

1-1997

The Validity of the Qualitative Measures of Teaching Performance Scale as a Measure of Teacher Effectiveness

J. Len Gusthart
University of Saskatchewan

Ivan M. Kelly
University of Saskatchewan

Judith E. Rink
University of South Carolina - Columbia, jrink@mailbox.sc.edu

Follow this and additional works at: https://scholarcommons.sc.edu/pedu_facpub



Part of the [Education Commons](#)

Publication Info

Published in *Journal of Teaching in Physical Education*, Volume 16, Issue 2, 1997, pages 196-210.

<http://journals.humankinetics.com/jtpe-contents>

© 1997 by Human Kinetics Publishers, Inc.

This Article is brought to you by the Physical Education, Department of at Scholar Commons. It has been accepted for inclusion in Faculty Publications by an authorized administrator of Scholar Commons. For more information, please contact digres@mailbox.sc.edu.

The Validity of the Qualitative Measures of Teaching Performance Scale as a Measure of Teacher Effectiveness

J. Len Gusthart, Ivan M. Kelly
University of Saskatchewan

Judith E. Rink
University of South Carolina

The purpose of this study was to determine the relationship between the Qualitative Measures of Teaching Performance Scale (QMTPS; Rink & Werner, 1989) and teacher effectiveness in producing student achievement. The QMTPS focuses primarily on variables related to teacher clarity and task presentation. Nine middle school generalist (classroom) teachers were asked to teach the volleyball forearm pass and serve over eight lessons as part of their normal curriculum. Students were pre- and posttested on the serve and forearm pass using the American Alliance for Health, Physical Education and Recreation (1969) volleyball tests. All lessons were videotaped and were coded using the QMTPS instrument. The relationship between the QMTPS total score and student achievement was significant for the forearm pass and for the serve. The authors concluded that the QMTPS was a valid measure of teacher effectiveness when the total QMTPS score for several lessons was used.

A variety of observational systems have been used in physical education in an attempt to understand the instructional process. Some observation systems have been designed to focus on the students, other systems to concentrate on the teacher, and others to examine both the student and the teacher in the instructional setting. Whatever the observation system, the link between the focus of the observation and student learning has contributed to our understanding of teaching and learning in physical education. For example, if discrete categories of an observation instrument are related to student achievement, then an assumption is made that teachers who demonstrate greater proficiency in these areas are more effective in producing student achievement. In this fashion, observation systems have been utilized as a proxy measure for student achievement and have become part of the literature informing teacher education (Rink, 1993; Siedentop, 1991).

In order to use observation systems with confidence to study, assess, or inform teaching, observation systems must be both valid and reliable. Validity in the context of an observation tool for teaching has been most associated with establishing a relationship between teaching performance and some measure of student achievement or learning. In physical education, validity studies of an observation instrument have been most associated with measures of student achievement

J. Len Gusthart and Ivan M. Kelly are with the College of Physical Education at the University of Saskatchewan, Saskatoon, SK Canada S7N 5C2. Judith E. Rink is with the Department of Physical Education at the University of South Carolina, Columbia, SC 29208.

(Silverman & Buschner, 1990; Silverman, Devillier, & Ramirez, 1991). Schempp (1987) and Metzler (1989) have identified the two most frequently deployed observational instruments in the study of teaching physical education as the Cheffers's Adaptation of Flander's Interaction Analysis System (CAFIAS; Cheffers, Mancini, & Martinek, 1980) and Academic Learning Time-Physical Education (ALT-PE; Metzler, 1979).

Silverman and his colleagues have recently conducted validity studies on both the CAFIAS and ALT-PE observation systems to determine their validity as process estimates of student achievement (Silverman & Buschner, 1990; Silverman et al., 1991). Both studies were conducted in a similar contextual situation and had similar achievement scores. Ten middle school classes were pretested, received instruction, and posttested on the volleyball serve and forearm pass. Achievement scores were calculated by posttest on pretest regression for each skill, with residual scores used for subsequent analysis as described by Silverman (1988). The classes were videotaped and analyzed with the CAFIAS or ALT-PE systems. Residual achievement scores were then correlated with the observation systems. The initial study of CAFIAS correlated CAFIAS raw categories and chains of behavior with achievement in the volleyball forearm pass and serve. For the forearm pass, no raw categories were related to achievement. For the serve explanation/demonstration and predictable student response were related to achievement. The conclusion was that the CAFIAS may not be valid in all situations and that to use the instrument in a universal situation may not be appropriate (Silverman & Buschner, 1990).

