The systematic observation of teacher behaviour in physical education: the design of an instrument

This item was submitted to Loughborough University's Institutional Repository by the/an author.

Additional Information:

- A Doctoral Thesis. Submitted in partial fulfillment of the requirements for the award of Doctor of Philosophy of Loughborough University.

Metadata Record: https://dspace.lboro.ac.uk/2134/13876

Publisher: © Lynne Spackman

Please cite the published version.
This item was submitted to Loughborough University as a PhD thesis by the author and is made available in the Institutional Repository (https://dspace.lboro.ac.uk/) under the following Creative Commons Licence conditions.

For the full text of this licence, please go to:
http://creativecommons.org/licenses/by-nc-nd/2.5/
<table>
<thead>
<tr>
<th>date due:</th>
<th>Loan Copy</th>
<th>Class Mark</th>
</tr>
</thead>
<tbody>
<tr>
<td>24 JAN 1989</td>
<td>28 JUN 1996</td>
<td>9 MAR 1999</td>
</tr>
<tr>
<td></td>
<td>27 JUN 1997</td>
<td>14 APR 2000</td>
</tr>
<tr>
<td></td>
<td>26 NOV 1997</td>
<td></td>
</tr>
<tr>
<td>- 2 JUL 1993</td>
<td>12 DEC 1997</td>
<td></td>
</tr>
<tr>
<td></td>
<td>14 MAR 1999</td>
<td></td>
</tr>
<tr>
<td>- 1 JUL 1994</td>
<td>23 FEB 1998</td>
<td></td>
</tr>
<tr>
<td>30 JUN 1995</td>
<td>20 MAR 1998</td>
<td></td>
</tr>
</tbody>
</table>

**LOUGHBOROUGH UNIVERSITY OF TECHNOLOGY LIBRARY**

**AUTHOR/FILING TITLE**

Spackman, L

**ACCESSION/COPY NO.**

012418/02

**VOL. NO.**

<table>
<thead>
<tr>
<th>Class Mark</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
</tr>
</tbody>
</table>

**MARK**

LeAN ClaP

**LEAN COPY**

LOAN 3 WKS + 3 UNLESS RECALLED SWANSEA

- 2 JUL 1993
- 1 JUL 1994
- 30 JUN 1995
THE SYSTEMATIC OBSERVATION OF TEACHER BEHAVIOUR IN PHYSICAL EDUCATION:
THE DESIGN OF AN INSTRUMENT

BY

LYNNE SPACKMAN

A Doctoral Thesis

Submitted in fulfilment of the requirements for the award of
Degree of Doctor of Philosophy
of the Loughborough University of Technology

June, 1986

c Lynne Spackman, 1986
CONTENTS

Abstract i
Acknowledgements iii
List of Tables iv
List of Symbols v
Publications vi
Preface vii

CHAPTER 1 INTRODUCTION

CHAPTER 2 PROCESS-PRODUCT RESEARCH SINCE 1960 6

2.1 Introduction 6
2.2 The demise of research in Physical Education 6
2.3 The changing focus in classroom research: The Interactionists 7
2.4 The development of observation techniques 11
2.5 Classification of Observer Systems 12
2.6 The structure of an Observer System 14
2.7 A Category Instrument 15
2.8 A Sign Instrument 21
2.9 A unit of analysis 22
2.10 A Rating Instrument 24
2.11 Alternative classroom research: The Anthropologists 26
2.12 Summary 27

CHAPTER 3 RESEARCH INTO THE DESIGN OF OBSERVATION SYSTEMS IN PHYSICAL EDUCATION 29

3.1 Introduction 29
3.2 The changing focus in Physical Education research 29
3.3 Interaction Analysis Systems: Boston University 34
3.4 Behaviour Analysis Systems: Ohio State University 38
3.5 Multi-Use Observation Instruments: Teacher's College, Columbia 48
3.6 The response from other countries 58
3.7 The identification of a research interest 60
3.8 Summary 62

CHAPTER 4 THE DESIGN OF AN INSTRUMENT: BOS 4 64

4.1 Introduction 64
4.2 The production of videotaped lecture material 64
4.3 Lessons learned during the filming phase 67
CHAPTER 7

IMPLICATIONS

7.1 Introduction
7.2 Implications for initial teacher education programmes in physical education
7.3 The Bailey Observation System: BOS 2
7.3.1 Early development
7.3.2 BOS 2 categories
7.3.3 Interval Recording Procedure
7.3.4 System Testing
7.3.5 Computer Use
7.4 In-Service Training

REFERENCES

BIBLIOGRAPHY

APPENDICES

A.1 Teacher Profile Form
A.2 Observer Training Diary
A.3 Observer Training Report
A.4 BOS 2 Computer Program
A.5 Computer Print-out
A.6 Test Tape PL2/1
A.7 Computer Program used to scale down samples
A.8 Results of BOS 2: Tables 10 - 15
ABSTRACT

Compared with research in classrooms, the teaching of physical education has not proved to be a popular research topic. The reasons for this apparent neglect can be seen in the lack of suitable research instruments to describe the teaching of physical education and the complex technical problems of recording lessons which take place out of doors in large, open spaces.

This study set out to design an instrument which could be used to describe the teaching of physical education in any setting (indoors and outdoors).

The investigation consisted of three studies. The first was the production of an instrument (BOS 1) with 42 categories which recorded every item of teacher behaviour in a lesson. The system was difficult to learn, required a great deal of time to analyse and was found to be unsuitable for live coding.

The second study set out to resolve the difficulties of BOS 1 and this was achieved by the production of a nineteen category system of teacher behaviour under the principal dimensions of Managing or Teaching, Direction and Mode, using a fifteen second interval recording system. The percentages for inter-observer agreement were 80% for reliability and 82% for believability which is high for an instrument that records lessons directly and does not use transcripts.
The third study was the use of BOS 2 on nineteen videotaped lessons provided by eleven teachers of physical education representing three local education authorities. Each tape was coded using the BOS 2 observation system providing 3399 samples of behaviour. Male and female teachers across all of the activities and age groups did not show any significant differences in their managerial and teaching behaviours examined collectively. The teachers were consistent in using a predominantly prescriptive teaching method, preferring to address the whole class simultaneously. These results which have significance for teachers engaged in changes in the curriculum requiring a range of teaching approaches, and their implications for research on teaching, are explored in depth.
I wish to thank my colleagues for their invaluable advice during the course of this research.

In particular, thanks go to Len Almond for his vision and tireless support;
to Rod Thorpe, Brenda Read and Bob Pigott for their patience in validating the instrument;
to the unnamed teachers without whose generosity there would be no instrument;
and to David Spackman, who never stopped believing in its value.
LIST OF TABLES

TABLE
1 Dimensions and Categories of Teacher Behaviour
2 The Modified System (BOS 2)
3 Group Coding Results
4 Group Coding of each unit
5 The Bailey Observation System 2
6 Sample Data from Reliability Test
7 Percentages of Agreement (Reliability)
8 Percentages of Agreement (Believability)
9 Percentages of Observed Teacher Behaviour
10 Percentages of category use by male teachers
11 Percentages of category use by female teachers
12 Percentages of category use by male teachers in gymnastics lessons
13 Percentages of category use by male teachers in games lessons
14 Percentages of category use by female teachers in games lessons
15 Percentages of category use by female teachers in gymnastics lessons
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BOS 1</td>
<td>Bailey Observation System 1</td>
</tr>
<tr>
<td>BOS 2</td>
<td>Bailey Observation System 2</td>
</tr>
<tr>
<td>T</td>
<td>Teaching Dimension</td>
</tr>
<tr>
<td>M</td>
<td>Managing Dimension</td>
</tr>
<tr>
<td>O</td>
<td>Other Dimension</td>
</tr>
<tr>
<td>Mo</td>
<td>Mode Dimension</td>
</tr>
<tr>
<td>D</td>
<td>Direction Dimension</td>
</tr>
<tr>
<td>01</td>
<td>First Observation</td>
</tr>
<tr>
<td>02</td>
<td>Second Observation</td>
</tr>
</tbody>
</table>
PUBLICATIONS

Part of this thesis has been published in the following Journals:-


For convenience throughout this text the terms 'he/his' have been used to refer to either sex.

Letters have been used in preference to names when describing participants in the study for the purpose of preserving confidentiality.
CHAPTER 1

INTRODUCTION

This thesis is concerned with the design and evaluation of a systematic observation instrument which may be used to describe teaching in physical education. The research grew out of the author's interest in the problems of observing and supervising trainee students of physical education.

The training of teachers depends critically on the supervision by the expert of the novice whose classroom performance is moulded, monitored and evaluated until such a point as he or she is given licence to practise alone. This supervision of the trainee in the school or in the workshop sessions makes heavy demands upon the observation skills of the tutor-supervisor or the co-operating teacher. The first question with which he is faced is 'what is the teacher doing?' and his lesson report is based upon an interpretation of what was seen together with likely recommendations as to what he might like to have seen. In this way the trainee is guided through the complexities of classroom life, avoiding some of the pitfalls and acquiring the basic teaching skills which are to be the tools of his trade.

There are problems associated with this form of supervision, however. The practised eye of the expert may be
able to anticipate difficulties and even to understand them, but are his observations always accurate and presented in the most meaningful way to the trainee?

Two supervisors observing the same lesson are unlikely to write identical lesson reports. Each makes reference to his own set of values, albeit unconsciously, and their collective experiences may differ markedly one from another. The result will be quite different anecdotal records of the lesson and the consequences potentially unhelpful to the trainee.

When no longer a trainee but a practising teacher, one can fairly assume that informal observation and anecdotal report writing are his only familiar tools of teacher evaluation. These will be scarcely adequate for the teacher intent upon educational advance for whom the role of an 'extended professional' is a goal. According to Stenhouse:

"The critical characteristics of that extended professionalism....seem to me to be:
The commitment and the skills to study one's own teaching;
The concern to question and to test theory in practice by the use of those skills.
To these may be added as highly desirable, though perhaps not essential, a readiness to allow other teachers to observe one's work - directly or through recordings - and to discuss it with them on an open and honest basis.
In short, the outstanding characteristics of the extended professional is a capacity for autonomous professional self development through systematic self-study, through the study
of the work of other teachers and through the testing of ideas by classroom research procedures."

(Stenhouse 1975, 144)

The skills of self-study and their application to the study of others are essential if the teacher is to be able to make informed decisions about teaching, learning and the content of the curriculum. Without such skills the teacher is left without an answer to such questions as 'What is happening in this lesson?' or risk making assumptions about what is happening and there is evidence to suggest that teachers do not judge accurately classroom behaviour (Hook and Rosenshine 1979).

Perceived problems in classroom observation and informal reporting for supervisors, trainees and practising teachers led the author to consider the alternatives available. Systematic observation appeared to offer one kind of a solution yet its use in physical education was virtually untried in Britain.

Examination of the literature revealed it to be a relatively new field of research internationally, with the majority of work being done in the United States only within the last decade. Educational research in classrooms appeared to be one step ahead of that which had been attempted in physical education although not by a wide margin, and it had been clearly favoured in Britain where the teaching of physical education seemed to be the poor relation.
The American literature reported the design of systematic observation systems used by researchers and teacher educators to provide descriptions of events in the lesson. Foci varied but included teacher behaviour and such a system, it was felt, might benefit the trainee and the practising teacher by yielding descriptions less subject to respondent bias and, therefore, more acceptable between fellow professionals.

The choice of systems focussing on Teacher Behaviour exclusively in physical education lessons was not extensive, and the language used to describe the dimensions and behavioural categories was obviously American, requiring translation if any such system was to be accepted for use in this country. In addition the videotaped lessons invariably were taken indoors and included types of activities not found in the physical education curricula in Britain. This had contributed, it was felt, to descriptive categories which were either not replicated or might not be given such importance in this country. It seemed likely that it might be necessary ultimately to design another observation system to overcome these difficulties, and with this target in view the literature of greatest relevance to instrument design was reviewed.

The thesis is divided into seven chapters. Chapter Two focusses on process-product research inside the classroom which developed rapidly in the 1960's and necessitated the design of techniques of observation.
Chapter Three reviews the contributions of leading researchers into the design and implementation of systems for use in physical education, and the selection of one of these to be used as a starting point for the organisation of material necessary for the design of the author's own observation instrument to be used for the description of Teacher Behaviour in physical education.

Chapter Four describes the methodology of instrument design which led to the production of the Bailey Observation System (BOS 1). The chapter traces in detail the vital phase of instrument validation through the process of observer training, and the decision to modify BOS 1 in order that it be made simpler and easier to use.

Chapter Five begins with the modifications to BOS 1 and its airing to fellow professionals which resulted in the final version known as BOS 2 being tested successfully for reliability.

Chapter 6 includes presentation of the main findings and a report on the data obtained through the use of BOS 2 to describe Teacher Behaviour in nineteen physical education lessons.

Chapter Seven concludes with the implications for use of the instrument and the necessary limitations of the study.
CHAPTER 2

Process - Product Research since 1960

2.1 INTRODUCTION

The systematic observation of classroom events owed its rise in popularity to the pioneering work of educationists in the 1960's, in particular. It was during this time that the 'system' became the tool of classroom researchers who were to pass on their skills to the relatively impoverished researchers in physical education. This chapter traces the development of their unique contribution and examines the influential tool of their trade: the observer system.

2.2 THE DEMISE OF RESEARCH IN PHYSICAL EDUCATION

The teaching of physical education has not proved to be a popular research topic in the last twenty years. Compared with research into classrooms it has attracted few funds and produced little published literature other than infrequent journal articles and unpublished doctoral theses. The effect of such apparent neglect is a dearth of material which might contribute to a theory of instruction in physical education.

The reasons for this apparent neglect may lie in the context of the activity. During physical education classes pupils do not sit neatly in rows of desks nor are they
expected to remain immobile for the duration of the lesson. Classes are often large in number, heterogeneous in organisation and they meet in an environment which is subject to change with conditions varying between wet, dry, hot or cold. The numerous activities which might be found on a physical education programme could take place in very different environments such as the swimming pool, the playing field, the gymnasium, the tennis courts and the dance studio. Frequently these are located on the periphery of school sites and so time, distance and the organisation of pupil location become features of lesson planning.

It is not difficult to speculate that the would-be researcher of events in the physical education lesson may have found the context too hard to handle. In addition he has not had a tradition of research in physical education to follow. Research funds have tended to be allocated to those aspects of the school curriculum perceived as 'core' subjects or for the promotion of basic literacy and numeracy skills for which there is considerable accountability. Critically, early research focussed on events outside of the classroom itself, preferring to examine variables such as intelligence quotients, the results of streaming and curriculum development. All this must have seemed a long way from activities in the gymnasium.

2.3 THE CHANGING FOCUS IN CLASSROOM RESEARCH: THE INTERACTIONISTS
In the sixties there began a new tradition in educational research which signalled a fresh direction for future projects. The classroom itself became the new focus as researchers recognised the need to understand and appreciate classroom events for the purpose of analysing classroom processes. Medley and Mitzel (1963), in reviewing this development, commented that

'The research worker limits himself to the manipulation or studying of antecedents and consequents... but never once looks into the classroom to see how the teacher actually teaches or the pupil actually learns.'

(Medley and Mitzel (1963)

With the shift from looking at variables outside of the classroom to processes inside the classroom went a change in emphasis from research into teacher effectiveness to research into teaching.

To understand the distinction between these two areas of research it is helpful to refer to a model for the study of classroom teaching (Dunkin and Biddle 1974 p.38). Altogether there are thirteen classes of variables suggested in this model which, for the sake of convenient review, will be reduced to four larger classes using the terminology suggested by Mitzel (1960).

Research into teaching teaching and teacher effectiveness has been concerned with PRESAGE, CONTEXT, PROCESS and PRODUCT variables.
PRESAGE variables concern the personal characteristics of teachers and would include their own formative experiences, teacher training experiences and teacher properties. These characteristics are examined for their effects on the process of teaching.

CONTEXT variables concern the conditions about which the teacher can do little and to which he must adjust. These would include such things as the pupil population, the school buildings, community expectations and the school budget.

PROCESS variables concern the observable activities of the classroom itself, namely what teachers and pupils do in the lesson.

PRODUCT variables concern the outcomes of teaching which may be a product of involvement in classroom activities either through teacher or pupil contact.

The early research into teacher effectiveness focussed primarily on the relationship between presage and product variables. It produced little of substance, as the Committee on Criteria of Teacher Effectiveness of the American Educational Research Association (1953) commented:

"The simple fact of the matter is that, after forty years of research on teacher effectiveness during which a vast number of studies have been carried out, one can point to few outcomes that a superintendent of schools can safely employ in hiring a teacher or granting him tenure, that an agency
can employ in certifying teachers, or that a
teacher-education faculty can employ in planning or
improving teacher-education programs."

(in Dunkin and Biddle, 1974)

Dunkin and Biddle (1974) summarise the reasons offered by
critics for the failure of this early research effort as
- a failure to observe teaching activities;
- theoretical impoverishment;
- use of inadequate criteria of effectiveness, and
- lack of concern for contextual effects.

For the purposes of this study, the first reason is the most
relevant. Indeed Dunkin and Biddle (1974) cited it as
perhaps being the most significant shortcoming of these
early studies for it meant that researchers simply did not
enter the classroom to look at the actual activities
therein.

Gage (1963) commented that to adopt such an approach treated
the classroom as if it was a 'black box' into which were fed
the characters and hardware and out of which came the
products which were measurable in the short or the long
term.

Entering the classroom to observe events as they happened
represented a considerable advancement in the study of
teaching. The move had the appeal of simple logic:-

"If teachers do vary in their effectiveness, then it must be
because they vary in the behaviours they exhibit in the classroom. To shed light on this point one must study classrooms......where the activity is."

(Dunkin and Biddle, 1974)

It had also the support of previously tried methods of research. Descriptive research is intended to describe and clarify behaviours so that the phenomenon of teaching can be understood. Smith and Meux (1962) justify a descriptive approach to studying teaching as follows:-

"If very little is known about a phenomenon, the way to begin an investigation of it is to observe and analyse the phenomenon itself. It must be observed, analysed and classified into its various elements. Until the factors which are involved in the phenomenon are understood and described, there is little likelihood that significant correlations, predictive or causal studies can be made. In other words, the state of knowledge about a phenomenon dictates to some extent the kind of enquiry of it which is appropriate."

(Smith and Meux, 1962)

2.4 THE DEVELOPMENT OF OBSERVATION TECHNIQUES

Hand in hand with the realisation that to study classroom phenomena was vital went the need to develop techniques of observation available to the researcher. As early as 1929, Olson introduced time sampling techniques - the recording of
certain categories of behaviour at specified time intervals - in order to cope with the volumes of descriptive data obtained when observing children. Much of this early work was sponsored by those responsible for the child development movement in America during the 1920's. The study of teaching gained a foothold at this time as research workers became interested in the contrasting methods of instruction used by the teachers and the search began for the more or less 'effective' teacher. With the change in emphasis to the study of classroom processes the growth in the number of studies was rapid. Early reviews of the literature concerning these studies include Flanders and Simon (1970), Simon and Boyer (1970) and Rosenshine (1971).

It was these studies which provided the initial stimulus for the development of systematic observation techniques. Observer systems became the tools to study 'dynamic, on-going interaction between people' (Simon and Boyer 1968) and in the American tradition, with its roots in behavioural psychology, became synonymous with 'interaction analysis'. According to Stubbs and Delamont (1976), 'Characteristically, research of this type involves using an observational system to reduce the stream of classroom behaviour to small-scale units suitable for tabulation and computation.'

(Stubbs and Delamont 1976)

2.5 THE CLASSIFICATION OF OBSERVER SYSTEMS
As the number of systems in use increased there began the inevitable attempts to classify them.

Simon and Boyer (1974) attempted to classify according to the focus of the system in use. Their list included the Affective, Cognitive, Psychomotor, Activity, Content, Sociological Structure and Physical environment as popular foci.

Dunkin and Biddle (1974) reviewed approximately 500 descriptive studies and classified them according to the focus or theoretical perspective of the instrument used, a selection of whose authors is offered here:

Classroom Climate
(Withall 1949; McCandless 1961; Smith and Hurdgins 1964; Flanders 1967)

Management and Control
(Kounin 1970; O'Leary and O'Leary 1971)

The Classroom as a Social System
(Bellack 1966; Gump 1967; Adams and Biddle 1970)

Knowledge and Intellect
(Taba 1966; Wright and Nuthall 1970)

Logic and Linguistics
(Rosenshine 1968; Hiller 1971)
The Sequential Patterns of Classroom Behaviour
(Bellack 1966; Resnick 1971; Lundgren 1972).

2.6 THE STRUCTURE OF AN OBSERVER SYSTEM

As the number of observation systems in use rose and the foci of the systems increased to account for the complex nature of classroom events, the basic characteristics of the systems remained the same.

