Effects of Holding Students Accountable for Social Behaviors During Volleyball Games in Elementary Physical Education

Craig A. Patrick
Pepper Ridge Elementary School

Phillip Ward
University of Nebraska

Darrell W. Crouch
Carlock Elementary School

This study investigated the effects of a semiformal accountability intervention (a modified version of the good behavior game) on the occurrence of appropriate and inappropriate social behaviors, and appropriate skill attempts during a 20-lesson volleyball unit. Participants were 67 students in Grades 4, 5, and 6. Following the collection of baseline data, students received intervention consisting of (a) differential awarding and removing of points for appropriate and inappropriate behavior, (b) public posting of team points, (c) the establishment of daily criteria, (d) a special activity for teams that met the criteria, and (e) an end-of-unit activity for teams that consistently met the criteria. A multiple baseline design across students showed that the intervention was effective in reducing inappropriate social behaviors and increasing appropriate social behaviors, but did not affect the number of correct volleyball skills performed. Results are discussed relative to task systems and social skills.

The promotion of socially responsible behavior in the form of moral character, conformity to social rules and norms, cooperation, and positive styles of social interaction has been a traditional and valued educational objective for American . . . public schools in almost every educational policy statement since 1848, being promoted with the same frequency as the development of academic skills. (Wentzel, 1991, p. 2)

Despite the implied equity between academic and social objectives in Wentzel's statement, there is a substantive difference between the number of studies that
investigate academic outcomes and those that investigate social outcomes. Nonetheless, social skills are commonly investigated in several literatures, including classroom management, social competence, and fair play. In each literature a shared objective for social skills is that students learn and apply the rules of a particular context. Because the contexts often differ, the type of social skill required varies (e.g., the skills needed to work together to complete a group assignment in a classroom are different from those needed to respond to point losses during volleyball games). Furthermore, the theoretical perspective through which social skills are viewed by a particular literature also influences the type and function of social skill observed (e.g., management versus fair play).

In the classroom management literature, the purpose of social skills is viewed primarily, but not exclusively, as contributing to classroom order by developing skills necessary to participate successfully in classroom events (Colvin & Sugai, 1988; Doyle & Carter, 1984; Soar & Soar, 1979). In physical education, several researchers have addressed the theme of what it takes to be a “member in good standing” in a class, often concurrent with their primary focus, and often from the perspective of preventative management (Hastie & Pickwell, 1996; Johnston, 1995; Oslin, 1996; O’Sullivan & Dyson, 1994). While the problem of inadequate social skills is acknowledged in physical education (e.g., Hellison, 1995; Sharpe, Brown, & Crider, 1995), there has been little effort to empirically examine efforts to remedy it.

In the social competence literature, the purpose of social skill development is to learn the rules not just of the classroom and the school, but to acquire skills that generalize beyond the classroom to other settings (e.g., home, after school, present and future work settings). It includes studies of moral citizenship and values education (Kohler & Fowler, 1985; Ostroky & Kaiser, 1995; Wentzel, 1991). In physical education, social competence studies have investigated self-responsibility for delinquency-prone youth (DeBusk & Hellison, 1989), moral development (Gibbons, Ebbback, & Weiss, 1995; Romance, Weiss, & Brockover, 1986; Weiss & Bredemeier, 1986), and values education (Chen, 1996; Ennis, 1992; Wandzilak, Carroll, & Ansorge, 1988).

