TEACHER FEEDBACK AND ACHIEVEMENT IN PHYSICAL EDUCATION: INTERACTION WITH STUDENT PRACTICE

STEPHEN SILVERMAN
University of Illinois at Urbana-Champaign, U.S.A.

LEAANN TYSON
Western Washington University, U.S.A.

JOHN KRAMPITZ
Concordia Lutheran College, U.S.A.

Abstract—The purpose of this study was to investigate the relationships between teacher feedback and achievement in physical education. Students \( N = 202 \) were pretested, received instruction, and posttested. An instrument was developed and validated to collect data on the amount and type of feedback received by individual students. Significant relationships were found for limited combinations where the type of feedback was positive, corrective, descriptive, or prescriptive. Various summed categories of prescriptive, descriptive, and corrective feedback showed significant part correlations with achievement when appropriate practice was a covariate. The results from this study suggest that relationships of feedback with achievement may be subject matter specific.

Teachers play many important roles in helping students in their classes learn. One teacher function is to monitor student progress and provide feedback. By providing appropriate feedback to the student, teachers hope that performance will be modified and the student will learn the material. Often teachers provide feedback, but do so based on intuition without systematic feedback plans.

Research provides some indication of the appropriate use of feedback for learning. Studies of learning in elementary and secondary school classrooms where the subject matter was mathematics, social studies, or language skills (Brophy & Good, 1986; Fisher, Berliner, Filby, Marliane, Cahen, & Dishaw, 1981; Good & Grouws, 1977) showed that specific, non-evaluative, task-relevant feedback is associated with increased student learning. Other authors (e.g., Rosenshine & Stevens, 1986) have made prescriptions based on this research, but have suggested that these prescriptions be used with caution since studies have occurred in only a few subject matter areas and the transfer of results to other subjects may not be valid.

Physical education is a subject matter in which transfer of results from classroom studies may be especially problematic. The objectives of physical education — learning of motor skills, kinesiological principles, and fitness — are different from those of many classroom subjects. In addition, during the learning of a motor...
skill students are actively engaged in movement that may complicate the teacher's task of providing feedback.

Based on the suggested importance of feedback in the learning process, one might conclude that ample research has been conducted in ecologically valid learning situations for all subject matters. Investigations of the role of feedback in the learning of a motor skill, however, have occurred almost exclusively in laboratory settings. Research on feedback, knowledge of performance, or knowledge of results (KR), the more common term in the laboratory motor behavior areas, typically occurs with one subject at a time learning a novel skill (often a fine motor skill) and with feedback provided by the experimenter or, more recently, by a computer (Magill, 1989; Schmidt, 1988). Although the results of laboratory studies (e.g., Magill & Wood, 1986; Salmoni, Schmidt, & Walter, 1984) generally support many findings of research that have focused on learning in classrooms, it also is inappropriate to extend these laboratory studies, where individual subjects are tested under highly controlled settings, to situations where students learn motor skill in much larger groups and under much less control.

While a few studies (de Knop, 1986; Eghan, 1988; Piéron, 1982) have had a component focus on teacher feedback and student achievement in gymnasia, the results have not provided any basis from which to draw conclusions. The most recent of these studies (Eghan, 1988) found a variety of nonsignificant relationships using two classes of university students learning a tennis skill. Although this might suggest no relationship exists between feedback and achievement in physical education, a study of time use and achievement (Silverman, Tyson, & Morford, 1988) found that practice time with feedback was strongly correlated with student learning.

Based on the collective results of research conducted in classrooms and motor learning laboratories and the relationships found in the Silverman et al. (1988) study, further research is needed to understand the role of teacher feedback as it relates to student achievement in physical education. Recently, prominent motor learning researchers (Christina, 1989; Magill, 1988) have emphasized the need to investigate the feedback—achievement relationship for learning sports skills. Therefore, if prescriptions provided to physical education teachers based on the results from classroom and motor learning research are to be valid for improving instruction, research must occur with a large subject pool in actual school settings.