Each system consisted of a predetermined set of categories, category definitions, ground rules and coding procedures for use by the observer. The observer could code the classroom events live or record the lesson for later analysis in order to obtain a description of the activity under observation. System designers attempted to ensure that whatever criteria they identified for the purpose of describing classroom events, these should be unambiguous and mutually exclusive. Observers using the system aimed to arrive at identical descriptions of the same event or behaviours so that at the end of the observation period, matched codings had been produced. This level of 'observer agreement' demonstrated the reliability of the system in use. Low levels of reliability could mean a high level of inference on the part of the observer which was the very thing that systems were purported to reduce.

In order to achieve high levels of reliability, observers needed to spend time training and becoming familiar with category identification and coding procedures. The training
process was one which had to be repeated at regular intervals for stability and consistency in the use of the system to be maintained. The extent of the training required, the means used for data collection, testing for reliability and the details of coding and analysis would all depend on the complexity of the system and the use to which the data were being put.

The majority of observation systems developed for use in the classroom were 'category', 'sign' or 'rating' instruments.

2.7 A CATEGORY INSTRUMENT

A category instrument consists of a set of behavioural categories which are both exhaustive and mutually exclusive. Each time an event occurs it is recorded by a tally or symbol which codes the behaviour into a predefined category. The size of the system will be determined by the designer and the purpose of the study but it will usually be limited as it is impossible to observe and record all classroom behaviours simultaneously. Continuous coding is made possible by the inclusion of a catch-all category called 'other'. Category systems are usually low-inference instruments requiring very little interpretation by the observer yet they contain items which are not always simple and obvious and could affect the value of the information they yield. For example, the two items 'teacher moves' and 'pupil moves' are clear but general and fail to tell the
investigator anything about the type of movement in the lesson. Categories such as 'teacher demonstrates', 'teacher participates' and 'teacher supports' tell the reader more about the reason for and the type of movement undertaken by the teacher in this example from a physical education lesson. The category 'pupil moves' could be broken down into natural units to include 'walks', 'runs', 'balances', 'vaults' and 'climbs', for example. Each category would require a definition for ease of identification although it should remain clear in which behaviour the subject is engaged. The majority of instruments produced by researchers were of this type because they were considered to be more sensitive and precise than other types and owed their origin to the much-used Flanders System of Interaction Analysis (FIAC).

The system was developed at the University of Minnesota between 1955 and 1960 by Ned Flanders and his colleagues. A full description of the development of the system with its coding procedures, means of tabulation and interpretation, may be found in the text 'Analyzing Teacher Behaviour' (Flanders 1970).

Flanders was concerned essentially with Teacher Education. For him the crucial issue was

'To decide how teachers and college students can explore various patterns of interaction and discover for themselves which patterns they can use to improve instruction'.

(Flanders 1970)

He considered that one answer could be for teachers to learn to use procedures of interaction analysis together with any
other appropriate techniques in order to provide them with information about their own teaching behaviour. His preference was for the study of verbal communication, following the traditions of those whose interests lay in analysing only the statements made by the teacher and pupil. He recognised that such a concept might seem to provide a narrow base for interaction analysis, for it ignored the features of the classroom and many aspects of verbal communication also which could not be classified readily. Nevertheless he maintained that capturing selected elements of classroom verbal communication did prove helpful in the analysis of teaching behaviour and to this end he devised his ten category system which employed the following headings:

1. Accepts Feeling
2. Praises or encourages
3. Accepts or uses ideas of pupils
4. Asks questions
5. Lecturing
6. Giving directions
7. Criticizing or justifying authority
8. Pupil-talk – response
9. Pupil-talk – initiation
10. Silence or confusion

From the ten categories it can be seen that seven are used to describe teacher talk, two for pupil talk, and to permit continuous coding at a constant rate, the last category termed 'silence or confusion' catches all other
Selection of these categories reflects the intent of the designer which was to provide a system which could be used for the analysis of initiation and response as characteristics of interaction. Using this system, an observer could estimate the balance between this initiative and response from an analysis of the time spent in teacher or pupil talk or silence. Closer examination of the categories reveals that seven of the ten categories discriminate among teacher statements with four reflecting an indirect influence exerted by the teacher (categories one to four) and three a direct influence (categories five, six and seven). The observer could describe more accurately this balance of classroom interaction by comparing the tallies opposite columns illustrating indirect influence as opposed to direct influence. In addition it was possible to investigate relationships between the balance of teacher statements and those expressed by the pupils, for example, above average use of categories five, six and seven was more likely to be associated with a higher incidence of category eight. Similarly Flanders considered that above average use of categories one, two and three was more likely to be associated with category nine. It was this kind of evidence which led Flanders to report that a teacher's verbal communication pattern is associated with pupil learning and pupil attitudes toward learning. One of the reasons why Flanders was able to make statements about 'balance', 'relationships' and 'patterns' was because of his use of a matrix for tabulating sequential pair frequencies of coded
behaviour which revealed a general picture of the interactions between the teacher and pupils. Acknowledged by Flanders to be time consuming and subject to coding error, it nevertheless greatly increased the amount of information obtainable from raw data and was a method of particular interest to researchers investigating patterns or sequences of behaviour in the classroom.

FIAC takes between four and twelve hours to learn to use accurately and with confidence. Familiarisation with the coding procedures includes practice at identifying the categories, coding them on paper and learning to do this retrospectively every three seconds. It is a system which can be used live or with recordings. Flanders used it to investigate work in classrooms during the 1950’s and his rule of 'two-thirds' was a product of looking at English, Mathematics and Social Studies in Elementary Schools. Two-thirds of the total time spent in classrooms was spent talking and two-thirds of that talk was teacher talk, spent in giving directions, lecturing or criticising. In addition he attempted to form teaching strategies which sought to promote pupil independence and self direction:

"For example, emphasizing teacher response when goals are not clear and then increasing teacher initiation after work gets underway appears to be a promising variation of teacher behaviour. Shifting back to emphasising teacher response when diagnosing learning difficulties may be another helpful strategy. There is a possibility that teacher response to pupil ideas is an essential feature of discourse which
reaches higher levels of logic during classroom discussion".  
(Flanders 1970)

Flanders' system has been more widely used than any other in observation studies. An educational researcher with degrees in chemistry, electrical engineering and educational psychology, he provided educationists with a tool to assist achieve the objectives of professional self-development. Teacher behaviour modification programmes following the traditions established by the behavioural scientists mushroomed in the last two decades and those using FIAC or modifications of FIAC outnumber any other. These systems are reviewed in Dunkin and Biddle (1974), Flanders (1970) and Cheffers (1980).
2.8 A SIGN INSTRUMENT

The sign instrument is not unlike a category instrument in that it consists also of a number of events or behaviours which are mutually exclusive and of low inference type. However the list is not an exhaustive one and the observer tallies a behaviour when or if it occurs usually within a specified time interval, and he has a long list of signs to look for during this period. The time period will vary according to who is making the observation and the purpose of the study itself. Teachers conducting their own observations in a primary school classroom would find it extremely difficult to engage in coding for longer than one or two minutes every five minutes or more, whereas observers who were not teaching could concentrate for the majority of a lesson. Sign systems are frequently long with the 'signs' for observers to note ranging across different categories of behaviour. In spite of their length, sign systems usually require fewer tallies than a category system because of the wider range of behaviours being observed and there may occur a long period during which little coding takes place.

Sign systems tend to be the tool of the classroom teacher rather than the educational researcher. Frequently the researcher, in trying to establish patterns of behaviour or quantifying occurrences of behaviour over an extended period, will turn to an exhaustive category system where the sequence of behaviour could be monitored and the changes in behaviour are noted each time they occur. With a sign system
the event is noted only once within the specified time period in spite of the fact that it may have lasted for all of that time period or re-occurred several times in the time period. This makes the sign system less sensitive and less flexible than the category system and accounts for its absence from major pieces of educational research.

2.9 A UNIT OF ANALYSIS

Whether using a category or a sign system, a designer has to decide upon the 'unit of analysis' which will be used for coding purposes.

Adams and Biddle (1970) identified four units of analysis used in the different types of systems:

Arbitrary Time Units - based on a specified, predetermined interval of time;
Naturally Occurring Units - based on the onset and termination of key events;
Analytic Units - these reflect the key concept defined by the designer;
Phenomenal Units - these are determined by a natural break in the flow of classroom events.

The choice of unit will be dependent on the purpose of the research, the nature of the observation instrument in use, the context of the observation and the type of data required.

Arbitrary time units provide a consistent time frame for
coding, lending an almost automatic character to the descriptive exercise. The observer does not have to decide when a behaviour occurs, he simply notes the behaviour occurring at the specified time interval. Behaviours are, therefore, sampled throughout the period of observation. The disadvantage lies in the fact that an incomplete picture of behaviours is compiled during the observation period. Nevertheless the aid of the mechanical 'alarm' to observers of the time interval may well be considered a distinct advantage and lead to greater accuracy in the coding of specified behaviours over a lengthy period of observation.

Naturally occurring units require the observer to select and record specified key behaviours or events as and when they occur. Whilst the flow of these especially selected occurrences is maintained, continuous accounts of all the classroom transactions are not provided.

Analytic units are the more popular choice for designers of category systems because they are based on the researcher's own perspective and are defined operationally. They are especially relevant for those systems which might focus on one aspect of teacher behaviour such as 'feedback' for example, and the researcher requires that the observer using the system codes only that behaviour when it occurs.

Sentences, questions and gestures are examples of phenomenal units. Their identification is usually reliable for observers are readily trained in their use and they have a
common vocabulary to describe them. Their drawback is that they tend to describe isolated acts rather than cycles or episodes for which the analytic unit is better designed.

2.10 A RATING INSTRUMENT

Requiring no unit of analysis and simpler in design than both the category and the sign instruments is the rating instrument.

Rating instruments may contain a list of events, behaviours or characteristics similar to both sign and category instruments but they differ markedly in one regard - the observer makes one judgement only, usually at the end of the period of observation. That judgement or assessment is an estimate of the degree, frequency or quality of specified items and normally a bipolar scale is used with a high rating at one end and a low rating at the other. Unlike both sign and category instruments, a rating is not a series of tallies indicating whether or not a behaviour occurred or how often it was repeated, but a retrospective judgement made by the observer often about a high inference variable such as 'teacher clarity' or 'pupil involvement'. Rating instruments utilise many different scales, a clear summary of which is provided by Deakin University (1981). In spite of the fact that rating instruments are notoriously inaccurate they have proved popular in educational research. Kerlinger (1973) cites many weaknesses which includes the
major charge that they are subject to respondent bias. Knowing the person under observation, for example, can prejudice a rating. The 'halo effect' can be especially evident in the classroom when those of good behaviour or previously high attainment can continue to score well regardless of changes in academic quality. In addition respondents differ markedly from one another, some rating consistently more leniently than others and yet others obeying the 'error of central tendency' by staying near the middle of the rating scale.

Their popularity has been due largely to their relative ease in construction compared with observation systems. Clearly their use serves a different purpose in providing a judgement yet the overall aim of providing feedback is not very different. For this reason the rating instrument has proved a useful tool in classrooms for completion by teachers, teacher educators, pupils and other observers.

For the researcher the chequered history of the rating scale has meant little serious use in the last decade. Prior to that time Ryans (1960) provided Education with the most extensive rating study yet published. He had observers rate teachers on eighteen, seven point global scales which included such items as 'partial-fair', autocratic-democratic', 'aloof-responsive' and 'harsh-kindly', and pupils on four scales which included 'apathetic-alert' and 'obstructive-responsible'. He was able to identify three apparently independent characteristics of teachers: teacher warmth or understanding, teacher
responsible or business-like classroom behaviour, and teacher stimulation or imaginativeness. These results, however, have been the subject of doubt (Dunkin and Biddle 1974; Solomon and Kendall 1979) and it would appear that the changes levelled earlier at this form of instrumentation render it impossible for consideration as a serious research tool.

2.11 ALTERNATIVE CLASSROOM RESEARCH: THE ANTHROPOLOGISTS

The development of instrumentation to record classroom events was the response to a perceived need to systematise observations and generate reliable data describing classroom processes. The move by researchers into the classroom produced more than one research method, however. The interactionists on the East coast of the United States were challenged by researchers further West and in Britain in particular where anthropological methods of enquiry had found favour. Initiators of this form of enquiry were to be found in social anthropology, psychiatry and participant observation research in sociology. There appears to be some confusion as to its terminology as well as to its origins (Stubbs and Delamont 1976) for it has been described variously as 'microethnographical' (Smith and Geoffrey 1968), 'naturalistic' (MacDonald 1970) and 'ecological' (Parlett 1969).

The key differences for the anthropologist include his role
as a participant observer who may talk with participants, conduct interviews and provide questionnaires. He will use field notes or anecdotal records, keep diaries or logs of events and be altogether more open-ended, informal and unsystematic than the system researcher. Whereas the latter starts with pre-ordained descriptive categories, the anthropologist uses a holistic framework and from this wide base will develop a new language to describe events.

Both traditions have their strengths and weaknesses which have been argued well elsewhere (Stubbs and Delamont 1976; Delamont 1973; Walker 1972). As far as their influence on our understanding of what happens in classrooms the value of both is beyond doubt. In terms of their use in Physical Education lessons, the interactionists have provided us with many more years of information than the researchers adopting the naturalistic approach which holds much promise but is only just beginning (Tousignant 1982).

2.12 SUMMARY

Research into the teaching of physical education owed its origins to the new focus in educational research in the 1960's which took the observer into the classroom. Process-product research necessitated the development of techniques of observation for the description of classroom events and this need was met with the rapid production of systems of observation. Category, sign and rating instruments proved to be the most popular with the
interactionist system produced by Flanders and his colleagues in the United States of America as arguably the most influential of its time. Later developments saw the birth of new approaches such as those of the anthropologist, yet it was the interactionist systems which were to be of particular use in research into teaching and behaviour modification programmes.
CHAPTER 3

Research into the design of Observation Systems in physical education

3.1 INTRODUCTION

Educational research was providing the physical educationists with a much-needed lead in the observation and description of classroom events.

In this chapter the response of the physical education researchers is examined, showing the clear lead taken by the universities of North America who quickly established their own traditions in the development of Interaction Analysis, Behaviour Analysis and Multi-Use Observation Systems. The rest of the world offered a smattering of support, with the British Isles contributing so little to this area of research that it appeared indifferent to its importance.

This dearth of research was to provide the impetus for the development of a system which could be used to describe teacher behaviour in physical education.

3.2 THE CHANGING FOCUS IN PHYSICAL EDUCATION RESEARCH

The apparent reputation of physical educationists for lagging behind their classroom counterparts would appear to be deserved. The advances in classroom research which took
the researcher into the 'black box' of the classroom were not to be echoed immediately in physical education classes for which the environment posed potential constraints. Perhaps it was because of these that research had traditionally centred on discrete, overt motor responses examined under laboratory-type conditions (Bahrick, Fitts and Briggs 1957; Pew 1966; Fleishman 1965; Fleishman and Rich 1963; Schmidt and White 1972; Adams and Creamer 1962; Caplan 1970).

In 1967, however, Elizabeth Bookhout (1967) produced her study of 'climate' in the gymnasium and this heralded the start of a new era of research into the teaching of physical education. There followed a series of studies which used systematic observation techniques for the recording and analysis of events in the gymnasium with a particular stress on teacher and pupil behaviour (Barrett 1969; Dougherty 1970; Gasson 1971; Nygaard 1971; Bosches 1972; Mancuso 1972; Cheffers 1973; Hughley 1973; Rife 1973; Hamilton 1974; Hupe 1974; Hurwitz 1974; Darst 1974; Boehm 1974; Dodds 1975; Rankin 1975; Taylor 1976; Rink 1979). The majority of these programmes of descriptive-analytic research which resulted in the development of a series of observation systems took place in American Universities where a team approach to research was encouraged. At Teacher's College, Columbia University in New York, graduates investigated the different dimensions of behaviour observable in the gymnasium using the same taped lesson material (Anderson 1975) and produced a series of doctoral dissertations in the process (Barrette
Echoing their colleagues in classroom research, physical educationists warmly welcomed this development which for the first time was attempting to answer the question 'What is happening in the physical education lesson?' Locke (1977) commented that observation instruments went straight to the heart of teaching - moment to moment events in the gymnasium. He noted the failure of past research to go to the source and to observe directly those events which were available for recording and measurement. Agreeing with Gage (1963) he said

"Researchers have treated the gymnasium as a black box, even when used for experiments on teaching. The standard drill has been to feed teachers, kids, hardware, curriculum or organisation in at one end of the gym and to observe various consequences at the other end......

The black box strategy was a disabling one. It prevented studies from generating even weak inferences about relationships among the interacting elements in the teaching process. If teacher behaviour is thought to be a factor of importance in determining the production of educational outcomes then there is only one way to study it and only one place to study it. That way is by observing teaching directly where it occurs, in the place where the action is, inside the black box".

(Locke 1977)
Physical Educationists faced the same cry for accountability as did their classroom colleagues. Teacher education programmes were required to increase their emphasis on basic research and many adopted a scientific approach for the study of the teaching-learning environment. The development of systematic observation techniques enabled the physical educationists to collect descriptions of behaviours, events and time utilisation, for example. Teachers were observed and analysed for the type of feedback they offered students, how often information was offered and whether they initiated moves or responded to students' initiatives. Correspondingly, students' behaviours were monitored also and data obtained to see whether they spent their time waiting, engaged in 'on-task' or 'off-task' activity or involved in different stages of game playing, for example. These kinds of investigations using observation systems were undertaken in the belief that such systems could be used to:-

describe classroom practices;
modify teacher behaviour;
provide a tool for the analysis of teaching;
give feedback about one's own teaching;
train student teachers;
discriminate between patterns of teaching;
determine relationships between classroom behaviours and student growth;
help project future teaching patterns.

(Batchelder and Cheffers 1976)

Gradually the teaching process was losing some of its
earlier mystique as it became apparent through the application of such systems that as a process it can be planned for, observed and described, analysed and assessed, evaluated and modified.

Foremost in these research endeavours were the physical education staff at the Universities of Boston (Cheffers), Columbia New York (Anderson) and Ohio State (Siedentop). Whilst other institutions have contributed significantly to the development of this type of work, these three Universities were responsible for the majority of studies produced in this field and many of their graduates took their skills to the other Universities which have subsequently developed these specialisms. It is also of interest that these three Universities developed the work along complementary lines which form natural divisions for the purpose of cataloguing and describing the work in each. Darst et al (1983) have provided a comprehensive text which analyses the most widely used observation systems in sport and physical education, some of which will be referred to here.
3.3 INTERACTION ANALYSIS SYSTEMS

The most widely used observation instrument reported earlier and used in both its original and adapted forms was the Flanders System (FIAC).

At Boston University John Cheffers developed the most popular interaction analysis system used in physical education when he modified FIAC to produce CAFIAS: Cheffers Adaptation of Flanders' Interaction Analysis System.

Cheffers (1973) pointed to three major limitations in FIAC which prevented its successful use in physical education settings:

- it was concerned only with verbal behaviour;
- it viewed the teacher as the sole teaching agent in the classroom,
- and it allowed only for the coding of the class structure when the whole of the class was functioning as a unit.

Critical features of physical education lessons include non-verbal activity, augmented feedback and different operational features in the lesson with children working individually, in pairs, small groups and teams for example, possibly in different physical locations and at different activities.

CAFIAS, therefore, was designed to record the original Flanders categories plus non-verbal behaviours, variations of the teaching agency, differential class structures and
greater sensitivity in student behaviours (Cheffers 1980). Ten non-verbal categories were placed alongside the Flanders original categories becoming the 'teen' equivalent of its verbal category. Additionally, a distinction was made between 'silence' and 'chaos' and also between censure and constructive criticism. CAFIAS, its ground rules and coding procedures may be found in Cheffers, Mancini and Martinek (1980). As for the Flanders system, coders may work in the live setting or use videotaped material and record a numerical symbol as a behaviour occurs. If a behaviour extends for a long period a time limit of three seconds is imposed. The data is summarised on to a matrix and frequency counts of each category are made and percentages and ratios calculated for each CAFIAS parameter together with the pattern of interaction between teachers and students, and among the students themselves. This three way interpretation of the data is one of the strengths of the system and the use of the matrix, in particular. It is possible to calculate the number of times a category is used and so to calculate the percentage of category usage in a lesson. Secondly it is possible to interpret the main CAFIAS parameters in percentages and ratios, for example, a teacher's verbal contribution, non-verbal contribution and total teacher contribution. This is made possible in each of the forty one parameters by adding together the tallies recorded in certain categories. Finally it is possible to establish patterns of interaction
between the teacher and students and among the students themselves. This is obtained through the use of the matrix where the density of tallies in cells reveals the predominant behaviours, their sequencing and their patterns of interaction.

To establish the reliability of the system Cheffers (1972) submitted cell rankings to Kendall's Coefficient of Concordance ($W$). In one comparison of the total matrices a ($W$) ranging from .60 to .81 was established. A second comparison of the ten main cells found a ($W$) ranging from .44 to .87. These comparisons were considered to be reliable at or beyond the .05 level of significance.

