Effects of Intervention on Differential Treatment of Boys and Girls in Elementary Physical Education Lessons

Raymond R. Dunbar
Yellowhead School Division, Alberta

Mary M. O'Sullivan
University of Victoria

This study examined the effects of verbal and graphic feedback on the distribution of teacher verbal behaviors (positive and corrective feedback, praise, desist, and questioning) and the teacher's use of student demonstrators during elementary coeducation physical education lessons. Data were collected over a 3-month period on two female nonphysical education specialist elementary teachers. A multiple baseline design was used to show the efficacy of the treatment. The results indicated that in baseline both teachers interacted with boys and girls inequitably on all variables. The intervention package and daily follow-up were influential in establishing more equitable teacher interaction patterns with boys and girls. The teachers' use of demonstrators was also distributed more equitably between boys and girls following the intervention.

Society is increasingly aware of the differential treatment of males and females in all facets of daily life. Despite the acceptance that sex-role messages pervade children's schooling experiences, little is known about the differences in day-to-day teacher-student interaction patterns in classroom or gymnasium settings.

Much of the effort to eliminate sex inequities in the public school system has dealt mainly with the development of nonexist curriculum materials and the provision of activities to raise teachers' awareness levels of inequitable practices. Unknowingly, many teachers provide differential treatment to boys and girls in both their formal and informal interactions with them (Brophy & Good, 1970; Serbin, O'Leary, Kent, & To-nick, 1973).

Brophy and Good (1974) reviewed the literature on student-teacher interactions and reported that "data from many different sources and educational levels agree in showing . . . the primary difference is quantitative; boys tend to have more interactions of all kinds with their teachers than girls" (p. 229). They concluded that boys and girls received quite different amounts and types of teacher feedback and that "the sex of the student and his behavior characteristics [are] important predictors of teacher-student interaction patterns" (p. viii).

Request reprints from Mary O'Sullivan, School of Physical Education, University of Victoria, Victoria, B.C., Canada V8W 2Y2.
There are conflicting findings about the extensiveness of the differential treatment of students in classroom settings. Some studies indicated that girls received more praise than boys (Galejs & Hegland, 1982; Lippitt & Gold, 1959) while others found that girls received less teacher praise and instructional assistance than did boys (Brophy & Good, 1974). Several studies indicated that boys received more teacher disapproval and criticism than did girls (Brophy & Good, 1970, 1974; Good, Sikes, & Brophy, 1973; Leinhardt, Seewald, & Engel, 1979; Lippitt & Gold, 1959; Meyer & Thompson, 1956). Spaulding (1963) reported that boys received a greater percentage of disciplinary action, and Etaugh and Harlow (1975) noted that boys were scolded more often than their behavior warranted.

Studies of teacher-student interaction in physical education have reported similar differential treatment in elementary and junior high physical education lessons (Griffin, 1981; Solomons, 1976). Griffin (1981) suggested that one of the goals of physical educators is "to provide maximum opportunities for successful participation in a variety of activities for all students regardless of their sex or ability level" (p. 17). With the present trend toward coeducational physical education classes, teachers must be more conscious of their interaction patterns and teaching style to ensure that all students have equal opportunity to participate and experience success. Griffin (1981) documented that having coeducational classes did not necessarily imply sex equality within those classes.

While the initial strategy in developing sex equity in the classroom is to make teachers aware of their established classroom behavior, teachers must be "coached" (Joyce & Showers, 1982) as they try to implement these new concepts of sex equity in their day-to-day interactions with students. Many of the coaching procedures outlined by Joyce and Showers (modeling, practice of the behavior, feedback, and reinforcement) have a long and successful history in applied behavior analysis. Bossert (1981) has argued persuasively that a definition of equity that relies solely on equitable teacher feedback to students is incomplete. Sex equity in classroom or gymnasium interaction must be defined as (a) equity in teacher feedback, (b) equity in the opportunity structure that girls and boys have to participate in the academic and nonacademic roles of the classroom or gymnasium, and (c) equity in students' perceptions of role access.

The purpose of this study was to determine if teachers' feedback and students' opportunity structures could be more equitably distributed to boys and girls in sex-integrated physical education lessons as a result of the following: making the teachers aware of their interaction patterns with students, discussing why and how they might establish an equitable gymnasium environment, monitoring their teaching, and providing them with feedback and reinforcement for gradual approximations toward the goals set for specific target behaviors. The establishment and maintenance of a more positive learning environment was implicit in setting targets for more equitable teacher interactions with boys and girls. The study's specific focus was to assess the effect of this intervention package on the distribution of teacher feedback (positive and corrective feedback, praise and desist statements) and the opportunities provided for boys and girls to answer questions and act as demonstrators in physical education.

