier et al., 1986), it was hypothesized that children in either of the two experimental groups would score significantly higher on the four indicators of moral development than those children in a control group. The hypothesized effects were expected by the end of the school year, with no basis for generating any hypotheses about what changes may have occurred at the mid-year data collection.

Method

Participants and Design

The sample comprised 204 male (n = 87) and female (n = 117) physical education students from nine intact classrooms. The participants in this study were enrolled in Grades 4, 5, and 6, which was consistent with the participants in the Gibbons et al. (1995) study. All students attended one of two elementary schools in a suburban area of a large Canadian city.

Three classrooms at each grade level were randomly assigned to one of the following experimental groups: (a) control—no teaching strategies from *Fair Play for Kids* implemented (n = 63); (b) social learning (SL)—social learning teaching strategies implemented by physical education teachers (n = 72); and (c) structural-developmental (SD)—structural-developmental teaching strategies implemented by physical education teachers (n = 69). The experimental protocol extended for 7 months of the school year (mid-October through mid-May), with measures administered prior to, in the middle of, and following the intervention.

The experimental intervention for all groups was incorporated into the prescribed physical education curriculum (British Columbia Ministry of Education [BCME], 1995). The elementary physical education curriculum (K–7) provides for instruction in physical activities from the following five movement categories: alternate-environment activities, dance, games, gymnastics, and individual and dual activities. This curriculum guide includes the recommendation that “no less than 15% of instructional time be spent in any one movement category” (BCME, 1995, p. 6). All groups in the study adhered to this recommendation. The physical activity content at each grade level was also consistent with the BCME (1995) curriculum guidelines, allowing for individual class variation due to scheduling and facility constraints. The knowledge, skills, and attitude learning outcomes within each of the five movement categories are grouped under the following three curriculum organizers: personal and social development, active living, and movement. The teaching strategies in *Fair Play for Kids* focus on learning outcomes within the personal and social development curriculum organizer. Time allocation for physical education for all groups was between 150 and 180 minutes per week. Students in the fourth-grade groups received physical education each day for 30 minutes. The fifth- and sixth-grade classes participated in four 45-minute physical education classes each week.

Description of Teaching Strategies

Teaching strategies for both experimental groups were selected by the researchers from the *Fair Play for Kids* resource manual. This manual was designed
to provide teachers with a variety of interdisciplinary teaching strategies that focused on the ideals of fair play and could be readily incorporated into a variety of subjects. This study incorporated the teaching strategies specifically designated for physical education lessons. The ideals of fair play included (a) respect for the rules, (b) respect for officials and their decisions, (c) respect for the opponent, (d) providing all individuals with an equal opportunity to participate, and (e) maintaining self-control at all times. It is important to note that although the teaching strategies for the two experimental groups differed, the thematic content of the strategies for both groups was based on the same five ideals of fair play. The authors further describe the teaching strategies in the manual as guided by the following learning processes: “(a) identifying and resolving moral conflicts, and (b) changing roles and perspectives” (Fair Play for Kids, 1990, p. i). These two concepts are highlighted throughout the manual in both the statement of objectives and instructions to teachers.

Selection of teaching strategies for the SL group was guided by the major concepts of modeling and reinforcement. The teachers in the SL group used the following teaching strategies on an ongoing basis: (a) role-modeling of the expected fair play behavior; and (b) verbal praise to reinforce appropriate fair play behavior. Examples of specific strategies used by teachers in this group included the implementation of a fair play awards program, role modeling by former Canadian Olympians, and a self-monitoring and peer-monitoring program based on an established code of fair play. The following are samples of several of the SL teaching strategies.

Compliment Cards. Students document examples of fair play displayed by their classmates. These examples are collected and examples are highlighted at the beginning of each week.

Fair Play Code. Students discuss ways they will put the ideals of fair play into action in their physical education class (e.g., we will cheer good plays by our opponents as well as our own team). This list is posted in the gym.

My Fair Play Agreement. Each student lists several fair play behaviors he or she will endeavor to exhibit during a particular activity. The student signs this list as a contract between himself or herself and other members of the class.