The validation study conducted with ALT-PE concluded that while some categories of the instrument such as intervals of skill practice were related to achievement, others such as total number of motor appropriate intervals were not significantly related to achievement. In general the authors concluded that the "validity of the ALT-PE system as a process measure of achievement can be partially substantiated" (Silverman et al., 1991, p. 319).

The results of the validation studies on CAFIAS and ALT-PE indicate that some constructs of the instruments were related to student achievement, and others were not. At the very least, further investigation is required into the validity of applying systematic observation systems in physical education.

A more recent observation system, Qualitative Measures of Teaching Performance Scale (QMTPS) has been developed by Rink and Werner (1989) to describe characteristics of teacher clarity and task presentation. The following describes the categories included:

1. *Clarity of task presentation*: Teacher's verbal explanation/directions communicated a clear idea of what to do and how to do it. This judgment is confirmed on the basis of student movement response to the presentation, and is relative to the situation.
 - Yes: Students proceeded to work in a focused way on what the teacher asked them to do.
 - No: Students exhibited confusion, questions, off-task behavior, or lack of intent to deal with the specifics of the task.
2. *Demonstration*: Visual information modeling desired performance executed by teacher, student(s), visual aids, or a combination of these.
 - Yes: Full model of the desired movement.

- Partial: Incomplete model of task performance exhibiting only part of the desired movement.
 - No: No attempt to model the movement task.
3. *Appropriate number of cues*: The degree to which the teacher presented sufficient information useful to the performance about the movement task without overloading the learner.
 - Appropriate: Three or fewer new learning cues related to the performance of the movement task.
 - Inappropriate: More than three new learning cues related to the performance of the movement, or none given when needed.
 - None Given: No attempt at providing learning cues was given.
 4. *Accuracy of cues*: The degree to which the information presented was technically correct and reflected accurate mechanical principles.
 - Accurate: All information present was correct.
 - Inaccurate: One or more instances of incorrect information.
 - None Given: No cues given.
 5. *Qualitative cues provided*: Verbal information provided to the learner on the process or mechanics of movement.
 - Yes: Teacher's explanation or direction included at least one aspect of the process of performance.
 - No: Teacher's explanation or direction included no information on the process of performance.
 6. *Appropriateness of student response*: The degree to which student responses reflected an intent to perform the task as stated by the teacher.
 - All: No more than two students viewed on the screen exhibited inappropriate responses.
 - Partial: Three or more students viewed on the screen exhibited inappropriate behavior.
 - None: No students exhibited appropriate behavior.
 7. *Specific congruent feedback*: The degree to which teacher feedback during activity was congruent (matched) the focus of the task.
 - Yes: More than two instances were evident of teacher feedback being congruent with the task.
 - Partial: One or two instances of congruent feedback were evident.
 - No: No congruent feedback was given.

Like the CAFIAS and ALT-PE instruments, the QMTPS has also been used to study instructional variables in relationship to student learning (Gusthart & Kelly, 1993; Gusthart & Sprigings, 1989; Werner & Rink, 1989). Both the Werner and Rink (1989) and Gusthart and Sprigings (1993) studies described student learning in relation to the variables of the QMTPS. Although direct relationships between the variables of the QMTPS and improvement in jumping and landing skills were not identified, the descriptive data suggest the possibility that student gains might very well be related to the quality of the teacher's task presentation.

The Gusthart and Kelly (1993) study prompted an investigation to determine if student learning in volleyball motor skill acquisition was related to teachers' instructional actions as described in QMTPS. The QMTPS was designed primarily to describe the quality of the teacher's task presentations. Task presentation is viewed as a teaching function having as its role to communicate to learners

what they are to do and how they are to do it. In this perspective, the individual and discrete behaviors of each category are only important as they contribute to the description of whether the function (communication in this case), was adequately performed (Rosenshine & Stevens, 1986). The instrument asks the question, "Do learners have a clear idea of what to do and how to do it when they are sent off to work on a task by the teacher?" Each of the categories describe an aspect of the teacher's ability to perform this function, the extent to which may vary in different contexts.