Cheffers' use of a well rehearsed system and its careful adaptation to include non-verbal behaviours in particular made it the popular choice of researchers during the seventies.

In the manual to the system (Cheffers, Mancini and Martinek 1980) the uses of CAFIAS are shown to extend to classroom subjects as well as to physical education lessons because of the expansion of the Flanders categories. Fifty four research projects are referenced between 1972 and 1980 which give testimony to its popularity.

CAFIAS itself has been subjected to modification, first of all by the designer in collaboration with a colleague, when a component was added to describe emotional episodes giving greater sensitivity in the non-verbal dimension (Cheffers and Mancini 1979).
Other changes included the MAINE modification (Cheffers, Archambault and Greene 1974) to measure specific behaviours in elementary school mathematics classrooms in southern Maine; the BAKE system (Batchelder and Keane 1977) for measuring college lecturing; the WOOD modification (Wood 1978) for measuring congruence with program objectives; ITBAS (Lewis 1978) which used both FIAC and CAFIAS to produce the Individualised Teacher Behaviour Analysis System; MODCAF (O'Donnell 1978) to examine the diversity of interaction patterns and levels of cognitive responses of small group discussions; PROXEMICAFIAS (Travis 1977) which is a system to look at proxemics and the pattern of classroom questions; an adaptation of CAFIAS for determining Fine and Gross Motor Activities (Hope 1978), and a Dyadic Adaptation of CAFIAS (DAC) for measuring Teacher and Individual Student Interaction (Martinek and Johnson 1978). In common with developments at other American Universities these studies were undertaken in part fulfilment of the doctoral dissertation. Although few were used in other research projects, collectively they contributed to a growing understanding and expertise in the use of observation systems at the time. Largely because of their origins and concern for the interactive dimension of lesson events they became known as Interactive Analysis Systems.
3.4 BEHAVIOUR ANALYSIS SYSTEMS: OHIO STATE UNIVERSITY

At Ohio State University, under the guidance of Daryl Siedentop, there developed a complementary interest in the use of observation systems which concentrated for the main part on one facet of the teaching-learning process, namely teacher behaviour.

The emphasis in the research programme at Ohio State was the development of systems which could be used for the study and modification of teacher behaviour. The most influential of these was the OSU Teacher Behaviour Rating Scale which developed from the earlier work of Breyer and Colchera (1971) and Pollack (1971). Unlike its name implied, the system, which was developed by Hughley (1973), was not designed to determine emphatically the effective teacher. It was considered that teacher effectiveness could be implied on the basis of the rate of occurrence of specific types of behaviours but that these rates should not be used to determine teacher effectiveness per se. The system was intended for use in applied behaviour analysis or behaviour modification work. Popular for use in the general supervision of student teachers, the system focuses on eight categories of teacher behaviour and uses an event recording procedure with five minute intervals. The eight categories include:

Category 1 'Input Teaching Acts' which includes all teacher behaviours which are directly related to learning and
function as a stimulus to the learner - for example asking questions, explaining, informing and providing guidance.

Category 2 'Managerial' which includes teacher initiated behaviours to establish order, relocate pupils, manage equipment and keep records.

Category 3 'Monitoring' which refers to watching pupils silently with no interaction evident.

Category 4 'No activity' caters for the non-interactive dimension of teacher behaviour.

Category 5 'Skill Attempt - Positive IF' refers to all positive verbal and non verbal reactions to a student's skill attempt.

Category 6 'Skill Attempt - Negative IF' on the other hand points to the negative verbal and non verbal teacher reactions to a student's skill attempt, which includes corrective feedback.

Category 7 'Positive Reaction to On-Task Behaviour' refers to the positive verbal and non verbal teacher responses to on-task behaviour which does not include skill attempts.

Category 8 'Negative Reaction to Off-Task Behaviour' refers to all negative verbal and non verbal teacher reactions to off-task student behaviour.
The system is relatively simple to use. Once the categories are learned for the purpose of easy identification the observer notes the time the lesson starts and for the first five minutes places a tally mark by the appropriate category under the first interval column each time a behaviour occurs during that first five minute period. A two minute rest period is recommended between the intervals. The time is noted at the end of the lesson to calculate the rate of behaviour in each category, the category tallies are totalled and divided by the number of minutes observed excluding the two minute rest periods.

In general terms the strategy is to increase the rates of behaviour in some categories and reduce them in others. Hughley (1973) considered it desirable to increase the rates of behaviour in categories five and seven and to reduce to zero categories three and four if the rates are substantial. He acknowledged that category one was difficult to comment upon except that certain kinds of those behaviours were both necessary and useful.

It is not possible to be precise in the use of this instrument for the somewhat critical reason that recommended behaviour rates do not exist.

It is worth taking a closer, critical look at this instrument for it raises issues which help to explain some of the criticism levelled at system designers and it might explain the developments in Ohio State's observation systems since 1973.
The selected categories necessarily reflect the philosophy of the designer. In this Rating Scale, the fact that one category can contain the majority of the input teaching behaviours which are not considered sufficiently important to distinguish one from another must make some kind of statement about the designer's view of their contribution to learning. This is confirmed when one sees that two whole categories describe the teacher's reactions to skill attempts by the pupil. Clearly the designer places skill acquisition as a high priority in physical education lessons and considers feedback from the teacher of paramount importance as a factor in pupil skill learning. It is interesting to reflect on how effective this strategy might be if, correspondingly, there is a reduction to zero in 'Monitoring' by the teacher. One might fairly judge that lack of time engaged in observing a pupil's skill attempts could lead to inappropriate feedback or no feedback at all. The placement of corrective feedback in a negative category is puzzling. The assumption here is that corrective feedback is not positive which cannot be a true reflection of corrective feedback which contains the information required for correction and usually a word of encouragement in addition. Using this system, this type of feedback would be split into two parts and tallied under categories six and five. This would not provide an accurate description of that teacher behaviour either in intent or in fact.

The placing of a tally each time a behaviour occurs notes the fact that it happened but not the length or the duration
of that behaviour. In one, five minute period therefore, it could be the case that four tallies totalled in both categories two and five might appear to be equal. Presumably the observer would hope to see many more tallies in category five than two but to describe these results as 'equal' would be inaccurate and meaningless. The four instances of positive feedback might have lasted a total of three minutes whilst the managerial behaviours could have lasted a total of forty seconds - or vice versa. To convert tallies into a rate per minute without knowing their original duration could, therefore, be considered to be misleading.

In spite of these criticisms it is acknowledged that this Rating Scale was not designed as a research instrument but as a simple system for use in the supervision of student teachers. Its simplicity should guarantee adequate levels of observer reliability, and Hughley reported levels of reliability ranging from 82% to 95% with undergraduate students learning to use the system achieving 85% reliability after only a few hours of practice. This is understandable because catch-all categories like one and two may lack discrimination but they help the novice in the early stages of system use, and event recording with built-in rest intervals is a sympathetic technique for the observer to handle.

This original system spawned many others which were used for the study of physical education teachers by researchers at Ohio State.
Siedentop and Rife, reported in Darst et al (1983), collaborated to produce DACOME-PE, a system of Data Collection for Managerial Efficiency in Physical Education. Focussing entirely on the management component of Physical Education classes it uses both an event and duration recording procedure to measure the managerial behaviour of teachers. Observers are required to record data for the entire class period.

The Managerial category noted in the Rating Scale is expanded in this system to include five types of managerial episodes, the length of the episode, the managerial behaviour, positive teacher reactions to student On-Task Behaviour and negative teacher reaction to student On-Task Behaviour.

This goes a long way to answering the earlier criticism of using 'catch-all' categories for such important behaviours. This system notes one aspect of a physical education teacher's behaviour in depth. By the same token the noting of the duration of the behaviour in this system makes comment on the totals in terms of the 'amount' or 'number' much more meaningful.

Also at Ohio State, Darst (1974), Hamilton (1974) and Boehm (1974) developed the OSU Rating instrument for use in their doctoral dissertations and ultimately to be used with student teachers in teacher training intervention programmes. Their developments of the original scale also showed a concern to provide a system which permitted greater
in-depth analysis along at least one dimension of behaviour.

Dodds (1975) acknowledges that her observation system 'STOP'—Student Teachers Observing Peers—is based on the previous work of Ohio State colleagues who developed the majority of the behaviour categories in the system. Designed to be part of a training intervention package, this instrument is intended for use by student teachers paired in the same setting in order that they can observe and provide feedback for each other. STOP consists of four, major categories of Teacher Behaviour: Instructional Input, Management, Skill Feedback and Social Behaviour Feedback. These familiar categories each contain more than one subcategory label. Additionally event recording, duration recording and placheck are utilised—placheck being a scan lasting approximately three to five seconds around the whole class. A full account of the procedures for use is reported in Darst (1983). This continuing expansion of the original system yields a greater quantity of more relevant data giving the student teachers for whom it is intended a relatively detailed account of their teaching behaviours within the outlined parameters.

Other notable students of Ohio State have produced observation systems ever increasing in their complexity (Stewart 1977; Olson 1979). Arguably the most influential of these has been that produced by Metzler (1979) who first used the variable of Academic Learning Time (ALT) conceptualised in the Beginning Teacher Evaluation Study.
The study was conducted in elementary schools during lessons in reading and mathematics in the early seventies. ALT was defined as the amount of time a pupil spends in relevant academic tasks at a high rate of success (Marliave 1977). Since Metzler's study a fully developed system known as ALT-PE complete with a coding manual has been available to researchers (Siedentop 1982). It reached this revised state after several studies had been conducted at Ohio State using this variable (Rate 1980; Birdwell 1980; Whaley 1980). Researchers in other Universities such as Texas, Liege, Victoria in British Columbia, Laval in Quebec and Georgia, USA also investigated its possibilities and it was as a result of this combined effort that the revised system was adopted for use in 1982.

In the coding manual, ALT-PE is conceptualised as a two-level hierarchical decision system making it somewhat more sophisticated than previous observation systems. In the first level of the system the observer has to decide on the context of the setting he is observing to determine whether the class is in general or subject matter content. The subject matter content is further divided into knowledge content and motor content between which the observer will need to choose.

In similar vein, in the second level of the system the observer has to decide on the learner involvement at an individual level to determine whether the individual is motor engaged or not motor engaged. The observer needs to know the goals of the activity in order that he can decide whether the learner is motor engaged in relevant,
goal-orientated activity otherwise he is considered to be not motor engaged for the purposes of coding.
Ultimately the observer is looking for the unit of ALT-PE which Metzler (1983) considers to be the best process measure of effective teaching in Physical Education to date.
To understand what this unit might be it is helpful to imagine an example of a class playing 3 V 3 Basketball games with the coach intervening on occasions and the target pupil actively engaged in the exercise and able to cope adequately in the game situation. The observer decides that the class are in a scrimmage situation (subject matter motor content) and that his target pupil is motor engaged in a motor appropriate situation. Finally he makes the decision that the level of difficulty for the pupil was also appropriate from the choices afforded by the system which include Easy, Medium or Hard. This final decision is of fundamental importance to the system. If the learner is achieving a high degree of success the activity must be sufficiently easy to grasp yet not too simple to underchallenge him. This is deemed to be motor appropriate and any observation sample in which motor appropriate is chosen for the second level decision becomes one unit of ALT-PE.
The options for measuring ALT-PE include interval recording, group time sampling and/or duration recording. These options, category definitions and coding procedures are all explained in detail in the coding manual (Siedentop, Tousignant and Parker 1982) and it is not proposed to reproduce them here.
ALT-PE has received much attention recently including the devotion of a whole Monograph to a review of its development and uses (Dodds and Rife 1983). It has been used primarily as a tool in descriptive research studies although its advocates remain convinced of its potential as a tool in teacher education. It is early days in its development yet its uses have grown rapidly accompanied by a rich source of reviews which fairly point out its limitations as well as extolling its strengths (Siedentop, Anderson and Metzler 1983).
3.5 MULTI-USE OBSERVATION SYSTEMS: TEACHER'S COLLEGE, COLUMBIA UNIVERSITY

During this same period of time on the east coast of the States in New York, students at Teacher's College, Columbia were also developing the use of systematic observation instruments. Theirs was a unique project. In 1971 the Data Bank Research Project in Physical Education was instigated for the purpose of developing and applying valid and reliable descriptive research systems to study teacher and pupil behaviour in physical education lessons. Co-ordinator of the project was Dr W.G. Anderson who, together with twelve doctoral students, collected eighty three videotapes of lessons conducted in elementary and secondary schools over a two-year period. These tapes provided the valuable raw material for both individual and collaborative studies in the field of system development and application. The systems which were developed focused on teacher and/or pupil behaviour and have subsequently been labelled 'Multi-Use Observation Instruments' (Darst et al, 1983).

One of the first systems to be developed was that by Fishman (1974) who chose to focus on augmented feedback provided by teachers for their pupils. For the purposes of the study, augmented feedback was defined as a teaching behaviour dependent upon the verbal or motor response of one or more pupils and intended to provide information related to the acquisition or performance of a motor skill. The system itself was a category system
selected, according to Fishman because it seemed appropriate for a recording instrument designed to focus upon a dimension of teacher behaviour (augmented feedback) which involved a temporal component of considerable importance. Additionally it was considered a more relevant choice because studies based upon previously developed theories had traditionally selected a category design (Medley and Mitzel, 1963) and feedback as a variable controlling performing and learning was well established. An analytic unit was selected in order that variables associated with the temporal component of augmented feedback could be retained. Coding, therefore, was based on the natural occurrence of units of teacher augmented feedback in six selected dimensions of the system described as 'Form', 'Direction' and 'Time' in the Methodological Dimension and 'Teacher Intent', 'General Referent' and 'Specific Referent' in the Substantive Dimension.

To use this recording instrument a tally was placed in one of the sub-categories belonging to each of the six dimensions whenever a unit of augmented feedback was observed. Six tallies, therefore, were placed for each coding.

In testing the use of this instrument by four teams of observers who analysed selected units of feedback, a mean level of overall agreement of 90.34% was achieved.

Tobey (1974) modified Fishman's system to distinguish between the positive, negative or neutral intent of the unit
of feedback and used the modified system to code teacher behaviour using the Data Bank tapes. Reliability figures for the system exceeded the .80 level. His analysis of eighty one P.E. lessons revealed 4392 instances of feedback.

At the same time, Hurwitz (1974) was developing a system to describe certain aspects of the P.E. Teacher's Role in the learning-activity selection process (TRI-LASP). A five second time unit was used to record the learning activity of the target student and the teacher's role in selection of the learning activity. The first set of categories used to describe the target student's learning activity focussed on the manner of behaviour, the broad content alternatives and the limited content alternatives identified by the designer. The second set of categories described the teacher's role in each of the preceding categories, identified by Hurwitz as 'encourager', 'identifier', 'predictor'.

To test the system Hurwitz selected 340 minutes of videotape and tested his four trained coders. Overall intercoder agreeability exceeded .90 and intracoder agreeability exceeded the .88 level.

Laubach (1974) was another member of the Columbia University doctoral team whose dissertation BESTPED focused on a student's use of time in Physical Education classes. The system contained four dimensions: Function, Mode, Content
and Time, and three forms of the system were produced each requiring more or less precision in the coding procedure and suitable for different users. A full description of these coding procedures may be found in Darst et al (1983). The system was tested using three coders and the figures achieved by using the percent of agreement method were 90.5% (inter coder reliability) and 88.8% (intra coder reliability).

Use of the system provided descriptions of how a student spent his time in Physical Education classes. The behaviour could be analysed to determine its function and substance, the duration and sequence of the behaviour and whether the student was mobile or not at the time. In common with other systems which focus on student behaviour, BESTPED permitted study of only one student at a time by any one observer.

One of the features of the Data Bank Project was the sharing of the work by pairs of researchers one of whom was the system designer and the other the user. As Tobey developed and used Fishman's system reported earlier, so Costello (1978) used Laubach's system to describe student behaviour in a sample of the videotaped classes. Costello selected twenty elementary classes for study and from these he focused on a sample of 193 students from a possible 469 for a total of 15 minutes per student. Full account of the results may be read in the summary printed in the Monograph which reported on the results of the Data Bank research team (Anderson and Barrette 1978) Whilst it is not intended to
reproduce all of the findings here, some of the results are worthy of comment for they illustrate the kind of quantitative data these descriptive studies produced and which have justified their continuing existence. For example, the 193 elementary school students were observed collectively for a total duration of 168,454 seconds. These students spent slightly less than two thirds (63.2%) of the total time they were observed in non-movement behaviour i.e. they spent a little more than one third (36.8%) of their time moving in P.E. classes. Furthermore a little more that a quarter (27.5%) of their time was spent in movement related to the accomplishment of various physical education objectives. The largest proportion of a student's time was spent in waiting either for a chance to play in a team or practice or in line for a turn to participate. In all, 35.4% of a student's time was spent this way. Costello attributed this high percentage to various factors which included the nature of the activity (e.g. gymnastics), periods of time spent in teacher talk and inadequate organisational ability of the teachers. In summarising the implications for teachers, Costello comments on both the more obvious ways which might be employed to reduce the amount of time a student spends in waiting and the less obvious, such as a shift in the teaching methodology used in teaching physical education. Giving a student greater responsibility for his own learning could be accomplished by using an alternative teaching style such as 'guided discovery' or 'problem solving' rather than the lecture method or didactic approach commonly used by P.E.
teachers. Costello speculated that such a change might improve opportunities to explore the environment in order to solve various motor problems, give information or engage in dialogue with the teacher and other students, and assist the teacher or other students in the class.

Morgenegg (1978) was also interested in the behaviour of students but chose to widen his observations and focus on the teacher also. His interest lay in the pedagogical moves in the classroom, a term used earlier by Bellack (1966), to describe the communication manoeuvres of teachers and pupils. Morgenegg adapted the original system to take account of the subject requirements of Physical education. A full description of the system dimensions and categories appears in the Monograph (1978).

The system was used to code forty physical education classes from the Data Bank tapes, of which twenty were from elementary schools and twenty from secondary schools. From his research he found that teachers dominate the pedagogical move labelled 'soliciting' (39.8%), children spend much time 'responding' (30%), and teachers then 'react' (15.7%). Whilst there was little variability in these roles from class to class, Morgenegg found that the ways in which these roles were exercised did vary considerably according to class activity, teacher personality and class size, in particular. Overall teacher moves are short in duration, lasting between two and three seconds, unless the teacher is 'structuring' a situation in which case he tends on average to take as long as twelve seconds. Teachers tended to
solicit directly rather than indirectly, a point noted by Costello (1978), and they tend to react positively twice as often as negatively. Speech dominates the means of communication used by teachers, yet pupils' responses were shown to be largely non-verbal. Additionally, pupils' reactions tended to be neutral rather than positive or negative. Morgenegg acknowledged that such results painted a particular picture of life in the gymnasium, noting that there seemed to be certain rules that govern the ebb and flow of communications and thus define the respective roles of teachers and pupils.

Director of the Data Bank team, Dr W.G. Anderson, produced his own system (1974) which was used to describe Teacher Behaviour. The system was developed over a two-year period and was intended to provide a professionally meaningful description of teacher behaviour. To achieve this objective the categories in the system were designed to classify teacher behaviour using the terminology and concepts that were familiar to the professional educator. Reliability tests yielded acceptable levels of agreement between coders for each of the major dimensions of the system ranging from 86.7% agreement to 99.6% agreement. Full descriptions of the system and its coding procedures may be found in Anderson (1974), Barrette (1977) and Darst et al (1983).

The system was designed for use by researchers to describe videotaped teacher behaviour. Its use will provide a
systematic account of the teacher's behaviour from the predetermined start of the lesson to the end. To use the system an observer would first decide the category in the main function dimension which best describes the teacher's behaviour at the time. He then codes the other categories in the next dimensions in relation to the function dimension. These other dimensions add substantially to the overall descriptive power of the system, going beyond the straightforward function description to include note of the function subscript (indicating whether or not the teacher carries out the whole of the function or shares it); the mode (means of communication); the direction (to whom the communication is directed), and the substance (the subject matter of the interaction).

The system was used by Barrette to code forty Data Bank videotapes which included twenty elementary schools and twenty secondary schools. The average duration of the P.E. lessons was approximately 33 minutes. The results showed that teachers are highly interactive, spending 94% of their class time performing interactive functions. Silently observing pupils accounted for at least 21% of a teacher's time, although this figure did not include any periods of observation which lasted for less than five seconds in length. Concurrent instruction was the most frequently occurring teacher behaviour (17.8%), closely followed by Observing (16.9%) and Officiating (14.1%). Teachers spent 37% of class time attending to instructional pursuits, a figure which rose to 75% if one included
Instruction Related Behaviours and Observing. These figures did not vary significantly between elementary and secondary teachers. In common with other reports, teachers communicated verbally almost exclusively and students seldom initiated verbal interaction. Talk about the substance of a lesson was absent and seldom did teachers question students. Demonstration was not a popular teaching technique, teaching aids were rarely used and written materials were almost never used.