Teaching strategies for the SD group focused primarily on dialogue and problem-solving tasks associated with the identification and resolution of moral conflicts and moral dilemmas. Examples included use of a selection of games with built-in moral dilemmas or challenges followed by a student-centered dialogue session designed to help students toward resolution of the moral dilemma or conflict, and tasks that require students to invent games to resolve preestablished dilemmas (e.g., inequitable play time or insufficient equipment). In each of the SD strategies, the role of the teacher was to facilitate the dialogue process to allow students to come to mutually agreed upon solutions. The following are samples of such SD strategies.

Our Secret Game. Team 1 will play the game by the rules. Team 2 will be allowed to have a “secret” set of rules (e.g., our team can pick up the ball from out of bounds; our team can pick up the ball with our hands as well as kick it). After the game is finished, a “let’s talk” period provides the opportunity for students to discuss the following questions: (a) Were the players on Team 1 unhappy? Why? (b) Was there conflict or confusion? (c) Was there potential for danger? and (d) What happens when there are not the same rules for everyone?
Invent a Game. Working in small groups, students are provided with a series of game problems (e.g., unequal playing time, no referee, rule infractions). Each group must design their game to solve their designated problems.

The Problem-Solving Running Shoe. This was a strategy utilized by the SD teachers on an ongoing basis to address naturally occurring moral conflicts between individuals (e.g., disagreements over rules), or incidents involving individual students (e.g., losing self-control during a game). A diagram of a running shoe was presented on a piece of paper, with specific areas marked in which students were asked to identify the problem, alternatives, consequences, and solution. Children involved in conflicts such as disagreements over rules were instructed to go to a designated area called the “listening bench” and fill out the running shoe. Peer counselors also utilized the problem-solving running shoe to help classmates involved in unfair play resolve the problem situation.

Teachers of the control group did not use any of the teaching strategies in the Fair Play for Kids manual. In addition, these teachers were asked to avoid any explicit reference to the definition of fair play as stated in the manual used by the teachers of the experimental groups. However, as fair play was a stated learning outcome in the physical education curriculum guide, it was not appropriate for fair play to be ignored. Control group teachers utilized several traditional rule-based strategies to address fair play during physical education class. These strategies included the formulation and implementation of rules for (a) safety, and (b) appropriate behavior during game play and skill practice. Formulation of these rules was the sole responsibility of the teacher, and did not involve discussion with or input from students. In addition, the teachers of the control groups used positive reinforcement for acceptable behavior and punishment for unacceptable behavior (e.g., time out, verbal reprimand). The preceding strategies were consistent with ones utilized by the teachers prior to the start of the present study.

Measures

The measures used to assess moral judgment, moral reason, moral intention, and prosocial behavior were consistent with the measures employed by Gibbons et al. (1995). The variables were assessed across 10 moral dilemmas common to the physical education setting. The dilemmas were based on behaviors identified by Horrocks (1979) that teachers could readily observe. The behaviors included teasing others, showing off, arguing with teammates, complaining, sharing equipment, disobeying rules of the game, “hogging” the ball, disputing officials’ decisions, not taking turns, and ignoring teammates’ suggestions for improving. An example of a dilemma would be, “When you lose games in PE classes, you wonder whether to complain.” A mean score was then calculated across the 10 dilemmas for each variable demonstrating adequate internal consistency, which is generally regarded as $r \geq .70$ (Nunnally, 1978).

Moral Judgment. These 10 items focused on the extent to which an individual thought it was OK to engage in each of the behaviors identified in the different dilemmas. An example of an item would be, “Do you think that it is OK to complain when you lose in PE class?” Students responded to a 5-point Likert-type scale ranging from it’s always OK (1) to it’s never OK (5). Cronbach’s (1951) alpha reliability coefficients for moral judgment at pre-, mid-, and postintervention ranged from .86 to .94.
Moral Reason. While the term moral reasoning reflects the structure of moral development, moral reason focuses on moral content or the rationale given for solving a moral dilemma (Gibbons et al., 1995). Ten items assessed the most important thing that an individual considered when deciding whether it was OK to engage in the behaviors identified in the dilemmas. An example of an item would be, “Which is the most important thing to consider when you decide whether it is OK to complain when you lose?” The response options were whether or not I would be punished (1), whether or not I wanted to get even with a classmate (2), whether or not it is nice (3), whether it's against the rules (4), and whether or not it's fair or right (5). Alpha reliability coefficients for moral judgment ranged from .90 to .94 across the three time periods.