Though similar to investigations of social competence, the fair play literature is specific to physical education. Within the past decade a small number of researchers have investigated social skills in the context of game play (Giebink & McKenzie, 1985; Grant, 1992; Sharpe, et al., 1995). Of particular interest are the studies conducted by Giebink and McKenzie (1985) and Sharpe et al., (1995). Giebink and McKenzie (1985), using a multitreatment reversal design (A-B-C-D-A), intervened on three behaviors during softball lessons: (a) compliment your teammates, (b) play fair, and (c) accept the consequences. The behaviors were assessed across baseline (A) and three experimental conditions: teacher instructions and praise for fair play behaviors (B), modeling of fair play behaviors (C), and a point system for fair play behaviors (D). All three interventions increased fair play behaviors and decreased inappropriate behaviors when compared to baseline. The behaviors developed in softball however, did not generalize to a new setting: recreational basketball games. Giebink and McKenzie (1985) then intervened in the recreation setting using a multitreatment reversal design (A-B-A-C) to compare baseline (A) with the teacher instructions and praise for fair play behaviors (B) and the point system (C). In the recreational basketball setting, inappropriate social behaviors decreased in both experimental conditions compared to baseline levels; however, fair play behaviors did not improve in any condition.
Sharpe et al. (1995) used an intervention designed to teach conflict resolutions and leadership skills in physical education and reported that these behaviors generalized to regular classroom settings. The Sharpe et al. (1995) study is particularly significant because of evidence of generalization of social skills to classroom settings, and because it provides an empirical validation of a social skills curriculum. At least three other social skills curricula have been developed to address the context-specific needs of children in sports: Fair Play for Kids (1990), Sport Education (Siedentop, 1994), and Teaching Responsibility Through Physical Activity (Hellison, 1995).

One conclusion from the above review is that, unless planned for and taught by the teacher, appropriate social skills often remain underdeveloped. If one accepts that improving social skills ought to be part of the functional curriculum, then teaching social skills becomes one of the tasks of teaching. Given this conclusion, the task system paradigm provides a useful framework to empirically investigate the improvement of social skills in education and physical education in particular. Though originally derived from the classroom management literature, the task system framework has great utility as a tool for investigating dimensions of classroom life. The major task systems in physical education are instructional, managerial, and social (Siedentop, 1991; Tousignant & Siedentop, 1983). Within a task system, tasks are defined and maintained by the effectiveness of the accountability used by the teacher. When there is no accountability, or when it is ineffective, task accomplishment may be incomplete, or the task may be modified by the student in such a manner as to change the intended outcome (Doyle, 1983). Holding students accountable for the accomplishment of social tasks is a key instructional procedure to ensure that such skills taught by the teacher are acquired by students.

This study was occasioned by a concern of the first author (an elementary school physical education teacher), who noted that during game play in volleyball (including applied tasks with modified rules), students in his classes were seldom encouraging and supportive of each other and that at times some students behaved inappropriately. He wanted to find a proactive strategy designed not only to reduce the occurrences of inappropriate behaviors, but to increase the occurrence of encouraging and supportive behavior. After some discussion, we decided to modify an intervention called the “good behavior game” (Barrish, Saunders, & Wolf, 1969) to meet the teacher’s goals. The good behavior game is a group contingency that typically operates as follows. A class is divided into at least two groups, and when any member of the group misbehaves, a point is marked against that group. At the end of a period of time, the group with the fewest points wins. Winning typically allowed the group members to engage in some special activity. In short, the group is held accountable for its members’ inappropriate behavior. In our discussions, we decided to modify the good behavior game to hold students accountable for both appropriate and inappropriate behaviors by adding points for appropriate behaviors and removing them for inappropriate behaviors. Furthermore, we decided that rather than have teams compete against each other, we would instead have them compete against a daily criterion. Thus, any and all teams that met the criterion would “win.”

In classroom studies of social competence, an implied outcome of social skill improvement has often been improved academic performance (Wentzel, Weinberger, Ford, & Feldman, 1990). One possible explanation for improved
achievement in the classroom is that with fewer inappropriate social skills less disruption and distraction occurs, which improves the opportunity to learn. In the present study, in addition to assessing the effects of the intervention on social skills, we were also interested in determining whether or not the number of successful forearm passes and overhead passes were affected as a result of the social skills intervention.

Three experimental questions guided our investigation during volleyball game play:

1. What is the effect of the modified good behavior game on the number of occurrences of inappropriate social behaviors?
2. What is the effect of the modified good behavior game on the number of occurrences of appropriate social behaviors?
3. What is the effect of the modified good behavior game on the number of successful forearm passes and overhead passes?

**Method**

**Participants and Setting**

Participants in the study were the students enrolled in three intact physical education classes and their physical education teacher at a suburban elementary school: a fourth-grade class consisting of 21 students (12 boys, 9 girls), a fifth-grade class consisting of 25 students (11 boys, 14 girls), and a sixth-grade class consisting of 21 students (11 boys, 10 girls). Parental consent for participation was obtained for all students. In addition, each student volunteered to participate in the study. The teacher was in his 10th year of teaching.