The researchers at Teacher's College, Columbia University provided those interested in Teacher Research with a rich source of information about 'What's Going On In Gym'. Their Data Bank of video tapes proved a valuable resource for not only their own doctoral students but other researchers - notable among them being Cheffers and Mancici whose work was reported earlier, and who were invited to analyse all of the tapes using CAFIAS.

It is interesting to note that whichever system was in use or by whom ever, the pattern of the results varied little. The teaching analysed belonged mainly to a traditional mode. Teachers taught directly and talked a lot; children listened and responded non-verbally. There was a lot of waiting around in evidence and little student involvement in their own learning. One might conclude that there was much indifferent teaching going on that was used in this sample of tapes yet this is refuted by those involved in the Data Bank project. The fact that the same tapes were being
analysed and reanalysed by the different researchers would account for much of the consistency in the findings even though different systems were in use. It is also important to note that the tapes were of indoor lessons only and therefore of a limited number and type of physical activity which would account for some of the findings.
3.6 THE RESPONSE FROM OTHER COUNTRIES

Material from Boston, Ohio and Columbia following their respective traditions of Interactionist, Behaviour Modification and Descriptive-Analytic Research, provided a wealth of information for those engaged in research in physical education and prompted corresponding research from universities in other countries.

In Belgium at the University of Liege, Pieron and his colleagues have produced many studies which reveal the influence of earlier American work. In conjunction with Mathy, Pieron has produced a bibliography of works which focus on the Study of Teaching Physical Education, Teaching Effectiveness and Teacher-Students Interactions and Related Subjects (Pieron and Mathy, 1981). Twenty six of these 429 works were produced at the University of Liege.

In Finland, Heinila adopted an Interactionist approach to design a system for describing Teacher-Pupil interaction in P.E. classes whilst Telama developed conceptual schemes of evaluation in P.E. Their published works are also reported in the Bibliography (Pieron and Mathy, 1981).

In Canada a team of researchers has emerged from the Universities of Victoria and Laval, Quebec, with an interest in coach behaviour and the use of systems to study teacher and student behaviour (Brunelle 1980; Tousignant 1982 and
In England the popularity of this type of research has been less marked. Fox (1976) devised a rating schedule for use in the P.E. lesson. Twenty one Teacher, Pupil and Context categories were devised for an observer to complete during a lesson.

Hallam (1978) produced a modified version of the Boydell Teacher and Pupil Records designed for the observation of educational dance and gymnastics in junior and middle schools.

Underwood (1976), following the Flanders tradition, produced an interaction analysis system for studying teacher behaviour in Physical Education using a total of nine categories.

Classen and McAleese (1976) turned their attention to the Primary School P.E. teacher in their design of the Primary Physical Education Teacher Analysis Record (P.E.T.A.R.) which is an instrument designed to record and analyse specific dimensions associated with the teacher - namely his movement and verbal/nonverbal behaviour during the lesson.

More recently Mawer (1981) designed an observation instrument which would describe the nature of teacher guidance behaviour in primary and middle school educational gymnastics lessons. Known as P.E.T.G.A.S.(Physical Education...
Teacher Guidance Analysis Schedule) it comprises a set of operationally defined categories of teacher guidance behaviour within the major dimensions of rejecting, focusing, accepting, questioning and organising. Mawer used the instrument to analyse categories of behaviour, patterns of teaching and styles of teaching. He identified seventy different forms of interaction which were classified into teaching 'moves' or 'cycles' according to length, direction, specificity and content.

3.7 THE IDENTIFICATION OF A RESEARCH INTEREST

Studying the teaching-learning processes in physical education classes in England using systematic observation techniques is still in its infancy. A small but nevertheless welcome start has been made by these researchers. It was partly this apparent apathy which prompted an interest in this type of research because the problems which had been apparent in the United States and which were outlined earlier seemed to be very evident in this country also. It was difficult, if not impossible, to establish what was actually happening in physical education classes because descriptions gathered systematically were not available. At the same time there were increasing calls for accountability and the need for research in Teacher Education both in Colleges and Universities just as there had been earlier in the United States.

Working with students on Initial Training courses and with
teachers In-Service prompted further interest in Teacher Behaviour as a focus for study. It seemed that recent calls for teacher assessment could not be met adequately until the means were available for providing accurate descriptions of teacher behaviour in the lesson. It seemed appropriate that such means were equally relevant for use by those inside the education system working in the schools and institutes of Higher Education as those outside of it looking in, requiring teacher profiles.

For this reason it was decided to focus on Teacher Behaviour exclusively and use a system which was professionally relevant to teachers and students alike which would provide the observer with accurate, meaningful descriptions of teacher behaviour in the P.E. lesson.

The most relevant system in use was that produced by Anderson and used by Barrette (1978) because it provided a broad description of the teacher's behaviour and was professionally orientated. Unfortunately it had been produced in the United States for use in American classes, and the sample was of their type of P.E. activities further restricted to indoor work only, which rendered it not entirely useful for the provision of descriptions of outdoor games work in Britain. As this aspect of work constitutes such a large proportion of our P.E. lessons then this was felt to be a major point of disadvantage.

The decision was made to use the Anderson system as a starting point to organise descriptions of teacher behaviour but to devise another system which would describe teacher behaviour during the teaching of games lessons. These would
be both indoor and outdoor games lessons with all the attendant problems of the necessary technology to film outdoors in inclement weather having to be resolved.

3.8 SUMMARY

By way of contrast with research in classroom education, physical education teaching lacked a research tradition and was operational in a work environment which posed constraints for researchers. This meant that little was accomplished until the late 1960's, by which time colleagues in the classroom could provide considerable guidance in the design and use of systematic observation instruments. American Universities took a lead in the initiation of programmes in descriptive-analytic research producing a series of observation systems capable of being used to describe the different dimensions of teacher-pupil behaviours in physical education.

At Boston University, Cheffers pursued the interactionist lead of Flanders by modifying his system in order that it may be used in the context of physical education and spawning numerous others as a result.

At Ohio State University, Siedentop preferred to concentrate system development on one facet of the teaching-learning process. His teacher behaviour systems were to be used principally in behaviour modification programmes for those pursuing a career in teacher education and, in common with Cheffers, the early work in the 1970's has led to numerous
influential systems in use today. At Columbia University in New York, Anderson and his colleagues developed Multi-Use Observation Instruments dating back to 1971 when the Data Bank Research Project in physical education was begun for the purpose of developing and applying descriptive research systems to study teacher and pupil behaviours. Contributions from other sources to the description of events in physical education lessons have been comparatively few in number and concentrated mainly in Canada and Europe where it has followed the lead provided by the Americans. Lack of experience and expertise in Britain in particular prompted an interest in the development of a system which might be used to describe teacher behaviour in games lessons, and the category system produced by Anderson in Columbia University, New York was selected as a starting point for the organisation of material necessary for instrument design.
CHAPTER 4

The design of an Instrument: BOS 1

4.1 INTRODUCTION

The decision having been taken to design an instrument which could be used to describe teacher behaviour during physical education (games) lessons, it was necessary to obtain relevant material which could be used for analysis and to begin the process of instrument design.

This chapter outlines the stages in that process from the initial attempts at filming material through to the modification of the first system made necessary by its complexity and consequent difficulties encountered in observer training. The training process is detailed in Appendix 2 and the Observer Training Manual (Bailey, 1981).

4.2 THE PRODUCTION OF VIDEOTAPED LECTURE MATERIAL

For the purpose of instrument design it was necessary to analyse videotapes of physical education lessons. Taped material was not available initially for this research was breaking new ground at the University where the Audio Visual Aids Department was in the process of re-organisation and it was necessary to wait until a new television unit was established before recording might begin. It was fortunate that the new unit saw the research as an opportunity to test procedures and new techniques and agreed to co-operate in
the programme.

It was decided to test the equipment on the University campus before taking it out to schools because recording out of doors was unknown territory for the television crew. Three lectures of indoor basketball, tennis and outdoor field hockey were recorded. The recording sessions provided the invaluable experience necessary for the production of good quality videotapes of games lessons in schools.

The equipment comprised an Hitachi camera FP3060 Single saticon tube with a 6/1 zoom lens. This was more sensitive than the standard vidicons which was of importance when filming in natural light. Sixty minute Sony U-Matic cassette tapes were used to record the lessons during which the teacher was fitted with a Shure radio microphone and an audio transmitter/receiver. A license for its use was obtained beforehand.

Two technicians were required to transport and operate the equipment excluding the researcher in order that she might be active as an observer during recording. A great deal of technical preparation was necessary prior to recording, the major points of which are noted here.

It was necessary to calculate the time for transportation of the equipment from the vehicle to the venue for recording and this varied considerably especially when playing fields were a distance from the main buildings. At some venues it was possible to leave much of the equipment in a van which could be driven to the edge of the field. Rigging time was approximately ten minutes for indoor recording and twenty minutes for outdoor recording. The temperature had to be
noted for this camera and video unit required a minimum working temperature of five degrees centigrade. It had to be switched on and left to reach its operating temperature and condensation allowed to evaporate. Whilst this time was dependent on humidity and air temperature it could be up to half an hour in extreme conditions which had to be added to preparation time.

New batteries had to be fitted to the radio microphone for each recording. In attaching the microphone to the teacher it was best done by slipping the unit into a pocket or alternatively the microphone could be clipped to a V-necked sweater, taped to the person or attached to a belt. The wearing of a nylon garment would interfere with reception through the build up of static electricity.

A flat surface was essential to support items of equipment and early location of this and power points was needed particularly for outdoor filming. Whilst it was possible to use a portable pack equipped with batteries, outdoor games lessons in this country are often over one hour in length which goes beyond the life of most power packs.

It was necessary to calculate the length of cable required for recording. Additionally it was found that the camera could not be operated in snow because of excessive glare nor in the rain because none of the equipment was waterproof although shields could protect the camera from a light shower, briefly. For indoor work unless supportive light was used it was better to place the camera with its back to a window and in a corner for wide angle vision. It was important to take care that cables did not intrude on to the
4.3 LESSONS LEARNED DURING THE FILMING PHASE

Lessons were learned regarding the setting up of this type of programme with colleagues. Such a programme requires a considerable commitment on behalf of technical staff because lesson recording requires at least a half day away from base. Staff need to know the full extent of this commitment before the programme starts. Flexibility is important because not all tape recording sessions may be successful, requiring extra visits perhaps to alternative venues. Technical staff may require practice before the tapes are of an acceptable standard and should the tapes require editing practice may also be required. Tape analysis will necessitate the use of viewing facilities. The member of staff taking the recorded lesson will need careful instructions and it is necessary to attach the microphone to the teacher well before the lesson begins. The lesson start may not be obvious so an agreement will have to be reached with the teacher as to when the first significant act will occur. This may be in the changing room and out of sight of the researcher therefore the teacher must know how to attach his own microphone and assemble it for use. It may be necessary to interrupt the lesson prior to the one being filmed in order to set up the equipment and possibly run over into the next one whilst packing away. All staff
would need to be appraised of this situation. Some staff perceive the use of cameras to be intrusive and such a research programme would need careful explanation to give confidence to participants. The telephone is not a successful medium for such explanations. Finally, time should be made available for showing results of recording to participating staff.

4.4 INSTRUMENT DESIGN (1):
CHECKING THE ANDERSON SYSTEM OF TEACHER BEHAVIOUR

The recorded basketball, tennis and hockey lessons were used to check the categories and definitions of the Anderson System.

It was anticipated that there would be some difficulties in using categories which might be relevant to Modern Dance, Parachute activities, Scooter activities, Wrestling and Relay Races for example, and less relevant for the analysis of teacher behaviour in games lessons in this country. Anderson's list of activities also included single lessons of weight training, Indian wrestling, Hula hoop, Rope jumping and Folk dance.

Perhaps it was partly due to this diverse spread of activities that the system overall was found to be very general in its application. Additionally some of the categories were scarcely used and other appeared misplaced under particular headings.

For example, 'Spotting': defined as 'Interacting with
students for the purpose of protecting them against injury, assisting them, and positioning oneself for the above' would be very relevant for gymnastics but not for games. Similarly 'Leading Exercises' which referred to 'counting cadence, regulating the performance of exercise, etc' might occur frequently in dance lessons but would be unusual in games work.

Anderson had chosen to place all those functions which identified the purpose of the teacher's interactive behaviour in one set of categories. These eight categories had been subdivided so that a total of eighteen categories described the teacher's behaviour. Examination of Category 1 'Preparing For Motor Activities' revealed a mixture of teaching and organising in the three subdivisions, namely 'Organising', 'Preparatory Instructing' and 'Providing equipment or Readying the Environment'. In similar vein Categories 2 through 6 contained a mixture of teaching and organising functions.

It was considered that there would be merit in organising these functions under two main dimensions of Management and Teaching. These two dimensions were quite distinctive with Management engaging the teacher in the organisation and control of a lesson for the purposes of facilitating teaching and learning and Teaching being the instruction of pupils for the purpose of achieving learning. Such a distinction for the teacher's behaviour would not stop the teacher instructing the pupils in an aspect of management or organisation for the purpose of the pupil achieving learning.
in the method of erecting equipment, for example, and this would clearly be Teaching as opposed to Management. Nevertheless such a distinction would serve the purpose of achieving clarity for the coder who has to make rapid decisions regarding the identification of categories and the placement of tallies on a coding sheet as well as provide a useful division in the data for the estimate of how a teacher spends his time in one or other of the two principal dimensions of behaviour.

The category 'Officiating' included the starting and stopping of activities which Anderson indicated had resulted in a higher percentage of tallies in this category than was warranted. Use of this category convinced the author that to include start/stop under officiating was misplaced in any case because it might have nothing to do with performing the duties of an official which was the principal function of an umpire or referee. Starting and stopping activities in physical education lessons can use a lot of time and therefore it was considered appropriate to place this behaviour in a category of its own.

In these early days of watching teacher behaviour, identifying Anderson's categories and determining their usefulness and appropriateness, it appeared that there was a dimension of teacher behaviour missing from the system. One of the hallmarks of a physical education teacher's behaviour which might distinguish that teacher from another is the mobility of the P.E. teacher. Even watching only three
lecturers one was struck by the differences in physical approach to the lecture. It was felt that an attempt should be made to capture this aspect of behaviour in any description a system might afford.

The dimensions Function Subscript, Mode and Direction were more straightforward and presented fewer problems from the point of view of identification or applicability. The category 'Listens' was open to debate simply because it could only be inferred and was rarely mutually exclusive. The Mode dimension also had no provision for the description of 'physical contact' other than purposeful manual assistance by the teacher. Observation of the three lecturers revealed frequent casual contact such as a pat on the back or hand on a shoulder. Also in this dimension it was felt that the use of a whistle in games work was so frequent as to merit a category of its own.

The major dissatisfaction in using this system lay in the fact that the teaching categories were very general and told the analyst interpreting the codings very little about the teaching method employed by the teacher. The word 'Instructing' was not broken down to provide a description of what the teacher was doing whilst engaged in instruction, only whether that instructing was preparatory, concurrent or intervening. It was considered that if a system was to be devised which might be used to provide a description of teacher behaviour then the teacher would need to know more about what he was actually doing whilst he was teaching.
In an attempt to answer this criticism the Anderson system was put on one side and the games lectures were studied to see if descriptions of teacher behaviour could be obtained which referred to teaching behaviour only. To do this, statements made by the lecturer were examined and a label pinned to each which best described that statement. For example, the statement "Well done - good shot" was the teacher's evaluation of the student's performance and he was offering the student feedback which was evaluative. In another example the teacher said "You took off your left foot" which was another example of feedback but this time the teacher was offering a description of the take-off to the student without evaluating its worth.

Analysing teaching behaviour this way proved to be a time-consuming task and fraught with difficulty. The chief problem lies in the fact that much can be inferred which might not actually be the case. The teacher's intent cannot readily be incorporated into a category system for the simple reason that no observer can ever accurately estimate and provide a description of a teacher's intentions. Initially the attempts to analyse the teacher's statements in these lectures contained numerous examples of high inference labels which had to be abandoned, such as teaching for 'technical appreciation' or 'consolidation', 'progression', 'correction' or 'transfer'. Whilst these may have been the intentions of the teacher at the time they could only have been confirmed retrospectively by the teacher. Perhaps more important for the purpose of
instrument design was the fact that they were descriptors of intent rather than descriptors of action by the teacher, a distinction which is vital in instrument design.

The other important factor was the necessity to arrive at a set of categories which were mutually exclusive. In other words there should be no overlapping of the categories so that if a coder wished to place a unit of behaviour in one category he would not be faced with a difficult choice of deciding between at least two categories. This would occur if the definitions of those categories showed similarity or if the coding procedures permitted the use of more than one category in the same dimension at the same time. For example the teacher can be observing and providing feedback simultaneously. Anderson got around this difficulty by stating that observation had to be silent and last for at least 5 seconds for it to receive a mutually exclusive tally. This showed that it was the predominant purpose of the interactive function at the time.

It was recognised that there was limited usefulness in attempting to finalise the design of the instrument through the study of three recordings of university lecturers' work because there were bound to be differences in content, method and context between these lectures and the school lessons due to be recorded. Nevertheless study of these tapes provided a useful beginning to the whole exercise.

As ideas and points of procedure were clarified so examples of teaching behaviours were put under descriptive headings.
A system of marked cards was used to contain these examples. A red card contained the category number, the title and the definition of the category. A yellow card followed giving examples from the tapes of the teachers' verbal or nonverbal behaviour which would be coded under this category. An example is given below:

**RED CARD**

Category 8
Title: Starts/Stops Activity
Definition: Gives instruction to start/stop activity
Notes/points of procedure: The instruction must be specific and stated, not implied (even if the activity stops or starts as a consequence) unless a whistle is blown. In this case the whistle takes the place of speech.

**YELLOW CARD**

Example: 'Stop!'  
'Hold it there!'  
'Right - off you go!'

Final whistle at the end of a game to conclude activity.

4.5 INSTRUMENT DESIGN (2):

SELECTING THE DIMENSIONS AND CATEGORIES OF TEACHER BEHAVIOUR

Prior to collection of the school recordings a total of forty three categories had been identified and placed in dimensions resembling those used by Anderson. Not all of these were given definitions and complete examples at that
stage because it was surmised that changes would be inevitable. A list of these dimensions and categories appears in Table 1.

Table 1

Dimensions and Categories of Teacher Behaviour

Dimension: Interactive Function

MANAGES
1 Organises People
2 Manages Equipment
3 Establishes/Enforces Codes of Behaviour
4 Officiates
5 Checks for Safety and Welfare
6 Administers Policy
7 Starts/ Stops Activity

TEACHES
1 Prescribes pupil response
2 Describes pupil response
3 Evaluates pupil response
4 Observes
5 Gives Information
6 Seeks Information

OTHER INTERACTIVE BEHAVIOUR
1 On- Task
2 Off- Task

Dimension: Function Subscript

1 Does
2 Shares
3 Delegates

Dimension: Mode

1 Speaks
2 Whistles
3 Gestures
4 Observes
5 Demonstrates
6 Uses Pupil Demonstration
7 Manually Assists
8 Contacts
9 Performs a Task
10 Participates
11 Uses Aids

Dimension: Direction

1 One
2 Group
3 Class
4 One plus either group or class
5 Other combination
6 Other person(s)

Dimension: Action
1 Mobile
2 Stationary

Dimension: Non-Interactive Function
1 On-Task
2 Off-Task
3 Absent from the scene
4 Inadequate VTR
4.6 THE PRODUCTION OF VIDEOTAPED LESSON MATERIAL

Arrangements were made to visit two secondary schools in the locality to videotape record a number of physical education lessons. The schools had adequate facilities for games work although neither possessed a sports hall which, it was felt, made them representative of the majority of secondary schools. The schools had sufficient numbers of staff to enable the filming of different teachers and their programmes contained several types of games.

The two schools which agreed to take part in the exercise were a mixed, Catholic Comprehensive school (11 - 18) and a mixed Comprehensive school (11 - 14). Both schools had playing fields, a hard court area and a gymnasium for games work.

A total number of nine staff agreed to have lessons recorded of whom three were students undertaking their final teaching practice as part of their post-graduate certificate in education.

Each participant plus the Headteacher had the purpose of the research programme explained. The draft observation instrument was shown to help with this explanation. The point was made strongly that the instrument enabled the observer to obtain descriptions of teacher behaviour and was not able nor designed to evaluate that behaviour. Each was informed that tape recordings and completed observation schedules would be available to participants should they wish to take advantage of this offer.
A timetable of visits was planned over a four week period during the months of March and April which provided poor weather for filming purposes. Whilst the programme did get interrupted several times it was felt that records of realistic games lessons in less than ideal circumstances had been obtained.

Fifteen videotaped lessons were suitable for use at the end of the filming period. Two of these developed intermittent speech faults but they were able to be used in part. The tapes included the recording of seven games, namely rugby football, netball, hockey, tennis, volleyball, basketball and badminton. These were the games on offer at these two schools during the Spring Term and were considered to be a typical selection. Of the four outdoor and three indoor games, one was taught on a field, three on the hard courts, one in the school hall and two in the gymnasium.

