A COMPARISON OF FOUR RATING SYSTEMS FOR THE EVALUATION OF PHYSICAL EDUCATION LESSONS

by

PETER RICHARD LIND

B.Ed., Massey University, 1980

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF ARTS
in the Faculty
of
Education

We accept this thesis as conforming
to the required standard

[redacted]
Dr B. Howe

[redacted]
Dr J. McLeish

[redacted]
Dr F. Martens

[redacted]
Dr A. Farquharson

UNIVERSITY OF VICTORIA

June 1982

Supervisor: Dr B. Howe

ABSTRACT

The purpose of this study was to compare three analytical observer systems with each other and an intuitive system of evaluation, using 24 videotaped physical education lessons selected for analysis. The physical education lessons were divided into four equal groups of volleyball, basketball, indoor soccer and badminton lessons.

These were analyzed by the author using the Academic Learning Time - Physical Education - Teacher Behaviour observation instrument developed at Ohio State University (1980), the McLeish Observation System (1975), and the Cheffers' Adaptation of Flanders' Interaction Analysis (1972). Each of the observer instruments measured the behaviours of the selected subjects, according to the pre-determined behavioural categories they theorised as defining effective physical education teaching. In addition, each of the lessons of the four groups of floor games were ranked for effectiveness on a scale from one to six, by six experienced physical educators.

The results showed that the three analytical observer systems were better able to reach common measurements of effective physical education teaching. The judges were not able to reach any agreement other than that which could be explained by chance.

Of the three analytical observer systems, the Academic Learning Time - Physical Education - Teacher Behaviour observation instrument was shown to be the most sensitive in measuring

the effectiveness of the 24 physical education lessons. When a factor analysis (based on selected categories from the the three observational systems) was computed, four factors were shown to account for 53.5 per cent of the variance. On the basis of this analysis, it was considered that a combination of the three scales, dropping duplicated and unsatisfactory variables, would be the most effective means of evaluating lessons in physical education.

Examiners: [redacted]

Dr B. Howe

[redacted]

Dr J. McLeish

[redacted]

Dr F. Martens

[redacted]

Dr A. Farquharson

TABLE OF CONTENTS

Page

Chapter

LIST OF TABLES

ACKNOWLEDGEMENTS

To the graduate students who allowed me to use the videotapes they collected - Mary O'Sullivan, Jaye Hickey, Monty Holding and Glenda Dhillon - thank you. I'm especially grateful to Glenda Dhillon for letting me use the judges' rankings that she collected.

To Barry, Mick and Dave, thanks for the cinnamon buns, coffee and suggestions.

To Dr Fred Martens, thank you for your helpful advice and assistance.

To Dr John McLeish, the long hours you spent in helping me grapple with the problems I faced in writing this thesis are much appreciated.

To the Howe family - Bruce, Jan, Carolyn and Jonathon - thank you for opening your home to Janis, Matthew and myself and making us feel very welcome. To Bruce, thank you for your wise counsel.

Janis, thanks for your support and patience.

CHAPTER ONE

INTRODUCTION

Context of the Problem

In the last decade, within British Columbia schools, there has developed an increased awareness of the role of physical education in the school syllabus, as evidenced by the revisions of the Elementary (1975) and Secondary (1980) Physical Education Curricula. Both curricula acknowledge that the teacher makes an important contribution to the learning of the student. Good, Biddle and Brophy (1975, p.3) support this stance:

> "There is...a small and consistent part of learning that is attributable to teaching and, more specifically, to the particular kind of teaching the student receives."

Assuming that this is true, the effectiveness of teachers of physical education is an important variable in evaluating a physical education programme.

Improving teaching continues to pose problems for theorists, researchers and practitioners. Gage (1963) stated that all research into the act of teaching prior to 1948 was suspect due to inadequate descriptions of the precise nature of the teacher variables. In the past, most evaluation had been

based upon subjective assessments and not upon scientific analysis (Cheffers, 1978.) Research on the effectiveness of physical education teachers using observer systems has only begun over the last decade (Locke, 1978.)

The current research in teaching effectiveness indicates there is no one method of teaching that invariably results in student learning. However, "a pattern of effectiveness that has a broad applicability within many contexts may be emerging" (Siedentop, Birdwell and Metzler, 1979, p.1.)

'Time on task', in terms of student time engaged on tasks relating to the objective of the lesson, has been considered as one of the most consistent and most significant correlators with student outcomes (Rosenshine and Berliner, 1978.) This process-product strategy has revealed some interesting findings applicable to the teaching of science, mathematics and reading. This strategy, however, does not seem to be appropriate for physical education, since it is difficult to find valid and reliable measures of student achievement in physical education (Locke, 1977.)

Using a process analysis, studies have been made of the behaviours and interactions of students and teachers during physical education lessons. Each observer system developed has attempted to measure the effect of different variables thought to be of

consequence for effective teaching and, or, student achievement (Cheffers, 1977). However, because of the lack of standardization and replication of studies using these observer systems, results have lacked supportive evidence and meaning (Locke, 1977). This indicates a need to use existing observer systems, in order, to develop a common base upon which empirical statements concerning the effectiveness of physical education teachers can be made.

Statement of the Problem

The purpose of this study was to compare the analyses of three analytical observer systems with each other and with an intuitive system of evaluation, using 24 videotaped physical education lessons specially collected for this purpose. The intuitive observer system involved six physical education experts, who ranked each of four types of games lessons, from the most effective to the least effective lesson, using a global rating. The study focused on the following questions:

1. What relationship exists between the three analytical observer systems?
2. What relationship exists between the analytical observer systems and the intuitive observer system?
3. What agreement exists between the judges using the intuitive method of evaluation?
4. What differences, if any, exist in the behavioural structure of the four types of games lessons?

CHAPTER TWO

REVIEW OF RELATED LITERATURE

This study is based on the assumption that teachers affect student learning. Recent studies support this premise (Bloom, 1976; Berliner et al, 1978; Petersen and Walberg, 1979.) The evidence of these studies indicates that, during the course of the year, up to 25 per cent of student learning can be traced to the actions of the teacher. Further, teacher influence is cumulative. Over the course of a student's schooling, teaching will have contributed importantly to the student's learning. This evidence gives a mandate for studying the effectiveness of teaching.

Research on teacher effectiveness (RTE) faces two major problems. Firstly, there is a need to identify what constitutes effective teaching. Secondly, there is a need for a reliable, valid and objective method of evaluating the effectiveness of teachers.

Mitzel and Medley (1963) identified three approaches to evaluating teacher potential and performance. Evaluation of the personal characteristics of the teacher, they termed the presage approach. The process approach, they described as evaluating the method of instruction and the interactive patterns between teacher and student. Finally, evaluating the student growth resulting from the teaching, they considered to be the product approach. A number of systematic observation instruments existed prior to the Mitzel and Medley paper. Notably, Withall (1949), Bales (1950) and Flanders (1955) had introduced

observer systems which could be used to describe classroom behaviours. Since the Mitzel and Medley paper, the development of systematic observation instruments has reached the thousands (Cheffers, 1978).

The Use of Observer Systems

Systematic observer systems have been used by some individuals for years. For example, professional basketball coaches have used charts to record players' performances for a considerable time. Within the teaching context, observer systems can give an accurate description of classroom interaction.

Batchelder and Cheffers (1976, p.8) state an observation instrument can be used to:

> (1) describe current classroom practices, (2) modify teacher behaviour, (3) provide a tool for the analysis of teaching, (4) give feedback about one's own teaching, (5) train student teachers, (6) discriminate between patterns of teaching, (7) determine the relationship between various classroom behaviours and student growth, (8) help in projecting future teaching patterns.

Classroom observation instruments have been defined in a number of ways. Gage (1969) and Rosenshine (1970) have divided such instruments into category and rating systems. Category systems have low-inference measures. They focus upon specific, observable and relatively objective behaviours. These events can be recorded as frequency counts. Rating systems have high-inference measures. The items require the observer to infer these from the lesson and, further, to infer the frequency of such behaviours. An item of the rating system may, for example, require the observer to infer whether the teacher displays "enthusiasm". The observer then may have to decide whether the

teacher displays this behaviour "consistently", "sometimes" or "never".

Cheffers (1978) divides observation systems into two main categories. Systems, such as anecdotal recording, which materialize after a series of observations have been made, are called inductive. Observations that are recorded within a predetermined formula are called deductive systems.

Siedentop and Olson (1978) have limited the definition of observation instruments to deductive or, more specifically, category or sign systems. A sign system divides the observation into intervals. For each interval, the observer decides which behaviour category best describes that interval. In a category system, the behaviour is recorded as it occurs (Duncan and Biddle, 1974.)

Having considered the types of observation systems that have been used, Coker, Medley and Soar (1980, p.132) report that RTE should now use low-inference systems "because of a conviction that high-inference measures...are not dependable measures of specified behaviours." Further, low-inference systems, by definition, require the researcher to make direct observations of instructional activities and to report the teacher and student behaviour systematically (Locke, 1977.)

Variables Considered to be Indications of Teacher Effectiveness

One of the most important figures in the development of objective observer systems, Flanders (1955), expanded the seven category observer system of Withall (1949) into a ten category system. This system was designed to study the effect of direct and indirect teacher influence upon classroom behaviour (Rosenshine and Furst, 1971.)

Flanders' influence can be observed in the Rosenshine and Furst (1971) review of 51 process-product RTE studies. The authors grouped their findings into 11 variables which they considered to represent the most significant and consistent results. They devised these variables in 1973 and, basing their revision on the same RTE studies, grouped the results under nine variables. These were:

> (1) clarity, (2) variability, (3) enthusiasm, (4) task oriented and/or business-like, (5) criticism, (6) teacher indirectness, (7) student oppurtunity to learn criterion material, (8) use of structuring comments, (9) multiple levels of questions.

They reported that the findings of the studies reviewed were tentative and unconfirmed. Most of the studies had centred on classroom climate and the directness and indirectness of the teacher.

Dunkin and Biddle (1974) reviewed 178 process-product RTE studies. They reported that:

> (1) Most of the studies had focused on the directness and indirectness of the teaching style and classroom climate.
> (2) No process-product data had been accumulated in the area of classroom management.
> (3) Very few observation systems had been used more than once or twice.
> (4) The process-product data gained from the reviewed studies was inconclusive and inconsistent.

Schalock (1979), after an extensive review of the literature, found that no consistent relationships had yet been found between what teachers know, or the characteristics they possess, and their teaching effectiveness. Berliner (1976) stated that, although there is considerable evidence that particular personality characteristics of teachers have a discernible

influence on student behaviour, there is little evidence that certain personality characteristics are more desirable than others for teaching in general. Furthermore, research has found no consistent relationship between particular teacher behaviours and teacher effectiveness. Brophy (1976) observed that effective teaching requires a synthesis of skills into complex patterns of behaviour. McLeish (1978, p.10) supports this view:

> "Effective teaching is a function of the discriminate use of a variety of operants. This is a dynamic process, not a mechanical procedure. Learning involves a multiple output of stimulus and reinforcement patterns, with responses by the subject. These are caught up in the interactive pattern generated by the response of the teacher and other participants."

Since 1973, RTE studies have stated a need to focus on two student variables, in particular. Firstly, the content covered by the students and, secondly, the amount of time the students are engaged on tasks related to the objective of the lesson (Rosenshine and Berliner, 1978.) No longer are teacher behaviours of primary concern, rather, it is believed that certain behaviours of the learner are more critical.

Time on Task

Bloom (1976), basing his research on earlier studies by Carroll (1963), stated that if the form of instruction and the amount of time allowed for study are made appropriate to the characteristics and needs of each learner, the majority of students will achieve mastery of a given subject. Therefore, the critical factor in the study of effective teaching is knowing whether the students are engaged in mastering skills and their progress towards achieving mastery of a particular

skill (Rosenshine and Berliner, 1978.)

Research into the teaching of elementary school mathematics, reading and science have yielded the most important results thus far. Significant relationships have been found between the content covered and student achievement gain (Husen, 1967; Shutes, 1969; Pidgeon, 1970; Chang and Raths; 1971; Rosenshine, 1971; Barr, 1973; Comber and Keeves, 1973; Walker and Schalfarzick, 1974; McDonald, 1975; Rosenshine, 1976; Armento, 1977). Only one study (Brown, 1969) reported in the literature did not find a significant relationship between these two variables.

Medley (1977) examined 289 empirical studies of teacher effectiveness. Teacher effectiveness was measured only in terms of student gain. He noted that effective teachers spent a greater amount of time on academic tasks than the ineffective teachers.

This work is supported by the findings of the Far West Laboratory for Educational Research and Development in the Beginning Teacher Evaluation Study (BTES) in California by Berliner and his associates (1972-1978). This has been one of the most extensive studies of teacher and student behaviour to date. Three of their findings were:

> (1) The amount of time the teacher allocated to instruction in a curriculum area was positively associated with learning in that content area.
>
> (2) The proportion of allocated time that students were engaged was positively associated with student learning.
>
> (3) The proportion of time that academic tasks provided a high success rate for a student was positively associated with learning.

Hall, Delquardi and Harris (1977, p.13) have found, as a result of their work in the Juniper Garden's Children's Project (JGCP), that:

> the basic element lacking in the homes and classrooms of the inner-city is not motivation per se, nor does it seem to necessarily be curriculum materials. Rather, it is quite possible that the major factor may be a lack of oppurtunity for children to make active learning responses.

From these recent studies a pattern is emerging. In order to maximize the successful learning responses of students, classroom management skills are very important (Good, 1979). Teachers need to provide an academic focus, clear goals for students, sequenced and structured learning tasks, and sufficient time for students to master a particular learning task before they are asked to proceed to another. In addition, they must provide immediate and task specific feedback (Tomlinson, 1981).

These findings are drawn from studies of the teaching of reading, mathematics and science. Very little process-product research has been conducted in other subject areas and with older students, to discover whether these findings can be generalised to all teaching (Rosenshine and Berliner, 1978; Schalock, 1981).

Research on Teaching Effectiveness in Physical Education

There have not been many scientific inquiries into effective physical education teaching (Locke, 1979). However, those that have been completed have revealed:

> (1) There is no general consensus as to which behaviours may be labelled as "effective" or "ineffective" in the

teaching of physical education (Cooper, 1972; Cheffers, 1977; Locke, 1977).

(2) There is no one single superior method for analyzing effective physical education teaching at present.

(3) There exists no common base upon which empirical statements concerning teacher effectiveness in physical education can be made because there has been no consistency in the use of measuring instruments and variables examined.

Locke and Nixon (1973) stated the need to analyze teacher behaviour while students engaged in learning movement skills. Rosenshine and Furst (1973) and Fishman and Anderson (1971) saw a need to gather descriptive data from the learning environment. Goldberger (1974) stated that regardless of the intents of teachers, it was their actual behaviour which the learners perceived and to which they reacted. In evaluating the BTES study, Berliner and Tikunoff (1976) stated that descriptive information helped the researcher to understand further the complexity of the teaching instruction and could help discriminate between "effective" and "less effective" teachers.