The categories of the QMTPS are theoretically based in both motor learning theory and research on the role of cognition in motor skill acquisition and classroom research in teacher clarity. Motor learning theorists have long recognized the role of cognition in learning motor skills (Lee, Swinnen, & Serrien, 1994; Magill, 1993; Schmidt, 1989). Fitts and Posner (1967) labeled the first stage of learning a motor skill the *cognitive stage*. It is in the cognitive stage that the learner formulates a motor plan for movement. The accuracy and appropriateness of the plan is affected by the quality of the information the learner receives on how to perform a skill. The QMTPS instrument includes several dimensions of the use of demonstration, the use of cues, and the use of a focused type of feedback. The use of cues for performance and the use of demonstration have been associated with higher levels of motor skill acquisition throughout the motor learning and pedagogy literature (Fairweather & Sidaway, 1994; Landin, 1994; Rink, 1994).

The instrument also includes a category describing the use of congruent feedback, which is probably the most controversial category in the instrument because of its conditional support in the motor learning and the pedagogy literature (Magill, 1994; Silverman, Tyson, & Krampitz, 1991). This specific category of the instrument is designed to identify when a teacher uses feedback that specifically reinforces the information given in the task presentation. This type of feedback to students creates accountability for the information used in the task presentation and keeps the learners and the teacher more narrowly focused (Rink, 1993). Although few strong relationships between motor skill acquisition and the use of teacher feedback have been established in the pedagogy literature, it is probably too early to remove the use of teacher feedback as an important teacher function.

Classroom research provides additional support for ways in which teachers can facilitate learning through the manner in which information is delivered to the learner (Brophy & Good, 1986; Doyle, 1990; Kennedy, Cruickshank, Bush, & Meyers, 1978). The categories of the QMTPS have been designed to be consistent with both the general information regarding teacher clarity, as well as the more specific information related to information processing.

Although QMTPS has been used in several studies, there has been no effort to determine the relationship between the data from the instrument and student achievement in a physical education setting. The purpose of the current study, therefore, was to investigate the relationship between the total score on the QMTPS instrument and student achievement in psychomotor skill development.

Method

Participants

Students ($N = 222$; 128 boys and 94 girls) from nine regular coeducational classes in nine local public schools and their normal teachers ($N = 9$) served as

participants for this study. The participants were eighth-grade students with class size ranging from 24–32. The teachers were selected by the Research and Development Division of the local school board with no researcher input or criteria. The policy of the school system was to share research projects equitably among the various schools. Nine teachers were nominated and contacted by the Superintendent of Instruction based upon no criteria related to the study. In accordance with local school board policy, informed consent was obtained. All teachers were certified and had a variety of experience and preparation. Teaching experience ranged from one to 16 years. Eight of the teachers were male and 1 was female. To facilitate data collection, 6 teachers and their classes participated in the study in Year 1, and the remaining 3 teachers and classes in the subsequent year.

Context

The teachers were generalist teachers in the Canadian system who were responsible for teaching several academic subjects. The school district supports the position that instruction in physical education should be provided by the generalist teacher up to the eighth-grade level. Specialists who teach only physical education are not supported.

In this school district, detailed lesson plans (Saskatoon Board of Education, 1988) are provided for teachers for each unit in the physical education curriculum. The plans for the volleyball unit at the eighth-grade level included objectives for the serve, forearm pass, and overhand pass. Specific learning activities and organizational arrangements were suggested and reflected a direct instruction approach and an expectation for highly active teaching. The plans were structured to provide detail on the warm-up, skill development and conclusion for each lesson. However, teachers were free to select additional learning experiences.

Each of the eight sessions was taught during the regular physical education period, which amounted to 150 minutes per week of instruction. One additional class period prior to instruction and one additional class period after instruction were used for the pre- and posttest. All teachers had adequate equipment. One teacher had a small space not considered adequate for the lesson objectives. All lessons (72) were videotaped using a camcorder positioned so that the total class could be seen. Teachers wore a cordless microphone to assure that all teacher verbal behavior could be heard.

Teachers were asked to concentrate and focus on the serve and forearm pass over eight lessons to maximize student improvement in the two skills. The proportion of time allocated to each skill was left to the discretion of the teachers. Teachers tended to follow the general format of the lesson plans provided and conducted their own warm up system and motivational sets for all lessons.

Instrumentation

Students were pretested and posttested on each of two volleyball skills. The criteria for selecting these items were that the skills were used in previous studies with this grade level and that they were part of the district curriculum. The American Alliance for Health, Physical Education and Recreation (AAHPER, 1969) test for the volleyball serve and forearm pass were used to measure student achievement to facilitate data comparisons with previous studies using these skills (French, Rink, & Werner, 1990; French et al., 1991; Gusthart, Kelly, & Graham, 1994;

Rink, French, Werner, Lynn, & Mays, 1992). All testing was conducted by a team of four research assistants who were trained in the test protocol by the skill test coordinator. The coordinator (the men's varsity volleyball coach at the university) was present at all testing sessions.