Moral Intention. Moral intention was assessed with 10 items that focused on the extent to which the individual thought he or she would engage in the behaviors identified in the dilemmas. An example of an item would be, “If you lose in future PE classes, what do you think you will do?” The 5-point Likert-type response format ranged from never complain (1) to always complain (5). The reliability coefficients for moral intention ranged from .88 to .94 across the three data collections.

Prosocial Behavior. An adaptation of Horrocks’s (1979) Prosocial Play Behavior Inventory (HPPBI) was used to assess prosocial behavior. This inventory asked teachers to rate each of their students on 10 behaviors commonly associated with fair play during participation in sport and games. An example of an item would be, “The child accepts defeat without complaining.” Each item was answered on a 4-point Likert-type scale ranging from not at all like the child (1) to very much like the child (4). Reliability coefficients for prosocial behavior were found to range from .90 to .92.

Procedure

Parents and students were informed of the general nature of the project and provided signed consent forms prior to student involvement in the intervention program. There was a 100% return rate of consent forms. Students and parents were informed that all responses were confidential.

Nine teachers from two different schools volunteered for the project as a result of an invited presentation about fair play in physical education offered by the first author (Sandra Gibbons). A brief explanation of the nature of the project, including the purpose of control groups, was provided as part of the presentation. All the teachers who volunteered were aware that they might be selected to teach a control group. All teachers were generalist elementary classroom teachers responsible for teaching their own physical education classes. However, in the 2 years prior to this study, these teachers had been involved in daily physical education curriculum implementation projects in their respective schools. In-service activities associated with these projects included participation in numerous workshops to increase teaching skills in several of the movement categories in the physical education curriculum. Thus, although the teachers were not physical education specialists, they had acquired more expertise in physical education than would be expected of a typical elementary classroom teacher.

In order to become familiar with the content and implementation procedures for the experimental teaching strategies, teachers of the experimental groups participated in a half-day workshop at the beginning of the school year. In addition,
Sandra Gibbons made monthly site visits to classes or phone calls to teachers in the experimental groups for the duration of the intervention period. These visits and phone calls were included to provide general monitoring of the project and availability for questions from the teachers. Dr. Gibbons held individual meetings with each of the teachers of the control groups prior to the start of the project. The purpose of these meetings was to clarify their assignment as a control group teacher and explain the teaching strategies associated with this group. Scheduling conflicts did not allow for a single group meeting of control group teachers; however, the same script was used to communicate expectations in all three individual meetings. It was not deemed necessary to hold a workshop for the control group teachers, as all three teachers had considerable experience with use of the control group teaching strategies. Dr. Gibbons also made monthly site visits or phone calls to the teachers of the control groups. All teachers were requested to avoid discussion with one another about the project for the duration of the intervention.

In order to provide a guideline for frequency and consistency of use, teachers in the experimental groups were advised to implement targeted fair play teaching strategies on a weekly basis. Weekly choice of specific strategies within each of the SL and SD groups was left to the judgment of individual teachers. This flexibility of choice was necessary to allow for differences in class schedules. Teachers were strongly encouraged to use all of the teaching strategies (15 for the SL group, 16 for the SD group) available to them during the course of the program. Each teacher kept a written record of the dates and situations in which a particular teaching strategy was implemented. These records indicated that all teachers had employed the fair play teaching strategies on a weekly basis, with frequencies ranging between 21 and 29. In addition, by the end of the intervention, teachers of the experimental groups had utilized the full range of assigned strategies across the physical activity units covered during the intervention. It should be noted that several of the teaching strategies in both experimental groups are designed for repeated use. For example, the problem-solving running shoe, a SD conflict-resolution strategy, was used on numerous occasions throughout the intervention period.