Researchers studying physical education teaching, therefore, have accepted the need for observer systems used in a natural physical education setting.

The Use of Observer Systems in Physical Education

Locke cited 11 studies between 1969 and 1975 using Flanders' Interaction Analysis Observation System or an

adaptation of this instrument. The Cheffers' Interaction Analysis System (CAFIAS, 1972) is such an adaptation. This process system is designed to describe behaviours in teaching or modifying human interaction. Cheffers expanded the Flanders' system to include more non-verbal behaviour categories. Further, it does not see the class teacher as the only teaching agency. The enviroment and other students can also act in this capacity. Cheffers (1980) cites 41 studies that have used CAFIAS. Locke (1977) concludes that there is no evidence to suggest that the correct type of data had been gathered from the Flanders' based studies. As a result, a number of researchers have developed other observation instruments (Lupien, 1970; Alder, 1972; Graham, 1973; Manson, 1973; Fishman, 1974; Hope, 1974; Anderson, 1975; Laubach, 1975; Rankin, 1975; Siedentop and Hughley, 1975; Gasson, 1976; Taylor, 1976; Morgenegg, 1978). However, because of the lack of standardization and replication of studies using these observer systems, results have lacked supportive evidence and meaning.

All the observation instruments cited are descriptive systems. Process-product strategy necessitates finding valid and reliable measures of student achievement. Thus far, physical educators have not been able to reach a consensus on what these should be (Nygaard, 1975; Siedentop, Birdwell and Metzler, 1979).

The Relationship Between 'Time on Task' and Physical Education Teaching

Locke and Nixon (1973) cited studies by Overstreet and Verducci which concluded that large physical education classes, where children spent a large percentage of the time lined up

waiting to perform an activity, resulted in lower skill achievement than where children spent more time performing the activity.

An early teacher effectiveness study in the area of physical education was done by Yerg (1977). Forty student teachers taught the cartwheel to 120 elementary students. Each teacher taught three students over a 20 minute period. Yerg found no significant skill achievement as a result of the teaching. Interestingly, each student on the average had only nine practice trials during the 20 minute period.

Anderson (1975) observed 83 elementary and secondary school physical education teachers teach a physical education lesson. His observer system allowed for three movement categories. Fifty-three per cent of the time the students were inactive. Thirty-five per cent of the time was spent on movement related to the skill being taught and 12 per cent of the time was spent on other movement, for example, lining up.

Costello and Laubach (1978), using the Behaviour of Students in Physical Education (BESTPED) System, observed 193 students in 20 different classes of varying grade levels. They found that:

1. 60.8 per cent of the time was spent was spent waiting for or listening to the teacher.
2. 15.3 per cent of the time was spent practicing motor skills.
3. 10.3 per cent of the time was spent playing games.
4. 3.6 per cent of the time was spent exercising.
5. 0.2 per cent of the time was spent on exploratory movement

Anderson and Barrette (1978), using Anderson's Descriptive System to observe 20 elementary and 20 secondary school physical education classes, concluded that the greater part

of the time was spent in the students listening to teacher talk, management activities or waiting to perform the motor activity.

Siedentop, Birdwell and Metzler (1979) have cited studies by Stewart (1977), Quaterman (1977) and Freedman (1978) that showed physical educators spend a great deal of time in management and instruction with relatively little time available for active practice by students.

Siedentop, Birdwell and Metzler developed the Academic Learning Time-Physical Education (ALT-PE) System, as a process approach for measuring teacher effectiveness of physical education lessons. Their system was based on the findings of the Beginning Teacher Evaluation Study (BTES) and the Juniper Garden's Children's Project (JGCP) which were noted earlier. These findings stated the need for students to make active responses and that the learning material should be structured so the student could perform the tasks with success. Further, the system should not involve expensive equipment and greatly increase the workload of the physical education teacher.

Metzler (1979) used the ALT-PE observation instrument to describe the physical education lessons of 33 classes, ranging from the elementary, junior and senior high school grades, of schools from Columbus, Ohio. He reported the following findings:

(1) Content-PE accounted for 73.6 per cent of the class time.
(2) Students were engaged 49.2 per cent of the observed intervals.
(3) ALT-PE occurred 26.8 per cent of all intervals.
(4) ALT-PE (motor) occurred 7.5 per cent of all intervals.

Metzler's findings indicate the need for increased opportunity

for students to make on-task motor responses in the physical education lessons observed.

Birdwell (1980) modified the original ALT-PE instrument so that teacher behaviours could be recorded simultaneously. Stage III of the BTES study had identified three teaching behaviours that appeared to be consistently associated with student achievement. These were the ability to diagnose and provide feedback on a student's performance and the ability to keep students on task. Birdwell intervened on the teaching behaviour of three in-service teachers in an attempt to modify specific teaching behaviours. She reported the following results:

> (1) There was a decrease in the mean percentage of student non-engagement in the physical education lessons of all three teachers.
> (2) There was an increase in the mean percentage of ALT-PE in the physical education lessons of all three teachers.
> (3) There was an increase in the mean percentage of teacher feedback in the lessons of the secondary level teachers.
> (4) There was a decrease in the mean percentage of managerial time in the lessons of the secondary level teachers.

The elementary level teacher was measured only on one teaching behaviour because he did not meet his class as often as the secondary level teachers.

Whaley (1980) intervened to influence the behaviours of physical education teachers and students of four classes of four separate schools, using the ALT-PE observations as graphic feedback. As a result of the interventions, he observed no significant change in the level of physical education content in the lessons. Any change that was noted in the physical education

content was the result of changes in the type of physical activity. Whaley (p.104) concluded that, rather than just supplying graphic feedback, future interventions should:

> develop management and organisational skills of the teacher so that students are provided more oppurtunities to respond.

Teacher Effectiveness Studies at the University of Victoria, B.C.

The School of Physical Education at the University of Victoria, British Columbia, has been involved in the study of effective physical education teaching for the past three years (1979-1982). The pioneer study was carried out by Bevan Grant. McLeish, Jackson and Howe then proceeded to supervise the collection of 104 physical education lessons videotaped on location in British Columbia elementary, secondary and adult recreation classes. These videotapes were part of individual studies by four graduate students — O'Sullivan, Hickey, Holding and Dhillon. An overall analysis of the 104 tapes was carried out by McLeish.

Grant (1979) videotaped 16 student teachers (eight of whom had specialist physical education training) teaching a physical education lesson to an intermediate grade class. He used Gasson's analytic instrument which described teacher-student interaction in terms of verbal and non-verbal behaviour and recorded the amount of time on task and the physical position of the teacher. Grant also used expert, experienced physical educators, using a global rating to rank the lessons from the most effective to the least effective. His findings indicated that there were no significant generic behaviour patterns amongst the most effective teachers, as rated by the experts. Grant also found that the most effective teacher had the class working on task for 31.6 per cent of the lesson. The least effective teacher had the

class working on task as little as 6.71 per cent of the time. Further, Grant found no significant difference between the group with physical education training and the group without this training. Grant recommends that further studies of this nature should use a larger number of judges, as it is possible that three judges will not perceive all the behaviours during their observation of the lesson.

O'Sullivan (1980) videotaped 16 physical education specialist secondary school teacher trainees teach a physical education lesson. Like Grant, she used three expert physical educators to rank the lessons from the most effective to the least effective. The ALT-PE observation instrument (noted earlier) was used to describe the student behaviour in the dance, basketball and soccer lessons. O'Sullivan analysed the relationship between the different physical activities, class size and sex of the teacher trainee with the judges' rankings and the student 'time on task'. There was a significant difference between each of the activities taught and the amount of student 'time on task'. This is supported by Whaley's (1980) findings (noted earlier). There were no significant differences between any of the other variables.

Hickey (1980) videotaped a physical education lesson of each of 12 matched physical education teachers selected on the basis of teaching experience, their sex and the grade level taught. These variables were compared with the observed student behaviour in each lesson, as recorded by the BESTPED observation instrument (noted earlier), and judged effectiveness. Teacher effectiveness was decided by exactly the same method as that of Grant and O'Sullivan. Hickey reported a

significantly greater amount of on-task movement in the female teachers' classes than the males' classes, but no other meaningful significant differences were found. Hickey noted the importance of using a common observation instrument for future research into effective physical education teaching and the need to develop a common data base.

Holding (1980) videotaped 32 adult swimming and tennis recreation classes. He was interested in examining the relationship between activity type, class size, age and sex of the teacher with student behaviour, as described by the ALT-PE analysis, and the judged effectiveness of the lessons. Holding used four experienced recreation teachers to rank the lessons from the most effective to the least effective. The only significant difference he reported was between the activities taught and the amount of time students spent on motor activity related to the objective of the lesson. A further important finding was that there was no significant difference between teaching six or twelve adults in a recreation class.

Dhillon (1982) examined the relationship between the global ratings and perceptions of 24 videotaped games lessons by six experienced physical educators and the ALT-PE observation instrument. She found no significant agreement between any of the judges' rankings, nor between their rankings and the ALT-PE behaviour variables. However, she found certain common features in the written comments on each lesson, among the three professors of physical education and, as a separate group, the three teachers of physical education.

McLeish (1981) carried out an analysis of the 104 videotapes collected by the afore-mentionned graduate students.

His study had four thrusts:

1. To compare the descriptions of four observation instruments (Flanders, Gasson, McLeish and ALT-PE) of the 16 videotapes of balance lessons collected by Grant (1980).
2. To make an analysis of the internal consistency of the ALT-PE observation instrument by comparing the ALT-PE analysis of each of the 104 videotaped physical education lessons.
3. To compare the relationship between the ALT-PE analysis and the judges' rankings of the videotapes of Grant (1980) and Holding (1981).
4. To compare the selected class and teacher characteristics of the 104 videotaped physical education lessons with the ALT-PE analysis.

McLeish reported that the Gasson and ALT-PE observation instruments had a higher degree of agreement with the judges' rankings of the lessons than the McLeish and Flanders observation instruments. He hypothesised this was because the former systems were designed to analyse physical education lessons specifically, whilst the other two were designed for general classroom use.

The ALT-PE instrument was consistent in measuring the amount of time the students spent on task and, conversely, the amount of time spent in management tasks and waiting. Further, the observation instruments (particularly the ALT-PE instrument) described the lesson in much more detail than the judges. The judges tended to emphasise the classroom climate or milieu of the lesson.

Of the class and teacher characteristics selected to be compared, the following results were reported:

1. There was no significant difference in teacher effectiveness because of the sex or the experience of the teacher.

2. The sex of the class (whether all male, all female or coeducational) did affect the nature of the lesson.
3. Teaching effectiveness, as measured by time-on-task, declined with an increase in class size.

McLeish noted (p.28) that future research into effective physical education should:

> control the nature of the activity in the lesson by limiting it to one kind only, for example, gymnastics.

Summary

This chapter has reviewed the recent literature on teacher effectiveness in education and physical education in particular. Research on teacher effectiveness (RTE) of physical education is relatively new. Studies have shown that the teaching act encompasses a wide range of behaviours and that no one single behaviour by itself has been consistently correlated with student learning.

It has been theorised that the majority of students can master academic skills provided that the amount of time allowed for study and the form of instruction is made appropriate to the characteristics and needs of the learner (Bloom, 1976). Recent RTE studies (BTES, 1972-1977; JGCP, 1974-1977), using a process-product strategy, have indicated that the amount of time students are engaged on task is significantly correlated with student outcomes. Based on these studies, it has been theorised that the mediating link between the behaviour of the teacher and student achievement in physical education is the amount of time engaged on tasks related to the objective of the lesson (Siedentop, Birdwell and Metzler, 1979).

Therefore, it would appear that effective physical education teachers are able to keep students engaged on task in appropriate activities.

CHAPTER THREE

RESEARCH METHODS

This chapter describes the method of subject selection, the variables considered, the method of data collection and the method of data analysis.

Subjects

A collection of 104 physical education lessons were videotaped on location in British Columbia elementary, junior and senior high school and adult recreation classes, as part of a three year study on effective physical education teaching at the University of Victoria, B.C. (1979-1982). These videotapes were made available to the researcher.

Since other studies had indicated that different types of physical activity resulted in significant differences in the amount of time spent in motor skill practice (Whaley, 1980; O'Sullivan, 1980; Holding, 1981; McLeish, 1981), 24 secondary school floor game lessons were selected from the original sample of lessons. These 24 floor-game lessons were divided into four groups consisting of six volleyball, six basketball and six indoor soccer lessons. They were chosen from

the larger pool of tapes entirely on the principles: (a) they were games of the kinds specified, and (b) the sound and vision were adequate for coding purposes.

Dependent Variables

The dependent variables of the study were:

1. The behaviour categories of the Academic Learning Time - Physical Education - Teacher Behaviour Observation Instrument (ALT-PE-TB). There are five major categories in this system of analysis. These are:
   (a) The setting category. There are six behavioural definitions possible to describe the teaching style being used in a given interval. They are: direct instruction, task, reciprocal, group, guided discovery, and problem solving.
   (b) The content category. There are 12 definitions possible to describe the content of the activities observed in a given interval. They fall into two main groups:
      (i) The content-general group includes the activities which do not have a physical education focus. These are: wait, transition, management, break and non-academic instruction.
      (ii) The content-physical education group includes activities that have a physical education focus: skill practice, scrimmage, game, fitness, other motor activity, knowledge and social behaviour.
   (c) The learner moves category. There are six definitions which describe the nature of student engagement in a given interval. They fall into two main groups:
      (i) The engaged group includes engaged-motor response, engaged- indirect and engaged-cognitive.

(ii) The not-engaged group includes not-engaged interim. not-engaged waiting and not-engaged off-task.

(d) The level of difficulty category. This reflects the observed student's ability to perform the assigned task. The three levels of difficulty are easy, medium and hard.

(e) The teacher behaviour category. There are 16 behaviours contained in this category. They are: lecturing, giving directions, listening, asking questions, answering questions, non-functional behaviour, monitoring, hustling, spotting, maintenance, modelling, feedback, behaviour praise, nagging, punishment, teacher participation and officiating.

(A detailed description of these categories is included in Appendix A).

2. The behaviour categories of the McLeish Observation Instrument. This system of analysis seeks to categorise all the possible behaviours of individuals in a social situation. McLeish recognises four broad dimensions of behaviour when two or more individuals are communicating in a common enviroment. These are:

(a) categories of reference: tacts and extended tacts;

(b) categories of action: mands and submissive autoclitics;

(c) categories of expressive behaviour: positive and negative affective behaviour;

(d) Thematic connections: echoics and intraverbals.

(A detailed description of the categories is included in Appendix B).

3. The behaviour categories of the CAFIAS Observation Instrument. There are six major categories in this system of

analysis. These are:

(a) Teacher verbal behaviours. They include praise, accepting and developing suggestions of the learner, questioning, giving facts or opinions, giving directions, and negative criticism.

(b) Teacher non-verbal behaviours. These are the corresponding non-verbal behaviours of the seven teacher verbal behaviours. These include actions demonstrating approval, demonstrating empathy, adopting a quizzical posture, drawing or writing or demonstrating activities and, finally, actions displaying disapproval.

