

Be Fair to Physical Education Teacher: Subject-Specific Approach to Teacher Evaluation

特為體育教師而設的專項教學表現評估

Bill Lam Chi Biu

Department of Physical Education,
Hong Kong Baptist University, HONG KONG

林志彪

香港浸會大學體育學系



Abstract

Studies done to validate most teacher evaluation systems depend on establishing the content validity of the observational instrument. The studies usually consist of a comparison between the general job description of teachers and the behaviors listed as items on the measurement tool as well as input from teachers on the importance of the behaviors included. This study attempted to assess the validity using the same two types of data, i.e. expectations for effective teaching and teacher input. However, the study took the validity issue one step farther by developing a list of effective teaching behaviors in one particular subject, namely physical education. The study asked for open-ended teacher input on what constitutes effective teaching in physical education rather than asking the teachers to respond to an established set of items.

摘要

要使教師教學表現評估制度成為有效，就必要使其評估制度所採用的視學工具擁有有效的量度內容。這方面的研究通常將視學工具裏的量度內容跟教師的一般工作內容作一比較，並搜集教師對視學工具中所提的教學行為所抱的意見。本研究正是以此類資料，評估「教學表現評估」的效度。作者還藉這比較提出一系列體育教師應有的高效教學行為。

Introduction

Research studies have provided a wealth of information regarding the characteristics of effective teaching (Brophy & Good, 1986; Medley, 1979). A comprehensive review of these researches has frequently been used as a foundation for teacher evaluation systems. However, such a foundation should be used with extreme caution. As McDonald and Elias (1976) found in the beginning Teacher Evaluation Study, what worked well in one subject or at one grade level or in one teaching context did not necessarily work well in another. The generalizability of findings from research on effective teaching is questionable.

Stodolsky's work (1984, 1988) documents behavioral differences between the teaching of mathematics and social studies in an elementary classroom and delineates the problems of assessing generic teaching behaviors across varied subject matter. Sergiovanni (1987) also rejects generic models that "seem to equate explicit

teaching with effectiveness" and calls instead for an approach that recognizes the ambiguity of professional work. Good and Mulryan (1990) accuse the education community of misusing the research on teacher effects and quote Brophy (in press), "Research on teacher effects has been seriously misused in many teacher evaluation and accountability programs developed by state departments of education and districts. Any such effort that in effect imposes a single lesson format on all teachers in all teaching situations is simply invalid, and cannot be justified by claiming that it is supported by research on teacher effects."

Validity Issues

In the field of measurement, validity has traditionally referred to the issue of whether a particular test or instrument measures what it is intended to measure. Recently, an emphasis has been placed on the fact that an instrument or test is not valid in and of itself but is valid only for particular purposes. Popham

(1981) explains that when speaking of the validity of a test or measurement tool we must recognize that technically it is preferable to discuss the validity of the interpretation of the results of the test in connection with the purpose for which the test is being used rather than discussing the validity of the test itself. This concept of validity shall also be applied to the field of personnel evaluation. In other words, the measurement procedures for evaluating staff should be chosen or developed and implemented on the basis of the described role and the intended use, so that inferences concerning the evaluation are valid and accurate. Thus, educators in the field of evaluation should question whether generic teaching behaviors delineated from effective teaching research really reflect the roles of all teachers and thus, their accuracy as the tools of evaluation.

Physical Education Teacher

In this study, the focus will be placed on physical education teacher. Historically, physical education in schools has often been regarded as “extra-curricular”, something marginal and peripheral as compared with the “core-subjects” like languages, mathematics, sciences, etc. As a result, not many people seemed to care about whether quality teaching did happen in the gymnasium. Physical education would probably be ranked number one for being taught by most number of non-specialists. A belief that most teachers would be capable of teaching physical education was prevalent in schools. This belief was based on the assumption that physical education lesson meant keeping the kids busy and happy in class activities. Recently, as people have become ever more health concerned and also realized that physical education and recreation have the potential to improve the quality of life, the pressure increased for having quality physical education program in schools. It becomes logical that physical education teachers are held accountable for effective teaching in order to have quality physical education program.