Method

Subjects and Setting

Two female elementary school teachers volunteered to participate in the study. Both were classroom generalists who taught physical education to their own classes. Sub-
ject A taught grade 2 and had 21 students (12 girls, 9 boys). Subject B taught grade 1 and had 22 students (11 girls, 11 boys). The subjects were told that the study was focusing on how teachers and students interact in the gymnasium. They were told that the students and teacher would be observed for several lessons and, during the course of the study, the researcher would meet several times with them to discuss recorded observations. Extreme care was taken not to refer to the dependent variables under study before they were to be intervened upon.

Independent Variable

The intervention package consisted of specific feedback from the researcher on observed verbal teacher behaviors. This feedback included verbal feedback and graphic representation of baseline data. Following a conference with the teacher at which time targets were set for her behavior, she was given daily feedback on her progress toward the end goal.

Dependent Variables

The teacher's verbal behavior was coded using the following categories:

1. Praise. A positive teacher verbalization to appropriate student conduct.
2. Desist. A teacher verbalization used in an effort to terminate student misbehavior.
3. Positive feedback. A positive teacher verbalization to student's skill behavior.
4. Corrective feedback. A judgment by the teacher on the incorrectness of a student's skill response(s).
5. Questioning. Teacher questions directed to students.
6. Demonstrations. Teacher requests of students to demonstrate a specific motor skill or motor task.

Experimental Design

As a result of the intervention, it was hoped that changes in teacher verbal behavior patterns would result in a more equitable distribution of teacher verbal behavior to boys and girls. The research design used to show the efficacy of the intervention package was the multiple baseline design across teacher behaviors (Hersen & Barlow, 1976). The data were graphed showing two replications of a two-tiered multiple baseline design (Hersen & Barlow, 1976). Intervention by the researcher occurred at staggered intervals on two dependent variables at a time for each subject to form a multiple baseline design across behaviors. In this study each pair of behaviors was intervened upon in sequence after relatively stable baselines had been established.

Data Collection

A series of 15 to 18 consecutive physical education lessons were videotaped (15 for Subject A and 18 for Subject B). The videotaping was conducted in the schools' gymnasium. The teacher was equipped with a cordless microphone and a small belt-attached transmitter for recording all verbal responses. Immediately prior to the study, the teachers were videotaped teaching five physical education lessons. This was to allow the teachers
and students to become accustomed to the camera equipment and the researcher's presence. Following this, baseline data on the dependent variables were collected for both subjects.

Each videotaped lesson was coded using a predetermined coding procedure and provided a frequency count for each dependent variable. All lessons were 23 minutes long. The content of the lessons varied from ball handling skills, track and field activities, and skipping for Subject A to skipping, ball handling skills, and gymnastic activities for Subject B. The lessons were coded by two graduate students. Both had previous experience teaching elementary physical education and were familiar with systematic observation and analysis of physical education teaching.

Results

Interobserver Agreement

Interobserver agreement was established prior to and during the course of the study by comparing the scores of two independent observers. Two methods of establishing interobserver agreement were used, one being more stringent than the other. The formula for establishing interobserver agreement for event recording is the smaller total of observations divided by the larger total and multiplied by 100 (Sulzer-Azaroff & Mayer, 1977). To ensure that the observers were recording the same incidence of behavior, observational periods were broken down into 10-second intervals. Not only did the observers have to tally the occurrence of the incident, they also had to tally it in the appropriate interval. The mean of the scored and unscored interval was used as a much more stringent estimate of the reliability of the data (Hawkins & Dotson, 1975). The average interobserver agreement for Subject A was 88.2% using the mean of the scored and unscored

Table 1

<table>
<thead>
<tr>
<th>Subject</th>
<th>Mean of scored and unscored intervals</th>
<th>Event-recording method</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>88.3%</td>
<td>93.1%</td>
</tr>
<tr>
<td></td>
<td>89.0</td>
<td>100.0</td>
</tr>
<tr>
<td></td>
<td>88.3</td>
<td>90.8</td>
</tr>
<tr>
<td></td>
<td>90.0</td>
<td>92.9</td>
</tr>
<tr>
<td></td>
<td>85.8</td>
<td>93.1</td>
</tr>
<tr>
<td></td>
<td>88.1</td>
<td>90.5</td>
</tr>
<tr>
<td>B</td>
<td>81.5</td>
<td>89.0</td>
</tr>
<tr>
<td></td>
<td>91.1</td>
<td>94.3</td>
</tr>
<tr>
<td></td>
<td>87.0</td>
<td>94.5</td>
</tr>
<tr>
<td></td>
<td>80.7</td>
<td>94.7</td>
</tr>
<tr>
<td></td>
<td>86.4</td>
<td>93.7</td>
</tr>
<tr>
<td></td>
<td>92.5</td>
<td>95.3</td>
</tr>
</tbody>
</table>
intervals, and 93.4% using the event-recording method. The average interobserver agreements for Subject B using the mean of the scored and unscored interval method and the event method were 84.2% and 93.5%, respectively. The interobserver agreement for each subject is shown in Table 1.