(c) Student verbal behaviours. These include entirely predictable student responses, student responses that require some measure of interpretation, student-initiated talk, and disorderly noise.

(d) Student non-verbal behaviours. These are the non-verbal behaviours which correspond to the four student verbal behaviours. These include entirely predictable movement responses, movement responses that require some measure of interpretation, innovative movement responses and students sitting doing nothing.

(e) The teaching agency. The teacher,or other students, or the enviroment can act as teaching agencies.

(f) Differential class structures. The class may be taught as a whole, or may be taught in groups.

(A detailed description of the categories is included in Appendix C).

4. Effectiveness of the lesson. After viewing the videotapes, six judges independently ranked the lessons of each of the four different activities — volleyball, basketball, indoor

soccer and badminton — for teaching effectiveness in six categories, from best to worst. The rating one (1) represented the most effective lesson and the rating six represented the least effective lesson.

Independent Variables

The independent variables of this study were the six tapes of four different floor-game activities:

1. volleyball;
2. basketball;
3. indoor soccer;
4. badminton.

The Lesson

The 24 floor-game lessons used in this study were part of the research on teacher effectiveness by the University of Victoria's School of Physical Education (1979-1982). These videotapes were selected from the samples collected by O'Sullivan (1981), Hickey (1981) and Dhillon (1982). The selection of tapes was made on the basis of clarity of the recording.

In making the videotapes, the following stipulations were made by the researchers. The subjects were invited to teach a physical education lesson to a junior or senior high school class which they normally taught. The lesson was to be in a gymnasium enviroment.

To avoid as much influence as possible by the researcher upon the teaching act, the selection of the lesson theme, the teaching method, the choice of activities and the use of specified equipment were left to the individual teacher.

In videotaping the lessons, the camera was stationed so

that the whole class could be viewed with minimal distraction. The teacher wore a cordless microphone in order to obtain a clear recording of his or her voice. This did not interfere with the physical aspect of the teaching act.

Each researcher used the following equipment:

1. a portable Sony videotape camera;
2. a half-inch Sony videotape recorder;
3. an Edcore cordless microphone, transmitter and receiver.

Ranking the Lessons for Teaching Effectiveness

Dhillon (1982) selected six judges, on the basis of their experience and knowledge of secondary school physical education teaching, to rank the six lessons of each of the four floor game activities. These rankings were made available to the researcher.

The judges viewed a representative sample of each of the lessons. Using a global rating, they independently ranked each of the lessons so that one (1) denoted the most effective lesson, whilst six (6) denoted the least effective lesson. The judges were not permitted to assign tied ranks.

The global ratings were based on subjective impressions of the quality of the lessons. The judges were encouraged to use only those criteria they normally used.

Behaviour Analysis of the Lesson

One observer coded all the behaviour dimensions of the videotaped lessons, as defined by the three observation instruments. The three systems each required an intensive training

period in a workshop situation. A trained expert in the ALT-PE-TB and McLeish observation systems provided the training and checked reliabilities on these and on the CAFIAS observation system.

The ALT-PE-TB Observation Instrument. The researcher had been involved in the original analysis of the 104 videotapes using the ALT-PE observation instrument by the University of Victoria, 1980-1981 (noted earlier). This had required participation in a training workshop conducted by an expert, who had received training at Ohio State University, where this instrument was devised. A high reliability coefficient between codings by three coders trained in the same workshop was obtained. The author of this dissertation then spent over 150 hours analysing videotaped physical education lessons using the ALT-PE observation instrument, as part of training for this study.

A further workshop was conducted by the same expert, to train the coder in the ALT-PE Teacher Behaviour Dimension. A high degree of agreement was reached between the instructor and the coder.

A scored interval (S-I) reliability measure was calculated to determine the reliability of the observer. This measure has been used in other studies by coders using interval recording systems (Metzler, 1979; Birdwell, 1980; Whaley, 1980). Scored interval reliability is recommended when there is more than a small number of behaviours being observed (Hawkins and Dotson, 1975).

The observer viewed a randomly selected videotape three times on three separate occasions over the period of the ALT-PE-TB analysis. (This occurred at the beginning, middle and end of

the analysis). Each interval recorded in any of the three viewings was analysed. If a behaviour was not scored on any of the videotape viewings, it was eliminated from the scored interval reliability calculation. The agreements and the disagreements for each behaviour interval were counted and reliability was calculated according to the following formula:

$$S\text{-}I = 100 \times \frac{\text{agreements}}{\text{agreements} + \text{disagreements}}$$

Scored interval reliability was scored among:

V_1 , V_2 , and V_3 .

(V_1 = first viewing; V_2 = second viewing; V_3 = third viewing)

An agreement was scored when a behaviour was recorded in the same interval at all three viewings, whereas, a disagreement was scored when a behaviour was not recorded at the same interval in all three viewings. (Refer to Table 1 for the reliability scores).

The McLeish Observation Instrument. After analysing the tapes using the ALT-PE-TB observation instrument, the coder took part in an intensive training workshop conducted by the originator of the system. A high degree of agreement was established between the instructor and the coder before the next step in the study was conducted.

A scored interval reliability measure (noted earlier) was again used to determine the reliability of the coder. Two videotapes were randomly selected and viewed three times by the coder on three separate occasions throughout the McLeish analysis. Reliability scores were calculated for each of the behaviours observed. (Refer to Table 2).

The CAFIAS Observation Instrument. Having completed the

Table 1
Scored-Interval Reliability Agreement
Percentages for the ALT-PE-TB Observation Instrument

Category	Reliability Scores			
	V_1	V_2	V_3	S-I
Management mode				
DM	6	7	6	94
DT	26	25	27	96
Instructional mode				
DKC	15	15	14	97
Skill practice mode				
TPM	14	15	14	97
TPNW	21	19	20	95
TPNI	1	1	1	100
Game mode				
TGI	2	2	2	100
TGNW	3	3	3	100
Fitness mode				
TFM	12	13	11	91
Difficulty level				
E	41	43	39	95
Teacher Behaviours				
Le	8	10	9	88
G	26	24	25	96
AQ	4	4	4	100
M	27	26	26	98
MT	1	2	1	75
Mo	3	3	3	100
N	12	11	13	91
F	20	21	20	98

V_1 - first viewing V_2 - second viewing V_3 - third viewing.

S-I refers to the scored-interval reliability agreement between the V_1, V_2, and V_3 percentage scores of each recorded interval.

Table 2

Scored-Interval Reliability Agreement
Percentages for the McLeish Observation Instrument

Category	Reliability Scores							
	Videotape A				Videotape B			
	V_1	V_2	V_3	S-I	V_1	V_2	V_3	S-I
A. Teacher Behaviours								
Mands:								
P.E. command	41	40	39	97	55	53	52	98
P.E. request	4	3	4	90	5	5	5	100
management command	46	47	48	97	9	10	9	96
management request	2	1	2	75	4	4	3	81
Tacts	1	1	1	100	5	5	6	93
extended tacts	2	2	2	100	7	8	8	91
Autoclitics:								
positive	1	1	1	100	7	7	8	95
dominant	1	1	1	100	4	4	5	92
submissive	1	1	1	100	1	1	1	100
Intraverbals	2	3	2	85	1	1	1	100
echoics	-	-	-	--	2	2	2	100
B. Student Behaviours								
Mands:								
P.E. request	1	2	2	75	-	-	-	-
Autoclitics:								
submissive	91	90	89	98	96	96	95	99
negative	8	8	9	96	4	4	5	92

V_1 - first viewing V_2 - second viewing V_3 - third viewing

S-I refers to the scored-interval reliability agreement between the V_1, V_2, and V_3 percentage scores of each recorded interval.

other two forms of analysis, the researcher was trained in the use of the CAFIAS observation instrument. A reliable consistency of recording was reached and maintained by repetitive practice using videotaped lessons.

Again, a scored interval reliability measure (noted earlier) was used to determine the reliability of the observer. Two videotapes were randomly selected and viewed three times by the observer on three separate occasions throughout the CAFIAS analysis. Reliability scores were calculated for each of the behaviours observed. (Refer to Table 3).

Recording procedures. Each of the 24 videotaped lessons was viewed and analysed using one of the three observation instruments. On completion the analysis was carried out again for the second system. On completion of the second analysis, the third system was used to yield a third series of codings. (Refer to Appendices A, B, and C).

After an analysis of the lessons by all three systems had been made, the frequency of each behavioural variable was determined. These scores were converted to percentages to standardize the data for statistical analysis. Those targeted behaviours which were not recorded, or represented a very minute part of the videotaped lessons, were eliminated prior to the statistical analysis.

Data Analysis

In the first place, a canonical correlation was calculated on the data to establish what kind of relationship existed between the ALT-PE-TB, the McLeish and the CAFIAS observation instruments (Hie, Jenkins, Bent, Hadlai Hull and Steinbrenner, 1975). This yielded the optimum correlation which could be obtained by

Table 3

Scored-Interval Reliability Agreement
Percentages for the CAFIAS Observation Instrument

Category	Reliability Scores							
	Videotape A				Videotape B			
	V_1	V_2	V_3	S-I	V_1	V_2	V_3	S-I
A. Teacher Behaviours								
Verbal:								
praise	3	3	3	100	3	2	3	75
empathetic response	1	1	1	100	3	3	2	75
asks questions	2	2	2	100	4	5	4	92
gives facts or opinions	12	11	10	90	3	3	3	100
demonstrates	1	1	1	100	2	2	2	100
constructive criticism	7	6	6	94	1	1	1	100
Non-verbal:								
uses whistle/claps for class attention	4	4	5	92	6	6	5	88
takes part in activity	-	-	-	--	11	10	12	90
B. Student Behaviours								
Verbal:								
predictable student answer	1	2	1	75	2	2	2	100
student-intiated talk	1	1	1	100	2	2	2	100
Non-verbal:								
practice drills	19	18	20	94	20	20	21	98
interpreting movement; games playing	27	28	26	96	31	31	30	98
C. Teaching Agency								
teacher	100	100	100	100	100	100	100	100
D. Class Organisation								
groups	54	54	56	98	32	31	30	96
whole class	46	46	44	97	68	69	70	98

V_1 - first viewing V_2 - second viewing V_3 - third viewing

S-I refers to the scored-interval reliability agreement between the V_1, V_2, and V_3 percentage scores for each recorded interval.

collapsing each lesson into a single score for each system.

Secondly, a multi-variate discriminant function analysis was performed to relate the judges' rankings of the videotaped lessons to the behaviour analyses of the three observation instruments. The purpose of this analysis was to discover whether the computer could calculate the different weights to assign to each of the variables of the three systems which would yield the same rankings of the lessons as the judges. If this operation was successful, then it would be possible to say which variables (for example, 'skill practice') influenced the particular judge and the relative importance of these variables (Hie et al, 1975).

Finally, a factor analysis was run on all the variables of the three systems together. This analysis enables us to establish how the 24 scores on the 38 variables can be explained in terms of a small number of factors or independent dimensions. The factors enable us to identify the ways in which these variables can be classified in meaningful ways. In other words, the factor analysis represents an attempt to discover which variables belong together and, more particularly, which combination of variables, if any, best measures teaching effectiveness.

CHAPTER FOUR

RESULTS

This chapter reports the results of the data analyses.

A Comparison of the Analyses of Each of the Three Observer Instruments

Since the videotaped lessons involved different lengths of time, it was necessary to convert the recorded interval scores to standard intervals. To do this, percentages were used for each of the three systems. From these percentage scores, three descriptions of each of the lessons were made.

1. The ALT-PE-TB Observation Instrument. The teachers spent an average of 24 per cent of the lesson in a managerial role, giving directions. Twenty-five per cent of the lesson was spent providing feedback to the students involved in physical activity, 17 per cent observing them and 14 per cent in an instructional role, lecturing.

   Several of the teacher behaviours were seen in a small number of lessons only, but, where these behaviours occurred, they accounted for a relatively high percentage of the particular lesson. For example, the behaviour of 'teacher participation' occurred in eight of the lessons and accounted for

less than five per cent of the teaching behaviours in six of these lessons. However, in badminton lesson six, it accounted for 32 per cent of the teacher behaviours.

In the ALT-PE dimension of the observation instrument, an average of 22 per cent was spent in management activity. Nineteen per cent of the lesson in instructional activity, 41 per cent of the lesson was spent in motor activity (skill practice, fitness activity or games playing) and, for 16 per cent of the time, the students were not engaged during motor activity tasks.

Ninety-eight per cent of the motor activity and the cognitive dimension of the 24 lessons was coded at the 'easy' difficulty level (see Appendix D).

Of the 44 behaviour variables defined in the ALT-PE-TB observation instrument, 14 behaviour variables were selected for analysis because they occurred more than two per cent of the time (see Table 4).

2. The McLeish Observation Instrument. The behavioural variables of the McLeish Observation Instrument, observed in the 24 videotaped lessons were totaled and averaged.

The predominant teacher behaviour was giving directions, or instructions. In the average lesson, 40 per cent of the time was spent giving commands related to the physical education content of the lesson; six per cent of the time was spent making requests related to the physical education content of the lesson; 19 per cent of the time was spent in giving commands unrelated to the physical education content; and two per cent of the time was spent in making management requests. An average of 13 per cent of the time was spent in the 'tact' behaviour dimension. The teachers positively reinforced the students 10 per cent of the time and spent four

Table 4

The ALT-PE-TB Analysis

Lesson	Behaviour Variables													
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
Volleyball														
1	19	15	16	4	10	1	3	32	14	22	9	46	1	0
2	20	12	23	0	15	5	6	16	11	29	13	35	1	0
3	28	22	3	5	0	4	13	25	22	38	2	9	0	16
4	13	24	38	6	0	0	0	18	15	13	26	30	9	0
5	14	6	21	3	13	4	14	25	6	16	14	39	0	19
6	16	15	4	7	3	1	4	13	15	16	17	36	0	0
Basketball														
1	23	24	17	8	0	9	15	4	22	25	36	8	2	0
2	36	21	9	6	13	1	2	13	13	30	16	19	14	0
3	25	16	26	0	12	3	13	4	13	35	32	9	3	0
4	11	14	31	5	0	0	0	9	8	11	8	51	6	0
5	25	25	25	2	11	0	0	12	14	26	11	29	11	0
6	34	22	23	0	10	0	0	11	20	46	8	14	2	0
Indoor Soccer														
1	30	26	14	0	13	2	2	13	26	25	12	11	5	7
2	13	11	16	16	0	9	27	8	6	12	31	13	0	32
3	26	24	28	5	8	0	0	8	20	29	26	14	5	0
4	27	19	19	3	18	0	0	16	20	31	0	41	3	0
5	25	35	8	6	0	0	0	25	30	22	16	16	0	0
6	18	21	0	0	21	8	23	9	14	24	15	4	2	35
Badminton														
1	16	11	34	1	15	5	2	16	8	26	22	31	5	0
2	21	31	21	7	0	8	0	12	16	21	17	23	15	0
3	32	14	14	0	12	0	2	25	8	26	28	20	3	0
4	17	11	24	4	2	10	6	26	7	19	40	24	4	0
5	24	10	32	0	9	0	0	25	6	24	0	54	4	0
6	11	20	12	5	20	9	6	28	6	10	22	22	3	0
Average Scores:	22	19	19	4	9	3	6	16	14	24	17	25	4	4

ALT-PE Dimension

A - Management time
B - Instructional time
C - Skill practice (motor)
D - Skill practice (indirect)
E - Fitness activity
F - Game (motor)
G - Game (indirect)
H - Non-engaged activity

Teacher Behaviours

I - Lecturing
J - Giving Directions
K - Monitoring
L - Feedback
M - Modeling
N - Officiating

per cent of the time listening to the students' comments and questions.