The author of this study argues that physical education is a very specialized area of education. The pedagogical skills required of a physical education teacher are different from those of teachers from other subject matters. It would be inappropriate to suggest that their respective professional competencies be judged by the same measure. Thus, the present study aims to find out that to what extent are effective teaching behaviors in physical education distinctive from the generic ones.

Educational Significance

The goal of teacher evaluation is to improve the quality of education for students through assisting teachers to realize their potential and to carry out their duties more effectively. Teacher evaluation also assists teachers in their professional

development and career planning and those responsible for taking decisions about the management of teachers. All these are based on the belief that if we can accurately and effectively assess teaching, teaching and learning will improve. Valid teaching evaluation must be based on appropriate criteria. The present study identifies appropriate teaching effectiveness criteria, which are distinctive to physical education. Therefore, the study is significant in its contribution to the identification of appropriate teacher evaluation criteria for physical education teachers, and thus, through valid teacher evaluation to better management of physical education teachers.

Methodology

Procedures

In order to collect evidence for the question “to what extent are effective teaching behaviors in physical education distinctive from the generic ones”, three procedures were used:

1. A literature review was conducted to delineate the generic effective teaching behaviors.
2. A literature review in physical education was conducted to develop a list of the effective teaching behaviors in physical education.
3. Semi-Structured interviews were conducted with physical education teachers. The interview responses were used to develop a list of effective teaching behaviors in physical education.

Samples for Semi-Structured Interviews

Subjects for the teacher interview were practicing physical education teachers from secondary schools in Hong Kong. Hong Kong has 430 daytime secondary schools and serves approximately 433,208 students. Among the 430 secondary schools, there are 40 government schools, 310 aided schools, and 80 private schools.

A stratified sample of one government, nine aided, and two private school physical education teachers was desired for the teacher’s interview.

The selection of physical education teachers was by no mean random but based on the fact that among these four hundred and some schools, some of which were traditionally supportive and strove to excel in physical education and sports. The information with regard to the school supportiveness to its physical education program and sports was either informed by veterans in the field of physical education or confirmed by the school performance in inter-school sports competitions. Twelve of these schools, one government, nine aided, and two private, were first selected.

The physical education teachers in these selected schools were first telephone contacted to ask of their willingness to participate in the study. The willing teachers in each of these selected schools and their principals then received letters describing the purpose and goals of the study and the nature of the participation being requested. The number of teachers who actually participated in the interviews was 12: 1 government, 9 aided, and 2 private.

The goal of conducting the interviews and the reviews of the literature was to develop lists of effective behaviors in physical education as well as a list of generic effective behaviors found in the effective teaching research. The generic list came only from the research literature. On the physical education list, behaviors found in the literature are noted with an "L"; those mentioned by the teachers are noted with a "T"; and those behaviors mentioned more than three times by the teachers are noted with a double asterisk. When behaviors were mentioned by both sources an "L" and a "T" are used. If the wording of the two sources was somewhat different, both phrases are included in the order denoted by the letter preceding them.

Rationale for Semi-Structured Interviews

Semi-structured interviewing methodology is one of the several main techniques used for data collection in qualitative research. In general, the procedures involve discussion based on a small number of broad, central questions. The questions are designed to elicit the participants' perceptions about area of interest, in this case effective teaching behaviors in physical education. In each interview, the interviewer uses a schedule and starts off with some close-ended questions and then moves on to the broad, central questions provided with prompts. The interviews are tape recorded, and are transcribed. The transcription constitutes the raw data, the analysis of which is the basis for inferences and conclusions about the area in question. The semi-structured interviews were designed to gather the views of practicing physical education teachers concerning effective teaching behaviors in physical education. The following seven questions were used in each interview:

1. In your school, how much time per week per class is allocated to physical education (PE)?
2. Is there any difference in terms of time allocated to PE between junior form (F. 1-3) and senior form (F. 4-7)?
3. What types of activities do you teach in the junior form PE lesson?
4. In your school, what governs which activities to teach in PE lesson?
5. What is/are your educational goal/goals in teaching junior form pupils PE?
6. In your opinion, what roles does PE play in your school? What are the educational objectives of PE in your school?