The purpose of this study was to investigate the relationships between teacher feedback (type, form, time, referent, number of students, and quality of teacher feedback) and student achievement in physical education. In addition, since other studies have shown practice to be strongly related to achievement (Silverman, 1990; Silverman et al., 1988), the relationships of summed feedback variables with achievement were examined when appropriate student practice was a covariate.

Method

This study used a large data bank that included over 200 students in 10 middle school/junior high school physical education classes and their teachers. Students were pretested, received instruction, and then were posttested on two volleyball skills. Instruction was videotaped for subsequent coding of teacher and student behavior. A systematic observation system for collecting data on teacher feedback in physical education was developed and pilot tested for use in this study. Feedback data were used to determine the relationships of various feedback patterns to achievement. In addition, data from a previous study with the data base (Silverman, 1990) focusing on student practice trials and achievement was used so that practice could be partialed out and relationships reexamined.

Subjects

Seven physical education teachers (4 men, 3 women ranging from a first-year teacher to a teacher with 8 years of experience) and their students were the subjects for the study. The teachers in this study were the regularly assigned teachers and participated in an orientation session outside of normal class time. Teachers were presented the guidelines for the study and then performed the skills tests that the students subsequently completed. Teachers also
had the opportunity to ask questions, and the procedures relative to their participation were clarified.

Students in the sixth, seventh, and eighth grades were the student subjects. Each student was enrolled in physical education as a normal part of their overall educational experience. Two hundred and two students who completed both the pretest and posttest and attended at least five of the seven classes were included in the study. This unit of instruction occurred during the regular physical education class and the subject matter would have been taught as a regular part of the curriculum. Parental consent was obtained prior to student testing.

Pretesting and Posttesting

Students were pretested and posttested on both skills to be taught. Both tests have been previously shown to be reliable and valid for this age group (Carlisle, 1982) and, in pilot studies, were shown to be practical, in that they could be completed in one class session with junior high-aged children in the school gymnasium setting. The pretest for the students took place in the class session prior to instruction, and the posttest on the day following completion of instruction.

Serve test. The American Alliance for Health, Physical Education, Recreation, and Dance volleyball serve test (American Alliance for Health, Physical Education, Recreation, and Dance, 1965) was used to determine skill in serving. Students stood in a designated area behind the service line and had 10 serving trials. The court was marked so that the subject received from 0 to 4 points for each serve depending on where the ball landed. The total number of points was summed and used for analysis.

Forearm pass test. The Brumback forearm pass test (Carlisle, 1982) was used to determine skill on the forearm pass. The test consisted of passing the ball above an 8 foot line and against a wall as many times as possible in 1 minute.

Instruction

Instruction took place during regular physical education classes for seven class sessions over a 2-week period. Each class was approximately 30 minutes in length. Teachers had complete teaching control over their classes. In order to investigate relationships with student achievement, no intervention was used to change the teacher’s style or methods. Teachers taught the class as they would under normal circumstances.

Instruction during this 2-week period was videotaped using a two camera split screen set-up so that the entire area of the gymnasium was recorded for subsequent coding. The videotape equipment was in the corner of the gymnasium. One camera was focused on the far end of the gymnasium using a telephoto lens and a second camera with a wide angle lens was focused across the near end of the gym. A split screen generator coordinated the two video signals and the resultant image had the elapsed time superimposed. This combined signal was recorded on a portable VHS video recorder. A monitor was attached to the system to determine the quality of the recording during taping. Teachers wore a portable microphone (weight approximately 1/2 pound) and the audio signal was recorded simultaneously with the video signal.

During instruction each student wore a numbered pinafore so they could subsequently be identified on videotape. Pinafores were assigned to students and they wore the same one each day.