The predominant class behaviour was following instructions or listening to the teacher. This accounted for 94 per cent of the average lesson. A further four per cent of the time was spent asking the teachers' questions related to the physical education content of the lesson.

Certain of the lessons had a high percentage of behaviours that were rarely, if at all, coded in other lessons. For example, as already noted, instances were coded of the teacher reinforcing the students for 10 per cent of the average lesson time. Yet, this behaviour occurred 29 per cent of the time in volleyball lesson five, and 34 per cent of the time in basketball lesson four. The average coded score of the teacher variable, 'physical education request', was five per cent of the lesson, yet, in volleyball lesson six, this behaviour was coded 45 per cent of the time (see Appendix D).

Of the 17 behavioural variables coded using the McLeish observation instrument, 10 variables were selected, seven of which were teacher behaviour variables. These behaviours accounted for more than three per cent of the lesson time, except for one of the student behaviour variables. Negative student response towards the teacher or the assigned task accounted for one per cent of the lesson time. However, this behavioural variable was included because it accounted for more than three per cent of the time in four of the lessons (see Table 5).

3. The CAFIAS Observation Instrument. After calculating the average percentages for each of the teacher behaviours measured by the CAFIAS observation instrument, it was found

Table 5

The McLeish Analysis

Lesson	Behaviour Variables									
	A	B	C	D	E	F	G	H	I	J
Volleyball										
1	50	7	18	2	5	8	3	97	3	0
2	39	6	8	11	8	11	3	97	3	0
3	30	6	17	10	10	16	5	94	6	0
4	33	2	30	11	11	6	5	96	3	0
5	38	3	12	5	9	29	1	97	3	0
6	18	35	18	3	2	10	13	85	3	0
Basketball										
1	27	1	31	6	15	5	5	95	3	0
2	48	4	23	4	4	6	4	96	4	0
3	36	0	28	6	12	5	1	97	2	1
4	38	1	16	1	6	34	1	99	1	0
5	33	4	29	8	8	12	1	98	2	0
6	50	5	9	5	7	7	1	96	0	4
Indoor Soccer										
1	47	7	19	5	2	8	7	93	6	0
2	37	7	22	5	10	9	4	88	5	7
3	44	4	30	2	2	9	3	97	3	0
4	41	4	41	1	2	5	1	95	1	4
5	28	6	19	16	9	3	5	90	4	4
6	63	4	1	9	10	6	4	92	9	0
Badminton										
1	47	8	13	2	5	9	7	92	7	0
2	34	9	3	18	13	14	0	98	2	0
3	30	8	20	9	11	4	4	95	5	0
4	36	4	31	5	7	11	4	95	4	0
5	53	10	10	1	4	7	3	93	3	4
6	48	7	17	8	2	9	8	93	7	0
Average Scores:	40	6	19	6	7	10	4	94	4	1

Teacher Dimension

A - PE Command
B - PE Request
C - Management Command
D - Tact
E - Extended Tact
F - Positve Autoclitic (Reinforcement)
G - Submissive Autoclitic (Listening)

Student Dimension

H - Submissive Autoclitic (Listening/responding to commands)
I - Requests/questions
J - Negative response

that 21 per cent of the lesson was spent giving directions, 12 per cent of the time was spent giving information, and five per cent of the time giving positive reinforcement. The other teacher behaviours accounted for no more than three per cent of the lesson.

The average percentages for each of the student behaviours, as defined by the CAFIAS system, were calculated. For 26 per cent of the lesson, the students were involved in practice drills. Twenty per cent of the lesson was spent in physical education tasks that required some interpretation by the students. Each of the other student behaviours accounted for less than three per cent of the lesson time.

On the average, the teacher acted as the teaching agent 97 per cent of the time, compared to three per cent when a student acted as the teaching agent. Each of the observed classes was organised into groups for instruction an average of 35 per cent of the time and spent 65 per cent of the time in whole class instruction.

On the average, the teacher acted as the teaching agency 97 per cent of the time compared to three per cent of the time, when a student acted as the teaching agency. Each of the observed classes was organised into groups for instruction an average of 35 per cent of the time and spent 65 per cent of the time in whole class instruction.

Again, certain of the CAFIAS behavioural categories were coded frequently in certain lessons, but when the average score was calculated, they represented a small percentage of the lesson. For example, in volleyball lesson six, the class was organised into groups for 82 per cent of the lesson, whereas, the average percentage for this behaviour was 35 per cent (see Appendix D).

From the 25 behavioural variables coded using the CAFIAS observation instrument, 14 variables were selected. Only one of these variables, the behaviour category of 'predictable student verbal response', was coded for less than two per cent of the lesson. This variable was included, as no other student verbal behaviour occurred as frequently (see Table 6).

Data Analysis. Canonical correlations were calculated to examine the relationships which existed between the ALT-PE-TB, McLeish and CAFIAS observation instruments. Because canonical correlations of plus one were obtained in all three comparisons, this alerted the researcher to the fact that they were purely mathematical artifacts, since such results would not be expected in educational research, where complete control of all experimental variables, as in laboratory research, cannot be expected. Further, the pattern of variables identified by the computer made no sense and bore no relationship either, to our expectations, or, to anything in the available related literature.

Since no meaningful relationships between the three observation instruments were found after performing a canonical correlation study, a factor analysis was calculated. The variables of the three systems were grouped together as one and the analysis calculated to discover if the behavioural variables could be reduced to a smaller number of factors to account for the relationship between the scores on the three observation instruments. Factor analysis is a method which helps to identify the different ways variables classify themselves together in actuality, regardless of how the authors of the systems might describe them.

Table 6

The CAFIAS Analysis

Lesson	Behaviour Variables													
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
Volleyball														
1	3	2	15	1	25	4	2	2	16	26	54	46	100	0
2	6	2	10	2	25	3	1	2	29	20	42	58	100	0
3	10	0	18	2	24	2	0	0	28	15	0	100	100	0
4	3	2	14	6	16	1	5	1	30	20	0	100	100	0
5	15	1	8	3	17	0	2	0	18	31	59	41	86	14
6	5	3	14	3	16	0	0	2	27	25	82	18	56	44
Basketball														
1	2	1	14	2	21	3	0	1	23	28	49	51	100	0
2	3	3	11	11	19	0	0	2	30	13	21	79	100	0
3	2	1	9	8	27	3	0	0	35	13	20	80	100	0
4	17	0	7	6	12	3	11	1	34	9	0	100	100	0
5	6	5	10	12	18	0	0	4	26	13	13	87	100	0
6	3	4	16	4	28	3	1	2	26	9	10	90	100	0
Indoor Soccer														
1	4	5	10	8	24	2	5	2	32	6	8	92	100	0
2	7	3	13	0	17	0	0	0	7	49	38	62	100	0
3	1	2	16	9	21	1	0	2	39	5	43	57	100	0
4	3	3	9	3	28	1	0	3	30	12	75	25	100	0
5	1	3	21	6	19	2	0	3	20	18	39	61	100	0
6	3	3	18	3	23	0	0	1	20	27	2	98	82	18
Badminton														
1	4	4	9	2	29	1	0	1	17	29	59	41	100	0
2	8	3	12	12	20	3	0	1	14	24	47	53	100	0
3	2	3	13	2	25	2	0	2	36	11	60	40	100	0
4	4	0	8	4	21	3	0	0	28	31	16	84	100	0
5	4	6	9	3	24	3	0	2	26	18	64	36	100	0
6	3	4	7	2	7	1	15	2	25	31	32	68	100	0
Average Scores:	5	3	12	5	21	2	2	1.5	26	20	35	65	97	3

Teacher Dimension

Verbal:

A - Praising
B - Questioning
C - Giving facts/opinions
D - Demonstrating
E - Giving directions/orders
F - Constructive criticism

Non-Verbal:

G - Teacher takes part in student activity

Student Dimension

Verbal:

H - Predictable student response

Non-Verbal:

I - Robot-like responses eg. practice drills
J - Interprets movement eg. game playing

Organisation

K - Class work in groups
L - Class work all together
M - Teacher as teaching agency
N - Student as teaching agency

Four factors accounted for 53.5 per cent of the total variance. Factor one, which could be described as a student-centred as compared to a teacher-centred approach, accounted for 17.6 per cent of the variance.

Table 7

Factor One

"Student versus Teacher-centred Approach"

System	Student-centred Behaviours	Saturation	System	Teacher-centred Behaviours	Saturation
CAFIAS	Movement interpretation	.69	McLeish	Class listening; following directions	-.77
ALT-PE-TB	Skill practice (indirect)	.68	ALT-PE-TB	Management time	-.51
ALT-PE-TB	Game (indirect)	.62	CAFIAS	Predictable student response (verbal)	-.51
McLeish	Teacher listening; responding to students	.57			
McLeish	PE request (teacher)	.56			

According to this factor, student-centred lessons, where the teacher maintained a low profile and provided the opportunity for students to participate in motor tasks and discussion, were opposed to lessons, where the teacher was the centre of attention and the students had a very passive role.

The ability to keep students actively involved on tasks directly related to the objective of the lesson has consistently distinguished between more and less effective teachers (Berliner and Tikunoff, 1976; Brophy and Everston, 1976; Mcdonald, 1978;

Graham and Heimerer, 1981). Therefore, effective physical education teachers are able to keep students actively involved in motor skill practice related to the objective of the lesson and provide feedback on the students' performance of these motor skills (Siedentop, Birdwell and Metzler, 1979). Further, Cheffers (1980) theorised that one of the characteristics of an effective teacher was the ability to develop a positive classroom climate.

These teacher behaviours are grouped together in factor two of the factor analysis. They are in contrast to teacher behaviours that have been suggested as characteristic of ineffective physical education teaching (Siedentop et al, 1979). This factor accounted for 13.3 per cent of the variance.

Table 8

Factor Two

"Effective Teaching"

System	Effective PE teaching	Saturation	System	Ineffective PE teaching	Saturation
ALT-PE-TB	Feedback	.81	ALT-PE TB	Non-engaged activity time	-.68
McLeish	Positive Reinforcement	.64	ALT-PE-TB	Instructional time	-.59
CAFIAS	Praise	.64	ALT-PE-TB	Management time	-.58
ALT-PE-TB	Skill practice	.46	ALT-PE-TB	Giving directions	-.51
ALT-PE-TB	Teacher participation	.45	CAFIAS	Gives facts or opinions	-.48

Factor three accounted for 12 per cent of the variance.

In this factor, the behavioural variables displayed in game situations, where the teacher acted as the referee and any student-teacher interaction was limited to that of an official-to-player nature, were grouped together. These were opposed to behaviours displayed in verbal interchanges.

Table 9

Factor Three
"Games Activity versus Talk"

System	Limited Verbal Interaction	Saturation	System	Free verbal Interaction	Saturation
ALT-PE-TB	Game (indirect)	.65	McLeish	Teacher listening; responding to students	-.61
ALT-PE-TB	Game (motor)	.55	McLeish	PE request	-.60
McLeish	Tact	.48	ALT-PE-TB	Feedback	-.47
CAFIAS	Class working together	.45	ALT-PE-TB	Asking questions	-.47

Factor four accounted for 10.5 per cent of the variance. Indirect teacher control was compared to direct teacher control. Lesson content, such as fitness activities and other kinds of behaviours connected to command-style teaching were compared to motor skill activity in a task-oriented setting (see Table 10).

A comparison was made of the variance associated with the variables of each of the three observation instruments on the four factors reported. The total variance associated with the ALT-PE-TB observation instrument was 42.77 per cent compared to 29.45 per cent associated with the CAFIAS instrument and 27.73 per cent associated with the McLeish instrument (see Table 11).

Table 10

Factor Four

"Direct versus Indirect Teacher Control"

System	Indirect Teacher Control	Saturation	System	Direct Teacher Control	Saturation
ALT-PE-TB	Game (indirect)	.45	McLeish	PE command	-.77
			ALT-PE-TB	Fitness	-.61
			CAFIAS	Gives directions	-.48

Table 11

Variance Associated with Each of the Three Systems

System	Factor 1	Factor 2	Factor 3	Factor 4	Totals
ALT-PE-TB	2.295 (12.55%)	2.576 (14.08%)	1.608 (8.79%)	1.345 (7.35%)	7.824 (42.77%)
McLeish	1.791 (9.79%)	.816 (4.46%)	1.220 (6.67%)	1.246 (6.81%)	5.073 (27.73%)
CAFIAS	2.158 (11.80%)	1.147 (6.27%)	1.233 (6.74%)	.849 (4.64%)	5.387 (29.45%)
Totals	6.244 (34.14%)	4.539 (24.81%)	4.161 (22.20%)	3.440 (18.80%)	18.284 (99.95%)

In order to bring out the comparisons between the three systems, the percentage variance was shown of each factor explained by each of the systems. These figures are, of course, derived from Table 11.

1. Factor 1: Teacher-centred versus student-centred teaching.

   Of this factor, the coding systems contributed to the variance as follows:

ALT-PE-TB 36.76% McLeish 28.68% CAFIAS 34.56%

2. Factor 2: Teaching effectiveness.

Of this factor, the coding systems contributed to the variance as follows:

ALT-PE-TB 56.75% McLeish 17.98% CAFIAS 25.26%

3. Factor 3: Games activity versus talk.

Of this factor, the coding systems contributed to the variance as follows:

ALT-PE-TB 38.64% McLeish 23.32% CAFIAS 29.58%

4. Factor 4: Direct versus indirect teacher control.

Of this factor, the coding systems contributed to the variance as follows:

ALT-PE-TB 39.10% McLeish 36.22% CAFIAS 24.68%

If the percentages are totalled in the other direction, an estimate of the meaning of the three different analytical systems is obtained, comparing each with the other two. (Allowance was made for the McLeish instrument which has only 10 measures compared to 14 measures for each of the other two instruments). The three coding systems were effective in accounting for the common variance of scores on the 38 variables, according to the ratios comparing the ALT-PE-TB instrument 100: to the McLeish instrument 91: to the CAFIAS instrument 69. This suggests that in attempting to measure effective teaching the ALT-PE-TB coding system is to be preferred to the other two.

A Comparison between the Analyses of the Observer Instruments and the Judges' Rankings

A discriminant function analysis was carried out to relate the judges' rankings with the three observation instruments. The purpose of this analysis was to discover the different weights to assign to each of the variables of the three

systems, to provide the maximum possible separation between the lessons of each of the four floor games assigned to each of the six ranks supplied by the judges.