Serve Test. The purpose of the serve test was to measure accuracy of execution for the underhand serve. A regulation volleyball court was marked so the student would receive 0–4 points for placing the ball in assigned areas. The student stood outside the serving line and in any legal court position. Two practice and 10 test trials were given. Total points were cumulative, with the maximum score of 40 points.

Forearm Pass. The purpose of the forearm pass test was to measure the student's skill in passing a legal forearm pass from the rear of the court toward the net. The student began by standing within a designated "passing area." A trained and skilled tosser from the test team was positioned at center court on the attack line. The tosser used a two-hand underhand toss to lob the ball so it would drop in the passing area. The student was to contact the ball with a forearm pass so it would pass over a rope elevated to 8 feet and drop into a target area. Tosses outside the "passing area" were repeated. Ten attempts were provided from the right side and 10 from the left. The total score was the sum of cumulative hits on target with a maximum score of 20. Two practice trials were provided from each side.

QMTPS

The QMTPS was used to collect instructional data from the videotapes. The instrument is divided into four sections: type of task, task presentation, student response appropriate to task focus, and specific congruent feedback. The first section (type of task) describes the nature of the movement task in terms of its contribution to content development (informing, refining, extending, and applying; Rink, 1993). This section is descriptive and, as such, is not used in the overall QMTPS score and was not used for this study. The remaining sections of the instrument (task presentation, appropriateness of student response, and specific congruent feedback) were used in the calculation of the overall QMTPS score used in this study. Seven categories describe these sections of the instrument.

The unit of analysis for the instrument was the individual movement task. After the teacher delivered each movement task, the observer stopped the video tape as students initiated practice of the task and made a judgment regarding the task presentation categories of the instrument. The tape was continued to observe the full practice of the task in order to code the appropriateness of the student responses and the specific congruent feedback categories. All coding procedures identified in the original instrument were followed.

Data were recorded in number of occurrences and converted to percentages of total tasks on a lesson (e.g., what percentage of the tasks delivered were delivered with a full demonstration?). Percentages are used because the number of total tasks delivered by a teacher varies from lesson to lesson and from teacher to teacher. The scoring system for each construct of the instrument was designed so that one category was more desirable. The most desirable categories reflected the knowledge base in the literature. That is, teacher clarity is associated with students working in a focused way rather than exhibiting confusion. Each teacher was given a percentage score for each category reflecting the percentage of tasks delivered with a score in the highest category. Percentages of responses in the most desirable

categories were summed and averaged across categories for a total QMTPS score representing the instructional function of task presentation.

Data Collection

All QMTPS data were collected by viewing the 72 videotapes. One investigator and a graduate assistant coded the videotapes according to the suggested QMTPS system (Rink & Werner, 1989). The procedure continued until the coders established interobserver agreement of .93 for all categories of the instrument on two successive lessons. Once actual coding of the instructional lessons commenced, observer drift between the coders was checked on a random sample of eight lessons (one for each of the teachers). For these lessons and all categories, the interobserver agreement was in excess of .90.

Data Analysis and Results

Three different sets of data were collected and analyzed. First, achievement data were analyzed to determine the significance of achievement gains from pre to posttest. Second, the achievement data were correlated with the QMTPS score of the teachers. Third, a descriptive analysis of the QMTPS data was conducted to more fully describe the performance of the individual teachers across lessons and in relation to each other.

Achievement Data

An initial analysis was conducted to determine if improved performance occurred over the unit. Residual performance scores were calculated for each student on the two motor skills (Table 1). These residual gain scores were based on the entire sample ($N = 222$ students) and obtained by regressing posttest scores on pretest scores for each of the two skills. This residual gain score analysis was suggested by Silverman (1988) and Silverman and Buschner (1990) because such scores control for beginning (entry) skill, are reliable, and are not subject to ceiling effects. Positive residual scores indicate that a student (or class) did better than predicted, based on the regression equation. Conversely, negative residual scores indicate that the student did worse than predicted by their pretest scores. To test change in performance for all 222 students combined across classes, a dependent t test was calculated on pretest and posttest scores for both skills. For the serve, $t(221) = 7.55$, $p < .001$, and the forearm pass, $t(221) = 5.02$, $p < .001$, students made significant gains from pretest to posttest.