Observation of Teacher Feedback

Instrument development. An adaptation of the Fishman and Toby (1978) augmented feedback observation instrument was developed and pilot tested for this study. Scholars in the field of research on teaching (Brophy & Good, 1986; Rosenshine & Furst, 1973) have stressed the necessity of developing instrumentation specifically designed to answer the questions being asked through the study. Definitions and a decision log were developed for each category and subcategory of the system. Extensive piloting indicated that the adaptation was appropriate to answer the questions posed and that high levels of interobserver agreement (reliability) could be reached and maintained.

The system consisted of six categories each with subcategories. In addition, since it was possible that either the pass or serve could be
the focus of the feedback, the skill being addressed also was recorded. The feedback categories were: (a) type, (b) form, (c) time, (d) referent, (e) number of students, and (f) quality. The various categories and subcategories allowed for the analysis of whether the type of feedback (as opposed to just the quantity) is an important variable in student achievement. An overview of the categories and subcategories with definitions is provided in Figure 1. Ex-

**SKILL**

*Forearm Pass (P)*—forearm pass is the focus of feedback

*Serve (S)*—Serve is the focus of feedback

**TYPE**

*Positive (P)*—feedback that praises or reinforces a movement

*Negative (N)*—feedback that tells a student the movement is incorrect

*Neutral (Ne)*—feedback that doesn’t praise, encourage, criticize, or correct

*Descriptive (D)*—feedback that describes an error without evaluation

*Prescriptive (Pr)*—feedback that describes how to make performance better

*Corrective (C)*—feedback that is a combination of descriptive and prescriptive feedback (the error is described and a correction is provided in a “business-like” fashion

*Affective (A)*—feedback intended as motivation or to improve attitude

*Comparative (Co)*—feedback providing comparisons to another motor skill

**FORM**

*Auditory (A)*—feedback provided orally

*Visual (V)*—feedback provided by demonstration

*Tactile (T)*—feedback provided by manual assistance

*Auditory-Visual (AV)*—feedback that is both auditory and visual

*Auditory-Tactile (AT)*—feedback that is both auditory and tactile

*Auditory-Visual-Tactile (AVT)*—feedback that is auditory, visual, and tactile

**TIME**

*Concurrent (C)*—feedback provided during skill performance

*Terminal (T)*—feedback provided after skill performance, but before another trial has been attempted

*Delayed (D)*—feedback provided some time after skill performance and after another trial has been attempted

**REFERENT**

*Whole Skill (W)*—feedback on the multiple components of skill performance

*Part of Skill (P)*—feedback on a particular part of skill performance

*Outcome (O)*—feedback on the resulting outcome of the skill attempt

**NUMBER OF STUDENTS**

*One (1)*—feedback directed to one student

*Two (2)*—feedback directed to two students

*Three (3)*—feedback directed to three students

**QUALITY**

*Good (G)*—feedback that is appropriate to the situation without identifiable errors

*Moderate (M)*—feedback that generally is appropriate, but where the teacher misses a small component or makes a small error in the explanation or demonstration

*Poor (P)*—feedback with many errors

Figure 1. Feedback coding categories, subcategories, and definitions.
Instances of teacher feedback to individual students or to small groups of students were coded using the observation system. Data were collected on a specially prepared coding sheet. The coder observed the feedback and determined which of the subcategories were reflected for each category. For example, after watching the teacher give feedback to a student, the coder might record that the feedback was provided to student number 19, the skill was the pass, that it was corrective feedback (feedback identifying the error and providing the needed correction), the feedback was auditory and visual, it was after the skill attempt (terminal), the teacher referred to the whole skill, only one student received the feedback, and the feedback provided was good.

To determine if follow-up feedback was a significant predictor of student achievement, coders indicated on the coding sheet when the teacher provided follow-up feedback after a new skill attempt. For instance, follow-up feedback was indicated when the teacher provided feedback, the student reattempted the skill, and then the teacher provided additional feedback.

Training coders. Two of the investigators completed all coding. Extensive training occurred before actual data collection. Coders completed the following during training: (a) discussion of the instrument; (b) coding videotapes together to identify unambiguous feedback patterns; (c) coding videotapes together to identify ambiguous feedback patterns; (d) coding separately with discussion of disagreements; and (e) actual coding. The various parts of the training procedure were repeated, and actual data collection did not occur, until coders reached a .90 level of interobserver agreement for three classes in a row.