Once again, correlations of plus one were found - this time between the judges' rankings and weighted scores on the 38 variables. This alerted the researcher to the fact that the discriminant functions were almost certainly mathematical artifacts. Furthermore, the pattern of variables identified by the computer, and their relative weights, made no sense and bore no relation either to our expectations of how the variables relate to each other or to anything in the available related literature about effective teaching.

In addition, an analysis of variance was carried out with each of the judges separately as the independent variable, to determine if there was any relationship between the judges' rankings and any one of the 38 behaviour variables. No such relationship was found.

The Agreement between Each of the Judge's Rankings

Each of the six judges successively ranked each of the six lessons in the four floor game activities from the most to the least effective, on the basis of their intuitive judgement. An average of 26 per cent agreement was reached between the judges in these rankings. This cannot be regarded as better than chance.

A Monte Carlo Simulation of this part of the enquiry was carried out. (Four sets of six cards, one to six, were shuffled and drawn, without replacements, to provide the rank order values for the six "judges"). This was done twice. The first time there was an average agreement of 21 per cent and, on the second simulation, there was an average agreement of 15.3 per

cent. These scores closely compare to the agreement of 26 per cent reached by the six expert physical educators (see Appendix E).

The Relationship between Each of the Four Floor Games

An analysis of variance was carried out to determine if any differences existed between the behavioural analyses of the four different activities represented in the four groups of floor games. No significant differences were found (see Table 12).

Summary

The raw scores of each of the behavioural analyses provided by the three observation instruments were converted to percentages in order to standardize the duration of the 24 lessons.

Two statistical analyses were carried out to examine what relationship existed between the three observation instruments. As a result of a factor analysis, four factors were identified which accounted for 53.5 per cent of the variance.

A comparison was made between each of the 38 behavioural variables of the three observation instruments and the rankings by each of the six judges. A discriminant function analysis indicated that no meaningful relationships existed. This finding was supported by the results of an analysis of variance.

In comparing the rankings which the judges allocated to each of the six lessons in the four groups of floor games, very little agreement was found. The average agreement was only 26 per cent. This is no better than chance.

Analysis of variance revealed no significant difference between the behavioural structure of the lessons in the four groups of floor games.

Table 12

Significance of the Average Differences
Between Activities on the 38 Behavioural Variables

Behavioural Variables	Significance of Differences Between Activities
ALT-PE-TB:	
Management time	0.306
Instructional time	0.163
Skill practice	0.450
Skill practice (indirect)	0.602
Fitness	0.869
Game	0.462
Game (indirect)	0.582
Non-engaged activity time	0.115
Lecturing	0.390
Giving directions	0.373
Monitoring	0.596
Feedback	0.213
Modelling	0.186
Officiating	0.119
McLeish:	
PE Command	0.700
PE Request	0.341
Management command	0.574
Tact	0.865
Extended tact	0.225
Positive reinforcement	0.484
Submissive listening (teacher)	0.408
Submissive autoclitic (class)	0.204
Class questions	0.143
CAFIAS:	
Praise	0.458
Questions	0.586
Demonstrates	0.272
Gives directions	0.820
Constructive criticism	0.461
Teacher participation	0.422
Predictable student verbal response	0.555
Practice drills	0.559
Interprets movement	0.310
Group organisation	0.522
Whole class organisation	0.287
Student teaches	0.285
Teacher is teaching agency	0.564

For statistical significance p should be $\leq .05$

CHAPTER FIVE

SUMMARY, CONCLUSIONS, AND SUGGESTIONS FOR FUTURE RESEARCH

The purpose of this study was to compare the behavioural analyses of three objective observation instruments with each other and with intuitive evaluations of 24 videotaped physical education lessons. The study focused on the following questions:

1. What relationship exists between the three objective observer systems?
2. What relationship exists between the analytical observer systems and the intuitive observer system?
3. What agreement exists between the judges using the intuitive method of evaluation?
4. What differences, if any, exist between the behavioral structure of the four types of games lessons?

The conclusions and discussion of the findings to each of these questions are presented in this chapter.

What Relationship exists between the Three Objective Observer Systems?

Two statistical analyses were reported which reveal the relationship between the three observation instruments.

No meaningful relationships could be identified between any of the observation instruments as a result of calculating canonical correlations.

A factor analysis was made. Four factors were found to account for 53.5 per cent of the total variance of the three coding instruments. These four factors grouped the behaviours in the following patterns.

In factor one, variables where the teacher was the centre of attention (such as giving directions) were grouped together. These variables were in contrast to variables where the student was the focus of attention (such as in skill practice). This factor accounted for 17.6 per cent of the variance.

Factor two grouped together variables associated with effective physical education lessons, in contrast to those considered to be characteristic of less effective lessons. Providing feedback to students as they engage in motor activity and generating a positive classroom climate have been identified in the literature as effective physical education teacher behaviours (Siedentop, Birdwell and Metzler, 1979; Cheffers, 1980). Spending a large proportion of the lesson in management activities and having students waiting about and otherwise not engaged in motor activity have been theorised as features of less effective physical education lessons (Siedentop et al, 1979). The second factor is identified, therefore, as the effectiveness factor and accounted for 13.3 per cent of the variance.

After calculating the variance associated with each of the three instruments on factor two, it was found that the most sensitive instrument in measuring the behavioural categories identified as representing effective physical education lessons

was the ALT-PE-TB instrument.

In factor three, a contrast was drawn between teacher and student interaction behaviours. Interaction patterns associated with referee-player roles were contrasted with interaction patterns associated with free discussion. This factor accounted for 12 per cent of the variance.

Factor four accounted for 10.5 per cent of the variance. Direct teacher control behaviours, characterised by command-style teaching, were compared to indirect teacher control behaviours, characterised by task-oriented teaching.

Within each of the four factors, behavioural categories of each of the observer instruments were included. This demonstrates that the three observer instruments agree in certain common measurements of the 24 videotaped physical education lessons. The total common variance explained is equal to 53.5 per cent. Thus 46.5 per cent of the variance is associated with factors specific to the 38 variables and to errors of measurement. It is a point that the patterns of relationship do not lend a great deal of support for Siedentop's views. However, this could well be that Cheffers explores one dimension of effective teaching while Siedentop explores the other (the affective domain in the first, the psychomotor domain in the second). McLeish's analysis seeks to comprehend both, but is not too successful in the case of these physical education lessons. (This is probably due to the fact that the McLeish instrument was devised to encompass the communication process in all its diversity and the physical education lessons are quite limited in this respect).

<u>What Relationship exists between the Analytical Observer Instruments and the Intuitive Observer System</u>?

Again two statistical analyses were calculated to discover what relationships existed between the judges' rankings of the lessons in each of the four floor game categories and the three observation instruments.

A discriminant function analysis was first calculated. It gave a weighted score to each variable of the three observation instruments to discover whether it was possible to match each of the lessons to the ranks assigned by each of the judges. Unfortunately, the patterns of relationships between the variables of the three observation instruments were totally meaningless and had to be judged as mere statistical artifacts.

To ensure that no meaningful relationship between the variables of the three observation instruments and each of the judge's rankings had been missed, an analysis of variance was made. No significant relationship was found with any of the variables. Therefore, it can be concluded that none of the six judges consistently used any one of the behavioural variables of the observation instruments in ranking the 24 videotaped lessons.

What Agreement exists between Each of the Judges using the Intuitive Observer System?

The judges achieved on average a 26 per cent agreement in their rankings of the physical education lessons. This agreement was not large enough to be explained in any other terms than by chance.

Because the judges showed little, if any, agreement in evaluating the effectiveness of the videotaped lessons, their rankings could not be used to validate the three observation instruments.

What Differences exist between the Behavioural Structure of the Four Types of Games Lessons?

An analysis of variance was carried out to determine if there was any difference in the structure of the four types of games lessons, as described by each of the three observer instruments. As far as these 24 videotaped lessons of volleyball, basketball, indoor soccer and badminton were concerned, there was no significant difference. Since they were games lessons set in a gymnasium, this was to be expected.

Conclusion

Of the four systems used to evaluate the 24 physical education lessons, the three analytical observer systems were able to reach common measurements of effective physical education teaching. The judges were not able to reach any agreement other than that which could be explained by chance.

Of the three analytical observer systems, the ALT-PE-TB instrument was the most sensitive and is to be preferred to the other two systems in attempting to measure effective physical education teaching. Perhaps some combination of the three scales, dropping duplicated and unsatisfactory variables, would be the most effective means of evaluating lessons in this area.

Suggestions for Future Research

As noted earlier in the review of related literature, there is a need to develop a consistency in the use of measuring instruments and the particular variables to be examined in physical education teaching effectiveness studies. Two variables — the amount of content covered and the time students spend engaged on academic tasks — have been consistently correlated

with student achievement in other subject areas. The ALT-PE-TB observation instrument is particularly sensitive to these variables. Furthermore, of the three coding systems, it was the most sensitive in measuring the behavioural variables identified by the factor analysis as characteristic of effective physical education teaching of the 24 videotaped lessons.

Therefore, future studies should use a single observation instrument, such as the ALT-PE-TB instrument, to develop a common data base upon which empirical statements concerning effective physical education teaching can be made.

A feature of the 24 videotaped lessons was the relatively small range of effective teaching, as evidenced by the difficulty the judges had in ranking them. Therefore, instead of sampling isolated lessons of a large number of physical education teachers, future studies should attempt to analyse a much smaller sample of teachers in a more intensive manner. This could take several forms.

Firstly, studies involving the analysis of lessons covering a complete unit of work using the ALT-PE-TB observation instrument could be made, to discover, for example, if the behavioural structure of physical education lessons change over the course of a unit.

Secondly, intervention studies, where the investigator targets certain behaviours for modification over a period of time based on the analysis of a coding system, such as the ALT-PE-TB instrument, could be made.

Thirdly, the development of units of physical education lessons which incorporated measures of student learning and which could be analysed being taught, would provide empirical

evidence either to support or reject the theory that the mediating link between teacher behaviours and student achievement in physical education is the amount of time students spend on tasks related to the objective of the lesson.

REFERENCES

Anderson, W.G. Videotape Data Bank. Journal of Physical Education and Recreation, 1975, 46: 31-34.

Anderson, W.G., & Barrette, G.T. What's going on in the gym: Descriptive studies of physical education classes. Motor Skills: Theory into Practice, 1978, 1: 25-38. (Monograph)

Armento, B.J. Teacher verbal cognitive behaviour related to student achievement on a social science concept test. Dissertation Abstracts International, 1977, 38: 3406A.

Barr, R.C. Instructional pace differences and their effect on reading acquisition. Reading Research Quarterly, 1973, 9: 526-554.

Batchelder, A.S., & Cheffers, J.T. CAFIAS: An interaction analysis instrument for describing student-teacher behaviours in different learning settings. Paper presented at the International Conference, Quebec, July, 1976.

Berliner, D.C. Impediments to the study of teacher effectiveness. Journal of Teacher Education, 1976, 27: 5-11.

Berliner, D.C., & Tikunoff, W.J. The California Beginning Teacher Evaluation Study: Overview of the ethnographic study. Journal of Teacher Education, 1976, 27: 24-30.

Birdwell, D. The effects of modification of teacher behaviour on the academic learning time of selected students in physical education. (Doctoral dissertation, Ohio State University, 1980). Dissertation Abstracts International, 1980, 51: 8022239.

Bloom, B.S. Human Characteristics and School Learning. New York: McGraw-Hill, 1976.

Brophy, J.E. Reflections on research in elementary schools. Journal of Teacher Education, 1976, 27: 31-34.

Brophy, J.E., & Everston, C.M. Learning from Teaching. Boston: Allyn & Bacon, 1976.

Carroll, J.B. A model of school learning. Teachers College Record, 1963, 64: 723-732.

Chang, S.S., Raths, J. The school's contribution to the cumulating deficit. Journal of Educational Research, 1971, 64: 272-276.

Cheffers, J. T. Observing teaching systematically. Quest, 1977, 28: 17-28.

Cheffers, J.T. Systematic observation in teaching. AEISP Yearbook, 1978, 2: 7-30.

Cheffers, J.T., Mancini, J., Martinek, T., Amidon, E., Batchelder, A., Locke, R., Nolen, W., Interaction Analysis: An Application to Non-verbal Activity (2nd ed.). Association for Productive Teaching, Boston, 1980.

Coker, H., Medley, D.M., Soar, R.S. How valid are expert opinions about effective teaching? Phi Delta Kappan, 1980, 62: 131-134.

Comber, L., & Keeves, J.P. Science Education in Nineteen Countries, New York: John Wiley & Sons, 1973.

Cooper, S.H. Hope for the future: A view of research in teacher effectiveness. Quest, 1977, 28: 29-37. (Monograph)

Costello, J., Laubach, S.A. Student behaviour. What's going on in the gym: Descriptive studies of physical education classes. In Motor Skills: Theory into Practice, Ed.s Anderson, W,G., & Barrette, G.T., 1978, 1: 39-47.

B.C. Elementary School Physical Education Curriculum Guide. Province of B.C., Ministry of Education, Victoria. Revised ed., 1975.

Secondary Physical Education Curriculum and Resource Guides Province of B.C., Ministry of Education, Victoria. 1980.

Dunkin, M.J., & Biddle, B.J. The Study of Teaching. New York: Holt, Rinehart and Winston, 1974.

Fisher, C.W., Berliner, D.C., Filby, N.N., Marliave, R.S., Cahen, L.S., Dishaw, M.M., & Moore, J.E. Teaching and learning in the elementary school. A Summary of the BTES, Far West Laboratory for Educational Research and Development. San Francisco, CA, 1978.

Fishman, S.E., & Anderson, W.G., Developing a system for describing teachers. Quest, 1971, 17: 9-16.

Flanders, N.A. The changing base of performance-based teaching. Phi Delta Kappan, 1974, 55: 312-315.

Gage, N.L. Handbook of Research on Teaching. Chicago: Rand McNally, 1963.

Gage, N.L. Analyzing Teaching Behaviour. Reading: Addison-Wesley, 1969.

Gasson, I.S. The development of an observational instrument to record selected teacher-pupil behaviours in primary school physical education. (Doctoral dissertation, Ohio State University, 1971). Dissertation Abstracts International, 1971, 24: 2702B.

Goldberger, M. Studying your teaching behaviour. Journal of Health, Physical Education and Recreation, 1974, 45: 33-36.

Good, T.L. Teacher effectiveness in the elementary school. Journal of Teacher Education, 1979, 30: 52-64.

Good, T.L., Biddle, B.J., & Brophy, J.E. Teachers Make a Difference. New York: Holt, Rinehart & Winston, 1975.

Graham, G., & Heimerer, E. Research on teacher effectiveness: A summary with implications for teaching. Quest, 1981, 33: 14-25.

Grant, B.C. The relationship between specialist teacher training and effective physical education teaching in the elementary school. Unpublished M.A. thesis, University of Victoria, 1979.