Each teacher taking part in the project was given a Teacher profile form to complete, a copy of which may be found in Appendix 1. The purpose of this was to check details of the participants' teaching experience and background training to ensure that the sample was not too similar as far as was possible. Excluding the student teachers the staff were all aged between twenty one and forty five years and had taught from between one and fourteen years. From the five female and four male teachers, five had taken the post-graduate route into teaching and four had taken Certificate courses at specialist Physical Education Colleges.
Notes were taken on each lesson which might be helpful to an observer in the future, providing the details of the class size and age of the pupils, the place, date and time of the lesson and brief notes on the lesson content together with the teacher's objectives for that lesson.
4.7 INSTRUMENT DESIGN (3):
CHECKING THE CATEGORIES OF TEACHER BEHAVIOUR

The videotaped lessons provided the material necessary for the checking of the proposed categories prior to the putting together of an Observer Training Manual in order that others might validate and use the System.

In order to check the proposed categories the videotapes were played with particular attention being paid to the teacher's verbal exchange with pupils. All behaviour had to be able to be coded with the minimum number of behaviours falling into the category 'Other Interactive Behaviour'. Only one aspect of the viewed teachers' behaviours appeared to warrant another category in the system. Perhaps it was a significant difference in the two types of educational establishments, i.e. the university and the school, that teachers did spend time seeking the attention of their pupils whereas lecturers did not. The category 'Seeks Attention' was added to the Managerial section of the Interactive Function Dimension.

As the Teaching categories represented those which differed markedly from any in the Anderson System they were looked at closely during this preliminary examination of the teachers' videotapes.

In assigning labels to these categories originally it was considered important to select a word which most accurately described the behaviour being displayed. Fishman (1974) had described an aspect of feedback as being 'prescriptive' and
this seemed a most appropriate label for much of a teacher's form of instruction, not only feedback intended to provide instruction for the subsequent performance of a motor skill. Teachers were frequently engaged in prescribing precisely what the response of the pupil(s) was to be, whether by direct command or less directly by using prompts, cues or suggestions. Flanders (1970) had termed this 'Giving Directions'. One of the important results of using this whole section devoted to describing the Teaching behaviours should be that a picture of the teaching style is obtained by the observer. This tendency for the teacher to prescribe for the pupil what should be done, when and how was an important aspect of behaviour to note in the instrument design.

Fishman's (1974) categories of feedback provided the starting point for considering which should be included, if any, in this system. The giving of feedback is such an important aspect of a P.E. teacher's behaviour that it was considered any system purporting to describe that behaviour would be incomplete without it. Fishman's study of teacher behaviour showed that the most popular form of feedback was Evaluative. Examination of the videotapes showed that this was undoubtedly the case with teachers frequently commenting on pupil performance and judging its worth with such statements such as 'Good kick' or 'Well done' or 'That was rubbish'. Even more subtle forms of evaluative feedback, such as 'Some of you are standing far too far back' or 'A lot of you made it even harder for yourself', nevertheless
informed the pupil of the value that the teacher had placed on his response. It was decided to include a category labelled 'Evaluates Response' in the Teaching section.

One of the other feedback categories which Fishman used was 'Descriptive'. This category described that type of feedback which a teacher offered to provide the pupil with an account of the performance of a motor skill. This was considered to be potentially important in the description of a teacher's style of behaviour. Unlike the direct approach which might use the shortest route to telling a pupil how to do a skill and getting him to do it, by providing the most direct forms of evaluative and prescriptive feedback, the offering of a description of that performance does allow the pupil an opportunity to reflect on his performance and self-correct if he is able. The use of the category in this way would be indicative of a 'guided discovery' or a 'problem-solving' approach, for example. It was decided to include this category labelled 'Describes Response' in the Teaching section.

It was decided not to include any other feedback categories in the system. Fishman had identified two others, namely Comparative and Affective, both of which accounted for very few tallies in her results. As this system was not intended to be a specialist feedback system but one which would accurately reflect the whole of a teacher's behaviour, it was considered that a reasonable balance should be struck in the selection of categories to describe that behaviour as a whole.
The category 'Observes' was readily identifiable from the tapes and was included in the Teaching section to take account of a teacher's silent watching of a pupil who was engaged in an appropriate learning activity deemed to be part of the lesson.

When the teacher was not being prescriptive, giving feedback or watching his pupils, he was often engaged in giving information to them. This varied from a fact about an activity, for example, to a long explanation of the Rules of Play associated with a game. Occasionally the information was quite elaborate, or it might provide a short introduction to the lesson content that day or even a retrospective summary at the end of the lesson. Such a category has been used in numerous systems to indicate when a teacher might be 'Informing' or 'Explaining' something to his class. In the Teaching section this was labelled 'Gives Information'.

By contrast the teacher was also observed to 'Seek Information' from the pupil(s). This was more than simply asking a question which could have been managerial in context. It indicated a teacher's questioning of pupils about matters relevant to lesson content to which a response was invited. Again this was considered an important aspect of a teacher's behaviour because it was indicative of a certain style unlike that of an overwhelmingly prescriptive teacher.
These categories accounted for almost all of the behaviour exhibited by the teachers whose lessons were videotaped. A 'catch-all' category labelled 'other interactive behaviour' was available to claim any other behaviours not accounted for by the system. Note would be taken of the use of this category for its overuse would indicate that inclusion of other categories might be necessary.
In order for any Observation System to be made valid it must be open to the scrutiny of fellow professionals. They are charged with the task of examining the categories, for example, and deciding if they agree with their selection, nomenclature and procedures for use. Above all they need to look at a system and check that it describes that which it purports to describe. In using the system they will need to be tested for reliability in its use to ensure that they can establish high levels of agreement.

Knowing that such procedures would be necessary it was decided to produce an Observer Training Manual at an early stage which would be used to train experts in its use and serve to test the validity of the system as a whole. It was recognised that there would be many changes to this text before the system was fit for use but that it would provide a useful start for the production of the final set of Ground Rules which would accompany the system.

In preparation for the production of an Observer Training Manual a number of stages had to be achieved.

The card index file of definitions and examples taken from the videotapes was completed to form the basis of a Glossary of Terms with Further Information for observers. Sections from some of the videotapes were transcribed to provide practice in the identification of the key Managerial
and Teaching categories.

A coding sheet was designed which listed all of the categories and provided boxes in which tallies would be recorded opposite the identified category.

Sections from some of the videotapes were coded on these sheets to provide illustration of their use plus identification of the categories.

In order to permit continuous coding, the unit for analysis had to be identified and the procedures for coding outlined. Adams and Biddle (1970) indicated that there were four units of analysis which could be chosen for use in descriptive-analytic systems and these were described in Chapter two. It was decided that the analytic unit was the most appropriate for this system, chosen to reflect the two key concepts of Teacher Behaviour being studied, namely Management and Teaching. A procedure was worked out for plotting the unit of analysis on to the coding sheet. At each coding, the digital time display would be noted, a content code ascribed to the lesson tape profile, the teacher's main function of either teaching or managing, whether the teaching was preparatory, intervening or concurrent, the function subscript, the teaching or managing act, the mode, direction and action of the behaviour.

Finally the Manual was prefaced with an explanation of the system and an indication of the training schedule necessary for learning to use the system (Bailey, 1981).

4.9 INSTRUMENT DESIGN (5):
OBSERVER TRAINING

Five University lecturers from the Department of Physical Education and Sports Science were selected by the supervisor of the research project for observer training. Each was invited to attend a preliminary meeting at which the project would be outlined.

During this and subsequent observer training sessions it was decided to keep notes and prepare a summary report which would be circulated to each observer. Additionally a diary of events noting the problems encountered by the observers would be kept by the author. This diary would not duplicate the session reports but it would give an account of the difficulties encountered by those in training and the author which related to both the management of the sessions and the observers themselves. Its principal purpose would be to inform readers and users of this system and potential designers of other systems of the many types of difficulties which can be encountered and which may be a necessary, if not a widely reported, part of system development.

The reports and the diary may be found in Appendix 2. A final Observer Training Report may be found in Appendix 3.

4.10 INSTRUMENT DESIGN (6):

SYSTEM MODIFICATION

Completion of this initial phase of system development
resulted in the decision to modify BOS 1 in order that it
could be made simpler, take less time to learn and so be
made available to a wider audience of users. Preliminary
observer training had resulted in modifications to the
Observer Training Manual and to the form of training
required in the learning of a system both of which would be
taken account of in the development of the modified system,
hereafter referred to as BOS 2.

4.11 SUMMARY

The process of instrument design was started by using the
Anderson system as a guide to selecting the categories of
Teacher Behaviour most relevant to teachers of Physical
Education.
Initially three videotapes of games lessons taken by
university lecturers were used to check the dimensions and
categories and to establish the requirements for filming
games lessons.
This was followed by filming fifteen games lessons taken by
nine teachers in two secondary schools in order that the
system could be refined and made ready for observer
training. The process of refinement included the selection
of dimensions and categories of teacher behaviour together
with definitions, a glossary of terms and the coding
procedures for system use. These were put into an Observer
Training Manual and the system was labelled the Bailey
Observation System 1 (BOS 1).
The observer training was undertaken finally by four
university lecturers and a diary was kept of the training sessions. As a result of the experiences encountered in training which are detailed in Appendices 2 and 3, it was decided to alter the system in order that it should be quicker to learn, easier to use and more readily available to a wider audience. The Manual for BOS 1 would serve as the basis for the development of BOS 2.
CHAPTER 5

The Modification of BOS 1

5.1 INTRODUCTION

This chapter details the continuation of the instrument design process beginning with the necessary modifications to the original system in order that it could be made easier to use. The result was a system with fewer than half the total number of categories and an interval recording procedure, making the process of observer training altogether more satisfactory and yielding test results acceptable for reporting purposes. Particular attention was paid to the method of system testing used in this study which is the most rigorous known to be available.

5.2 INSTRUMENT DESIGN (7):

REDUCING THE DIMENSIONS AND CATEGORIES OF TEACHER BEHAVIOUR

Arising from the use of BOS 1, the decision was taken to develop the system in order that it should be simpler to learn, easier to use and available to a wider audience.

(The first, important step was to reduce the number of categories. This would mean that the observer had fewer decisions to make in the course of coding.)
In order to do this it was necessary to decide which to remove from the system and which to leave in. (It was important that the system still described the major dimensions of Teacher Behaviour and that any pruning should not result in dilution to the point where descriptions obtained from use of the system were meaningless.)

It was obvious that the selection of the teaching and managing Acts provided the chief description of the teacher's behaviour and that at least a selection from these categories must remain. In similar vein without knowing the direction and mode of the teacher's communication any description of the behaviour in a lesson would be incomplete.

Less essential was the knowledge that the teacher was mobile or not, and having used that dimension it became clear that the procedures for its identification and coding were less than satisfactory. The degree of mobility could not be described adequately without further categories being added, thereby making the observer's job even more difficult.

Use of the Function Subscript categories with BOS 1 had been almost entirely in favour of 'Does' i.e. the teacher did all of the managing or teaching himself. It was not felt necessary therefore to include these other, underused categories which left the dimension with only one category in it. The decision was made to delete this dimension.

The content code and the type of teaching code (preparatory, intervening or concurrent) were considered peripheral to a description of behaviour compared with the remaining dimensions. Notes on the lesson content could accompany a
lesson description independently, if required. Whilst the type of teaching was more relevant it was adding to the complications of coding considerably. The Anderson System made more of a feature of this dimension at the expense of describing the teaching act, in particular, and it was felt that observers wanting that discrimination would be better advised to use the Anderson System.

Having reduced the number of dimensions to four of principal interest - Manages, Teaches, Direction and Mode - the task was now to 'collapse' some of the categories within those dimensions to further reduce the number of decisions which had to be made in any one observation. It was recognised that this would lead to some loss of discrimination in the descriptive power of the system but this, it was hoped, would be compensated for by increased reliability in its use and an extended number of users.

Table 2 shows the initial attempt to produce a system with fewer categories in the four dimensions. The letters and numbers in brackets refer to the dimension and category in BOS 1 subsumed under the new dimension.

Table 2

The Modified System (BOS 2)

MANAGES
1. People (C1,C3,C4)
2. Equipment (C2)
3. Task (C5,C8)
4. Policy (C6,C7)

TEACHES
5. Prescribes (D1)
6. Gives Feedback (D2,D3)
7. Observes (D4)
8. Informs (D5,D6)

DIRECTION
9. One (F1)
10. Group (F2)
11. Class (F3)
12. Other (F4,F5,F6)

MODE
13. Audible (E1,E2,E11)
14. Silent (E4)
15. Physical (E3,E5,E6,E7,E8,E9,E10,E11)
16. Audible-Physical (Combination of above)
17. Silent-Physical (Combination of above)

18. Other interaction
19. Non-Interaction
20. Inadequate record

The most serious loss using this modified version was the teaching act 'Seeks Information' which had been put under the collective category 'Informs'. After initial trials using this system this was felt to be a mistake as well as a misnomer and the category was reinstated with 'Gives Information' also described separately.

The category 'Other' placed under Direction was considered unnecessary. It was scarcely used and also the definition contained reference to persons unconnected with the lesson. The main interest lay with the teacher and the pupils she was teaching as part of the lesson, so it was decided to exclude this reference from lesson coding.

5.3 INSTRUMENT DESIGN (B):

INTERVAL RECORDING PROCEDURE
At this point the system had twenty categories with no more than three to be selected during one observation. The number of categories had been reduced by more than half. It would have been possible to have continued using the same unit of analysis and coding procedures as for BOS 1 but this was considered undesirable. Continuous coding was tiring, led to errors and depended largely on lessons being videotaped. It was decided that for an observer to sample teacher behaviour regularly throughout the lesson would be adequate and so an interval recording was used. A fifteen second interval was chosen because it allowed very adequate time for an observer to look, decide and code without being too rushed which had previously led to errors being made. The digital clock markings on the videotapes permitted selection of the intervals at 00, 15, 30, and 45 seconds. Live coding could be done using a 'bleep' emitted from a timer every fifteen seconds which could, if preferred, be transferred on to a tape recorder for the observer's use.

5.4 INSTRUMENT DESIGN (9):
EXPOSING THE SYSTEM

The modified system, hereafter referred to as BOS 2, was first aired at a conference of experts in the education service (Leeds Carnegie, 1982).

In one hour the system was explained briefly, coding sheets were distributed to delegates and edited sections of games
lessons were shown for coding purposes. Most of the time was taken with identifying the categories. Definitions of the categories had been circulated but not learned beforehand. At the end of the very brief training session the delegates attempted coding seven units of behaviour.

Nineteen delegates attempted to code the seven units of behaviour. They did it without assistance and used the digital clock in order to code at 15 second intervals. Each delegate's coding was checked against the Designer's coding of the same units. Using the formula:

\[
\text{Agreements} \times 100 \over \text{Agreements} + \text{Disagreements}
\]

applied to each dimension and to the total number correct out of three for each coding, a percentage of agreement was obtained for the group as a whole. Table 3 shows the results of the coding expressed as a percentage.

Table 3

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Average % of Agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teaches/Manages</td>
<td>60.89%</td>
</tr>
<tr>
<td>Direction</td>
<td>84.95%</td>
</tr>
<tr>
<td>Mode</td>
<td>78.19%</td>
</tr>
<tr>
<td>Overall</td>
<td>75.43%</td>
</tr>
</tbody>
</table>

Table 4 shows the results of the coding of each unit expressed as a percentage.
Table 4

Group coding of each unit

<table>
<thead>
<tr>
<th>UNIT</th>
<th>T/M</th>
<th>D</th>
<th>M</th>
<th>x%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100%</td>
<td>100%</td>
<td>94.7%</td>
<td>98.2%</td>
</tr>
<tr>
<td>2</td>
<td>84.2%</td>
<td>89.5%</td>
<td>89.5%</td>
<td>87.7%</td>
</tr>
<tr>
<td>3</td>
<td>36.8%</td>
<td>73.7%</td>
<td>68.4%</td>
<td>59.6%</td>
</tr>
<tr>
<td>4</td>
<td>57.9%</td>
<td>63.2%</td>
<td>73.7%</td>
<td>64.9%</td>
</tr>
<tr>
<td>5</td>
<td>57.9%</td>
<td>89.5%</td>
<td>63.2%</td>
<td>75.4%</td>
</tr>
<tr>
<td>6</td>
<td>63.2%</td>
<td>78.9%</td>
<td>78.9%</td>
<td>73.7%</td>
</tr>
<tr>
<td>7</td>
<td>26.3%</td>
<td>100%</td>
<td>78.9%</td>
<td>68.4%</td>
</tr>
</tbody>
</table>

KEY: T/M Teaches or Manages
      D Direction
      M Mode
      x% Mean % of agreement per unit of behaviour

It had not been planned to test the audience after such a brief introduction to the system. The units which they attempted to code followed on from those they had been looking at earlier and had not been especially selected for coding. This resulted in only a few categories appearing in the test units rather than a spread of different categories. The results of this impromptu trial were interesting, nevertheless, and felt to be worthy of comment.

Overall the percentage of agreement was quite high at 75.43% in spite of the very short time spent looking at categories and even less time spent in explanation of coding procedure. The lowest agreement was for the
Teaching/Managing categories which are those requiring the longest time for accurate identification. Part of the explanation for the relatively low agreement figure overall and for two of the unit Acts in particular is the lack of practice at coding procedure rather than Act identification. Many of the delegates had coded the Act which occurred immediately prior to or just after the exact interval. Mistakes in the Direction dimension were largely as a result of coding procedure also although the high figure of 84.95% was very encouraging in spite of these errors. The percentage in agreement on the Mode dimension would have been higher overall but for the result of Unit 5 where delegates missed the physical demonstration which was accompanied by speech. The minority coded speech only.

It would be inappropriate to exaggerate the importance of a very small and impromptu trial such as this. It served a useful purpose however, reminding the designer that time must be concentrated on Act identification and learning the definitions at an early stage, and that coding procedures are equally important to learn thoroughly if reliability is to be achieved. Additionally it had been encouraging to note that the system had some potential for ready use, for the audience of lecturers, advisory staff and HMI had shown a keen interest in the exercise and were learning quickly to use the moderately difficult system.

Comments made at the conference and the experiences gained in explaining such a system to a knowledgeable audience
combined to produce a Manual for BOS 2 (Bailey, 1982).

Two alterations were made to the System itself. Under Mode, the Silent-Physical category was deleted because it was duplicated by the category 'Physical' if used alone. The definition of the category 'Physical' was altered to make it clear that it was a silent use of demonstration or contact. Secondly the categories were renumbered.
5.5 BOS 2

Table 5 shows the final version of BOS 2.

Table 5

The Bailey Observation System 2

<table>
<thead>
<tr>
<th>MANAGES</th>
<th>TEACHES</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. People</td>
<td>5. Prescribes Response</td>
</tr>
<tr>
<td>2. Equipment</td>
<td>6. Gives Feedback</td>
</tr>
<tr>
<td>3. Task</td>
<td>7. Observes</td>
</tr>
<tr>
<td></td>
<td>9. Seeks Information</td>
</tr>
<tr>
<td></td>
<td>10. Other Interaction</td>
</tr>
<tr>
<td></td>
<td>11. Non-Interaction</td>
</tr>
<tr>
<td></td>
<td>12. Inadequate Record</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>DIRECTION</td>
<td></td>
</tr>
<tr>
<td>13. One</td>
<td></td>
</tr>
<tr>
<td>14. Group</td>
<td></td>
</tr>
<tr>
<td>15. Class</td>
<td></td>
</tr>
<tr>
<td>MODE</td>
<td></td>
</tr>
<tr>
<td>16. Audible</td>
<td></td>
</tr>
<tr>
<td>17. Silent</td>
<td></td>
</tr>
<tr>
<td>18. Physical</td>
<td></td>
</tr>
<tr>
<td>19. Audible-Physical</td>
<td></td>
</tr>
</tbody>
</table>

The Manual contained an explanation of the system, coding procedures, category definitions and guidance for Observer Training. A computer program was designed which enabled the coder to put the numbers of the identified categories directly into the computer which would provide an analysis of category use. The computer program is reproduced in
Appendix 4. An example of the print out showing an analysis of category use may be found in Appendix 5. The computer program contributed to the overall speed and efficiency of BOS 2 very considerably.

5.6 INSTRUMENT DESIGN (10):

  OBSERVER TRAINING

The system was, at this stage, in a more advanced state prior to Observer Training than BOS 1 had been. This was in an attempt to improve accuracy at an early stage in training.

Two observers agreed to learn the use of the System and be tested for reliability. One of these was an observer who had learned the use of BOS 1 but the other observer was new to the system. Both were lecturers in physical education and would be considered as experts in the field. The observer new to the system was intending to use the system in his analysis of physical education lessons in local secondary schools. Originally the system had been designed for the description of games lessons only, but now it was felt that it might be possible to use it to describe other activities in the physical education lesson.