Interobserver agreement. Interobserver agreement was calculated with scored subcategory classifications and separately for summed category data using repeated measures ANOVA to obtain an intraclass correlation reliability coefficient (Winer, 1971). This method has been found to be appropriate and conservative for category data from systematic observation of teaching (Frick & Semmel, 1978). Both subcategory and summed data interobserver agreements coefficients were calculated because each level of data would be used in subsequent analysis. For research on teaching it is appropriate to calculate interobserver agreement/reliability for each level of data used in analysis (Evertson & Green, 1986).

During actual data collection, 10% of the class sessions were randomly selected and coded for interobserver agreement. All interobserver agreement checks, calculated for subcategory and summed data, were .90 or higher.

Data collection. Each session of each class was observed by a trained coder to collect feedback data. The coder observed the class, watching the teacher, and recording each instance of feedback provided during the class. Since instruction was on videotape, the coder had the option of stopping the tape to record information on the coding sheet. Coders also could rewind the tape to review part of the instruction before recording feedback data.

Observation of Student Practice

In previous studies with this data base, variables related to student practice have been found to have strong significant positive relationships with achievement (Silverman, 1990; Silverman et al., 1988). In addition, research conducted in field settings does not have the control that is possible in laboratory studies. Since practice is such a strong variable in predicting achievement, relationships over and above those found for practice variables were important to examine to fully understand feedback—achievement relationships. While it might be appropriate to examine other variables as covariates, individual student appropriate practice has been found to have the strongest relationships with achievement. Other covariates are likely to obfuscate the questions being asked here. A systematic observation instrument was designed and validated to provide data on the quality and quantity of student practice trials. An overview of the instrument and data collection are presented here. Readers are referred to the previous report (Silverman, 1990) for additional information.
Each practice trial by an individual student was categorized as either *appropriate* (a good practice trial) or *inappropriate* (a difficult or poor trial) for the following: (a) *serve without ball* — when the movement is practiced without a ball; (b) *serve with ball but without net* — when the student attempts the skill without attempting to get the ball over a net; (c) *serve with ball and net* — when the student attempts to get the ball over the net; (d) *forearm pass without ball* — when the movement is practiced without a ball; (e) *forearm pass with ball* — when the student attempts the skill by passing to him or herself or to another student; (f) *forearm pass with ball against wall* — when the movement is practiced against a wall similar to the testing conditions.

**Data collection.** Each student was observed in each of the seven class sessions. Due to the time-intensive nature of collecting data on each student in each class, data collection took approximately 900 hours over a 12-month period. Coders were trained with similar procedures to those used for the feedback instrument. Throughout the actual data collection 82 random observer agreement checks were performed. All interobserver agreement checks were .87 or higher with a mean of .93.

**Data Analysis**

**Achievement data.** For each skill, a residual score was calculated for each student as the measure of student achievement. For the entire sample a posttest on pretest regression equation was determined. Based on this equation predicted scores were calculated. The actual posttest score minus the predicted posttest score was the achievement score adjusted for entry skill level (the residual gain score). Therefore, if a student did better than predicted the student would have a positive residual score and if worse than predicted a negative residual score. Residual gain scores were selected because they partial out pretest skill level, are reliable, uncorrelated with entry skill, and not subject to ceiling effects.

**Feedback data.** For each student and each combination of skill, category, and subcategory data were summed for each class session. Unique subcategory feedback variables were created by summing skill, type, form, time, and referent combinations across the seven class sessions. In addition, since there were many subcategories, variables also were collapsed into logical units for subsequent analysis. For example, all feedback that was either of a good or moderate quality was summed into one subcategory and used in analysis. The computing of combined subcategories and the summing of unitary subcategories provided for maximum flexibility in determining statistical relationships.