Hall, R.V., Delquardi, J., & Harris, J. Opportunity to respond: A new focus in the field of applied behavioural analysis. A paper read at the Conference of the Midwest Association for Behavioural Analysis, Chicago, May, 1977.

Hickey, B. The inter-relationship of teacher experience, student behaviour and effective teaching in secondary school physical education. Unpublished M.A. thesis, University of Victoria, 1980.

Hie, J., Jenkins, B., Bent, J., Hadlai Hull, K.L., & Steinbrenner, D. (Ed.s), Statistical Package for the Social Sciences. (2nd ed.). New York: McGraw-Hill, 1975.

Holding, M.V. Teaching effectiveness in adult physical education recreation classes. Unpublished M.A. thesis, University of Victoria, 1981.

Husen, T. International Study of Achievement in Mathematics: A Comparison of Twelve Countries. New York: John Wiley & Sons, 1967.

Locke, L. Research on teaching physical education: New hope for a dismal science. Quest, 1977, 28: 2-16.

Locke, L. Learning from teaching. In Theory into Practice J.J.Jackson (ed.). Victoria: University of Victoria, 1978.

Locke, L., & Nixon, J.E. Research of teaching physical education. In Second Handbook of Research into Teaching, Travers, R.M., (ed.). Chicago: Rand-McNally, 1973.

McDonald, F.J. Report on Phase II of the Beginning Teacher Evaluation Study. Journal of Teacher Education, 1976, 27: 39-42.

McLeish, J. Learning in groups: Facilitation and inhibition processes. Bulletin of the British Psychological Society, 1976, 29: 7-15.

McLeish, J. Effective teaching: A new analysis. British Journal of Teacher Education, 1978, 4: 215-222.

McLeish, J. Effective teaching in physical education. Unpublished manuscript, University of Victoria, 1981.

McLeish, J. The Development of Modern Behavioural Psychology. Calgary: Detselig Enterprises Ltd., 1981.

Medley, D. Teacher Competence and Teacher Effectiveness. Washington, D.C.: American Association of Colleges for Teacher Education, 1977.

Medley, D., & Mitzel H. The scientific study of teacher behaviour. In Theory and Research in Teaching, Bellack, A.A., (ed.). New York: Teachers College, 1963.

Metzler, M.W. The measurement of academic learning time in physical education. (A doctoral dissertation, Ohio State University, 1979). Dissertation Abstracts International,

1979, 48: 7924660.

Nygaard, G. Interaction analysis of physical education classes. Research Quarterly, 1975, 46: 351-357.

O'Sullivan, M. A descriptive analytical study of student teacher effectiveness and student behaviour in secondary school physical education. Unpublished M.A. thesis University of Victoria, 1980.

Peterson, D.L. Direct instruction reconsidered. In Research on Teaching: Concepts, findings and implications, Peterson, D.L., & Walberg, W.H., (ed.s). Berkeley, CA: McCutchan, 1979.

Rosenshine, B.V. Evaluation of classroom instruction. Review of Educational Research, 1970, 40: 279-300.

Rosenshine, B. Classroom interaction. In The Psychology of Teaching Methods, 75th Yearbook of the National Society of Education, Chicago, 1976.

Rosenshine, B.V., & Berliner, D.C. Academic engaged time. British Journal of Teacher Education, 1978, 4: 1, 5-16.

Rosenshine, B.V., & Furst, N. Research in teacher performance criteria. In Research in Teacher Education, Smith, B.O., (ed.). Englewood Cliffs, N.J.: Prentice-Hall Inc., 1971.

Rosenshine, B.V., & Furst, N. The use of direct observation to study teaching. In Second Handbook of Research on Teaching, Travers, R.M., (ed.). Chicago: Rand McNally, 1973.

Schalock, D. From research to practice: The dilemma for teacher education. A paper read at the Second Annual Teaching-Learning Colloquium, University of Alabama, Birmingham, August, 1981.

Siedentop, D., Birdwell, D., & Metzler, M. A process approach to measuring teacher effectiveness in physical education. Paper read at the AAHPER Research Symposium, New Orleans, March, 1979.

Siedentop, D., & Hughley, C. Ohio State University teacher behaviour rating scale. Journal of Physical Education and Recreation, 1975, 46: 45, 112-116.

Simon, A., & Boyer, E., (ed.s). Mirrors for Behaviour: An Anthology of Observation Instruments. Philadelphia: Research for Better Schools, 1970.

Skinner, B.F. About Behaviourism. New York: Alfred A. Knopf, 1974.

Whaley, G. The effect of daily monitoring and feedback to students and teachers on ALT-PE. (Doctoral dissertation, Ohio State University, 1980). Dissertation Abstracts International, 1980, 51: 8022240.

Walker, D.F., & Schaffarzick, J. Comparing curricula. Review of Educational Research, 1974, 44: 83-112.

Yerg, B.J. Relationships between teacher behaviour and pupil achievement in the psychomotor domain. (A doctoral dissertation, University of Pittsburgh, 1977). University Microfilms, 1977, No. 77-21, 229.

APPENDIX A

THE ALT-PE-TB OBSERVATION INSTRUMENT

THE ALT-PE-TB OBSERVATION INSTRUMENT
DECISION LOG

Three students are randomly selected from the physical education class. These students and the teacher are observed systematically throughout the entire lesson

The lesson is divided into equal measures of time. The coder observes the behaviour of one of the selected students and the teacher for six seconds, before the observed behaviour is recorded over the next six seconds. Accuracy in this procedure is ensured by the use of a cassette recorder equipped with an earjack. Recorded cues tell the coder exactly when to 'observe' and when to 'record'.

The teacher's behaviour is coded at every interval. However, only one of the selected students is coded at each interval. To ensure that no one student is observed more frequently than the others, they are coded in an exact order. For example,

interval 1	student 'A' and the teacher are observed
interval 2	student 'B' and the teacher are observed
interval 3	student 'C' and the teacher are observed
interval 4	student 'A' and the teacher are observed
interval 5	student 'B' and the teacher are observed
interval 6	student 'C' and the teacher are observed

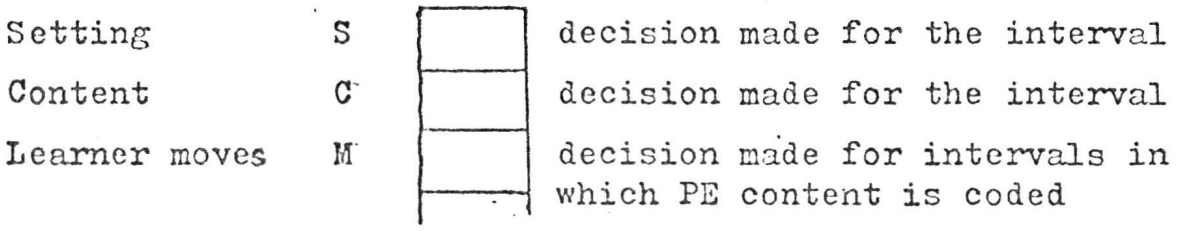
and so on.

Five major category decisions are made in coding the

behaviours observed. These are:

1. Setting. The setting categories describe the basic format for instruction within the lesson. These categories reflect Mosston's (1966) spectrum of teaching styles.
2. Content. The content categories are divided into two main groups. Those that reflect a non-academic focus and those that reflect a physical education content. Student time in physical education related content is the first level of the ALT-PE-TB instrument.
3. Learner involvement. The third decision concerns the behaviour of the learner. There are three categories that reflect differing forms of learner engagement and three categories that reflect non-engagement. Whether the student is actually engaged in physical education related activity is the second level of the ALT-PE-TB instrument.
4. Difficulty level.If the coder observes the particular student engaged in activity, the decision is made concerning the level of difficulty of the activity. This decision is based on the success rate of the learner. This is the third level of the ALT-PE-TB instrument.
5. Teacher behaviour. The final decision requires the coder to determine which behavioural definition best describes the teacher's activity during the observed interval.

An individual interval on the coding sheet looks like this:



Difficulty

Teacher Behaviour

decision made for intervals in which PE content and learner engagement is coded.

decision made for the interval

A detailed description of the ALT-PE-TB behavioural dimensions follows.

SETTING: Assign a setting and content code for each interval.

Direct Instruction (D)	Teacher controls instruction, focus, and pacing of instruction (includes attendance checks, announcements, other preliminary classwork, etc.).
Task (T)	Instruction defined by task - multiple station and/or multiple task.
Reciprocal (R)	Students in pairs for instruction and feedback to each other.
Group (G)	Same functions as reciprocal with larger group.
Guided Discovery (GD)	Teacher leads students toward predetermined goal through series of sequenced prompts.
Problem Solving (P)	Teacher controls instruction through sequenced problems in which alternative solutions are possible.

CONTENT-GENERAL

Wait (W)	Period when student has performed the necessary activity (after transition or management) and is

	waiting for the next. (Waiting in line or for the next turn to practice is coded on the learner move level, not on the content level.)
Transition (T)	Time devoted to class business that is related to instructional activity (includes managerial activities related to instruction, such as equipment change, lining up, selecting teams, etc.).
Management (M)	Time devoted to class business that is unrelated to instructional activity, such as taking attendance.
Break (B)	Intentional periods of no activity to rest students, drink water, etc.
Non-academic Instruction (N)	Activities which fall outside the narrow domain of focused instruction, such as rapport building.
CONTENT-PE	
Skill Practice (P)	Direct participation in drills and other activities in which the primary goal is individual skill development.

Scrimmage (S)	Controlled group practice in which instruction and feedback are frequent.
Game (G)	Practice under game conditions.
Fitness (F)	Repetitive activities for fitness development such as calisthenics, running laps, weight lifting, etc. Also warm up and cool down activities such as stretching.
Other Motor Activity (O)	Motor activity unrelated to specific goals of the day's instruction.
Knowledge Focus (K)	Activities in which teacher is giving verbal instructions or demonstration about skill, fitness, historical information, strategies, rules, etc. as the focus.
Social Behavior (B)	Activities in which social behavior, attitudes, etc. are the focus.

LEARNER MOVES: Assign a learner moves code for every interval in which one of the content-PE codes is assigned.

Engaged, Motor Response (M)	Student is performing a skill including supportive motor

	involvement such as screening in basketball, backing up in baseball, moving to block in volleyball.
Engaged, Indirect Participation (I)	Student is in an activity but is not directly involved with the immediate action such as right fielder during pitch, basketballer who does not move down court during a fast break, servicing another player by spotting, feeding balls, or other supportive activities unless that is the main focus of the instruction.
Engaged, Cognitive (C)	Cognitive involvement related to instruction, such as listening, questioning, verbal responding, or thinking about the activity (as in problem solving).
Not Engaged, Interim (NI)	Any non-instructional activity that is a natural part of the practice activity (such as changing equipment, changing sides of a court, retrieving the ball, etc.).

Not Engaged, Waiting (NW)	Time during activity when student is waiting for help or waiting to participate again. (Student does not have an opportunity to respond, such as in line or a substitute or waiting for equipment to be repaired, etc.).
Not Engaged, Off Task (NO)	Periods when student is inappropriately disengaged from the practice, including socializing, daydreaming, misbehaving and failing to respond when given the opportunity.

DIFFICULTY LEVEL: Assign a level of difficulty for every interval in which one of the engaged codes are entered on the learner moves level.

Easy (E)	Few errors are made and student performs appropriately with little effort, experiencing success frequently. In game or scrimmage situations the student anticipates properly, moves and responds to situations appropriately and fulfills the roles demanded by the game.

Medium (M)	Any performance that is other than easy or hard. When in doubt, code the response as medium.
Hard (H)	Many errors are made and student appears to be unable to perform appropriately, experiencing lack of success frequently. Chances of success are not much better than luck.

TEACHER BEHAVIORS: Assign a teacher behavior for every interval.

Lecturing (LE)	Teacher gives facts or opinions about content or procedures. Teacher may lecture to one or many students.
Giving Directions (G)	Teacher directs a student or group of students to perform a task. The direction may be verbal or non-verbal.
Listening (L)	Teacher listens to a student's question or response. Teacher may be listening to one student or a group of students and must be silent for the full interval.

Asking Questions (AQ)	Teacher asks questions about content or procedures with the intent of obtaining a response. Rhetorical questions should not be coded in this category. Questions may be individual or group oriented.
Answering Questions (Q)	Teacher answers a student or group of student's questions.
Nonfunctional (NF)	Teacher is engaged in behavior unrelated to the content or procedures of the practice. Talking to a visitor or an observer are prime examples.
Monitoring (M)	Teacher observes the practice without reacting verbally to the behaviors of individuals in the practice. The teacher's eyes must be directed toward at least one individual in the practice to code in this category. Watching a practice game is an example.
Maintenance (MT)	Teacher is engaged in activities that are indirectly related to the practice objectives. These

include such behaviors as checking attendance, putting away or handing out equipment, or administering first aid.

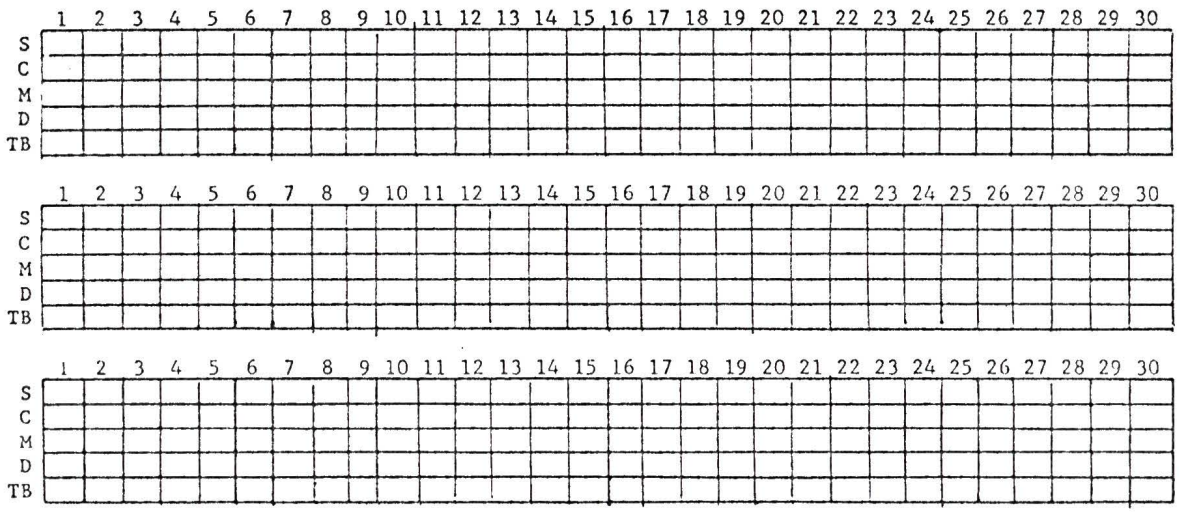
Hustles (H) — Teacher uses verbal statements or gestures to activate or intensify the motor performance of the student(s). Examples include such statements as "Run, run, run," "Move," "Go, go." Encouraging clapping of the hands would be considered a hustle. The tone of the voice and general level of enthusiasm are extremely important in this category. Do not mistake these statements and gestures as negative or corrective skill feedback statements.