Observer A had not used the system for over a year. A total of three hours was spent in revision of the BOS 2 categories and the coding procedures.
Observer B was given a Manual to read before training began. Training sessions were spread over a two week period and consisted initially of category identification with frequent reference to the definitions in the Manual. Coding procedures were begun earlier than with previous observers partly because this observer was spending free time practicing the use of the system and progressing more rapidly. During training the observer operated the tape machine himself.

All of the lessons learned during earlier training proved to be of enormous benefit. A total of eight hours was spent in training and no major problems were encountered. The Teaching Acts proved the most difficult to get right consistently and the distinction between observing a group or class was additionally a problem when it was not always clear on the tape. Points of clarification were added to the Manual during training.

Both observers were trained independently partly through necessity and partly through choice. Observer A had previous experience whereas Observer B had none. They worked in different institutions 100 miles apart which would have made joint sessions difficult. Perhaps most important, however, was the fact that observers for BOS 1 had shown their preference for individual tutorials and so this was chosen from the beginning in this training round.

Whilst the two observers were not able to discuss their training with each other, independently each was able to
discuss the categories, definitions and coding procedures with the designer. Observations were practised on videotape which had previously been coded by the designer for the purpose of comparison. Both the designer and observer practised simultaneously coding directly from videotape and a tape recording was made of Observer B's training in order that a record could be kept of disagreements and how they were resolved. This was not felt to be necessary for Observer A as this Observer was familiar with the original system and had been trained in its use.
5.7 INSTRUMENT DESIGN (11):

RELIABILITY TESTING

The final stage in training is the testing for reliability.

Little guidance is given in the literature regarding the form that reliability testing should take with the result that researchers devise their own. Hupé (1974) tested four observers using 6 short 20 minute tapes and during the testing the observers were allowed to talk to one another. Hurwitz (1974) also used four observers who looked at two sets of video recordings and they were provided with typescripts of the lessons they were coding. Catelli (1979) trained one observer who was tested on one tape plus a typescript lasting 29 minutes and 18 short segments from 29 tapes plus typescripts totalling 23 minutes i.e. 52 minutes of coding overall.

Such discrepancy in the structuring of these and other reliability test sessions led the designer to select her own structure for the test procedure. It was considered important that the observers should be allowed to code a complete lesson as this was what they would do if using the system to gather data to be used in research work. In addition it was felt that if an error was made which was one of consistent misinterpretation then this was more likely to show over the period of one whole lesson's coding. It was recognised that fatigue might contribute to inaccuracy when coding a whole lesson also, but this too was considered an
important and necessary feature of coding which should be tested. Four sets of results were considered adequate for the purpose of comparison, and to ensure that the test situation was as near realistic as possible to live coding, the observers were not supplied with typescripts of the lessons. Each observer coded the lesson(s) independently and alone.

Observer A was given a previously unseen tape recording of a gymnastics lesson to analyse in full. Observer B, who was new to the BOS 2 system, was given two tapes to analyse in full – one of the same gymnastics lesson and a swimming lesson. Neither was observed by the other during testing.

Whilst interobserver agreement is a measure of accuracy in the use of the system between observers, a test of stability is a measure of the consistency of an observer's coding over a period of time. This was considered to be a vital feature of the use of the system by the designer if the data obtained through its use were to be discussed with confidence.

Two tapes, one of a netball lesson and one of a badminton lesson, were selected for recoding one year after they had been initially coded. They were not selected randomly but because they were distinctive for different reasons. Both had caused initial problems in training – one because of the sheer volume of talk by the teacher and the other had taxed the coder procedurally due to the unusually high number of
instances of Teacher Observation in the lesson. Selecting difficult tapes together with a very long period of time between coding sessions it was felt would be a rigorous test of stability.

In order to calculate the results of the reliability testing, note was taken of research done by Hawkins and Dotson (1975) who were very critical of methods used to calculate interobserver agreement scores. They pointed out in their report that the method of calculation most widely used by independent observers was the interval-by-interval (I-I) method which uses every interval of recording in its calculation of interobserver agreement. They stated that there are three independent sources of error in obtaining accurate and objective data:

1. The definition of behaviour given to the observer by the experimenter may be vague, subjective or incomplete.
2. The observer may be poorly trained, unmotivated, or otherwise incompetent.
3. The behaviour may be difficult to detect because of its subtlety or complexity, because of distractions, or because of other factors obstructing the observing process. They maintained that it is not simply the accuracy and objectivity of the data themselves that need to be assessed, but also the "believability" or validity of the experimental effect. In their assessment of the I-I method they collected data and performed other analyses to assess the adequacy of the method in serving three functions: 1) as an index of how precise, clear, objective and complete the definition is; 2)
as an index of how competently the observer is recording; and 3) as an index of the believability of the experimental effect reported. The results of their assessment showed the gross unreliability of I-I scores as an index of definition adequacy, observer competence or believability of experimental effects due to the fact that I-I scores are highly subject to influence by the rate or duration of the behaviour being recorded. Their recommendations were intended to solve the problem of I-I reliability scores being affected by the rate of behaviour and secondly to assist with the problem of the believability of experimental effect. Their recommendations are summarised below:

A. Scored-Interval (S-I) agreement

In S-I agreement, all intervals in which neither observer scored the behaviour as occurring are ignored in calculating agreement scores. Only an interval in which both observers recorded the presence of the behaviour is counted as a behaviour. The authors pointed out that S-I reliability scores also had serious limitations if used alone. First they constituted a very stringent test of agreement used alone, and also at low frequencies of behaviour they became very variable. In order to get over these difficulties they recommended that not one but two reliability scores be calculated, one for the occurrence of the behaviour and one for its non-occurrence. They termed this kind of agreement Unscored-Interval agreement.
B. Unscored-Interval (U-I) agreement

In U-I scores an agreement is counted only when both observers recorded that the behaviour did not occur. A disagreement is counted when one observer recorded the presence of the behaviour and the other did not. Those intervals in which both observers scored the behaviour are ignored.

The purpose of using both types of scores is that they compliment each other. The authors considered that the limitations of both used independently should be eliminated or greatly ameliorated. It was further considered that averaging the two scores reduces the problem of variability which would still be evident if the two sets of scores were presented independently. They pointed out that the lower limits of the mean of S-I and U-I would seem more tolerable that those of I-I.

C. Total-Interval (T-I) agreement

In calculating T-I agreement the observer divides the total number of intervals in which one observer saw the behaviour into the number in which the other observer saw the behaviour, dividing the larger into the smaller and multiplying by 100. The authors concluded that this provides a much better assessment of the believability of the experimental effect than do other agreement scores because
it employs the same statistic as that used in presenting the experimental effect: the number - or percent - of intervals in which the behaviour was seen.

The authors concluded that a change to using S-I and U-I reliability scores offered much more accurate representation of the objectivity and accuracy of interval data. They felt that when combined with additional safeguards aimed specifically at assessing the believability of the experimental effect, these reliability measures should improve the methodology of a significant portion of applied behaviour analysis in education.

It was acknowledged that to apply such stringent tests of reliability would undoubtedly reduce the levels of agreement which might be obtained using I-I reliability scores or even S-I scores independently yet it was felt that the arguments used by Hawkins and Dotson were persuasive. By the same token it was acknowledged that the results would not be entirely comparable with data obtained by other researchers who had not used three sets of scores in their testing for reliability.

An example of the analysis of Test PL2/1 may be found in Appendix 6 and the methods employed to arrive at the results are summarised below.

5.8 METHODS USED IN TESTING FOR RELIABILITY
Stage 1

The category codings entered on the computer print-out were put on to a separate sheet in numerical form. Three sets of numbers were entered for each observation - one for the Managing or Teaching or Other Dimension; one for the Direction Dimension and one for the Mode Dimension. As two sets of codings were being compared for each test, the codings were arranged side by side in order that the observations could be compared readily and their agreements or disagreements made apparent. An example from the data using test tape PL2/1 from the first and second observation by the designer may be found in Table 6.

Table 6 Sample data from reliability test

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>02</td>
<td>01</td>
<td>02</td>
<td>01</td>
<td>02</td>
</tr>
<tr>
<td>T/M/O</td>
<td>T/M/O</td>
<td>D</td>
<td>D</td>
<td>Mo</td>
<td>Mo</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>15</td>
<td>15</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>15</td>
<td>15</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>15</td>
<td>15</td>
<td>(19</td>
<td>(16)</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>15</td>
<td>15</td>
<td>(16</td>
<td>(19)</td>
</tr>
</tbody>
</table>

KEY:
01 - 1st observation
02 - 2nd observation
T/M/O - Teaching, Managing or Other
D - Direction
Mo - Mode
Numbers correspond to the category numbers in the BOS 2 System.

Disagreements were circled as illustrated for ease in identification.

Stage 2
To calculate the number of S-I agreements, each category from 1 through 19 was examined down all of the columns under the different Dimensions to establish when a pairing was made. From the example in Table 6 it can be seen that there are 2 S-I agreements for both categories 5 and 8, 4 for category 15 and 2 for category 16.

Stage 3

To calculate the number of S-I disagreements, the process was repeated but this time the number of times a category was sighted and not paired with the same category was totalled. From Table 6 it can be seen that category 16 scores two disagreements as does category 19.

Stage 4

To calculate the number of U-I agreements i.e. the number of times that the observer recorded that the behaviour did not occur, this was calculated by 1) adding together the S-I agreements and disagreements 2) deducting their total from the number of intervals or observation points 3) dividing by the resultant figure by the the same total plus the number of U-I disagreements which is the same figure as the S-I disagreements.

Using the data from Table 6 as the example, the calculation of U-I agreements for category 16 would be as follows:

S-I Agree (2)  S-I Disagree (2)  Total (4) U-I
Agree(4-4=0-2=0)

This can be checked in this short example by simply adding the number of times in the four intervals that category 16 did not occur, and as can be seen, it was observed in all four intervals either in both sets of observations or in one of them. In the analysis of the full tape there were 186 intervals to be checked which made the use of this simple formula both helpful and necessary if mistakes were to be avoided.

Stage 5

To calculate the T-I believability of the experimental effect the total number of intervals in which the first set of observations reported the behaviour were divided by the total number of intervals in which the second set of observations reported the behaviour, with the lower figure always being divided by the higher figure. This is the same as saying S-I Agreements and S-I Disagreements for Observer 1 are to be divided by S-I Agreements and S-I Disagreements for Observer 2.

In the data in Table 6 using category 16 as the example once more, the S-I agreements totalled two. In interval three, Observer 2 noted a category 16 disagreement and in interval 4 Observer 1 noted a category 16 disagreement. Both sets of observations recorded the use of category 16 a total of 4 times which would give a percentage of 100% i.e. $\frac{4-4}{4} \times 100 = 100\%$
Stage 6

The raw data was now available and it remained for the percentage of agreement to be calculated using the formula:

\[
\text{Agreements} - \frac{\text{Agreements} + \text{Disagreements}}{\times 100} = \% \text{ of Agreement}
\]

This was done for S-I, U-I and T-I calculations.

Stage 7

The S-I and U-I percentages were averaged to give the average percentage of agreement for each category. The categories contributing to a Dimension e.g. 1 through 4 for Managing, were averaged to give the average percentage of agreement for each Dimension of the System. Finally the category averages were totalled and the Mean calculated for the System as a whole for both Reliability and Believability.

5.9 RESULTS OF RELIABILITY TESTING

The results of the reliability testing are shown in Tables 7 and 8.
Table 7
PERCENTAGES OF AGREEMENT (Reliability)

<table>
<thead>
<tr>
<th>TEST TAPE</th>
<th>OVERALL %</th>
<th>M %</th>
<th>T %</th>
<th>D %</th>
<th>Mo %</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPM</td>
<td>75.48%</td>
<td>75.64%</td>
<td>73.15%</td>
<td>76.90%</td>
<td>78.46%</td>
</tr>
<tr>
<td>GPF/1(T1)</td>
<td>81.73%</td>
<td>80.18%</td>
<td>84.30%</td>
<td>82.04%</td>
<td>80.55%</td>
</tr>
<tr>
<td>GPF/1(T2)</td>
<td>69.68%</td>
<td>68.08%</td>
<td>70.11%</td>
<td>76.48%</td>
<td>75.96%</td>
</tr>
<tr>
<td>GPF/3(T3)</td>
<td>73.14%</td>
<td>68.84%</td>
<td>74.63%</td>
<td>81.32%</td>
<td>75.99%</td>
</tr>
<tr>
<td>OVERALL%</td>
<td>75.01%</td>
<td>73.19%</td>
<td>75.55%</td>
<td>79.19%</td>
<td>77.74%</td>
</tr>
<tr>
<td>LH61</td>
<td>77.65%</td>
<td>75.25%</td>
<td>81.57%</td>
<td>77.29%</td>
<td>74.91%</td>
</tr>
<tr>
<td>PL2/1</td>
<td>90.62%</td>
<td>94.31%</td>
<td>92.09%</td>
<td>87.18%</td>
<td>89.05%</td>
</tr>
<tr>
<td>OVERALL%</td>
<td>84.14%</td>
<td>84.78%</td>
<td>86.83%</td>
<td>82.24%</td>
<td>81.98%</td>
</tr>
</tbody>
</table>

OVERALL RELIABILITY
Management Dimension 79%
Teaching Dimension 81%
Direction Dimension 81%
Mode Dimension 80%

BOS 2 RELIABILITY 80%

Table 8
PERCENTAGES OF AGREEMENT (Believability)

<table>
<thead>
<tr>
<th>TEST TAPE</th>
<th>OVERALL %</th>
<th>M %</th>
<th>T %</th>
<th>D %</th>
<th>Mo %</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPM</td>
<td>80.85%</td>
<td>66.2%</td>
<td>77.26%</td>
<td>92.83%</td>
<td>87.11%</td>
</tr>
<tr>
<td>GPF/1(T1)</td>
<td>81.79%</td>
<td>91.11%</td>
<td>78.95%</td>
<td>83.80%</td>
<td>73.30%</td>
</tr>
<tr>
<td>GPF/1(T2)</td>
<td>73.02%</td>
<td>70.15%</td>
<td>69.42%</td>
<td>82.15%</td>
<td>70.35%</td>
</tr>
<tr>
<td>GPF/1(T3)</td>
<td>77.77%</td>
<td>75.78%</td>
<td>78.21%</td>
<td>88.56%</td>
<td>68.51%</td>
</tr>
<tr>
<td>OVERALL%</td>
<td>82.07%</td>
<td>75.81%</td>
<td>75.96%</td>
<td>84.60%</td>
<td>74.82%</td>
</tr>
<tr>
<td>LH61</td>
<td>85.16%</td>
<td>67.66%</td>
<td>88.29%</td>
<td>91.36%</td>
<td>93.33%</td>
</tr>
<tr>
<td>PL2/1</td>
<td>93.85%</td>
<td>89.04%</td>
<td>96.24%</td>
<td>94.54%</td>
<td>95.57%</td>
</tr>
<tr>
<td>OVERALL</td>
<td>89.51%</td>
<td>78.35%</td>
<td>92.27%</td>
<td>92.95%</td>
<td>94.45%</td>
</tr>
</tbody>
</table>

OVERALL BELIEVABILITY
Management Dimension 77%
Teaching Dimension 81%
Direction Dimension 89%
Mode Dimension 81%

BOS 2 BELIEVABILITY 82%
5.10 DISCUSSION OF TEST RESULTS

Two major errors were responsible for the low percentages of agreement in sections of these test results. The first was one observer whose analysis of the gymnastics lesson which was compared with that of two other observers revealed a misinterpretation of the context in which the Teacher Behaviour was taking place. For example, if the teacher was telling the pupils to take up their positions in a part of the gymnasium, the observer had interpreted the instruction narrowly to mean 'Prescribes Response' when, in fact, the context of the instruction was that of Management not Teaching and the behaviour should have been identified as 'Organises People'. This observer had received the minimum of training which would appear to have been inadequate for the purposes of learning to use the system with accuracy. The error, which appeared consistently throughout the coding of the taped lesson, had not been evident in the training process. The observer had not found it possible to code the lesson until some weeks after the training had been completed and this time lag may have contributed to the lower percentages than were anticipated. But for the figures of 70% and 73% associated with this observer it can be fairly assumed that the overall result would have been higher.

The second point of disagreement which was noted in all of the test results to greater or lesser degree was not one of
context or category disagreement but one of procedural dispute. In an interval coding procedure, it is critical that for the purposes of comparing coding and realising agreement, the coders at least code the same behaviour. The problem for any observer of teacher behaviour is that the teacher shifts his behaviour very frequently so that when using an interval coding procedure to organise the observations, it is of paramount importance that the observers land on precisely the correct second to note a behaviour before it changes. The alternative is to watch behaviour over a prescribed period of time and for the observer to note the most dominant behaviour retrospectively. This was not favoured in BOS 2 because of the potential for disagreement over what might constitute a dominant behaviour and the necessity to build into such a system a hierarchy of behaviours. In the Test Tape SPM it was noted that the principal difficulty lay in the fact that one observer was coding 'Observes' whilst the other had coded the next verbal behaviour. As a result all three entries were invariably in disagreement which lowered the agreement percentage overall. These two points accounted for the majority of the disagreements between observers. The first was peculiar to the one observer as it was not replicated elsewhere and the second led to some clarification in the Manual regarding coding procedure. It was noted that increased time should be spent on coding procedure when training potential users of the system.

The other points of disagreement were less consistent in
their reporting. It was noted that some confusion existed between categories 5 and 8, 'Prescribes' and 'Informs' which should clarified in the Manual and in the training process. Similarly there were demonstrations missed (category 19) and these were invariably those which included part of an action to help the learner rather that the whole action. The observers occasionally confused the direction of a teacher's period of observation, in particular when the teacher was standing at the side of a class and looking one way. This should be coded at that second as observing a group (category 14) rather than the whole class (category 15) even though the period of observation might last for a minute and eventually have taken in the whole class.

The tests for stability undertaken by the designer yielded the highest percentages of agreement. There were some procedural changes from the first set of observations which were bound to result in changes one year later, but these accounted for almost all of the points of disagreement between the two codings.

Determining the level of agreement necessary between observers before a system and those users may be deemed to be reliable is a necessary feature of instrument design. "In behavioural observation research, using interobserver agreement calculation techniques, a criterion of 80% is typically required before observers can begin to collect data that are used for research."

(Siedentop 1983, p.264)

The test results taken overall achieved this recommendation
yet it must be said that this figure might be considered high for a moderately difficult system which underwent such stringent reliability tests. Note was taken of the statement by Hawkins and Dotson (1975) which said that greater tolerance might be afforded to the lower limits of the mean scores achieved using the reported method of calculating reliability. Taking this into account it was felt that 75% might be an acceptable minimum level of reliability for any use of BOS 2 which was tested in a similar manner. This figure would rule out the use of the system by the observer whose overall scores did not exceed 73% yet permit its use by the other two observers whose reliability scores averaged 81%. The average percentage of agreement achieved in the tests for stability was 85% and it was possible, therefore, to present the results of system use by the designer with confidence.
5.11 SUMMARY

The modifications of BOS 1 included a reduction in the number of dimensions and categories used to describe teacher behaviour and a change to an interval recording procedure.

The modified system was shown to an audience of experts in physical education at a conference in 1982 and as a result the new system (BOS 2) was prepared for observer training. A Manual, which included a guide for observer training and a computer program enabling the codings to be analysed, was made ready for the training of two observers. During the course of training the system was able to be used to code behaviour in gymnastics and swimming lessons equally well as in the games lessons for which it was designed initially.

After a period of training the two observers and the designer were tested for reliability and stability in the use of the system. The most stringent tests available were applied using the scored-interval, unscored-interval and total-interval methods. The overall inter and intra-observer reliability tests averaged 80% reliability and 82% believability or validity which were considered adequate for system use and the reporting of results.
CHAPTER 6

The results of using BOS 2

6.0 INTRODUCTION

This chapter details the use of BOS 2 to describe teacher behaviour in 19 lessons of physical education. The main findings are presented followed by a statistical analysis of the data and a discussion of the results.

Eleven teachers of physical education representing three Local Education Authorities in England provided nineteen videotaped lessons for analysis. Eight of these lessons were taught by women and eleven by men, and the activities included netball, hockey, badminton, gymnastics, swimming, rugby football, volleyball and basketball.

Each tape was coded using the BOS 2 Observation System. A total of 3399 samples of behaviour were taken with each lesson providing between 122 and 232 samples depending on length. In order that meaningful statistical comparisons could be made between teachers the category totals were scaled to the smallest sample using the computer program listed in Appendix 7. As the sample of teachers contained some who taught twice as opposed to once, the lesson data belonging to those teaching twice were totalled and then scaled to the smallest sample size in order to avoid any one
teacher's style unduly influencing the results.

From the data the category Totals, Mean, Standard Deviation and Percentage of lesson time spent engaged in the different teacher behaviours were calculated.

2 X 2 Chi Square tests were applied to the data in order to establish whether there were significant differences in the observed behaviours of male and female Physical Education teachers in general, when teaching games as opposed to gymnastics in particular, and by the same teachers but teaching different activities.