**Practice trial data.** For each student, each category of practice trials was summed across the seven class sessions. Categories then were summed so the number of appropriate, inappropriate, and total trials for each skill could be used in statistical analysis. For this report, only the total number of summed appropriate trials for each skill was used as a covariate in analyses.

**Statistical analysis.** For each skill, the various summed units of feedback variables were correlated with individual residual achievement using the Pearson product-moment correlation technique. The procedure of correlating residual scores with other variables is similar to calculating part correlation coefficients. The technique used here was chosen because it is slightly more powerful than using increments of $R^2$ for significance in part correlation and for consistency of analysis among studies with this data base. Based on research hypotheses, a one-tailed test was used to determine significance.

In addition, statistical power analysis was completed for correlation coefficients using the procedures described by Kraemer and Thiemann (1987). Because of the ambiguity of previous results examining feedback—achievement relationships in physical education, the procedure was conducted to determine both the power and Type II error for the correlation coefficients.

As stated, relationships of feedback variables with achievement adjusting for appropriate practice also were of interest. Part (semi-partial) correlation coefficients were calculated.
Feedback and Achievement

with summed feedback for each subcategory of feedback for both skills. In addition, the subcategories of descriptive, prescriptive, and corrective were summed for all two-way and the three-way combinations and these variables and total feedback were used to calculate part correlation. Incremental and total $R^2$ were calculated for posttest on pretest, appropriate practice trials, and the feedback where significance was found.

Results

Descriptive Data

The typical student received feedback 28.39 [standard deviation $(SD) = 28.81$] times summed across the seven class sessions. This was distributed as 14.55 times $(SD = 14.57)$ while practicing the pass and 13.85 times $(SD = 14.90)$ while practicing the serve. The number of instances of feedback directed to individual students ranged from 0 to 237 with all but 23 students receiving feedback less than 50 times and 4 students receiving no feedback. Only 6 students received feedback 100 or more times.

When examining the feedback subcategories, most feedback was either positive or prescriptive, was auditory, occurred at the completion of the skill (terminal), and addressed either part of the skill or the whole skill. Virtually all feedback was coded as good. Table 1 provides means and standard deviations for each type, form, time, and referent variable.

As expected based on the number of combinations, the subcategory combinations (skill—type—form—time—referent) resulted in very

Table 1

<table>
<thead>
<tr>
<th></th>
<th>Pass</th>
<th>Serve</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>3.27 (3.99)</td>
<td>2.43 (4.15)</td>
</tr>
<tr>
<td>Negative</td>
<td>0.46 (1.14)</td>
<td>0.31 (1.08)</td>
</tr>
<tr>
<td>Neutral</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Descriptive</td>
<td>0.75 (1.34)</td>
<td>1.14 (1.73)</td>
</tr>
<tr>
<td>Prescriptive</td>
<td>3.57 (4.60)</td>
<td>1.92 (2.77)</td>
</tr>
<tr>
<td>Corrective</td>
<td>1.04 (1.40)</td>
<td>1.05 (1.66)</td>
</tr>
<tr>
<td>Affective</td>
<td>1.17 (1.81)</td>
<td>0.83 (1.71)</td>
</tr>
<tr>
<td>Comparative</td>
<td>0.00</td>
<td>0.04 (0.36)</td>
</tr>
<tr>
<td><strong>Form</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Auditory</td>
<td>7.88 (9.14)</td>
<td>5.20 (7.62)</td>
</tr>
<tr>
<td>Visual</td>
<td>0.00 (0.07)</td>
<td>0.00 (0.06)</td>
</tr>
<tr>
<td>Tactile</td>
<td>0.00</td>
<td>0.01 (0.14)</td>
</tr>
<tr>
<td>Auditory—visual</td>
<td>1.87 (2.35)</td>
<td>2.24 (3.81)</td>
</tr>
<tr>
<td>Auditory—tactile</td>
<td>0.40 (0.96)</td>
<td>0.16 (0.51)</td>
</tr>
<tr>
<td>Auditory—visual—tactile</td>
<td>0.08 (0.30)</td>
<td>0.10 (0.36)</td>
</tr>
<tr>
<td><strong>Time</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Concurrent</td>
<td>2.56 (3.94)</td>
<td>0.88 (1.96)</td>
</tr>
<tr>
<td>Terminal</td>
<td>7.33 (8.27)</td>
<td>6.63 (9.62)</td>
</tr>
<tr>
<td>Delayed</td>
<td>0.38 (0.98)</td>
<td>0.20 (0.67)</td>
</tr>
<tr>
<td><strong>Referent</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Whole skill</td>
<td>4.13 (5.74)</td>
<td>3.42 (6.58)</td>
</tr>
<tr>
<td>Part of skill</td>
<td>4.93 (6.27)</td>
<td>3.11 (4.38)</td>
</tr>
<tr>
<td>Outcome</td>
<td>1.20 (2.33)</td>
<td>1.17 (1.95)</td>
</tr>
</tbody>
</table>
low means for any specific combination. Most subcategory combinations were not coded for either skill. The same combination was the most frequently coded (positive-auditory-terminal-whole skill) for both skills. It occurred an average of 1.52 times ($SD = 2.70$) for the pass and 1.94 times ($SD = 3.54$) for the serve summed over the seven class sessions.