7. By drawing on your own experience in teaching PE, let us discuss what does effective PE teacher do in his/her classroom instruction?

Prompts used to guide the discussion in this central question:

- a. Plan and preparation (What and how do you prepare for a PE lesson?)
- b. Progression of lesson (What are the procedures of a typical lesson of yours?)
- c. Instructional strategies (Illustrate by means of an example how would you teach your pupils a certain skill in an activity.)
- d. Interaction and feedback (How do you communicate with and give feedback to your pupils)
- e. Classroom management (Describe what you think as good classroom management which includes organizing and controlling the following four aspects:
 - i. Equipment
 - ii. Space available for activity
 - iii. Time
 - iv. Student and their misbehavior)

Data Analysis

Once the separate literature reviews were completed, comparison was made between the list developed for physical education and the list of generic effective behaviors which was developed. Items on the physical education list which were not on the generic list were considered distinctive to physical education (S).

Audiotapes of the interviews were transcribed word for word by the author. As the interviews were all conducted in the mother tongue, which is Chinese, of the respondents, the transcripts were in Chinese and were analyzed based on Chinese. The results were then translated into English and were verified by a second researcher. Names of specific people and places mentioned in the transcripts were removed to protect confidentiality. The transcripts were analyzed by examining the responses to the questions posed and identifying any additional major themes which emerged. The data were categorized into a list effective teaching behaviors mentioned by teachers in interviews with special designations beside behaviors which were mentioned by three or more teachers.

Overall Analysis and Synthesis

This was essentially a qualitative study and inferential statistics were not used. The analysis for this study attempted to synthesize the multiple sources of data as they relate to the questions which were asked. The results were synthesized into general conclusions about the extent to which effective teaching behaviors are distinctive in physical education. The list of effective practice in physical

education developed through the literature review and teacher interviews was compared with the list of generic behaviors developed from the literature. Behaviors found on physical education list which were not on the generic list were considered distinctive to physical education.

Results

The goal of conducting the teacher interviews and the reviews of the literature was to develop a list of effective teaching behaviors in physical education as well as a list of generic effective behaviors found in the effective teaching research. The lists which have been produced are shown below. The generic list came only from the research literature. On the physical education list, behaviors found in the literature are noted with an "L"; those mentioned by the teachers are noted with a "T"; and those behaviors mentioned by more than three teachers are noted with a double asterisk. When behaviors were mentioned by both sources an "L" and a "T" are used. If the wording of the two sources was somewhat different, both phrases are included in the order denoted by the letters preceding them.

Lists of Effective Behaviors

Generic Behavior (Literature Based)

- Provide a climate that is warm but task-oriented
- Select material at an appropriate level of difficulty
- Move through curriculum material at a brisk pace
- Present material in small steps insuring student mastery
- Maximize student opportunity to learn / keep students on task
 - Effectively manage student behavior
 - Minimize time spent in non-academic tasks such as transitions, preparation, etc.
- Have appropriately high expectations for student mastery
- Take responsibility for teaching and student learning
- Allocate time to academic objectives rather than another type of objective
- Be able to effectively diagnose student needs and prescribe appropriate activities
- Teach or supervise most of the time
- Present information through lecture and demonstration with elaboration through feedback rather than relying on curriculum to do it
- Give instruction and practice examples before assigning independent practice / convey information in brief presentations followed by recitation or practice opportunities
- Monitor student progress
- Teach again when necessary
- Structure learning for the students
- Present information clearly