**Subject A**

The collective data for Subject A were graphed illustrating a multiple baseline design across behaviors (see Figure 1). Teacher questioning and use of demonstrations

![Figure 1](image-url)
were the first variables intervened upon. During baseline, Subject A on average directed 61% of her questions to boys and 39% to girls, and called upon boys to demonstrate 2.5 times and girls .75 times per lesson. Following intervention, the teacher on average directed 55.3% of her questions to boys and 44.7% to girls. Boys were called upon to demonstrate on average 2.1 times and girls 1.8 times per lesson.

The distribution of the teacher's positive and corrective feedback was the focus of the second intervention. During baseline, the boys received on average 70.3% of the teacher's positive feedback and 63.1% of her corrective feedback, and the girls received 29.7% and 36.1%, respectively. Following intervention, the boys received 54.2% of the positive feedback statements and 54.6% of the corrective feedback, and the girls received 45.8% and 45.4%, respectively. Table 2 presents the average responses to boys, girls, and the whole class during baseline and intervention phases.

Table 2
Summary of Data for Subject A

<table>
<thead>
<tr>
<th>Variables</th>
<th>Baseline responses/lesson</th>
<th></th>
<th>Intervention responses/lesson</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Boys</td>
<td>Girls</td>
<td>Boys</td>
<td>Girls</td>
</tr>
<tr>
<td></td>
<td>Freq.</td>
<td>%</td>
<td>Freq.</td>
<td>%</td>
</tr>
<tr>
<td>Questions</td>
<td>6.4</td>
<td>61</td>
<td>4.1</td>
<td>39</td>
</tr>
<tr>
<td>Demonstrations</td>
<td>2.5</td>
<td>78</td>
<td>.7</td>
<td>12</td>
</tr>
<tr>
<td>Positive feedback</td>
<td>7.1</td>
<td>70</td>
<td>3.0</td>
<td>30</td>
</tr>
<tr>
<td>Corrective feedback</td>
<td>7.5</td>
<td>63</td>
<td>4.2</td>
<td>37</td>
</tr>
<tr>
<td>Praise</td>
<td>.6</td>
<td>68</td>
<td>.3</td>
<td>33</td>
</tr>
<tr>
<td>Desist</td>
<td>5.0</td>
<td>65</td>
<td>2.6</td>
<td>35</td>
</tr>
</tbody>
</table>

*Note. * = not intervened upon

Subject B

The collective data for Subject B were graphed, illustrating a multiple baseline design across behaviors (see Figure 2). Teacher questioning and use of demonstrations were intervened upon first. During baseline, Subject B on the average (per lesson) directed 69.6% of her questions to boys and 30.4% to girls, and called upon boys to demonstrate 2.6 times and girls to demonstrate 1.2 times. Following intervention, the teacher on average directed 53.2% of her questions to boys and 46.8% to girls. Boys were called upon to demonstrate an average of 1.1 times, and the girls 1.2 times per lesson.

The distribution of the teacher's desist and praise statements was the focus of the second intervention. During baseline, on the average per lesson, the boys received 69.1% of teacher desists and 64.3% of the praise statements, and the girls received 30.9% and 35.7%, respectively. Following intervention, the boys received 58.8% of the desist and 38% of the praise statements per lesson, and the girls received 43.2% and 62%, respectively. Table 3 presents the average responses to boys, girls, and the whole class during baseline and intervention phases.
Discussion

The purpose of the study was to determine the effects of an intervention on the distribution of teacher feedback, questions, and use of demonstrators to boys and girls in elementary physical education lessons. The results suggest that it is possible to modify teacher-student interaction patterns and establish a more sex-equitable gymnasium environment. The treatment effect in this study is shown on the graphs, not as an acceleration or deceleration of the frequency of specific variables but with the two data points coming together and tracking each other following intervention.
Table 3
Summary of Data for Subject B

<table>
<thead>
<tr>
<th>Variables</th>
<th>Baseline responses/lesson</th>
<th></th>
<th>Intervention responses/lesson</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Boys</td>
<td>Girls</td>
<td>Boys</td>
<td>Girls</td>
</tr>
<tr>
<td>Questions</td>
<td>3.2</td>
<td>70</td>
<td>1.4</td>
<td>30</td>
</tr>
<tr>
<td>Demonstrations</td>
<td>2.6</td>
<td>68</td>
<td>1.2</td>
<td>32</td>
</tr>
<tr>
<td>Praise</td>
<td>1.4</td>
<td>64</td>
<td>.7</td>
<td>36</td>
</tr>
<tr>
<td>Desist</td>
<td>13.2</td>
<td>69</td>
<td>5.8</td>
<td>31</td>
</tr>
<tr>
<td>Positive feedback</td>
<td>7.2</td>
<td>63</td>
<td>4.1</td>
<td>37</td>
</tr>
<tr>
<td>Corrective feedback</td>
<td>5.1</td>
<td>60</td>
<td>3.2</td>
<td>40</td>
</tr>
</tbody>
</table>