**Correlational Analysis**

*Power analysis.* The statistical power analysis indicated high power for the correlation coefficients calculated for this study. Power was greater than .85 for all correlations (therefore $\beta < .15$). Because of the high power, if correlations existed there was a high probability they would be found.

**Correlations with achievement.** For both skills, the summed total of all feedback received by students was not related to residualized student achievement. Based on the research hypotheses, for the summary categories only a few significant relationships were found. For the pass, feedback that focused on outcome as the referent was related to achievement ($r = .12, p < .05$). Three relationships were found for the serve. Feedback that was positive ($r = .15, p < .05$), occurred as auditory and tactile ($r = .19, p < .01$), or focused on the outcome of the skill ($r = .16, p < .05$) related to student achievement.

When the unique type-form-time-referent feedback combinations were correlated with achievement, relationships were found for both skills. Positive-auditory-concurrent-whole feedback was positively related to achievement for both skills. A few other combinations with positive feedback also were significant. When the feedback was coded as either descriptive, prescriptive, or corrective a number of correlations were found for subcategory combinations for both skills. In addition, two combinations of affective feedback were related to achievement for the serve. Table 2 provides an overview of significant correlations for combination feedback variables.

**Table 2**

<table>
<thead>
<tr>
<th>Feedback combination</th>
<th>Pass</th>
<th>Serve</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive/auditory/concurrent/whole</td>
<td>.19**</td>
<td>.23**</td>
</tr>
<tr>
<td>Positive/auditory/concurrent/part</td>
<td></td>
<td>.14*</td>
</tr>
<tr>
<td>Positive/auditory/concurrent/outcome</td>
<td>.13*</td>
<td></td>
</tr>
<tr>
<td>Positive/auditory/terminal/whole</td>
<td></td>
<td>.12*</td>
</tr>
<tr>
<td>Positive/auditory/terminal/outcome</td>
<td>.21**</td>
<td></td>
</tr>
<tr>
<td>Descriptive/auditory/concurrent/whole</td>
<td></td>
<td>.15*</td>
</tr>
<tr>
<td>Descriptive/auditory/concurrent/part</td>
<td>.12*</td>
<td>.14*</td>
</tr>
<tr>
<td>Descriptive/auditory/concurrent/outcome</td>
<td>.13*</td>
<td></td>
</tr>
<tr>
<td>Descriptive/auditory-tactile/terminal/outcome</td>
<td>.20**</td>
<td>.15*</td>
</tr>
<tr>
<td>Prescriptive/auditory/concurrent/outcome</td>
<td>.16**</td>
<td></td>
</tr>
<tr>
<td>Prescriptive/auditory/terminal/part</td>
<td>.12*</td>
<td></td>
</tr>
<tr>
<td>Prescriptive/auditory/delayed/outcome</td>
<td>.12*</td>
<td></td>
</tr>
<tr>
<td>Prescriptive/auditory-visual-tactile/terminal/whole</td>
<td>.18**</td>
<td></td>
</tr>
<tr>
<td>Prescriptive/auditory-visual/tactile/terminal/part</td>
<td>.14*</td>
<td></td>
</tr>
<tr>
<td>Corrective/auditory/terminal/part</td>
<td>.12*</td>
<td></td>
</tr>
<tr>
<td>Corrective/auditory/terminal/output</td>
<td>.13*</td>
<td></td>
</tr>
<tr>
<td>Corrective/auditory-visual/terminal/whole</td>
<td></td>
<td>.13*</td>
</tr>
<tr>
<td>Affective/auditory/concurrent/outcome</td>
<td>.16*</td>
<td></td>
</tr>
<tr>
<td>Affective/auditory-visual/tactile/terminal/outcome</td>
<td>.15*</td>
<td></td>
</tr>
</tbody>
</table>