Percentages were obtained for the use of the nineteen categories of Teacher Behaviour under the principal Dimensions of Managing or Teaching together with the Direction and Mode of the observed behaviour and these results may be found in Table 9. All activities except the swimming fell under the classification of games or gymnastics lessons and the category use by male and female teachers in these two types of lessons were compared. Tables 10, 11, 12, 13, 14 and 15 showing the percentages of category use by male and female teachers in gymnastics and games lessons may be found in Appendix 8.
RESULTS

6.1 SUMMARY OF MAIN FINDINGS

6.1.1 The Dimensions of Teacher Behaviour

From the sample, teachers of Physical Education taught for 67% of the available teaching time and were engaged in Management for 27% of the remainder of interactive lesson time. They used predominantly a prescriptive teaching method, preferring to address the whole class simultaneously by means of verbal communication only.

There were significant differences in the management and teaching of games as opposed to gymnastics by male Physical Education teachers (p<.001) and by female Physical Education teachers (p<.05), and significant differences were also recorded in the management and teaching of games by male as opposed to female teachers (p<.05).

In gymnastics lessons male teachers spent 32% of the time managing as opposed to teaching for 63% of the time compared with a management time of 20% in games lessons and a teaching time of 70%. Much of the remainder of the available time was spent changing and showering by the pupils.

Female teachers spent longer in management behaviour in both
gymnastics and games lessons at 35% and 27% respectively and less time engaged in teaching behaviour at 60% and 67% respectively.

Teaching profiles were compared by means of 2 X 2 Chi-Square Tests available on a computerised statistical package (Angseesing, 1980).

Three male teachers differed significantly in their observed managerial and teaching behaviours during gymnastics as opposed to basketball lessons. No other teachers exhibited changes in style in their teaching between one lesson and another even when the class and/or the activity were different. Applying the same 2 X 2 Chi-Square test, male and female teachers across all of the activities and age groups did not show any significant differences in their managerial and teaching behaviours examined collectively.

6.1.2 The Categories of Teacher Behaviour

Teachers of gymnastics spent longer engaged in managing equipment than did teachers of all other activities. As a whole the teachers spent longer organising pupils than either giving the pupils feedback, or asking them content-related questions or giving them information about their work. The teachers showed a marked preference for telling pupils how to do an activity rather than demonstrating how it should be done.
During their teaching of games and gymnastics, male and female teachers differed significantly in their use of silent observation (p<.001), the time spent talking (p<.001), their use of prescription (p<.05) and the frequency with which their attention was directed to groups (p<.005) and individuals (p<.05) as opposed to the whole class.

6.2 REPORT OF THE DATA

6.2.1 The Percentages of Observed Teacher Behaviour

Table 9 The Percentages of Observed Teacher Behaviour (overall)

<table>
<thead>
<tr>
<th>CATEGORY</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>MANAGEMENT</td>
<td></td>
</tr>
<tr>
<td>Organises People</td>
<td>13.5%</td>
</tr>
<tr>
<td>Manages Equipment</td>
<td>7.5%</td>
</tr>
<tr>
<td>Manages Task</td>
<td>4.8%</td>
</tr>
<tr>
<td>Manages Policy</td>
<td>0.9%</td>
</tr>
<tr>
<td>TEACHING</td>
<td></td>
</tr>
<tr>
<td>Prescribes Response</td>
<td>24.3%</td>
</tr>
<tr>
<td>Gives Feedback</td>
<td>8.3%</td>
</tr>
<tr>
<td>Observes</td>
<td>15.8%</td>
</tr>
<tr>
<td>Gives Information</td>
<td>12.6%</td>
</tr>
<tr>
<td>Seeks Information</td>
<td>6.2%</td>
</tr>
<tr>
<td>Other Interaction</td>
<td>2.8%</td>
</tr>
<tr>
<td>Non-Interaction</td>
<td>1.7%</td>
</tr>
<tr>
<td>Inadequate Record</td>
<td>1.6%</td>
</tr>
<tr>
<td>DIRECTION</td>
<td>24.2%</td>
</tr>
<tr>
<td>One</td>
<td></td>
</tr>
<tr>
<td>Group</td>
<td>32.8%</td>
</tr>
<tr>
<td>Class</td>
<td>39.8%</td>
</tr>
<tr>
<td>MODE</td>
<td></td>
</tr>
<tr>
<td>Audible</td>
<td>67.8%</td>
</tr>
<tr>
<td>Silent</td>
<td>18.7%</td>
</tr>
<tr>
<td>Physical</td>
<td>0.6%</td>
</tr>
<tr>
<td>Audible-Physical</td>
<td>10.9%</td>
</tr>
</tbody>
</table>
It can be seen that the teachers spent over 96% of the coded time performing interactive functions, a figure which compares closely with that of 94% recorded by Anderson and Barrette (1978). It is not possible to make exact comparisons between the other results from this study and that conducted by Anderson and Barrette because the two systems contained different sets of categories of teacher behaviour, nevertheless some of the dimensions of behaviour are comparable and provide a basis on which comparisons may be made.

6.2.2 The Management Dimension

Summing the percentages of time the American teachers spent in the different categories which bear a close resemblance to the Managerial Categories in the BOS 2 system, it would appear that the American Teachers in the Anderson and Barrette study spent approximately 23% engaged in those acts in which their English counterparts spent 27%.

Using BOS 2 data the majority of this time (21%) was spent in organising pupils and managing equipment. Task management accounted for the remainder with less than 1% spent in dealing with such things as absence, uniform or taking registers.

Female teachers spent 17% of games lessons organising pupils and 3% managing equipment. In gymnastics this changed to 12%
of the time organising pupils and 19% managing equipment.

Male teachers spent 12% of the time organising pupils in games lessons and 4% managing equipment but these figures altered to 10% and 18% respectively in gymnastics lessons.

6.2.3 The Teaching Dimension

The time spent teaching was comparable to that in the Anderson and Barrette study although the Observation time of the American Teachers was higher at 21% as opposed to 16% by the teachers from England.

Significant differences in category use by male and female teachers in games and gymnastic lessons were reported in two teaching categories:

Observes $X^2(1, N = 1730) = 17.22, p < .001$
Prescribes Response $X^2(1, N = 1730) = 6.47, p < .05$

Silent observation of pupils working was more likely to be engaged in by male teachers in both gymnastics and games lessons. In gymnastics lessons, male teachers spent approximately 16% of their time observing pupils silently ($n = 488, x = 19.4, \text{on-1} = 4.1$) whereas female teachers spent approximately 8% of their time silently observing pupils working ($n = 244, x = 9.36, \text{on-1} = 3.13$).
In games lessons, male teachers spent 14% of their time engaged in the silent observation of pupils working (n = 610, x = 17.5, on-1 = 5.03). Female teachers spent 15% of their time observing work silently in games lessons (n = 488, x = 18.6, on-1 = 19.4). As the results indicate whilst the final percentage may be similar in the games lessons, individual differences for the female staff were high. One teacher had coded seventy one instances of observation compared with the lowest figure of five by another teacher ranging to the next highest of thirty three.

The most frequently selected teaching behaviour was that of Prescribes Response in which teachers were engaged for 24% of the time. This would appear to be a broader category than those reported by other researchers who may have used such labels as 'lecturing', 'soliciting', 'structuring' or similar with the result that meaningful comparison is not possible.

The use of prescriptive teaching behaviour adopted by the teacher to tell the pupils precisely what they are to do, when and how, was preferred by male teachers compared with female teachers in gymnastics lessons and games lessons. Male teachers spent 24.4% of their teaching time in games lessons prescribing pupil responses (n = 610, x = 29.7, on-1 = 2.53) and 23.1% of the time in gymnastics lessons (n = 488, x = 28.1, on-1 = 10.05). Female teachers spent 23.3% of their teaching time in games lessons prescribing pupil responses (n = 488, x = 28.5, on-1 = 4.71) and 20.3% of the
time in gymnastics lessons (n = 244, x = 24.7, on-1 = 0.75).

Freedman (1978) studied the teaching behaviours of both student and in-service teachers and reported that approximately 5% of these behaviours focussed on questioning pupils about their work. BOS 2 percentages support this figure showing that approximately 6% of teacher behaviour was spent in seeking information from the pupils.

The female teachers spent more time asking pupils about their work in gymnastics lessons (11%) than did their male counterparts (5%).

Numerous studies have focussed on feedback of different kinds and those which report quantitively on augmented feedback— that which is provided by an external source to the learner— report figures which differ widely either in rate per minute or in percentages of lesson time. (Stewart, 1977; Fishman and Tobey, 1978; Freedman, 1978; Pieron, 1982). These can vary as much as from 10% to 25% with the lower end of the scale comparable to the BOS 2 figure of 8%.

Within games lessons, male teachers spend 17% of the teaching time giving information to their pupils compared with 14% in gymnastics lessons. A significant difference was noted between this giving of information to the whole class in games as opposed to gymnastic lessons \( \chi^2(1, n = 732) = 4.34, p < .05 \).
6.2.4 The Mode Dimension

Significant differences in category use by male and female teachers in games and gymnastic lessons were reported in two categories:

Audible $\chi^2(1, N = 1730) = 16.68$, $p < .001$
Silent $\chi^2(1, N = 1730) = 12.43$, $p < .001$

The Anderson and Barrette research also highlighted the overwhelming frequency with which teacher talk was used either as an exclusive mode or in combination with other mode categories.

Audible communication was significantly higher for the female staff than the male staff. In gymnastics lessons the female staff preferred audible communication exclusively for 76% of the time ($n = 244$, $x = 92.8$, on-1 = 8.6). The male teachers recorded a figure of 65% ($n = 488$, $x = 79.6$, on-1 = 9.7).

In the games lessons the female teachers used audible communication exclusively 72% of the time ($n = 488$, $x = 88.2$, on-1 = 19.5) whereas the male staff communicated audibly for 61% of the time ($n = 610$, $x = 74.2$, on-1 = 19.8). In these games lessons the alternative mode selected by the male teachers which accounted for the difference was that of 'Audible-Physical' i.e. demonstration with talk, which was used for 15% of the time by the males as opposed to 7% by the female teachers. In gymnastics the male
teachers used demonstration for 11% of the time compared with the female teachers who selected the demonstration mode for 8% of the time.

The work of Anderson and Barrette also revealed the few occasions on which teacher demonstrations were in evidence, citing slightly less than six times per class period. BOS 2 results show that demonstration was used approximately 7 times per lesson.

6.2.5 The Direction Dimension

Nearly 40% of the teachers' time was spent attending to the whole class throughout the interactive period as opposed to 24% which was directed at individuals or nearly 33% towards groups. The form of organisation preferred by the teachers used in the BOS 2 sample was group work, where the pupils were split into pairs or larger groups such as teams for play or practice purposes.

Significant differences in category use by male and female teachers in games and gymnastic lessons were reported in two categories:

Group $X^2(1, N = 1730) = 10.04, p < .005$

Individual $X^2(1, N = 1730) = 4.58, p < .05$

In games lessons, male teachers directed their attention to individuals, groups and the whole class for 22.4%, 30.6% and
40.5% of the time respectively. The standard deviations for these figures were 13.7, 22.6 and 22.2 respectively. For the female teachers in games lessons, their attention was directed to individuals for 28.1% of the time (on-1 = 7.7), to groups of pupils for 33.1% of the time (on-1 = 9.9) and to the whole class for 36.3% of the time (on-1 = 11.5).

The figures varied more widely in the gymnastics lessons with female teachers spending longer engaged with individual pupils (25.9%) than did the male teachers (20.2) and less time with groups (23.3%) than did the male teachers (25.1%). The female teachers directed their attention to the whole class for longer (48.8%) than did the male teachers (43.1%). Once again the standard deviations for the male teachers were greater overall than for the female teachers (14.3, 9.74, 19.4 for individuals, groups and the class respectively) as opposed to 9.7, 3.7 and 12.9 respectively.

6.3 INDIVIDUAL TEACHER DIFFERENCES

6.3.1 Teacher A

Of the three male teachers whose lessons were significantly different, Teacher A showed the most marked difference in Management and Teaching behaviours $X^2(1, n = 368) = 12.6$, $p<.001$). Considerably longer was spent in management behaviour in the gymnastics lesson (28%) compared with the basketball lesson (8%), in verbal communication exclusively
in gymnastics (68%) compared with basketball (35%) and in demonstrating basketball skills (30%) compared with gymnastic skills (10%).

6.3.2 Teacher B

This male teacher also revealed marked differences in Management and Teaching behaviours \( \chi^2(1, n = 329) = 10.38, p<.005 \). Longer was spent managing the gymnastics lesson (33%) than the basketball lesson (17%), in directing attention to the whole class in basketball (56%) than in gymnastics (23%) and in giving information about the basketball activity (23%) rather than the gymnastics (12%). Demonstrations occurred more frequently in the games lesson, being recorded at 18% compared with 12% in the gymnastics lesson.

6.3.3 Teacher C

Teacher C's Managerial and Teaching behaviours were different in the basketball and gymnastics lessons \( \chi^2(1, n = 331) = 5.49, p<.05 \). In common with the other teachers longer was spent engaged in management in the gymnastics lesson (40%) than in the basketball lesson (24%). 33% of the management in the gymnastics lesson was concerned with the arrangement of apparatus. The length of time engaged in management in this lesson meant that relatively little time
was spent teaching but from the data it was again noticed that whereas 14% of the mode selection in basketball was for demonstration it was as low as 6% for gymnastics. Feedback offered in basketball was offered for 12% of the teaching time and for 3% of the corresponding time in gymnastics. Individuals were given attention for 4% of the communications in gymnastics compared with the whole class for 56% and these figures compared with 12% and 52% in basketball.

Comparisons made of the other teachers in the sample who taught more than one lesson whether they were of the same or of different activities revealed no significant differences at the 0.5 level. The figures were as low as .42, .47, .073, .0045, .24 and .000077.

6.4 DISCUSSION

Teachers can expect to spend approximately 10% more time engaged in managerial behaviour in a gymnastics lesson than in games lessons in general. In particular the management of apparatus makes demands on a teacher's time which is deducted from the available teaching time. From the sample of lessons studied no obvious strategies for reducing this time were observed with teachers directing operations verbally for each group, assisting personally where necessary, checking the apparatus before use and with all of the groups waiting until every piece was ready before work commenced. Whilst it is unlikely that management time could
ever be reduced to levels apparent in games lessons and
corners should not be cut at the possible expense of safety,
it might be possible for teachers to use apparatus diagrams
for groups illustrating what pieces are to be put where on a
regular basis. This would save explanation time and permit
everyone to begin getting out their apparatus at the same
time. Once the pupils know what they are doing, have had
their apparatus checked and are not in the way of others, it
should be possible for them to begin work. Additionally the
reorganisation of apparatus stores to permit ease of access
might hasten the process of management.

Physical Education teachers show a marked preference for
prescriptive teaching over all other behaviours. By
selecting this behaviour which is the hallmark of didactic
teaching or coaching behaviour, the teacher is responsible
for the vast majority of the decisions in the Physical
Education lesson. Prescriptions are given as to the task to
be followed, when and for how long it is to be done and how
it is to be accomplished in the majority of cases. Rarely is
choice given either for the nature of the task to be pursued
or for the method of pursuit and the practice time
available. One of the reasons for this apparent state of
affairs is due to the fact that the material used by the
teachers in this sample centred on the acquisition of motor
skill rather than on the development of understanding or
attitudes towards physical education by the pupils. In
short, the lessons appeared to be more concerned with
physical training than with education of the physical
through the physical with the data reminiscent of that obtained by Tharp and Gallimore (1976) who studied the teaching behaviour of UCLA basketball coach John Wooden. In the case study they recorded 2,326 acts of teaching by Wooden which they classified into 10 categories of leadership behaviour. They reported that 50% of Wooden's behaviour communicated information about what to do and how to do it.

Another factor which prompted this observation was the very short period of time given by teachers to asking pupils questions about their performance or about themselves and how they might feel about their involvement in the activity. Those questions which were asked tended to be concerned with knowledge of results following engagement in competition.

By the same token feedback constituted a very short part of lesson time. To some extent this might have been attributable to the fact that a fifteen second interval system will miss on occasions instances of short, evaluative feedback such as 'Well done, John'. It is, nevertheless, indicative of the fact that teachers do not engage in dialogue with pupils about their work for any length of time. They do not tend to watch a performance, comment upon it at any length, correct when necessary and stay to watch the next attempt with follow-up comment. The Physical Education teacher tends to flit from one pupil or group to the next commenting upon immediate impressions and moving on before a next attempt is made, a behaviour noted by Mawer
(1982). Few instances of feedback convey information about a performance but rather satisfy a perceived need for general approval.

Pausing to elaborate and give information or explanation about an activity is not a frequent occurrence either. Such explanations as there are tend to be delivered to the whole class rather than to groups or to individuals.

Observation of pupils working would appear to be the second most important teacher behaviour for most teachers and it is typically followed by further prescription or feedback, sometimes both. The length of the period of observation varies considerably and there would appear to be no general pattern among teachers. The more spread out the pupils, such as over two rugby football fields for example, the more likely they are to be left without teacher intervention for longer periods during practice and the longer the teacher will observe them working.

Most communications made by the teacher are audible, with the vast majority consisting of talk only. In a subject which is predominantly concerned with movement of one kind or another this is perhaps surprising. One of the reasons would seem to be that the Physical Education teacher presents a variety of activities and might feel less than physically competent in some and therefore unable to use demonstration as a teaching aid. It was noticeable that the three male teachers with the different profiles in
basketball and gymnastics were obviously competent in the practical demonstration of ball skills whereas there was no sign of obvious gymnastics competence. For the same reason there was very little use of the category which might have described the demonstration of skills silently - what little demonstration there was the teachers preferred to comment upon at the same time as giving the performance.

There would seem to be very little logic behind the preference of Physical Education teachers to direct most of their attention in lessons to the whole class first, then to the groups and finally to the individual child for less than 25% of the time. The reason would appear to be that material is organised for all pupils in the same way even if the lesson is gymnastics and some groups may, of necessity, be using different apparatus. This form of organisation does not take account of individual differences or to put it another way, there is an implicit assumption that all pupils are at the same stage of learning and learn equally well the same way as one another. Clearly this cannot be the case but there was no evidence of mixed ability teaching in this study. Interestingly the form of organisation used for work did not dictate the direction of a teacher's communication as much as might have been thought. Most tasks were the same for everyone and pupils were paired or put into larger groups for their practice yet the majority of communications were still addressed to the whole class, therefore it would appear to be the nature of the task rather than the form of organisation which dictates the direction of a teacher's
communication. This tendency leads to a great deal of teacher intervention in a lesson with all of the pupils having to stop what they are doing whilst the teacher talks to them.

The teachers in the sample showed a remarkable level of consistency in their selection of behaviours regardless of the activity, environment, age range or their own personal attributes including their training. It was difficult not to escape the conclusion that the physical education stereotype cited by Hendry (1973 and 1978) was very much in evidence in 1985.
6.5 SUMMARY

BOS 2 was used to describe teacher behaviour in 19 lessons of physical education taken by eleven teachers from three Local Education Authorities.

From the data the category Totals, Mean, Standard Deviation and Percentages of lesson time spent engaged in the different teacher behaviours were calculated. Statistical comparisons were made between teachers to establish whether there were significant differences in the observed behaviours of male and female teachers in general, when teaching games as opposed to gymnastics in particular, and by the same teachers but teaching different activities.

The main findings were summarised and the data reported and discussed.

There was little distinction between the teachers in spite of differences in personal characteristics, training, the activity being taught, the pupils or the school environment. Didactic teaching was the preferred teaching method with verbal communication by the teacher dominating the lessons which were organised for the teaching of the whole class engaged simultaneously in practising the same task. Few questions were asked, few explanations given, and demonstrations were a rarity. Teachers spent most of their teaching time prescribing to the pupils what was to be done, when, where and how, and their next preferred behaviour was
to watch it being done. Such was the level of similarity and consistency in their behaviours that it was concluded that a stereotype did exist in the teaching profession.
CHAPTER 7

IMPLICATIONS

7.1 INTRODUCTION

This study sought to contribute to research in teacher education through the design of a systematic observation instrument which could be used to provide a description of teacher behaviour in the physical education lesson. Due to the fact that the research was breaking new ground in this country, it was resolved to write a comprehensive account of the stages necessary in instrument design together with any pitfalls encountered, in order that others may be assisted with the methodology in future research.

Using an American observation system as a preliminary guide, a 19 category system with a 15 second interval recording procedure was designed and subsequently tested by two observers. Figures of 80% and 82% were obtained for reliability and validity respectively, and the system was used to obtain descriptions of teacher behaviour in 19 lessons of physical education.

As the principal purpose for the design of the instrument was that it may be used to influence the work in teacher education programmes, the implications arising from the main findings reported in Chapter 6 will be discussed in the context of initial teacher education programmes in physical education. The BOS 2 instrument will be examined and
recommendations made for its use and further development.