The four dependent measures were assessed before, in the middle of, and following the intervention period. For each of the three data collections, teachers from all three groups completed a copy of the HPPBI for each student. The other three dependent measures were administered to the students during an extended morning classroom period. Researchers supervised the students as they completed the questionnaires and addressed any questions raised by the students. The teachers were not present during administration of the measures.

Results

Analyses were conducted using the class \((N = 9)\) and student \((N = 204)\) as the unit of analysis. The class-level analysis should be considered the primary analysis in this study given that the classes, not the students, were randomly assigned to conditions, and that the treatment was administered to classes as opposed to individuals.\(^1\) Horn (1985), however, has discussed the merits of analyzing

---

\(^1\) The authors would like to acknowledge that this idea was forwarded by one of the reviewers.
both between-group (class) and within-group (student) differences. In essence, between-group differences take into account the classroom setting, while within-group differences take into account individual variation. Using both the class and student as the unit of analysis was consistent with the investigation conducted by Gibbons et al. (1995) and allowed for comparisons across the two studies.

The bivariate correlations among moral judgment, reason, intention, and behavior were first inspected to determine if multicollinearity ($r \geq .70$) was a problem. There was evidence of multicollinearity and so univariate, as opposed to multivariate, analyses were employed. Four repeated measures analysis of variance (RM ANOVA) were conducted using first the class and then the student as the unit of analysis with the alpha level adjusted ($p < .0125$) to compensate for the multiple analyses. Each of the indicators of moral development (judgment, reason, intention, and behavior) served as the dependent variable, while group (control, SL, SD) and time (pre-, mid-, postintervention) served as the independent variables. Following a significant interaction, simple main effects were conducted to determine if the groups were significantly different at each time period, and where there were differences, a Student-Newman-Keuls (SNK) post hoc test was conducted to ascertain which groups were significantly different from each other.

**Class-Level Analyses.** The group by time interaction was significant for moral judgment: $F(4, 10) = 38.3, p < .0001$; for moral reason: $F(4, 10) = 12.4, p < .001$; and for moral intention: $F(4, 10) = 31.2, p < .0001$; but not for prosocial behavior: $F(4, 10) = 4.2, p < .030$. There were no significant differences at pre-intervention for moral judgment, reason, or intention. At mid-intervention, the SL and SD groups scored significantly higher than the control group on moral judgment. At postintervention, the SL and SD groups were significantly higher than the control group on moral judgment and intention. Also at post-intervention, the SD group was significantly higher on moral reason than the SL and control groups. Means and standard deviations using class as the unit of analysis appear in Table 1 for each of the variables at the three time periods.

Effect sizes were calculated to determine the meaningfulness of treatment effects. The SD group (selected because it was consistently different from the control group) was compared to the control group on each of the four dependent variables at postintervention. Effect sizes were calculated using the following formula: $ES = (M_{SD} - M_{control}) / SD_{control}$. The values were 10.69 for judgment, 4.86 for reason, 6.15 for intention, and 3.02 for behavior. Moderate effect sizes range from .41 to .70, while large effect sizes are greater than .70 (Thomas, Salazar, & Landers, 1991). Therefore, meaningful treatment effects were found to exist with all four dependent variables when comparing the SD and control groups.

**Student-Level Analyses.** The group by time interaction was significant for each of the following variables: moral judgment, $F(4, 402) = 180.5, p < .0001$; moral reason, $F(4, 402) = 271.0, p < .0001$; moral intention, $F(4, 402) = 140.9, p < .0001$; and prosocial behavior, $F(4, 402) = 36.6, p < .0001$. There were no significant group differences at pre-intervention for any of the four variables. At mid- and postintervention, the SL and SD groups scored significantly higher than the control group on moral judgment, intention, and behavior. Also at mid- and postintervention, the SD group was significantly higher on moral reason than the SL and control groups. The postintervention effect sizes comparing the SD and control groups were 1.57 for judgment, 1.69 for reason, 1.50 for intention, and 0.88 for behavior. These values suggest that the group differences for each of the
Table 1  Results for Variables at Three Time Periods Using Class as the Unit of Analysis