Physical education classes were held daily for 20 minutes in Grade 4, and daily for 30 minutes in Grades 5 and 6. The lessons were conducted in half of the school gymnasium, in an area approximately the size of one basketball court (90 x 50 ft). During each of the 20 lessons in the volleyball unit, 10 minutes was allocated for game play with modified rules. Students in each class were grouped into four teams of 5–6 students. The students remained in these teams for the duration of the study. Teams typically played against each other on a rotated schedule.

**Data Collection Procedures**

Three classes of behavior were measured: (a) the number of appropriate social behaviors per class, (b) the number of inappropriate social behaviors per class, and (c) the number of correct forearm passes and sets per class. Appropriate and inappropriate social behaviors were further subdivided into three categories: physical acts, verbal statements, and gestures committed by students. The following list presents the behaviors, definitions, and examples of each category:

**Appropriate Social Behaviors**

- **Physical:** Physical contact between students that is supportive in nature or that is a response to good play (e.g., high five, pat on the back, handshake).
- **Verbal:** Statements made by students that are supportive in nature or that are a response to good play (e.g., “good job,” “good try,” and “way to go”).
In coding an instance of an appropriate behavior, we made a judgment regarding an observed contingent relationship between the appropriate behavior and the events that preceded it. In addition, a separate subcategory labeled “false acts” was included to record instances where students used an appropriate behavior that was not contingent upon some success or effort by team members but occurred in the presence of the teacher merely to earn a point. For example, during a break in the game and as the teacher passes by, one student turns to another and says “well done.”

Data were collected via videotape for a 5-minute block of the 10-minute game for all 20 lessons of the unit. “Interactions during game play” was selected as the unit of analysis, because the teacher had observed the most inappropriate acts during this phase of the lesson. Two games occurred concurrently. Data collection was limited to 5 minutes of each game, due to equipment limitations and also to standardize the observation interval. Data collected for appropriate and inappropriate social behaviors were limited to the sensitivity of the camera’s microphone and the lens of the camera. Because of the need to capture the verbal comments of the students, the camera was placed to the side and at an angle to the court. As such, it was quite obtrusive. The school however, was a regular site for student teachers who were supervised and videotaped by university personnel. The practice of videotaping was therefore a common event in the school and in these classes in particular. Students were informed that the camera would be used to help the teacher make judgments about their performance during the volleyball unit.

The forearm pass and overhead pass skills were selected because they represented the content of the instruction that preceded games for each class and were the most frequently used skills in the game. The physical education teacher for the class determined the criteria for correct performance of the volleyball skills.

Independent Variable

The independent variable consisted of five components. Each will be discussed in turn.
Differential Awarding and Removing of Points. During the 10 minutes of


game play, the teacher moved between the two games and awarded points to teams


when members demonstrated appropriate behaviors. Points were removed from


the team score if the teacher observed instances of inappropriate behavior. A “false


act” also resulted in a lost point.

Public Posting of Team Points. During scheduled breaks in the game, stu-


dents recorded the points that were awarded for appropriate behavior and/or lost
due to inappropriate behavior on a wall poster under their team name.

The Establishment of Daily Point Criteria. At the beginning of each class,


the teacher established a criterion for each group to meet or exceed during the
daily game. On the first day of the intervention, the teacher established a criterion


that was 10 times that of the teams’ baseline. With the exception of the first day of
the intervention, teams were required to meet or exceed the previous day’s perfor-
mance, or a criterion established by the teacher in the case of an occasion where
there were an exceptional number of points accrued on the previous day.

Daily Special Activity. Teams that met the daily criterion played an addi-
tional three minutes of game play each lesson. Teams that did not meet the crite-


rion were not awarded the special game time. Any and all teams that met the criterion
were awarded the special game time.

Special End-of-Unit Activity. A special end-of-unit lunchtime game was
provided for the two (or more, if they were equal) teams in each class that met the
daily criterion most often. Thus, though a team may have exceeded the criterion
each day, in order to participate in this lunchtime game, the criterion had to be
consistently met over the duration of the unit of instruction.