- Be enthusiastic
- Use appropriate questioning
- Probe incorrect responses
- Provide immediate, specific feedback
- Provide opportunities for review and practice and monitor closely

Physical Education (Literature and Teacher Interviews)

- L – provide a supportive climate
 - T – be open to hearing and dealing with students' outside of class needs and problems **
 - T – be patient **
 - T – deal with affective needs of students in class, e.g. sportsmanship, enjoyment of sport **
 - T – be enthusiastic
 - T – set proper environment
 - T – motivate students **
- L – match instruction to group or individual
 - T – deal with multiple fitness and intellectual levels
 - T – diagnose class needs
 - T – monitor progress
- L – insure high time on task
 - T – be organized
 - T – get and maintain students attention
 - T – manage large groups of students **
- L – present tasks clearly
- L – provide time for practice on task / skills
- L – provide frequent, specific feedback
 - T – provide feedback, especially positive feedback **
- L & T – set expectations for students
- T – manage logistics of physical education, e.g. locker rooms, equipment **
- T – get students to follow directions
- T – be creative
- T – be flexible
- T – stress psychomotor skills
- T – set goals

Results of Comparison between the Lists

Physical Education with Generic

In order to answer the question "To what extent are effective teaching behaviors distinctive in physical education?", a comparison between the physical education list (PE) and the generic list (G) was made. A behavior on the physical education list (PE) which was not found on the generic list (G) would be considered physical education specific (S).

The following table shows the number of behaviors from the physical education list which belongs in each category.

Table 1. Comparison of Effective Behavior Lists

	PE
G	13 (52%)
S	12 (48%)
Total	25

For the subject physical education, the number of behaviors categorized in to each cell is given. This means that G-PE contains behaviors which were listed on the physical education list and on the generic list. Behaviors listed in cell S-PE are behaviors which were listed on the physical education list and not on the generic list. As shown in the table, approximately half of the behaviors in physical education are categorized as distinctive. (S-PE = 12 or 48%)

Each of the Lists produced by this comparison is reported below.

G-PE – Behaviors found on physical education list and generic list.

1. PE: match instruction to group or individual / G: select material at appropriate level of difficulty
2. PE: manage large group of students / G: effectively manage student behavior
3. PE: get and maintain students attention / G: effectively manage student behavior
4. PE: insure high time on task / G: minimize time spent on non-academic tasks
5. PE: be organized / G: minimize time spent on non-academic tasks
6. PE: manage logistics of physical education / G: minimize time spent on non-academic tasks
7. PE: set expectations for students / G: have appropriately high expectations for student mastery
8. PE: diagnose class needs / G: effectively diagnose needs
9. PE: present tasks clearly / present information through lecture and demonstration
10. PE: monitor progress / G: monitor student progress
11. PE: be enthusiastic / G: enthusiastic
12. PE: provide frequent, specific feedback / G: provide immediate specific feedback
13. PE: provide time for practice on tasks and skills / G: provide opportunities for review and practice

S-PE – Behaviors which were listed on the physical education list and not on generic list

1. Provide a supportive climate
2. Be open to hearing and dealing with students' outside of class needs and problems **

3. Be patient
4. Deal with affective needs of students in class, e.g., sportsmanship, enjoyment of sport
5. Set proper environment
6. Motivate students
7. Deal with multiple fitness and intellectual levels
8. Get students to follow directions
9. Be creative
10. Be flexible
11. Stress psychomotor skills
12. Set goals

Discussion

The following discussion is based on the conclusion drawn from the data reported in the study, but will extend beyond it to other issues that the conclusion suggests. Two basic issues will be discussed: 1) the purpose of teacher evaluation and 2) the assessment of generic processes applied differently in specific contexts.

Purpose of Teacher Evaluation

Many evaluation systems attempt to serve both accountability objectives and improvement concerns and fail to adequately serve either. Practically, the vast majority of teachers are not evaluated for accountability purposes after the beginning year(s) of teaching, but are assessed in order to recommend staff development activities to improve or enhance their instruction. Also, few, if any, dismissal decisions are ever based solely on formal evaluation data. Therefore, a teacher evaluation system with the purpose of providing feedback to teachers for improvement, not of accountability was proposed.