*p < .05.

**p < .01.
Part correlations with achievement. Part correlations for individual summed subcategories were not significant. The part correlation of total feedback and achievement was significant for the pass ($r_{12.3} = .12, p < .05$), but not for the serve. For all summed combinations of descriptive, prescriptive, and corrective feedback, significant part correlations were found for the pass with the three-way combination having the highest value ($r_{12.3} = .24, p < .01$). A similar trend was found for the serve; however, when descriptive and corrective were summed the relationship was not significant. All categories with significant part correlations are presented in Table 3.

Analysis of incremental $R^2$ showed that for both skills nearly half the variance (pass = 44.84%; serve = 49.33%) in posttest scores could be attributed to entry level (pretest scores). When appropriate practice trials were added to the equation approximately 70% of the posttest variance was accounted for both skills (pass = 69.54%; serve = 71.78%). The addition of feedback variables to the equation resulted in small, but significant increases in $R^2$ for those variables where part correlations were significant. The incremental $R^2$ analysis for both skills is presented in Table 4.

Discussion

Descriptive Data

Most students received relatively little skill-related feedback (about four times each class) during instruction. Although the standard deviations indicate a lot of variability in the data, both overall means and individual combinations of feedback were low. It is difficult to compare these data with other studies (e.g., Eghan, 1988; Fishman & Toby, 1978) since the other studies were conducted with different or a variety of motor skills. It is interesting, however, that the Fishman and Toby (1978) study showed that much of the feedback was negative and the data from this study indicates the teachers gave little negative feedback. This discrepancy could be a consequence of differences in the coding system or that the teachers for this study were specifically selected because of their instructional focus on motor skill learning, while the teachers in the Fishman and Toby sample were part of an early, large descriptive data bank.

Correlations with Achievement

Given the emphasis placed on feedback in prescriptions of teaching, perhaps the most notable result of this study is that total feedback by itself did not relate to student achievement. This occurred even though the power of the analysis was very high. Although this may seem like an unlikely result, it is important to note that to make comparisons with classroom and motor learning research, these analyses focused on feedback in isolation of other process variables. A recent study, with a relatively small data base, also found that feedback in physical education classes did not relate to achievement (Eghan, 1988). A previous in-

Table 3

<table>
<thead>
<tr>
<th>Feedback combination</th>
<th>Pass</th>
<th>Serve</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total feedback</td>
<td>.12*</td>
<td>.01</td>
</tr>
<tr>
<td>Prescriptive + corrective</td>
<td>.13*</td>
<td>.13*</td>
</tr>
<tr>
<td>Descriptive + corrective</td>
<td>.12*</td>
<td>.10</td>
</tr>
<tr>
<td>Descriptive + prescriptive</td>
<td>.14*</td>
<td>.14*</td>
</tr>
<tr>
<td>Descriptive + prescriptive + corrective</td>
<td>.24**</td>
<td>.22**</td>
</tr>
</tbody>
</table>