Correlation With Achievement. For a second analysis, the mean residual score for each class for both skills was calculated to provide an overall change in performance measure from pretest to posttest. The class mean residual score was chosen as a measure not only because it has been used in previous studies as a measure of learning of volleyball skills (e.g., Silverman, 1988; Silverman & Buschner, 1990) but also because the mean of a group is a more stable and reliable indicator of performance than scores of individual students. Pearson product-moment correlations were calculated between the class mean residual score of each skill and the total QMTPS score for the teachers. The class was selected as the unit of analysis because the QMTPS score is obtained for the teacher on class performance. Previous research (Werner &

Table 1 Mean Pretests, Posttests, and Residuals on Pretest R^2 for Motor Skills

Class	<i>n</i>	Serve				Forearm pass			
		Pretest	Posttest	R^2	Residual	Pretest	Posttest	R^2	Residual
1	26	21.31	26.04	.39	+3.03	7.73	9.19	.19	+1.48
2	26	15.04	20.96	.13	+2.83	4.12	6.31	.11	+0.83
3	23	12.61	15.87	.59	-0.36	4.13	6.61	.45	+1.12
4	25	20.00	21.96	.45	-0.0	3.08	4.76	.11	-0.08
5	19	8.63	11.84	.63	-1.29	3.63	4.74	.69	-0.44
6	28	19.61	22.75	.40	+1.06	6.75	6.29	.37	-0.82
7	26	13.54	14.62	.55	-2.34	4.69	4.96	.21	-0.87
8	19	16.11	18.74	.21	-.022	4.73	5.89	.15	+0.03
9	30	9.8	11.27	.55	-2.78	3.83	4.30	.17	-1.01
Total	222	15.32	18.34	.52	0.00	4.80	5.90	.31	0.00

Rink, 1989) had indicated that it was the total QMTPS score over a period of instruction that was critical in discriminating the effective teacher, not necessarily the use of single components in isolation.

This approach is also consistent with the notion of teaching functions rather than discrete behaviors being critical discriminators of teacher effectiveness (Berliner, 1987). The p values accompanying each correlation are one-tailed since positive associations were expected between each pair of variables. These expectations were confirmed. Alpha level was set at a more relaxed level of .10. Justification of this higher level was based upon the field-based nature and uncontrolled aspects of instruction. Given the small sample size ($N = 9$ classes), the power of the statistical test will be low even for large effects (Cohen, 1988). The correlation between teacher QMTPS score and class mean residual score for the pass ($r = .77$) was significant ($p < .008$). The correlation between teacher total QMTPS score and class mean residual score for the serve ($r = .50$) was significant ($p < .08$). The correlation between mean class residual scores for the serve and pass ($r = .73$) was also significant ($p < .01$).

Descriptive Data. The scores each teacher received for each category of the QMTPS, as well as individual scores for each lesson were reviewed to determine patterns important to understanding the relationship between individual QMTPS scores and the total QMTPS score used in this analysis. Mean and standard deviations for each teacher for each category across the lessons are reported in Table 2.

Teachers were ranked with each other for each category based on their score for that category of the instrument (Table 3). The three teachers who were the most effective in improving performance were also ranked in the top three for most individual categories of the QMTPS. Teacher 1, for instance, scored in the top three in five of the seven categories. Teacher 2 was ranked in the top three for six of the seven categories. Similarly, those teachers who did not produce as much achievement as expected were ranked lower for most categories. The two least effective teachers scored in the top three for only one and two of the categories. However, each teacher who was considered most effective had one or two categories that might be considered a "weakness." Likewise, each teacher who was not as effective in this study had one or two categories of strength. The strong but inexact relationship between each of the categories is support for using the total score of the instrument.

A descriptive analysis of variability across lessons for a particular category and a particular teacher was also done using the standard deviation across lessons (Table 2) and reviewing individual lessons for each teacher. Most teachers who were strong in a category would have one or two lessons in which they would score low, and most teachers who were weak in a category would have one or two lessons in which they scored considerably higher. Often that variability could be explained by specific events occurring on a particular day such as teachers not demonstrating because they wore a suit to school and did not change.