<table>
<thead>
<tr>
<th>Variable</th>
<th>Preintervention</th>
<th>Midintervention</th>
<th>Postintervention</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Moral judgment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>3.39</td>
<td>0.08</td>
<td>3.41</td>
</tr>
<tr>
<td>Social learning</td>
<td>3.54</td>
<td>0.17</td>
<td>3.78</td>
</tr>
<tr>
<td>Structural developmental</td>
<td>3.45</td>
<td>0.13</td>
<td>3.91</td>
</tr>
<tr>
<td>Moral reason</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>3.15</td>
<td>0.30</td>
<td>3.18</td>
</tr>
<tr>
<td>Social learning</td>
<td>3.33</td>
<td>0.34</td>
<td>3.34</td>
</tr>
<tr>
<td>Structural developmental</td>
<td>3.26</td>
<td>0.27</td>
<td>3.83</td>
</tr>
<tr>
<td>Moral intention</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>3.33</td>
<td>0.13</td>
<td>3.32</td>
</tr>
<tr>
<td>Social learning</td>
<td>3.47</td>
<td>0.23</td>
<td>3.76</td>
</tr>
<tr>
<td>Structural developmental</td>
<td>3.40</td>
<td>0.15</td>
<td>3.88</td>
</tr>
<tr>
<td>Prosocial behavior</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>3.32</td>
<td>0.16</td>
<td>3.34</td>
</tr>
<tr>
<td>Social learning</td>
<td>3.48</td>
<td>0.15</td>
<td>3.59</td>
</tr>
<tr>
<td>Structural developmental</td>
<td>3.40</td>
<td>0.15</td>
<td>3.62</td>
</tr>
</tbody>
</table>

four variables were both statistically significant and meaningful. Means and standard deviations using the student as the unit of analysis appear in Table 2 for each of the variables at the three time periods.

**Discussion**

The purpose of this study was to examine the effectiveness of teaching strategies grounded in either social learning theory or structural-developmental theory on the moral judgment, reason, intention, and prosocial behavior of elementary-aged children. Results suggest that physical education students benefited from the SL and SD teaching strategies in comparison to students not exposed to these teaching strategies. From a general perspective, these results support the assertion that moral development is not an automatic outcome of the physical activity these children experienced over the seven-month period of the study, but rather requires the implementation of specific teaching strategies in order to effect positive change. Consistent with the findings reported by Gibbons et al. (1995), these results provide support for the teaching strategies included in *Fair Play for Kids*, a practical and readily accessible resource for teachers.

Results of the present study partially supported the hypothesis that children in either of the two experimental groups would score significantly higher on all four indicators of moral development than those children in the control group. Moral judgment, moral intention, and prosocial behavior scores (for the student-level analyses) were significantly higher for children in the SL and SD groups
Table 2  Results for Variables at Three Time Periods Using Student as the Unit of Analysis

<table>
<thead>
<tr>
<th>Variable</th>
<th>Preintervention</th>
<th>Midintervention</th>
<th>Postintervention</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Moral judgment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>3.39</td>
<td>0.70</td>
<td>3.41</td>
</tr>
<tr>
<td>Social learning</td>
<td>3.52</td>
<td>0.53</td>
<td>3.75</td>
</tr>
<tr>
<td>Structural developmental</td>
<td>3.44</td>
<td>0.62</td>
<td>3.91</td>
</tr>
<tr>
<td>Moral reason</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>3.12</td>
<td>0.78</td>
<td>3.16</td>
</tr>
<tr>
<td>Social learning</td>
<td>3.34</td>
<td>0.65</td>
<td>3.34</td>
</tr>
<tr>
<td>Structural developmental</td>
<td>3.26</td>
<td>0.70</td>
<td>3.83</td>
</tr>
<tr>
<td>Moral intention</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>3.34</td>
<td>0.73</td>
<td>3.34</td>
</tr>
<tr>
<td>Social learning</td>
<td>3.45</td>
<td>0.59</td>
<td>3.73</td>
</tr>
<tr>
<td>Structural developmental</td>
<td>3.39</td>
<td>0.68</td>
<td>3.88</td>
</tr>
<tr>
<td>Prosocial behavior</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>3.32</td>
<td>0.59</td>
<td>3.34</td>
</tr>
<tr>
<td>Social learning</td>
<td>3.47</td>
<td>0.42</td>
<td>3.58</td>
</tr>
<tr>
<td>Structural developmental</td>
<td>3.40</td>
<td>0.51</td>
<td>3.62</td>
</tr>
</tbody>
</table>