Experimental Design and Procedure

A multiple baseline design across classes (Cooper, Heron, & Heward, 1987)
was used to assess the efficacy of the modified good behavior game in holding
students accountable for the targeted social behaviors. In single-subject designs,
judgments about internal validity are made on the basis of visual analysis of changes
in the data as a consequence of changing experimental conditions (which includes
the removal or introduction of a baseline). The multiple baseline design uses a
time-lagged strategy to assess internal validity when changes in the data path (level
and trend) plotted on the first tier occur at the point of intervention, without changes
occurring in the underlying tiers. This effect, when reproduced in Tier 2, and in
particular in Tier 3, increases confidence that changes in the dependent variable
are in fact due to the presence of the independent variable (see Cooper et al., 1987,
for a more detailed explanation).

Baseline. During baseline, students played the game of volleyball.

Intervention. On Day 1 of the intervention, the teacher took 10–15 minutes to
(a) explain the rules of the good behavior game, (b) have the students put the poster on
the wall with their team names marked on it, (c) allow a short rehearsal where the
teacher awarded points for good behavior, and (d) establish the daily criterion. Later
during that lesson and for the remainder of the unit, points were awarded or removed
contingent upon the targeted behaviors during game play. Due to the time-lagged
strategy of the multiple baseline design, Grade 5 received the intervention first, fol-


lowed by Grade 4 and then Grade 6. We made the decision to intervene in this order
based on the stability and trend of the data paths of each class.
Interobserver Agreement

Judgments of correct and incorrect performance of volleyball skills made by the teacher were compared to those of a second trained observer (another physical education teacher in the school) to determine the percentage of interobserver agreement. The second observer was not directly involved with the study’s implementation and had been trained using direct observation and video recordings to a criterion of three sessions at 80% or higher prior to the start of the study. Both observers coded the dependent variables from the videotape independent of each other. Interobserver agreement (IOA) was assessed on 50% of the sessions distributed across baseline and intervention phases (typically, every other day). The IOA percentages were calculated using a trial-by-trial method, by dividing the number of agreements by the number of agreements plus disagreements and multiplying by 100. Mean IOA percentages for social behaviors were: Grade 5, 85% (range = 73–93%); Grade 4, 85% (range = 74–97%), and Grade 6, 87% (range = 70–98%).

Interobserver agreement was also conducted on 25% of the sessions (two baseline and three intervention sessions for each class) for the correct performance of the forearm pass and overhead pass. The means for correct volleyball skills were calculated similarly for social behaviors (i.e., trial-by-trial) and were: Grade 5, 93% (range = 91–95%); Grade 4, 89% (range = 81–96%); and Grade 6, 92% (range = 90–95%).

While the IOA means for social behaviors in each grade lie in the mid-80s, the range of the IOAs for each grade indicate there was at least one occasion per class (in Grade 4 there were two occasions) where IOA agreement scores were in the 70s. In contrast, the IOA means for the forearm pass and overhead pass were quite high with a small range. The difference in variability in the ranges and level of agreement between the IOAs for social behaviors and volleyball skills may be an artifact of the difficulty of coding the less obvious social behaviors versus the more overt volleyball skills. This problem of lower reliability for social skills has been reported elsewhere (Dugan et al., 1995).

Procedural Integrity

During both baseline and intervention, in order to standardize instruction, the teacher was instructed to (a) only stop a game to deal with managerial problems (e.g., arguments over the score fights) or to allow the points accrued during the intervention to be posted on the wall chart, (b) restrict his feedback, and (c) maintain his monitoring (movement around the court perimeter) of games.

Results

Four primary dependent measures (inappropriate and appropriate social behaviors, and correct and incorrect skill trials) were totaled for each day and plotted. In addition, the mean and range for each variable during baseline and intervention were calculated. Also of interest, and totaled for each day, were two secondary variables: the number of false acts and the days where teams reached their criterion level of points.