Teacher 3 in this study achieved a great deal in one skill, but not in the other. That teacher tended to be an exception, since most teachers weak in one skill were also weak in the other or vice versa. Teacher 3 also had the highest total QMTPS score. Upon further investigation of the context in which the teaching occurred, it became clear that this teacher did not have the space to teach the serve effectively using a full court. It is possible that, had this teacher had the facilities to teach the serve effectively, this teacher would have been effective across both skills. Because

Table 2 Mean Percentage and Standard Deviations for QMTPS

	Teacher 1		Teacher 2		Teacher 3		Teacher 4		Teacher 5		Teacher 6		Teacher 7		Teacher 8		Teacher 9	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Clarity	96	7.51	89	7.42	88	13.07	90	6.59	89	12.07	100	0.00	61	14.00	83	17.32	98	4.63
Demonstration	52	23.32	32	16.36	38	23.03	13	12.67	33	29.83	12	13.04	36	18.38	42	19.36	14	16.37
Appropait number of cues	36	10.14	48	13.56	44	14.27	34	12.97	41	14.31	35	21.72	20	25.00	15	10.67	22	8.96
Accuracy of cues	58	13.92	59	17.45	81	25.19	55	26.38	77	17.10	46	15.63	53	20.78	34	19.54	33	20.78
Qualitative cues provided	34	22.06	45	21.18	75	22.64	54	19.56	70	17.09	27	14.43	23	19.41	41	12.76	30	26.91
Appropriate student response	98	4.63	85	8.62	68	13.19	61	18.48	64	15.21	98	4.63	59	25.49	61	19.48	98	4.63
Specific cognitive feedback	65	10.89	81	12.07	83	15.37	77	19.74	58	6.73	54	14.56	63	20.45	72	16.79	70	9.98
Total QMTPS	63	7.62	65	8.31	70	10.87	55	9.07	62	10.26	53	5.72	45	10.16	50	9.16	53	9.04

Table 3 Teacher Rank on QMTPS Variables

QMTPS variable	Teacher								
	1	2	3	4	5	6	7	8	9
Clarity	3	5	7	4	6	1	9	8	2
Demonstration	1	6	3	9	5	8	4	2	1
Appropriate number of cues	4	1	2	6	3	5	8	9	7
Accuracy of cues	3	4	1	5	2	7	6	8	9
Qualitative cues provided	6	4	1	3	2	8	9	5	7
Appropriate student response	1	4	5	7	6	2	9	8	3
Specific congruent feedback	6	2	1	3	8	9	7	4	5
Total QMTPS score	3	2	1	5	4	6	9	8	7

this teacher did score so high on the QMTPS, the relationship between the QMTPS and the serve would also have been stronger.

No statistical analysis was done on the separate categories of the QMTPS. The size of the n ($N = 9$) and the interrelationships between these categories would have made such an analysis inappropriate. The relationships between the categories and the relationships between the categories and the whole QMTPS are better understood when each of the categories is considered. Each category will be described in terms of the range and the relationship between the category and the effectiveness of the teachers.

The teacher clarity category had a range of 61–98%. Actually, all of the teachers scored above 83% in this category with the exception of one teacher whose score was 63%. This would seem to indicate that this particular category was not discriminating for most of these teachers. The teacher who did score at the lowest level was ranked 8th out of the 9 teachers in terms of student achievement for both skills.

The teacher demonstration category had a range of 12–52%, which means that most of the teachers presented tasks to learners without a full demonstration. Although 2 of the teachers who achieved the most in terms of student learning were ranked in the top 3 of this category, one was not. Although 2 of the 3 teachers who achieved the least were ranked in the bottom three of this category, 1 was not.

The category number of appropriate cues ranged from 14 to 48% for the 9 teachers. Again, teachers who produced the most achievement tended to be ranked at the top for this category, and teachers who produced the least achievement were ranked near the bottom for this category. There was not a direct relationship however.

Scores for the accuracy category ranged from 33 to 81%. This category seemed to have the same relationship to achievement as the categories cited above, with one exception. One of the least effective teachers scored 77% in this category. This teacher scored low in appropriate number of cues, as well as use of qualitative cues. The preceding most likely indicates that the teacher was accurate because little information was presented to the learners on how to do a task.

The use of qualitative cues ranged from 23 to 75%, one of the largest ranges for any of the categories. This category seemed to be the least predictive of student achievement of all of the categories, which may mean that the use of cues is important but the qualitative nature of those cues may not have to be identified.

Appropriate student responses was another category of the QMTPS that was mostly predictive of student achievement but had exceptions at either end. The range for this category was 59–98%. Five teachers had scores below 70%, and 3 teachers had scores of 98%. One of the teachers who scored 98% was ranked last in both skills in terms of student achievement.

The range for teacher congruent feedback was 54–83%, with many of the most effective teachers having the highest scores, but the least effective teachers not necessarily having the lowest scores.