compared to the control group, but moral reason scores were only significantly greater than the control group for children in the SD group. This finding parallels the one other investigation conducted in the physical education setting that has examined the effectiveness of implementing structural-developmental teaching strategies. Romance et al. (1986) found significant increases in the moral reasoning scores of fifth grade children following an 8-week structural-developmental intervention. Many of the SD teaching strategies in the present study are consistent with those utilized in the Romance et al. (1986) study. Although there are differences in how the dependent variables were conceptualized in the two studies, the similar positive impact of the SD teaching strategies is noteworthy.

The differential impact reported in the current study with regard to the SL and SD teaching strategies provides a contrast to the results presented by Bredemeier et al. (1986). Bredemeier et al. examined the effectiveness of SL and SD teaching strategies utilized in a summer sports program with children aged five to seven years, and found both teaching strategies to be equally successful in producing significant increases in measures of moral intentionality and moral reasoning. The different findings across the two studies could be due to any number of methodological inconsistencies, such as the context, age of participants, duration of intervention, or variables assessed. Subsequent research will be necessary in order to better understand what circumstances maximize the effectiveness of various teaching strategies.

Use of Rest’s (1984) four-component model of morality to frame the dependent variables in the present study provided a productive structure for the examination of any potential differential effects of SL and SD teaching strategies on several aspects of moral development. Results of this study indicated a significant difference
in effects between SL and SD teaching strategies on the measure of moral reason. The SD group showed a significant increase on moral reason compared to both SL and control groups. This finding provides some insight into the learning processes associated with the different teaching strategies. Whereas social learning teaching strategies primarily utilize aspects of modeling and reinforcement to encourage moral development, structural-developmental strategies concentrate on the processes necessary to both identify and resolve moral dilemmas or conflicts. By their very nature, structural-developmental teaching strategies require students to question, reflect, dialogue, and negotiate toward the resolution of a moral dilemma. One would expect that these actions associated with the process of resolution would have a greater impact on moral reason than the less interactive social learning strategies. Thus, it follows that use of social learning strategies alone may not provide the opportunity for the type of interaction and reflection afforded by the structural-developmental strategies. However, although SL teaching strategies were less effective in promoting positive change on the moral reason measure, these strategies were equally as effective as SD teaching strategies on the other dependent variables. This result provides teachers with a promising selection of effective moral development teaching strategies to choose from when they are planning their physical education lessons. We encourage future studies of moral development to continue to examine the various processes associated with both social learning strategies and structural-developmental strategies.

There was some variation in the time at which significant group differences were noted for each of the four variables. For the class-level analyses, by mid-intervention only moral judgment scores were significantly higher for the treatment groups compared to the control group, yet by post-intervention one or both treatment groups scored significantly higher than the control group on moral reason and moral intention as well as moral judgment. Furthermore, there were still no statistically significant group differences on prosocial behavior by the end of the intervention—a finding that is consistent with the Gibbons et al. (1995) study—although the effect sizes suggested meaningful treatment effects with this variable. For the student-level analyses, group differences evident at the mid-point of the intervention remained unchanged by the end of the school year on all four variables. It appears that weekly use of the teaching strategies, as was the case in the present study, has the potential to effect positive change in a relatively short period of time. In combination, these findings highlight the importance of repeated assessments in order to more fully understand the nature of change with regard to moral development. In addition, little is known about the stability of the recorded changes, so the question of how long the effects remain if the teaching strategies are discontinued needs to be addressed.

In conclusion, this study investigated theory-based teaching strategies on multiple measures of moral development. The systematic implementation of effective teaching strategies proved to be beneficial for all involved. Students experienced positive growth; teachers embraced strategies that were readily incorporated into existing lesson plans; and researchers gained additional insights about the nature of change with certain moral indicators. Clearly, such investigative efforts need to be ongoing in the future if the goal of promoting moral development is to be realized.

References