Balance between Generic and Subject-Specific Approach

Generic approach to teacher evaluation was developed in an effort to be objective and fair, therefore, served well for accountability objectives. However, this study has shown that applying only this type of instrument omits the assessment of critical behaviors teachers consider crucial to effectiveness in their discipline. As the proposed purpose of teacher evaluation system is to improve instruction through the provision of helpful feedback, a balance between the assessment of generic teaching processes and the varied application of those processes in specific teaching contexts must be maintained. Teacher involvement in the development of such system is therefore suggested. One way is by using teachers whose teaching experience was matched as closely as possible with that of the evaluatee in subject matter, age level, etc. as evaluators. These teachers would, of course, be trained in the use of the system and at least two teachers would team up for each evaluation to increase reliability and

reduce bias. In matching evaluator and evaluatee, a valid evaluation is possible without establishing an intricate hierarchy of generic / specific behaviors. It requires that the "matched" evaluators identify the appropriate subject specific behaviors under the generic processes as they implement the evaluation in each specific teaching context; then the system itself needs not to supply lists of those specific behaviors. The impact of the proposed system on the cost of evaluation would have to be investigated.

This study and the proposed evaluation system have important implications for teacher and administrator training. Teacher involvement in peer observation and/or evaluation needs to be presented in the teacher education program in a manner that begins to define teaching as a profession that requires teachers to play important roles outside their own classes. Future administrators need to be knowledgeable about the overall instructional approach and trained to evaluate that approach as well as to collaborate with subject matter experts in the evaluation of the application of the processes in a variety of contexts. Both groups could benefit from the simplicity of this overall process approach for structuring knowledge about the instructional process.

Conclusion

The overall finding of the present study was that physical education is a very specialized area of education which requires of a physical education teacher a very different set of pedagogical skills. An examination of the physical education literature and teacher input from interviews showed that many of the effective teaching behaviors in physical education were distinctive to this particular field. Twelve behaviors (48%) were identified as distinctive to instruction in physical education. Among these 12 behaviors, one dealt directly with the teaching of the relevant subject matter, i.e., "stress psychomotor skills". Others were mainly concerned with creating a positive learning environment, e.g., "motivate students", "provide a supportive climate", "be patient", "provide feedback, especially positive feedback" and "deal with affective needs of students in class".

It is clear that the behaviors mentioned by the teachers and the physical education literature were not only specific to the teaching of the subject, but also stressed the importance of creating the right affective environment for learning.

References

- Brophy, J.E., & Good, T.L. (1986). Teacher behavior and student achievement. In Wittrock, M. (Ed.), *Handbook of Research on Teaching* (3rd ed). New York, Macmillan.
- Good, T.L. & Mulryan, C. (1990). Teaching ratings: A call for teacher control and self evaluation. In Millman, J. and Darling-Hammond, L. (Eds.), *The New Handbook of Teacher Evaluation: Assessing Elementary and Secondary Teachers*. Newbury Park, CA: Sage Publications.
- McDonald, F. (1976). Report on Phase II of the beginning teacher study. *Journal of Teacher Education*, 27, 39-42.
- Medley, D.M. (1979). The effectiveness of teachers. In Peterson, P.L. & Walberg, H.J. (Eds.), *Research on Teaching: Concepts, Findings and Implications*. Berkeley, CA: McCutchan Publishing Corporation.
- Sergiovanni, T.J. (1987). *The Principalship: A Reflective Practice Perspective*. Newton, MA: Allyn & Bacon.
- Stodolsky, S. (1988). *The Subject Matters: Classroom Activity in Math and Social Studies*. Chicago: University of Chicago Press.
- Stodolsky, S. (1984). Teacher evaluation: The limits of looking. *Educational Researcher*, 13, 11-22.