Modeling (MO) — Teacher demonstrates a skill behavior for one or many students. If verbal instruction accompanies a modeling behavior, prioritize the modeling for coding purposes. Showing a student how to perform a skill properly would be a modeling behavior.

Officiating (O)	Teacher is actively refereeing a sport or game. Verbal behavior which entails rules, regulations or judgements about the sport or game should be coded officiating.
Behavior Praise (B)	Teacher provides positive or supportive statements or gestures to a behavior episode unrelated to motor skill performance. Examples include "Tremendous class, you lined up beautifully," or "I like the way you are sitting without talking, Judy."
Nags (N)	Teacher verbally or nonverbally scolds a student or group of students for undesirable behavior or skill attempts in a low intensity manner. Nags also include pleas to the class that they return to on-task behavior that go unheeded by the students. Examples of nags are "I told you to get in line," "Listen up, listen up," "Quit it," "Didn't I say to stop chattering."

DATE _____ CLASS ACTIVITY _____ TAPE NO. _____ CODER _____

DURATION _____



Setting	Content-Gen.	Content-PE	Learner Moves	Diff. Level	Teacher Behaviours
Dir. Instr. (D)	Wait (W)	Skill Practice (P)	Engaged:	Easy (E)	Lecturing (Le)
Task (T)	Transition (T)	Scrimmage (S)	Motor Rgs. (M)	Medium (M)	Giving direct. (G)
Reciprocal (R)	Management (M)	Game (G)	Indirect (I)	Hard (H)	Listening (L)
Group (G)	Break (B)	Fitness (F)	Cognitive (C)		Ask Questions (AQ)
Guid. Dis. (GD)	Non-academic	Other Motor (O)	Not Engaged:		Answer Ques. (A)
Prob. Sol. (P)	Instr. (N)	Social Behav. (B)	Interim (NI)		Non-functional (NF)
		Knowledge (K)	Waiting (NW)		Monitoring (M)
			Off-task (NO)		Maintenance (MT)
					Hustles (H)
					Modeling (Mo)
					Officiating (O)
					Behav. Praise (B)
					Nags (N)
					Teach. Partic. (TP)
					Feedback (F)
					Spotting (S)

ALT-PE-TB OBSERVATION SYSTEM
CODING SHEET

APPENDIX B

THE McLEISH OBSERVATION INSTRUMENT

THE McLEISH OBSERVATION INSTRUMENT

This system of analysis seeks to categorize all the possible behaviours of individuals in a social situation. It is based on Skinner's (1957) analysis of communicative behaviour and this coding system was originally used to analyse the behaviour of people in learning groups.

As in the other two observation systems, codings are based on a functional analysis of overt behaviours. Each overt act is seen as a combination of verbal and non-verbal components. These acts are coded as a whole and according to the specific function they perform in the teaching-learning situation.

McLeish recognises four broad dimensions of behaviour when two or more individuals are communicating in a common enviroment.

The first dimension includes categories of reference. The individual may refer directly to the immediate environment or to specific aspects of it. For example, the individual may refer to the event that occurred in the previous 10 seconds in that environment. Or, the individual may make some clear extension from the immediate environment, establishing some relevant connection from the immediate circumstances to circumstances outside this immediate reference. For example, relating what someone is saying at that particular point in time, to what was said the previous day. Finally, the individual, may make some reference to himself as the context for

responding,or reflecting, on the enviroment. These behaviours are coded as tacts or extended tacts (see definitions).

The second dimension includes categories of action. In this context, the individual may be controlling, or be under the control, of one or more other individuals. Thus, the teacher giving directions (in control) to the attendant class (under control) is an example of this category. These behaviours are coded as mands and submissive autoclitics.

The third dimension includes categories of expressive behaviour. This group is characterised by the positive and negative responses of the individual to the enviroment, or to things said or done in the group.

The final dimension includes communication responses that serve as thematic connections between the interaction of a group in a given situation. For example, a response such as "That reminds me..." is coded as an intraverbal, which is one of the two behavioural definitions in this dimension.

Adaptations were made to the system for the gymnasium setting. The mand category was divided into four parts to distinguish between references to physical education content and references to management activity (see definitions).

The interactive pattern between the teacher and student or students was observed at all times. Every overt act in this situation was coded. This was possible because the lessons were on videotape and allowed the coder to control the pace of the lesson. That is, the coder could stop and rerun any part of the lesson to ensure accurate coding, if it was necessary.

Definitions of the McLeish Instrument

Category	Definition of behaviour	Example
MANDS:		
PE command (1a)	Giving directions related to the physical education content. Making demands.	"Keep your eyes on the ball."
PE request (1b)	Questioning the students on physical education related content.	"Where is the opposition going to be?"
Management command (1c)	Giving directions/making demands unrelated to the physical education content.	"Go get changed."
Management request (1d)	Making requests unrelated to the physical education content.	"Is Jon absent today?"
TACT (2)	A response to something actually present.	"Here is the bat."
Extended tact (3)	A response associated with property of an event 'here and now'.	"It is like a morgue here."
ECHOIC (4)	A response which repeats a previous behaviour.	A group laugh or repeated phrase.
INTRAVERBAL (5)	A response, not echoic, but with a thematic connection.	"That reminds me..."
AUTOCLITICS:		
Dominant (6)	Behaviour that calls attention to the speaker or what he/she is saying.	"Now"; pausing for effect.
Informative (7)	Behaviour which clarifies response, without emotion.	"On the other hand..."
Negative (8)	Behaviour which indicates non-acception, rejection.	"I doubt that!" Disruptive behaviours.

Category	Definition of behaviour	Example
AUTOCLITICS:		
Positive (9)	Behaviour which indicates a warm acceptance.	Smiling, laughing "Very good!"
Submissive (10)	Following directions, listening, accepting behaviour.	Performing required assignments.

Example of a Coding Sheet

	T.	C.		T.	C.		T.	C.
1.	1a	10	11.			21.		
2.	1a.	10	12.			22.		
3.	3	10	13.			23.		
4.	10	1b	14.			24.		
5.	1a	10	15.			25.		
6.	1a	10	16.			26.		
7.	1a	8	17.			27.		
8.	8	8	18.			28.		
9.	1c	10	19.			29.		
10.	1c	10	20.			30.		

APPENDIX C

THE CAFIAS OBSERVATION INSTRUMENT

THE CAFIAS OBSERVATION INSTRUMENT

The CAFIAS instrument was designed to record the original Flanders categories and, in addition: non-verbal behaviours, variations of the teaching agency, differential class structures and greater sensitivity in student behaviours. Like the original Flanders Interaction Analysis System, CAFIAS is predominantly a process system designed to describe behaviours in teaching or modifying human interaction. This instrument primarily measures the classroom climate and the directness or indirectness of the teacher.

The lesson was divided into equal measures of time. The coder observed the class for three seconds and then recorded all the behaviours observed in this interval for three seconds. Whenever the class split into group work, the coder observed the interaction between the teacher and whichever group the teacher was working with.

The Behavioural Definitions of the CAFIAS Instrument

Coding Symbols

Teacher
Environment (E)
Student (S)

Categories	Verbal	Relevant Behaviors	Nonverbal
2-12	2		12
	Praises, commends, jokes, encourages	Face:	Smiles, nods with smile, (energetic) winks, laughs
		Posture:	Claps hands, pats on shoulder, places hand on head of student, wrings student's hand, embraces joyfully, laughs to encourage, spots in gymnastics, helps child over obstacles.
3-13	3		13
	Accepts, clarifies, uses, and develops suggestion and feelings by the learner.	Face:	Nods without smiling, tilts head in empathetic reflection, sighs empathetically.
		Posture:	Shakes hands, embraces sympathetically, places hand on shoulder, puts arm around shoulder or waist, catches an implement thrown by student, accepts facilities.
4-14	4		14
	Asks questions requiring student answer.	Face:	Wrinkles brow, opens mouth, turns head with quizzical look.
		Posture:	Places hands in air, waves finger to and fro anticipating answer, stares awaiting answer, scratches head, cups hand to ear, stands still half turned towards person, awaits answer.
5-15	5		15
	Gives facts, opinions, expresses ideas, or asks rhetorical questions.	Face:	Whispers words inaudibly, sings, or whistles.
		Posture:	Gesticulates, draws, writes, demonstrates activities, points.

Categories	Verbal	Relevant Behaviors	Nonverbal
6-16	6		16
	Gives directions or orders.	Face:	Points with head, beckons with head, yells at.
		Posture:	Points finger, blows whistle, holds body erect while barking commands, pushes child through a movement, pushes a child in a given direction.
7-17	7		17
	Criticizes, expresses anger or distrust, sarcastic or extreme self-reference.	Face:	Grimaces, growls, frowns, drops head, throws head back in derisive laughter, rolls eyes, bites, spits, butts with head, shakes head.
		Posture:	Hits, pushes away, pinches, grapples with, pushes hands at student, drops hands in disgust, bangs table, damages equipment, throws things down.
8-18	8		18
	Student response that is entirely predictable, such as obedience to orders, and responses not requiring thinking beyond the comprehension phase or knowledge (after Bloom)	Face:	Poker face response, nod, shake, gives small grunts, quick smile.
		Posture:	Moves mechanically to questions or directions, responds to any action with minimal nervous activity, robot like.
	EINE (8\)		EINETEEN (18 \)
eine (8\) & eineteen (18\)	Predictable student responses requiring some measure of evaluation and synthesis from the student, but must remain within the province of predictability. The initial behavior was in response to teacher initiation.	Face:	A "What's more, Sir" look , eyes sparkling.
		Posture:	Adds movements to those given or expected, tries to show some arrangement requiring additional thinking; e.g., works on gymnastic routine, dribbles basketball, *all game playing.*

Categories	Verbal	Relevant Behaviors	Nonverbal
9-19	9		19
	Pupil-initiated talk that is purely the result of their own initiative and that could not be predicted.	Face:	Interrupting sounds, gasps, sighs.
		Posture:	Puts hands up to ask questions, gets up and walks around without provocation, begins creative movement education, makes up own games, makes up own movements, shows initiative in supportive movement, introduces new movements into games not predictable in the rules of the games.
10-20	10		20
	Stands for confusion, chaos, disorder, noise, much noise.	Face:	Silence, children sitting doing nothing, noiselessly awaiting teacher just prior to teacher entry, etc.

Additional Ground Rules

1. It is possible to have three types of teaching agency in a physical education lesson.
   a. The classroom teacher (T).
   b. Students doing the teaching (S). In this case, the letter 'S' is placed beside the appropriate tally whenever this is observed.
   c. The environment (E). An example of this agent is a film. When this teaching agent is being used the letter 'E' is placed beside the appropriate tally.
2. Whenever the teacher is talking and demonstrating at the same time, the observer codes the verbal symbol and encircles it.
3. Whenever a change in class structure takes place, either the symbol 'W' (whole), 'P' (part), or 'I' (not influencing) is placed beside the relevant code symbol. In this way it is possible to code percentages of time spent with the class working as a whole or in groups or independent of immediate

teacher influence.

4. In order to clarify the use of 6, CAFIAS adopts the following recommendation. When directions are being given, only the executive part of the command is coded as 6. The other information-giving section of the statement is coded as 5.
5. There is a difference between helpful criticism and criticism intended to destroy or punish. The former is coded as a 7 followed by a 2 and the latter is coded simply as a 7.

TALLYING WORKSHEET

Teacher _____ Observer _____
Class _____ Date _____
School _____ Time Start _____ Stop _____

1. 20W	21. 5	41. 8	61. 18\	81. 18\
2. 5	22. (5)	42. 2	62. 18\	82. 18\
3. 5	23. (5)	43. 4	63. 18\	83. 18\
4. 5	24. (5)	44. 8	64. 18\	84. 18\
5. 5	25. 5	45. 2	65. 18\	85. 6
6. 5	26. 5	46. 4	66. 18\	86. 18
7. 5	27. 5	47. 8	67. 18\	87. 5
8. 5	28. 4	48. 2	68. 18\	88. 5
9. 5	29. 8	49. 6	69. 18\	89. 5
10. 5	30. 2	50. 18	70. 18\	90. 5
11. 5	31. 4	51. 18\	71. 18\	91. 5
12. 5	32. 8	52. 18\	72. 18\	92. 5
13. 5	33. 2	53. 18\	73. 18\	93. 5
14. 5	34. 4	54. 18\	74. 18\	94. 5
15. 5	35. 8	55. 18\	75. 18\	95. 5
16. 5	36. 2	56. 18\	76. 18\	96. 5
17. 5	37. 4	57. 18\	77. 18\	97. 5
18. 5	38. 8	58. 18\	78. 18\	98. 5
19. 5	39. 2	59. 18\	79. 18\	99. 5
20. 5	40. 4	60. 18\	80. 18\	100. 20

APPENDIX D

THE PERCENTAGE SCORES ON THE THREE OBSERVATION INSTRUMENTS

Table 13

Percentage Scores of the ALT-PE Dimension

LESSON	ALT-PE BEHAVIOURS																				
	A	B	C	D	E	F	G	H	I	J	K	L	M	O	P	Q	R	S	T	U	V	W
Volleyball																						
1	2	17	0	0	0	0	0	16	4	11	19	0	10	0	1	3	2	0	15	97	0	3
2	0	20	0	0	0	0	0	23	0	5	3	0	15	0	5	6	6	2	12	100	0	0
3	1	26	1	0	0	0	0	3	5	3	12	0	0	0	4	13	6	4	22	99	0	1
4	0	13	0	0	0	0	0	38	6	4	14	0	0	0	0	0	0	0	24	96	0	4
5	0	12	1	0	0	0	0	21	3	6	5	0	13	0	4	14	12	2	6	99	0	1
6	0	16	0	13	22	5	2	4	7	2	4	0	3	1	7	4	1	0	15	100	0	0
Basketball																						
1	1	22	0	0	0	0	0	17	8	1	1	0	0	0	9	15	2	1	24	100	0	0
2	1	34	1	0	0	0	0	9	6	3	8	0	13	0	1	2	1	0	21	99	0	1
3	3	20	2	0	0	0	0	26	0	1	0	0	12	1	3	13	2	1	16	97	0	3
4	0	11	0	0	0	0	0	31	5	3	6	0	0	0	0	0	0	0	14	100	0	0
5	0	25	0	0	0	0	0	25	2	4	7	0	11	1	0	0	0	0	25	100	0	0
6	3	29	2	0	0	0	0	23	0	3	5	2	10	1	0	0	0	0	22	99	0	1
Indoor Soccer																						
1	1	28	1	0	0	0	0	14	0	0	12	0	13	0	2	2	1	0	26	99	0	1
2	1	12	0	0	0	0	0	16	16	1	5	0	0	0	9	27	1	1	11	100	0	0
3	6	20	0	0	0	0	0	28	5	0	8	0	8	0	0	0	0	0	24	100	0	0
4	2	25	0	0	0	0	0	19	3	10	6	0	18	0	0	0	0	0	19	99	0	1
5	5	20	0	0	0	0	0	8	6	7	14	0	0	0	0	0	0	0	35	100	0	0
6	5	12	1	0	0	0	0	0	0	0	0	0	21	0	8	23	8	1	21	99	0	1
Badminton																						
1	0	16	0	0	0	0	0	34	1	2	8	1	15	0	5	2	2	3	11	100	0	0
2	1	20	0	0	0	0	0	21	7	5	7	0	0	0	8	0	0	0	31	100	0	0
3	6	26	0	0	0	0	0	14	0	1	22	0	12	0	0	2	2	0	26	100	0	0
4	1	16	0	0	0	0	0	24	4	7	9	0	2	0	10	6	6	4	11	100	0	0
5	4	20	0	0	0	0	0	32	0	14	10	1	9	0	0	0	0	0	10	100	0	0
6	0	11	0	0	0	0	0	12	5	6	12	0	20	0	9	6	3	7	11	99	0	0