Note. * = not intervened upon

The baseline data supported many of the research findings in the regular classroom setting for older grade levels (Brophy & Good, 1974). Boys were questioned and asked to demonstrate more frequently than girls. It is unclear what sex role messages are conveyed to students through this differential assignment of academic tasks. The effect of this hidden sex role curriculum has received little attention in the physical education literature.

The baseline distribution of skill feedback (positive and corrective) to boys and girls concurred with Spaulding’s (1963) earlier findings in a classroom setting where boys received a greater number of teacher contacts of all types than did girls. Despite the increasing awareness of the importance of sexual equality and equal opportunity, the baseline data in this study tend to suggest that awareness does not automatically result in more sex-equitable classroom environments. The more opportunity for practice with feedback, follow-up, and support in the field subsequent to the intervention conference, the more likely that significant behavior change can be accomplished. This notion of “coaching” teaching behavior has been successfully documented in the staff development literature (Joyce & Showers, 1982; McLaughlin & Marsh, 1978). The intervention used in this study that followed the coaching principle was influential in establishing more equitable distributions of the four variables intervened upon for each subject. While positive and corrective feedback were not intervened upon with Subject B, Figure 2 shows a much more equitable distribution after lesson 9. It seems that the teacher generalized what she learned at the first intervention conference about her behavior with boys and girls (demonstrations and questions) to her distribution of corrective and positive feedback to boys and girls.

Generalization did not occur for Subject A until after the second intervention conference, when there was a reversal in the trend of desists to students. More data would be needed to draw any conclusions from this, however (see Figure 1). The generalization of the teachers’ equitable behavior to variables not directly intervened upon was a positive side effect of the study. Such inductive effects reflect a tremendous cost effectiveness, as some teacher behaviors were improved upon without direct intervention.

Teacher reaction to student conduct in baseline was very similar to earlier studies; very little teacher praise and high rates of teacher desists with boys desisted twice as fre-
quentiy as with girls (Brophy & Good, 1970; Good, Sikes, & Brophy, 1973; Leinhardt, Seewald, & Engel, 1979; Meyer & Thompson, 1956). During the intervention phase, these variables were intervened upon with Subject B. Following intervention there was an increase in the total number of praise statements directed to the students, from 2.18 statements per lesson in baseline to 12.5 per lesson during intervention. These statements became more equitably distributed also (see Table 3). Desistive behavior was very high and differentially applied by Subject B during baseline lessons, with 13.2 and 5.8 desists to boys and girls, respectively. Following intervention, the total number of desists decreased and the distribution was more equitably distributed, with 5.5 and 4.2 to boys and girls, respectively. The purpose was not to increase the desists to girls but to establish a more positive learning environment by encouraging an appropriate balance between teacher desist and praise responses to boys and girls.

The results of this study with grade 1 and 2 students indicate a major discrepancy in the distribution of feedback to boys and girls, with the boys consistently receiving a higher proportion of all teacher feedback. Jackson and Lahaderne (1967) emphasized the cumulative effects of even small differences. When calculated in yearly totals, differential feedback and opportunity to respond may have immense impact on students' behavior and attitudes toward physical activity. Brophy and Good (1974) argued that inappropriate or ineffective teaching is due in large part to teachers' lack of awareness of their teaching behavior patterns. In talking with the subjects, it was revealed that both had thought they had similar interaction patterns with boys and girls. The intervention package made the subjects aware of the types and distribution of their feedback to boys and girls. The intervention then provided an effective way of quantifying the teachers' behavior and, using graphic representations of the data, communicating this information to them. The daily feedback, reinforcement, and suggestions for teaching in a more equitable way were an important component of the intervention. It was not possible to tease out the differential effects of daily feedback and teacher-researcher conferences.

It would also be appropriate to look at how contextual differences (dance, games, gymnastics) affect teacher-student interaction patterns. Some classroom studies have suggested that differential treatment of boys and girls may be varied across subject areas and activities (Bossert, 1981; Doyle, 1977). The effects of student behavior on teachers' differential treatment patterns have received little attention in the general classroom or physical education research literature (Bossert, 1981). Clearly, boys and girls enter the gymnasium with different behavior patterns and attitudes (Griffin, 1984, 1985). The consequences of these sex-linked behavior differences in the physical education setting deserve further investigation.

References