Discussion

The results of this study support the use of the total score on the QMTPS as a proxy measure of teacher effectiveness for the two skills studied. The correlation between the total QMTPS score and the volleyball forearm pass and serve was significant. Several factors related to this data need discussion.

The context in which the study was conducted facilitated the investigation of task presentation as a holistic concept related to teacher effectiveness. All teachers had as their objective the improvement of the two skills studied and were accustomed to teaching to objectives. The teachers in this study were generalist teachers in the Canadian system who were also responsible for teaching other academic subjects. There was a high expectation for teachers to do a good job of teaching, particularly in volleyball, which was a highly valued subject in the curriculum. Curriculum and instructional lesson plans, which the teachers were expected to use, were provided to teachers as part of district policy. The lesson plans contained objectives and rather explicit help in conducting the learning experiences of the unit. All groups in this study improved from the pretest to the posttest over the eight lessons except for one teacher in one skill. The data set as a whole reflects a range of effective teachers, but did not include ineffective teachers. It is unclear how the results of this study would have been affected with less effective teachers and less focused teaching.

All teachers approached the content using direct instruction. The categories of the QMTPS are intuitively appropriate for content approached in this manner. However it is also likely that clarity in teaching is generic to all teaching styles. Previous research using the QMTPS suggests that this might be so (Werner & Rink, 1989). Continued investigation of the relationship of the categories of the instrument with teacher effectiveness using different approaches to content is certainly warranted.

Performance (measured in pre- to posttest residuals) on the forearm pass and the serve were highly correlated in this study ($r = .73, p < .01$). Given the differences between these skills and the lack of control evident in this setting, differences are to be expected. It is also likely that the fact that the teacher who had the highest score on the QMTPS did very well in producing student learning in the set but not in the serve played a role in these differences. As previously mentioned, this teacher did not have the facilities to teach the serve effectively. Both of these skills, as tested, required force production abilities and most likely have more in

common than they do differences when practiced to this level. Had instruction in the forearm pass had as its objective a more open use of the skill, other factors than just task presentation, such as content development, would probably have had more influence on the amount of student achievement in the forearm pass.

The results of this study support the use of the total QMTPS score over several lessons and not any single category of the instrument or any single lesson as a measure of achievement. More effective teachers have weak categories and less effective teachers have strong categories, but more effective teachers have higher total scores and less effective teachers have lower total scores. The total score looks more at the total function of presenting information to learners rather than any single behavior.

This research is consistent with classroom research and other teacher effectiveness studies done in physical education, as well as motor learning studies that have strongly supported teaching variables related to task presentation, teacher clarity, the use of demonstration, and the role of feedback in learning (see Rink, 1994; Silverman, 1991). Likewise, the individual teacher profiles created by the instrument graphically demonstrate the importance of looking at teaching behavior in terms of the function it serves rather than in terms discrete behaviors (Rosenshine & Stevens, 1986).

The relationship established in this study was between the variables of the QMTPS and two volleyball skill tests. A retention test was not done in this study, limiting our understanding of the relationship between student achievement and student learning. Our knowledge of the ability of the subjects to transfer the use of the motor skills tested to a game situation is also limited by the nature of the validity of the skill tests used.

No single teacher behavior or function can represent the total teaching act. Teaching is more complex than that. Knowledge about individual functions however can make the total teaching act more effective. The QMTPS is yet another important piece of information that should help teachers be more effective and should provide direction for teacher education. Additional research on context should help us understand even more the relationship between these categories and teacher effectiveness in different settings, with different content, and with different approaches to content.

References

- American Alliance for Health, Physical Education and Recreation. (1969). *Volleyball skills test manual*. Washington, DC: Author.
- Berliner, D. (1987). Simple views of classroom teaching and a simple theory of classroom instruction. In D. Berliner & B. Rosenshine (Eds.), *Talks to teachers* (pp. 93-110). New York: Random House.
- Brophy, J., & Good, T. (1986). Teacher behavior and student achievement. In M. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 328-375). New York: Macmillan.
- Cheffers, J., Mancini, V., & Martinek, T. (1980). *Interaction analysis: An application to nonverbal activity* (2nd ed.). St. Paul: Association for Productive Teaching.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.) Hillsdale, NJ: Erlbaum.
- Doyle, W. (1990). Themes in teacher education research. In W.R. Houston (Ed.), *Handbook of research on teacher education* (3rd ed., pp. 3-24). New York: Macmillan.