*p < .05  
**p < .01
Investigation with this data base (Silverman et al., 1988) has shown, however, that class time spent practicing with the teacher providing feedback to students was positively related to achievement. As seen with the data here, it is likely that practice is the key variable and feedback serves to modify practice for the individual student. Although total feedback did not relate to achievement, a few categories of feedback did. The only category that related to achievement for both skills was feedback directed to the outcome. This result parallels motor learning research suggesting that knowledge of results (KR) is related to performance (Salmoni et al., 1984). It is perplexing, however, that this was the only category found for both skills, particularly since it is commonly suggested that feedback on form is as important as feedback on outcome (Siedentop, 1983). As a consequence of practice, intrinsic feedback occurs automatically from seeing the outcome. It is impossible to separate the outcome seen visually from teacher outcome feedback. It may be that the teacher feedback focuses the KR that is occurring and, perhaps, is more powerful.

It should be reemphasized that the achievement tests used here tested outcome. Student form in performing the skill was not tested. Time limitations in the school environment make this type of testing difficult, if not impossible, for a large number of students. Much of the feedback that occurred, however, focused on various aspects of form. Since the ultimate goal of improving form in this type of activity is to improve outcome, students whose form improves are likely to have better and more consistent outcomes.

When combination categories were correlated with achievement three trends emerged. First, for a small number of categories where the type of feedback was positive or affective, feedback related to achievement. This fits with many prescriptions that suggest positive feedback should be increased in physical education instruction (e.g., Siedentop, 1983). Perhaps the positive feedback motivates students to continue practice, permitting a greater number of total appropriate practice trials, instead of functioning to directly change skill performance. It also is possible that when students correctly perform the skill during trial and error that positive feedback focuses their practice.
Second, auditory feedback was a part of most combinations that related to achievement. This is logical since the form of the overwhelming majority of feedback was auditory.

A third trend that emerged was that categories with corrective feedback or its two subparts — descriptive and prescriptive feedback — often were related to achievement. This occurred for both skills and represented most significant subcategory combinations.

For both skills, combinations of descriptive, prescriptive, and corrective feedback were related to achievement when partialing out appropriate student practice. Again, feedback seemed to focus student practice. These were the only summed subcategories other than total feedback for the pass where significant part correlations were found. Total feedback for the pass probably was significant because most pass feedback occurred in these categories. In many ways the descriptive, prescriptive, and correlational feedback categories are similar to specific, nonevaluative, task-relevant feedback that previous research has shown relates to achievement in classrooms (Brophy & Good, 1986; Fisher et al., 1981; Good & Grouws, 1977; Rosenshine & Stevens, 1986). The part correlation and combination category results found in this study fit with those from classroom research and with the recommendations of physical education teaching methods textbook authors (Mosston & Ashworth, 1986; Siedentop, 1983).

An additional trend that emerged was that incremental increases of $R^2$ were significant, but small. The increases in $R^2$ were about 1%. This may cause some to discount how meaningful these feedback variables are in increasing achievement. It may be that, for motor skill learning, pretest and appropriate practice are the best predictors of posttest scores, but that other aspects of instruction may increase achievement despite a small window of variance in the posttest scores. In addition, if models such as those proposed by Rosenshine and Stevens (1986) are correct, feedback may work to modify practice and make it more appropriate for individual students.

The results of this study suggest that feedback focused on skill outcome or feedback that is corrective, descriptive, prescriptive, or positive may relate to achievement. The results, however, are not conclusive and support Magill’s (in press) contention that the direct role of augmented feedback in learning motor skill may be overstated. It probably is not appropriate for teacher educators and supervisors to provide prescriptions for teaching that focus exclusively on how to deliver feedback. The results of this study were that various process variables were interrelated and that perhaps a single variable should not be considered in isolation of other variables or factors. On this basis it seems that prescriptions for teaching should be based on the complex nature of the task, not on one unitary dimension.

References


Submitted 27 July 1991
Accepted 13 March 1992