The first question addressed in this study was “What is the effect of the modified good behavior game on the number of occurrences of inappropriate social
Table 1  Means and Ranges for Appropriate and Inappropriate Social Behaviors

<table>
<thead>
<tr>
<th>Grade</th>
<th>Baseline</th>
<th>Intervention</th>
<th></th>
<th>Baseline</th>
<th>Intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>Range</td>
<td>$M$</td>
<td>Range</td>
<td>$M$</td>
</tr>
<tr>
<td>Grade 5</td>
<td>12</td>
<td>8–16</td>
<td>102</td>
<td>35–184</td>
<td>25</td>
</tr>
<tr>
<td>Grade 4</td>
<td>10</td>
<td>2–16</td>
<td>121</td>
<td>53–196</td>
<td>23</td>
</tr>
<tr>
<td>Grade 6</td>
<td>12</td>
<td>6–20</td>
<td>135</td>
<td>100–153</td>
<td>25</td>
</tr>
</tbody>
</table>

behaviors?" As shown in Table 1, mean baseline measures were 25, 23, and 25 for Grades 5, 4, and 6, respectively. During intervention, the means dropped to 3, 2, and 1. Visual inspection of the graphed data in Figure 1 indicates that the change in level was immediate and was maintained throughout the intervention for each class.

The second question addressed in this study was "What is the effect of the modified good behavior game on the number of occurrences of appropriate social behaviors?" Mean baseline measures for appropriate behaviors were 12, 10, and 12 for Grades 5, 4, and 6 (see Table 1). When the intervention was implemented, the means rose to 102, 121, and 135. Visual inspection of the graphed data in Figure 1 indicates that the change in level was immediate and increased throughout the study for Grades 5 and 4, and was relatively stable after Day 1 of the intervention for Grade 6.

There were few false acts observed. No false acts were observed in Grade 5. In Grade 4, on Day 16, four false acts were committed by the same student. In Grade 6, on Day 11, two acts (by different students) occurred, and on Day 16, one false act was observed. Data were collected on the number of days a team in any class did not meet its established criterion. For Grades 5 and 6, there were no instances when teams in either class failed to meet the criterion. In Grade 4, there were two occasions (two separate teams) when the criterion was not met.

The final question addressed in this study was "What is the effect of the modified good behavior game on the number of successful forearm and overhead passes?" Figure 2 displays the number of correct and incorrect forearm passes and overhead passes performed by members of each class. The vertical dotted line indicates when the good behavior game intervention occurred for each class. Changes in the level or trend of the data at and following that point in time would indicate that the intervention targeted on the social behaviors influenced the volleyball skill performances as well. Visual inspection of the graphed data in Figure 2 indicates that there were no concurrent changes for any of the classes at the time of the social skills intervention. Data paths for correct and incorrect volleyball skill performances maintained their trends.

Discussion

The results of this study show that the modified version of the good behavior game was effective in reducing inappropriate social behaviors and increasing appropriate social behaviors, but did not affect the quality of students' skill attempts.
During the baseline in each class, more inappropriate social behaviors occurred than appropriate social behaviors. This finding was also obtained by Giebink and McKenzie (1985) in both the softball and basketball settings they investigated. The finding is also consistent with the conclusions from the classroom management, social competence, and fair play research that suggest that appropriate social skills often remain underdeveloped unless planned for and taught by the teacher.

The good behavior game represents one strategy that can be used by teachers to teach social skills in physical education. Focusing on the group rather than
on specific individuals in the good behavior game allows those individuals who have fewer social skills to be in the presence of peers who can model correct behaviors and who are present at the time of an inappropriate behavior to discourage it. One possible negative outcome of such contingencies is that an individual within the group may be unfairly punished by members for the inappropriate behavior if it prevents the group from achieving the daily criterion. In this study,
there were only two occasions in the entire study where teams did not meet their established criterion and therefore did not play in the daily special event. Furthermore, no instances were observed where one member of a group was unfairly treated by other group members. One explanation for this may lie in how the criteria were established. Each team’s criterion was based on the previous day’s score of that team (i.e., meet or exceed the previous day’s performance) and was not a mean of the class or an arbitrary judgment made by the teacher.

Although students were not interviewed to assess their enjoyment and sense of the intervention, the teacher reported that during the intervention, the students seemed to relax and there was less pressure to do well. He suggested that this was because it was less likely that a student would be ridiculed if a bad play was made during the intervention, and because playing the game under the good behavior game conditions was fun.