DIRECT SETTING

Management time:
- DM - A
- DT - B
- DW - C

Instructional time:
- DKC - T

GROUP SETTING

Skill practice:
- GPM - D
- GPI - E
- GPNI - F
- GPNW - G

TASK SETTING

Skill practice:
- TPM - H
- TPI - I
- TPNI - J
- TPNW - K
- TPNO - L

Fitness:
- TFM - M
- TFNW - O

Game:
- TGM - P
- TGI - Q
- TGNW - R
- TGNI - S

DIFFICULTY LEVEL
- Easy - U
- Medium - V
- Hard - W

Table 14

Percentage Scores of the TB Dimension of the ALT-PE-TB Instrument

LESSON	TEACHER BEHAVIOURS																
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
Volleyball																	
1	14	22	2	2	1	0	9	0	2	1	0	1	2	0	4	0	46
2	11	29	1	2	2	0	13	0	1	1	0	1	3	0	2	0	35
3	22	38	2	4	0	0	2	2	1	0	16	2	0	0	0	0	9
4	15	13	1	1	0	2	26	0	0	9	0	1	0	0	0	0	30
5	6	16	1	1	1	1	14	1	0	0	19	0	1	0	0	0	39
6	15	16	2	5	0	0	17	7	0	0	0	0	0	0	1	0	36
Basketball																	
1	22	25	1	2	1	1	36	1	1	2	0	1	0	0	0	0	8
2	13	30	2	0	1	2	16	1	1	14	0	0	0	0	0	0	20
3	13	35	1	1	0	4	32	0	0	3	0	1	2	0	0	0	9
4	8	11	0	1	0	0	8	0	0	6	0	2	0	0	14	0	51
5	14	26	1	3	2	1	11	0	0	11	0	0	0	0	3	0	29
6	20	46	1	3	0	2	8	0	2	2	0	1	2	0	1	0	14
Indoor Soccer																	
1	26	25	1	5	2	0	12	1	1	5	7	0	1	0	4	0	11
2	6	12	1	4	0	0	31	0	0	0	32	0	0	0	0	0	13
3	20	29	1	1	0	4	26	0	0	5	0	0	0	0	0	0	14
4	20	31	0	2	0	0	0	1	1	3	0	1	0	0	0	0	41
5	30	22	4	2	0	4	16	0	1	0	0	0	5	0	0	0	16
6	14	24	1	3	1	1	15	0	0	2	35	0	0	0	0	0	4
Badminton																	
1	8	26	1	2	1	1	22	0	1	5	0	1	0	0	0	0	31
2	16	21	0	3	0	0	17	0	2	15	0	0	2	0	0	0	23
3	8	26	1	1	0	0	28	1	0	4	0	0	9	0	0	0	22
4	7	19	0	0	0	1	40	4	0	4	0	0	1	0	0	0	24
5	6	24	1	1	3	1	0	1	1	4	0	0	2	0	0	0	54
6	6	10	0	4	1	0	22	0	0	3	0	0	0	0	32	0	22

A — Lecturing
B — Giving directions
C — Listening
D — Questioning
E — Answering questions
F - Non-functional
G - Monitoring
H - Maintenance
I - Hustling
J - Modeling
K - Officiating
L - Behaviour praise
M - Nagging
N - Punishment
O - Teacher participation
Q - Feedback

Table 15

Percentage Scores of
the McLeish Instrument

Lesson	Behavioural Variables															
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
Volleyball																	
1	50	7	10	2	2	5	8	2	1	3	1	1	2	1	97	0	0
2	39	6	8	3	11	8	11	1	8	3	0	2	1	2	97	0	0
3	30	6	17	2	10	10	16	1	1	5	2	0	5	1	94	0	0
4	33	2	30	1	11	11	6	1	0	5	0	0	2	1	96	0	1
5	38	3	12	1	5	9	29	0	1	1	1	0	2	1	98	0	0
6	18	35	18	1	3	2	10	0	0	13	0	0	2	1	85	0	8
Basketball																	
1	27	1	31	2	6	15	5	1	2	5	5	0	1	2	95	0	2
2	48	4	23	6	4	4	6	0	0	4	1	0	2	2	96	0	0
3	36	0	28	1	5	12	5	4	4	1	2	1	1	1	97	1	0
4	38	1	16	1	1	6	34	0	1	1	0	1	0	1	99	0	0
5	33	4	29	2	8	8	12	0	1	1	1	1	1	1	98	0	0
6	50	5	9	4	5	7	7	5	4	1	1	2	0	0	96	4	0
Indoor Soccer																	
1	47	7	19	3	5	2	8	1	1	7	1	0	5	1	93	0	1
2	37	7	22	2	5	10	9	2	2	4	0	0	1	4	88	7	0
3	44	4	30	3	2	2	9	0	1	3	1	1	1	2	97	0	0
4	41	4	41	2	1	2	5	0	1	1	2	0	1	0	95	4	0
5	28	6	19	0	16	9	3	5	1	5	1	2	2	2	90	4	2
6	63	4	1	1	9	10	6	0	1	4	1	0	7	2	92	0	0
Badminton																	
1	47	8	13	4	2	5	9	2	1	7	1	1	5	2	92	0	0
2	34	9	3	3	18	13	14	1	1	0	1	3	1	1	98	0	0
3	30	8	20	7	9	11	4	1	3	4	2	1	3	2	95	0	0
4	36	4	31	0	5	7	11	1	0	4	0	1	2	2	95	0	1
5	53	10	10	2	1	4	7	3	4	3	2	0	2	1	93	4	0
6	48	7	17	0	8	2	9	1	0	8	0	0	5	2	93	0	1

TEACHER BEHAVIOURS

A - PE command
B - PE request
C - Management command
D - Management request
E - Tact
F - Extended tact
G - Positive affect
H - Negative affect
I - Dominant autoclitic
J - Submissive autoclitic
K - Intraverbals
L - Echoics

STUDENT BEHAVIOURS

M - PE request
N - Management request
O - Submissive autoclitic
P - Negative autoclitic
Q - Tact

Table 16
Percentage Scores of the
CAFIAS Instrument

LESSON	BEHAVIOURAL VARIABLES																				
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
Volleyball																						
1	3	1	2	15	1	25	0	4	2	2	0	1	0	2	0	16	26	0	54	46	100	0
2	6	0	2	10	2	25	0	3	0	2	0	0	0	1	0	29	20	0	42	58	100	0
3	10	0	0	18	2	24	0	2	0	0	0	1	0	0	0	26	15	0	0	100	100	0
4	3	0	2	14	6	16	0	1	0	1	0	2	0	5	0	30	20	0	0	100	100	0
5	15	5	1	8	3	17	0	0	0	0	0	0	0	2	0	18	31	0	59	41	86	14
6	5	0	3	14	3	16	0	0	0	2	2	3	0	0	0	27	25	0	82	18	56	44
Basketball																						
1	2	0	1	14	2	21	4	3	0	1	1	0	0	0	0	23	28	0	49	51	100	0
2	3	2	3	11	11	19	3	0	0	2	0	1	0	0	0	30	13	2	21	79	100	0
3	2	1	1	9	8	27	0	3	1	0	0	0	0	0	0	35	13	0	20	80	100	0
4	17	0	0	7	6	12	3	3	0	1	0	0	0	11	0	34	9	0	0	100	100	0
5	6	0	5	10	12	18	5	0	0	4	1	0	0	0	0	26	13	0	13	87	100	0
6	4	0	4	16	4	28	0	3	2	3	0	0	0	1	0	26	9	0	10	90	100	0
Indoor Soccer																						
1	4	0	5	10	8	24	0	2	0	2	2	0	0	5	0	32	6	0	8	92	100	0
2	7	0	3	13	0	17	0	0	2	0	2	0	0	0	0	7	49	0	38	62	100	0
3	1	0	2	16	9	21	3	1	0	2	0	0	0	0	0	39	5	1	43	57	100	0
4	3	5	3	9	3	28	0	1	1	3	1	1	0	0	0	30	12	0	75	25	100	0
5	1	0	3	21	6	19	3	2	3	3	0	0	1	0	0	20	18	0	39	61	100	0
6	3	0	3	18	3	23	0	0	0	1	0	1	0	0	1	20	27	0	2	98	100	0
Badminton																						
1	4	3	4	9	2	29	0	1	0	1	0	1	0	0	0	17	29	0	59	41	100	0
2	8	0	3	12	12	20	3	3	0	1	0	0	0	0	0	14	24	0	47	53	100	0
3	2	0	3	13	2	25	1	2	0	2	2	1	0	0	0	36	11	0	60	40	100	0
4	4	0	0	8	4	21	0	3	0	0	0	1	0	0	0	28	31	0	16	84	100	0
5	4	0	6	9	3	24	1	3	3	2	1	0	0	0	0	26	18	0	64	36	100	0
6	3	1	4	7	2	7	0	1	0	2	0	1	0	15	0	25	31	0	32	68	100	0

TEACHER DIMENSION

Verbal:

A - Praise
B - Accepts; uses student's ideas
C - Questions
D - Gives facts or opinions
F - Gives directions
M - Confusion
H - Constructive criticism
I - Negative reinforcement

Non-verbal:

E - Demonstrate
N - Teacher participation
O - Blows whistle
G - Claps hands/whistles & verbalises instructions

STUDENT DIMENSION

Verbal:

J - Predictable response
K - Response requiring some interpretation
L - Student-initiated talk

Non-verbal:

P - Practice drill
Q - Interpretative movement
R - Silence, students doing nothing

ORGANISATION

S - Working in groups
T - Working as whole class

TEACHING AGENCY

U - teacher
V - Student

APPENDIX E

MONTE CARLO SIMULATION SCORES

Table 17

Monte Carlo Simulation I

Lessons	Rank Orders						
	Judge 1	Judge 2	Judge 3	Judge 4	Judge 5	Judge 6	Agreement
Volleyball							
1	4	5	4	2	2	5	3
2	6	4	1	6	4	2	2
3	5	3	5	3	3	6	3
4	3	1	3	1	6	3	4
5	1	6	2	5	5	4	1
6	2	2	6	4	1	1	2
Basketball							
1	4	5	5	2	4	2	3
2	6	6	1	6	6	6	9
3	5	4	4	1	1	4	4
4	3	3	3	3	5	3	10
5	1	1	2	4	2	1	4
6	2	2	6	5	3	5	2
Indoor Soccer							
1	5	4	5	2	5	2	4
2	6	5	4	6	1	4	2
3	4	6	1	3	2	6	1
4	2	1	3	1	4	3	2
5	1	3	6	4	6	1	2
6	3	2	2	5	3	5	3
Badminton							
1	2	1	5	2	5	2	4
2	1	5	4	6	4	4	2
3	5	6	6	1	2	6	3
4	6	4	3	4	1	1	2
5	4	3	1	3	3	5	3
6	3	2	2	5	6	3	2

Possible agreement = 4 x 6 x 15 = 360

Actual agreements = 77

Percentage agreement = $\frac{77 \times 100}{360}$ = 21%

Table 18

Monte Carlo Simulation II

Lessons	Rank Orders						
	Judge 1	Judge 2	Judge 3	Judge 4	Judge 5	Judge 6	Agreement
Volleyball							
1	2	4	5	6	5	1	1
2	6	1	4	4	3	6	2
3	4	2	3	2	6	5	1
4	5	5	6	3	2	4	1
5	3	3	2	1	1	2	3
6	1	6	1	5	4	3	1
Basketball							
1	5	1	3	4	6	5	1
2	6	3	6	3	2	2	3
3	2	2	4	6	3	6	2
4	3	6	2	1	4	1	1
5	1	4	5	2	1	4	2
6	4	5	1	5	5	3	2
Indoor Soccer							
1	4	1	3	4	3	5	2
2	6	3	4	6	4	1	2
3	5	2	6	3	6	2	2
4	2	5	2	2	2	3	6
5	3	6	5	1	1	6	2
6	1	4	1	5	5	4	3
Badminton							
1	1	1	6	6	3	5	2
2	5	4	2	4	4	2	4
3	4	5	3	3	6	3	3
4	3	3	4	2	2	6	2
5	6	6	1	5	1	1	4
6	2	2	5	1	5	4	2

Possible agreements = 4 x 6 x 15 = 360

Actual agreements = 55

Percentage agreement = $\frac{55 \times 100}{360}$ = 15.3%

VITA

Surname: LIND Given Surname: PETER RICHARD

Place of Birth: ELIZABETHVILLE, ZAIRE Date of Birth: July 31, 1954

Educational Institutions Attended, with Dates of Entering and Leaving:

PALMERSTON NORTH TEACHERS' COLLEGE, PALMERSTON NORTH, NEW ZEALAND	1973 to 1975
MASSEY UNIVERSITY, PALMERSTON NORTH, NEW ZEALAND	1973 to 1980
UNIVERSITY OF VICTORIA, B.C.	1980 to 1982

Degrees, Diplomas, Etc., Awarded, with Dates and Names of Institutions:

Trained Teachers Certificate	1977	Palmerston North Teachers' College, N.Z.
Diploma in Teaching	1977	Massey University, N.Z.
B.Ed	1980	Massey University, N.Z.

Honors and Awards:

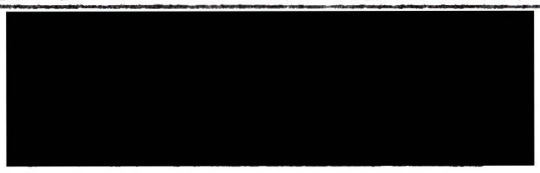
Publications:

PARTIAL COPYRIGHT LICENCE

I hereby grant the right to lend my thesis or dissertation (the title of which is shown below) to users of the University of Victoria Library, and to make single copies only for such users or in a response to a request from the library of any other university, or similiar institution, on its behalf or for one of its users. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by me or a member of the University designated by me. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Title of Thesis

A COMPARISON OF FOUR RATING SYSTEMS FOR THE EVALUATION OF PHYSICAL EDUCATION LESSONS.

Author 

PETER R. LIND

JUNE 1982