- Fairweather, M., & Sidaway, B. (1994). Implications of hemispheric function for the effective teaching of motor skills. *Quest*, **46**, 281-298.
- Fitts, P.M., & Posner, M. I. (1967). *Human performance*. Belmont, CA: Brooks/Cole.
- French, K., Rink, J., Rickard, L., Mays, A., Lynn, S., & Werner, P. (1991). The effects of practice progressions on learning two volleyball skills. *Journal of Teaching in Physical Education*, **10**, 261-274.
- French, K., Rink, J., & Werner, P. (1990). Effects of contextual interference on retention of three volleyball skills. *Perceptual and Motor Skills*, **71**, 179-186.
- Gusthart, J.L., & Kelly, I.W. (1993). Teachers' instructional variables in volleyball and students' improvement in motor skill. *Perceptual and Motor Skills*, **76**, 1015-1024.
- Gusthart, J., Kelly, I., & Graham, T. (1994). Student's learning of volleyball skills. *Perceptual and Motor Skills*, **79**, 200-202.
- Gusthart, J.L., & Springings, E. (1989). Student learning as a measure of teacher effectiveness in physical education. *Journal of Teaching in Physical Education*, **8**, 298-311.
- Kennedy, J.J., Cruickshank, D.R., Bush, A.J., & Meyers, B. (1978). Additional investigations into the nature of clarity. *Journal of Education Research*, **72**, 3-10.
- Landin, D. (1994). The role of verbal cues in skill learning. *Quest*, **46**, 299-313.
- Lee, T., Swinnen, P., & Serrien, D. (1994). Cognitive effort and motor learning. *Quest*, **46**, 328-344.
- Magill, R. (1993). *Motor learning: Concepts and applications*. Dubuque, IA: Brown.
- Magill, R. (1994). The influence of augmented feedback on skill learning depends on characteristics of the skill and learner. *Quest*, **46**, 314-327.
- Metzler, M. (1979). The measurement of academic learning time in physical education (Doctoral dissertation, The Ohio State University). *Dissertation Abstracts International*, **40**, 536A. (University Microfilm No. 8009314)
- Metzler, M. (1989). A review of research on time in sport pedagogy. *Journal of Teaching in Physical Education*, **8**, 87-103.
- Rink, J. (1993). *Teaching physical education for learning* (2nd ed.). St. Louis, MO: Mosby.
- Rink, J. (1994). Task presentation in pedagogy. *Quest*, **46**, 270-280.
- Rink, J., French, K., Werner, P., Lynn, S., & Mays, A. (1992). The influence of content development on the effectiveness of instruction. *Journal of Teaching in Physical Education*, **11**, 139-149.
- Rink, J., & Werner, P. (1989). Qualitative Measures of Teaching Performance Scale. In P. Dart, D. Zakrajsek, & P. Mancini (Eds.), *Analyzing physical and sport instruction* (2nd ed., pp. 269-276). Champaign, IL: Human Kinetics.
- Rosenshine, B., & Stevens, R. (1986). Teaching functions. In M. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 376-391). New York: Macmillan.
- Saskatoon Board of Education. (1988). *Physical education resource material Grade 7 and 8*. Saskatoon, SK, Canada: Author.
- Schempp, P.G. (1987). Research on teaching in physical education: Beyond the limits of natural science. *Journal of Teaching in Physical Education*, **6**, 109-110.
- Schmidt, R. (1989). *Motor control and learning: A behavioral perspective* (2nd. ed.). Champaign, IL: Human Kinetics.
- Siedentop, D. (1991). *Developing teaching skills in physical education*. Mountain View, CA: Mayfield.
- Silverman, S. (1988). Relationship of selected presage and context variables to achievement. *Research Quarterly for Exercise and Sport*, **59**(1), 35-41.
- Silverman, S. (1991). Research on teaching in physical education: Review and commentary. *Research Quarterly for Exercise and Sport*, **62**(4), 352-364.

- Silverman, S., & Buschner, C. (1990, April). *Validity of Cheffers adaptation of Flander's interaction analysis system*. Paper presented at annual conference of American Educational Research Association, Boston, MA.
- Silverman, S., Devillier, R., & Ramirez, T. (1991). The validity of academic learning time-physical education as a process measure of achievement. *Research Quarterly for Exercise and Sport*, **62**(3), 319-325.
- Silverman, S., Tyson, L., & Krampitz, J. (1991). Teacher feedback and achievement in physical education. *Teaching and Teacher Education*, **8**, 333-344.
- Werner, P., & Rink, J. (1989). Case studies of teacher effectiveness in second grade physical education. *Journal of Teaching in Physical Education*, **8**, 280-297.