The finding that volleyball skills remained unaffected by the intervention is interesting. Social competence studies suggest that social competence and achievement in classrooms may be causally related (Wentzel, 1991). If so, then improved volleyball skill performance during the intervention would have occurred concurrent with the introduction of the intervention. It did not, which suggests either (a) that the social and psychomotor response classes are unrelated (i.e., there is no causal link, as Wentzel suggests), or (b) that the baseline conditions in this study for social skills were not sufficiently “inappropriate” to influence the performance of volleyball skills. One explanation for this might be that the inappropriate behaviors deemed undesirable by the teacher and researchers did not functionally affect student achievement. Another explanation is that the intervention did not produce conditions that were positive enough to influence skill achievement. This seems quite unlikely given the level of appropriate social skills demonstrated. It is more likely that the positive environment did not influence the skill performance. Similar findings have been reported in classroom studies (Soar & Soar, 1979). Soar and Soar (1979), in their studies of classroom climates, found little difference in achievement gains in classrooms where the climate was either neutral or positive, but they found negative correlations in classrooms where the climate was negative.

The previous discussion notwithstanding, the development of social skills does not require the rationale of improved student learning. As Wentzel (1991) noted, the development of social skills and social competence as educational objectives was a component of “every educational policy statement since 1848” (p. 2). There are three challenges for researchers in the area of social skills training. First, strategies designed to improve social skill competence need to be empirically assessed. Second, such assessment should demonstrate that changes resulting from social skills training and interventions can be maintained in the setting. Third, researchers need to assess the generalization of social skills to new settings. In the present study, while the efficacy of a strategy designed to improve social skills was assessed, maintenance and generalization were not. In physical education, few studies have assessed the effects of maintenance and generalization of social skills. Giebink and McKenzie (1985) found that social skills training did not generalize from a softball setting to a basketball setting. Sharpe et al. (1995), however, did report generalization of conflict resolution skills from the gymnasium to classroom settings. These mixed findings, relative to the generalization of social skills in physical education, reflect a larger problem of generalizing social skills from one setting to another. In a review of social skills in preschool settings, the
most extensively and rigorously studied setting for social skills, Chandler, Lubeck, and Fowler (1992) noted that "generalization and maintenance have been particularly difficult to obtain in applied research with peer interactions and young children" (p. 416). Future studies in physical education should try to assess the maintenance and generalization of social skills and, in particular, the generalization of social skills to classrooms and playgrounds, as well as to different units of instruction.

Framing social competence as skill development allows social skills to be viewed in the same manner as cognitive and psychomotor skills. In this study, doing so allowed social skill to be examined within the tasks systems or ecological paradigm (Doyle 1979, 1986). In this paradigm, a task, a social skill in the present case, can be treated as a dependent variable, and some form of accountability (e.g., teacher monitoring, public posting, peer mediation) can be used as an independent variable to assess and improve task accomplishment. The task systems framework has allowed researchers in physical education to investigate different dimensions of teaching and learning. For example, researchers have (a) examined the specific systems in operation in physical education (e.g., Silverman, Kulina, & Crull, 1995; Tousignant & Siedentop, 1983); (b) examined the congruence between tasks stated by the teacher and the actual tasks performed by students (e.g., Jones, 1992; Lund, 1992); and (c) examined the ecology of the social system that operates in different physical education contexts, such as camp, dance, and sports settings (e.g., Hastie & Pickwell, 1996; Hastie & Saunders, 1990); and (d) most recently, used task accomplishment as a dependent variable and accountability as an independent variable to improve student achievement (Crouch, Ward, & Patrick, 1997; Ward, Smith, & Sharpe, 1997).

The present study's use of the good behavior game to increase socially appropriate behavior and reduce socially inappropriate behaviors extends the research on tasks and accountability in at least two ways. First, it assesses the efficacy of the good behavior game as a type of semiformal accountability. Second, the study represents an initial step toward expanding the empirical base of the task systems paradigm (specifically, tasks and accountability) into domains other than the psychomotor—in this case, the affective domain. In addition, the study represents one of a small number of studies that empirically validate strategies designed to improve social skill competence in physical education. We hope that additional research efforts along these lines will lead to other strategies that will focus on the affective domain in physical education.

References


---

**Acknowledgments**

Special thanks are due to Mary O’Sullivan and Bill Murphy for their comments on the manuscript.