7.2 IMPLICATIONS FOR INITIAL TEACHER EDUCATION PROGRAMMES IN PHYSICAL EDUCATION

From the sample observed, the teaching of physical education in schools had all the hallmarks of physical training for skill acquisition to the exclusion of anything else. The vast majority of decisions were taken by the teacher, and these included the type of task set, the location of the task, the duration and organisation of the task, with whom the pupil was to work and the form of accountability which was operating. For the purpose of skill development, the pupil was entirely dependent upon the teacher for augmented feedback. Additionally, the teacher managed the choice and distribution of equipment or apparatus to be used in the lesson.

Not surprisingly the teachers in the sample showed a clear preference for practice and command style teaching in their lessons, all of which showed a marked similarity to one another. The pupils were not normally invited to speak although they were not necessarily prevented from doing so, and they usually had a clear idea of what was expected of them. The lessons ran with the smooth precision of a military drill.

It would appear from the sample that these teachers had been trained in direct or didactic teaching and were well practised in the use of the contributory teaching skills.
Training would necessitate identification of sound management practice and the repetitive practice of those skills of both teaching and management which would lead to a smooth-running organisation. Adopting this style teachers would need to be good observers of pupil responses, accurate diagnostics of motor behaviour, clear speakers, possess sound leadership qualities and preferably be accomplished demonstrators. Their form of teaching requires of the pupils that they listen well and copy or translate accurately.

Practice and command style teaching both have a part to play in the teaching of physical education, and some would say a very important part. The fact, however, that the descriptive records obtained through the use of BOS 2 demonstrated the indulgence of teachers in prescriptive behaviour for such a large portion of the time and usually with the majority of the class simultaneously, must be a cause for concern among teacher educators. So well practised were the teachers in this sample that they did not alter their style even when teaching different pupils and activities, and it is not hard to surmise that the teachers did not possess alternative teaching styles.

Overuse of command and practice styles deny the pupils the opportunity to make decisions, create their own work and be more responsible for their own learning. The pupils in this sample were necessarily teacher-dependent during the observed lessons. During teacher training it would seem imperative that the trainee be familiarised with alternative teaching styles, a move which would produce other beneficial
effects.

In the sample, the favoured form of organisation was either the whole class working simultaneously at the same task or they worked in groups which varied in size according to the team required in the game or the amount of apparatus provided in the gymnasium. Familiarity with alternative teaching styles should encourage the teacher to be more flexible in the forms of organisation chosen as the context for learning, with these perhaps chosen as a priority rather than, as one suspects, out of expediency.

By the same token, examination of different teaching styles which focus on the ways children learn, might encourage evidence of mixed ability teaching - a feature which was totally absent from the sample observed. This in turn should encourage teachers to speak more with individual children and provide more frequent and longer examples of informative feedback. Practice at talking with individuals might also encourage the teacher to invite a verbal contribution from the pupil, which not only could assist in building good relations and a warm climate in the gymnasium but also contribute to the notion of language across the curriculum for all pupils.

Through the regular use of BOS 2 trainee teachers could build personal teaching profiles which could be compared during the teaching of different activities, with the various age groups and at different stages in their
training. As with the teachers in this sample, the data will reveal preferences in the chosen behaviour and the mode and direction of that behaviour. Depending on the interpretation of the data, a behaviour modification programme could then be implemented and followed by the obtaining of another record which could be compared with the earlier descriptions.

For the purposes of profiling, it would not be necessary to use the whole of the system at once. A descriptive record need only be taken of the teaching as opposed to the managerial behaviours, for example, and the mode and direction dimensions may also be used independently. A trainee who wishes to move from a predominantly prescriptive form of teaching behaviour exercised with the whole class could obtain records of his teaching using the teaching and direction dimensions only until such a point as they reveal more frequent use of other behaviours with his attention divided appropriately among class members.

The lack of teacher demonstration was noted among the sample in this study. There was ample opportunity for demonstration in these teacher-led lessons, and it must be concluded that the teachers concerned were incapable of performing the skills adequately or occasionally did not feel the need to demonstrate. One of the reasons given in the teacher education establishments for spending a considerable amount of the training time on practical studies is the necessity for the
teacher to be able to use this powerful tool of demonstration. The catch phrase 'a picture is worth a thousand words' is one which every trainee undoubtedly hears, yet it would appear that he might not be paying a great deal of attention. If that is the case, should the institutions spend so long on practical studies at the expense of subject or methodology studies? If the trainee is losing the demonstration habit, could it be that this aspect of his mode of teaching is inadequately monitored during training?

Use of BOS 2 to check and encourage the use of demonstration would also serve the purpose of advising the trainee on the amount of talking he is doing in a lesson. Trainee teachers are frequently heard to complain of sore throats and loss of voice during teaching practice in schools and the reason is not hard to find.

Whilst the teaching behaviours are undoubtedly important, the managerial behaviour of the teacher of physical education can make or mar a lesson. Too long in management is at the expense of learning time, and use of BOS 2 provides a clear, quantitative measure of the percentage of a lesson spent in management and differentiates between one kind of managerial concern and another. A great deal of attention is paid in the first two years of teacher training to the acquisition of managerial skills, yet it is a skill which is rarely monitored systematically. It appears doubtful that practising teachers have considered ways in which management time in gymnastics lessons in particular
might be reduced, yet the evidence from the sample suggests that this should be a priority.

Obtaining reliable descriptions of teacher behaviour could be invaluable for the trainee who has had to rely entirely on the subjective impressions of tutors and cooperating teachers for such feedback. Traditionally the skills of teaching would appear to have been passed on by word of mouth from those licensed to teach, and the evidence obtained from this sample would suggest that unsystematic observation and recommendation breeds dependence on a master model.

Such is the stereotype which emerges from this background, that it gives rise to real concern for the future of physical education teaching in this country. The physical education curriculum in many parts of the country is witnessing change - change in both content and in methodology. The objectives would appear to be shifting from predominantly skills promotion to developing knowledge, understanding and attitudes as evidenced by the keen interest in Health-based physical education and Teaching Games For Understanding currently being expressed by teachers and administrators alike. It would appear to be an almost hopeless cause for developers to challenge teachers and teacher educators to adopt models for change which demand flexibility, sensitivity, new knowledge and above all, a change of attitude towards the subject and the pupil, if those same people demonstrate the stereotype characteristics of the P.T. Instructor.
It would seem imperative that changes in the training courses be introduced immediately, and the inclusion of work which provides a 'mirror of behaviour' might go some way towards remedying the long term problem.

7.3 THE BAILEY OBSERVATION SYSTEM (BOS 2)

7.3.1 Early Development

The design of BOS 2 took a total of two years from inception to completion. The first eight months were spent obtaining material and devising professionally relevant dimensions and categories which could be validated by experts in the field of physical education. The process of monitoring this validation necessitated the production of an observer training manual and the keeping of a diary in order that progress of the observers could be checked as well as the development of the instrument itself. The comments and suggestions of the observers as they studied videotapes of teachers in action and used the categories in the system to describe what they saw and heard, served to modify the category definitions until each observer could independently accept their validity.

A system for use in the field, however, must be more than a valid collection of behavioural categories. As a complete package it must be manageable, and at the end of the first eight months it was decided that BOS 1 was both too long and it necessitated far too great a commitment on the part of observers for it ever to be used regularly by those for whom
it was originally intended. It was considered that these
beginnings were a necessary stage in instrument design and
had clearly demonstrated the need for simplicity and
versatility in any observation system for it to be a useful
tool in the education service. The essential features of BOS
1 were retained in the development of BOS 2 without too many
compromises, and the result was a much neater, simpler
system which required less than a full-time commitment from
its users.

7.3.2 BOS 2 Categories

Extensive use of the instrument has been sufficient to
convince observers that the major aspects of teacher
behaviour in physical education are covered by the system.
The category 'Other' has been scarcely used and that would
have been the first sign of inadequate design had it
collected behaviours unable to be categorized.

It must be accepted that for five categories to adequately
describe all of the aspects of teaching behaviour there must
be some lack of discrimination. BOS 2 shares this complaint
with the vast majority of systems which attempt to cover
more than one aspect of teacher behaviour such as management
as well as teaching. The alternative is to produce a system
which focusses, for example, on an aspect of teaching
behaviour such as Feedback and design a system which
provides the observer with descriptions of that behaviour
alone. The moment the observer stands further back and looks
at classroom events through a wider angle lense there is an inevitable loss of discrimination. Implicit in the design of BOS 2 is the acceptance of the wider role of the physical education teacher which accommodates a considerable amount of management as well as teaching in a lesson and the inclusion of this dimension led inevitably to a reduction in the number of teaching categories which an observer could use.

Mention has already been made of the fact that it would be possible for an observer to use the dimensions independently if desired. By the same token if the number of decisions required of an observer were to be reduced by such a strategy then it would be possible for the number of remaining categories to be increased if greater detail within a dimension was required. For example, category 5 'Prescribes Response' includes two major aspects of prescription given by the teacher. These could be divided to show 'Prescribes Task' i.e. telling the pupil WHAT is to be done, and 'Prescribes Method', i.e. telling the pupil precisely HOW the response is to be made. Examples belonging to these two categories respectively would be: 'Do a handstand followed by a forward roll', and 'Make sure that your legs remain extended for as long as possible'.

The second example should not be confused with further elaboration or extended explanation which might provide the pupil with information about how best to maintain a balanced position - the giving of information which should be coded
under category 8.

It is acknowledged that some confusion remains evident in the use of categories 5 and 8 by observers. The problem would appear to be one of some similarity in the message on occasions given by the teacher which only becomes more clearly different as the verbal interaction unfolds. The difficulty for the observer is that the 15 second signal encourages a prompt response which might, on occasions, be best made after a pause to listen as the content of the interaction unfolds rather than immediately. This advice needs to be given much earlier in the training programme than has been the case in the past.

The strategy of expanding categories could be used to extend category 6 'Gives Feedback' if the observer wanted to distinguish between one type of feedback and another. In a similar vein, analysis of the tapes showed that teachers seek information about results rather more often than information of another kind such as how a pupil is getting on, for example, and this category 9 could be divided. Observers requiring further information about the way pupils work with one another may prefer to add another category to the direction dimension to bridge the gap between an individual and a group, which by definition could be any number between two and short of the whole class. An obvious addition would be 2's in order that work in pairs could be monitored.

It should be stressed that any alterations would require the
provision of adequate definitions which could be validated and tested for reliability in their use by other experts. Fundamental alteration to a system is unwise on an ad hoc basis for then it becomes a different system and as such must be tested in the same way that the original system was tested. Data from an untried system would be unacceptable for purposes of publication and comparison with other data.

7.3.3 Interval Recording Procedure

The change from coding using a naturally occurring unit i.e. coding a new behaviour each time it happened, to using a 15 second interval procedure was one of the most significant changes in the development of BOS 2. Coding every behaviour as it happened was very tiring and led to inaccuracies. One alternative which was considered was to code for periods of time such as five minutes followed by a break for a similar period, but this was rejected because of the unnatural division of the lesson into segments which could alter the teaching profile of the physical educationist very considerably. For example, periods of management in a lesson can begin and end in five minutes, as can introductory activities which may reveal a different style of teaching to longer, extended practice periods. Should these lesson sections coincide with the observer's rest period then the teacher's profile would alter significantly.

The decision to select 15 seconds as opposed to 5 or even 10, was to enable the observer to look, listen and record
without undue haste and the risk of inaccuracy. It was hoped that this period would give the observer confidence and was arrived at through repeated practice at recording and putting the record on coding sheets in the three dimensions by the designer. 15 seconds permitted three decisions to be made, committed to paper and for the observer to look up and locate the teacher with approximately 4 seconds to spare. It is felt that well-practised observers could reduce this interval to 10 seconds which would permit a greater number of samples to be taken of the teacher's behaviour in the lesson.

It is bound to be a weakness of interval recording procedures that not every behavioural act in a lesson is coded. Researchers who might wish to obtain descriptions of a teacher's behaviour in order to write a detailed individual profile may consider it necessary to use a 3-second interval such as is used in the Flanders System. This is tantamount to continuous coding and permits one decision to be arrived at and committed to paper before the next observation is made. If only two dimensions of BOS 2 were being used then this might be possible but it is highly unlikely that an observer could adequately cope with three. It would not be recommended for trial other than with a videotape of the lesson which would permit use of the 'pause' button.

There is little doubt that the majority of errors made by the observers in the process of training and testing were procedural rather than errors of interpretation. This is a
A direct result of using an interval recording procedure which demands that the observer codes the observed behaviour evident at the point the interval bleeps or is displayed on the monitor. Teacher behaviours do not always land precisely on arbitrary time units, and it was often the case that a new category of behaviour would begin or end on the 'boundary' of a second interval, e.g. between 30/31 seconds. Decisions by two observers to code the first behaviour or the second behaviour were frequently at odds with one another and the result was an apparent mismatch in the observed behaviour. It should be noted, however, that although this practice served to reduce the reliability of the inter and intra observer scores, it should not be seen as invalidating the observations in any way. Both observers correctly identified that the behaviours occurred but at marginally different times.

An alternative to this method of timing the coding which has been used elsewhere is to look at the behaviour over a short period, e.g. 5 seconds, and then code the behaviour. The major problem with this strategy is that teacher behaviours change so frequently that it would be quite reasonable to have three behaviours occurring within a 5 second interval and then the observer has to select which behaviour he will record. This has led to designers writing into their systems a hierarchy of behaviours which observers have to use when they note more than one behaviour. Such a practice was rejected in the design of BOS 2 because of the high level of inference and designer influence it was considered operated implicitly in these systems.
7.3.4 System testing

Testing for validity and reliability would appear to be an area of controversy.
In the first instance it is not unusual to find that studies do not report the means whereby figures were obtained and used in the calculation of interobserver agreement scores. Secondly the figures themselves are not provided so it is impossible to verify the results or calculate the means by which they were arrived at.
Thirdly, and contentiously, varying means have been used to calculate for reliability which differ significantly in their stringency.
In testing BOS 2 for reliability and believability or validity of observation, note was taken of considerable criticism which had been levelled in the past at studies such as this one. The means used to test the system were carefully articulated in order that other researchers may use them as a blueprint for their own or simply be able to accept the data obtained through BOS 2 use with confidence. It was acknowledged that in so doing it was likely that the scores would be lower than is sometimes obtained in testing other systems, nevertheless the figures of 80% and 82% reported earlier were viewed optimistically. It would be hoped that other researchers might follow this example and that the comparisons between the reliability of one system and another might become more real than speculative.
7.3.5 Computer use

Adding to the speed and efficiency of the system has been the writing of appropriate software for computer use. Coding can be done directly on to a computer keyboard and a hard copy obtained of the raw data together with a summary of the category use within seconds of completing the coding. This facility extends the potential use of observer systems enormously. Not only does it take the drudgery out of post-coding calculations, but it opens new doors for the form of analysis which can be applied to the raw data. A preliminary examination of the BOS 2 data, for example, has revealed that patterns or cycles of behaviour are evident through the use of some categories more than others and in relation to other categories. Means of extracting these patterns from the raw data using a main frame computer are currently in progress.

The ready access to computers in universities, polytechnics, colleges and schools as well as videotape equipment, must signal a new advance in work of this kind. North American researchers have had the benefit of this technological advance for longer than colleagues in this country which might partially account for the relatively slow progress of work in this field to date. It can be anticipated that the availability of new technology might encourage greater enthusiasm to pursue this type of research in the future.
7.4 IN-SERVICE TRAINING

BOS 2 was designed and used for the purpose of describing teacher behaviour in physical education lessons. The implications arising from interpretation of the data would suggest changes might be necessary in the training programmes of students on initial teacher education courses in physical education and it has been suggested that use of systematic observation techniques could do much to remedy anomalies in the training.

The problems for serving teachers remain acute. Behaviour modification might prove harder for those who have practised their teaching for considerably longer yet their need could be seen to be all the greater. It is recommended that serving teachers should be encouraged to examine their teaching behaviour systematically with the aid of instrumentation such as BOS 2. Administrators should be encouraged to mount In-Service courses which focus on the use of such instrumentation and advisors sensitised to the problems of observer training support which the teachers would need. Teachers should be assisted in the drawing up of a 'contract' with a fellow colleague which enables both teachers to observe each other and share their observations and interpretations with confidence. Such a contract would serve to remove the perceived threat of 'supervision' by a senior colleague which might be used inappropriately and also encourage the individual teacher to adopt the role of researcher in his or her own classroom.
It is to be hoped that the design and use of systematic observations systems in this country might flourish in the wake of continued pressure from abroad and after the lead which has been taken in this project.

The project itself was unique. No other fully tested system, purpose built for the description of teacher behaviour in physical education in Britain has been produced. It comes at a time when changes in the physical education curriculum and in systems of accountability put pressure on those in education to mount courses and provide the means for change. It would seem imperative that a prelude to change should be the provision of a means whereby reliable descriptions of what is happening in our physical education lessons becomes available. Systematic observation instrumentation could supply this demand and well as providing a technique for monitoring ongoing events in the physical education lesson.
REFERENCES
(referred to in the text)


Physical Education. Summer, Monograph 1.


Training, Research and Education. New York: Wiley.


Hallam, V. (1978). Systematic observation of some aspects of
educational dance and educational gymnastics. Unpublished M.Ed. thesis: University of Leicester.


Martinek, T.J. & Johnson, S. (1978). Expectant levels of achievement effects on interaction patterns and self-concept
in elementary age children. Research Quarterly, 50, 1, p. 60.


Stubbs, M., & Delamont, S. (1976). Explorations in Classroom


BIBLIOGRAPHY
(additional texts consulted)


Dunkerton, J. (1981). *Should classroom observation be*
quantitative? In Educational Research, 23, 2, February.


Karafin, G.R. (1973). Discussion of Considerations for Selecting or Developing an Observational System. Classroom
Interaction Newsletter, 8, 2, pp 15-32.


de Janeiro.


APPENDIX 1

TEACHER PROFILE

Ref. No.

SCHOOL

NAME Male/Female

No. of years teaching experience:

Age Group:
Student...
21-25... 26-30... 31-35...
36-40... 41-45... 46-50...
51-55... 56-60... 61-65...

Details of Professional Training:

Any other information:
APPENDIX 2

OBSERVER TRAINING DIARY

Session 1: May 20th 1981

REPORT

A copy of the Manual was given to each of the observers. A brief explanation was given regarding the nature of the research programme, the system and the role of the observer in the training and testing programmes. The practical commitment required of each observer was investigated to enable those with heavy teaching loads to assess whether or not it was possible to undertake such work.

The dates and times of the next two sessions were fixed:
June 2nd at 8.45am in P507
June 3rd at 9.00am in P507

Observers were asked to read the Manual and bring any further queries to the next session.
Length of session: Half an hour.

DIARY

The principal concern of two out of the five observers chosen for this training was that of the time commitment each would have to make. These sessions have begun at a very busy time on the academic calendar for the observers. They were given a week to familiarise themselves with the Manual largely to determine whether or not they could make this
commitment to both the training and testing sessions.

Session 2: June 2nd 1981

REPORT

Training began today.

1. CORRECTION
A correction was made to D3 on page 4 of 'Further Information' in the Manual. The final sentence should read 'Categorise the brief response in D6 and the evaluative comment in D3'.

2. DISCUSSION
The observers were asked to comment on the Manual.

2.1 The use of 'Observes' in E MODE dimension was raised.
The point was made that this was not considered to be a method of communication. This was acknowledged to be the case yet it did describe the means of carrying out the Act, in this case D4, Observes Response.

ACTION - To reconsider definition of E MODE.

2.2 The use of Public/Private address in F DIRECTION dimension.
The point was raised that the system does not differentiate between a teacher's contact with one pupil being for his ears alone or made with the intention that others should be party to the communication. It was pointed out that intention could be difficult to determine accurately. It would require very careful definition and Ground Rules to distinguish, for example, between teacher-talk directed towards one which some or all could hear if they chose to
do, and teacher-talk directed towards one which no-one else should hear. It might be possible to do this, however, if it was felt to be of sufficient importance.

ACTION - To leave open the possibility of inclusion at a later date, if it is felt to be desirable.

2.3 The use of empathising comments did not have a separate category.

It was explained that this aspect of verbal behaviour had been carefully considered and had been included in a very early consideration of category design. It had been removed because of very infrequent use and now such comments appear under H1.

3. LIVE CODING OF EDITED TAPE

3.1 Identifying Managerial Acts

The observers were shown the edited tape 0/1 illustrating examples of Managerial Acts. These had been coded and could be followed on the prepared coding sheet.

3.2 Identifying E Mode

The same section of tape was re-run to show examples of Mode.

3.3 Discussion

3.3.1 C8 'Starts and Stops Activity'

Observers were asked to comment on C8 'Further Information'. It was felt easier to record all instances of starting and stopping activity as C8 even when the means used to do so by the teacher might appear to be better described by another category, e.g. C1.

ACTION - Delete the main paragraph from the Manual under C8 beginning 'Occasionally the teacher stops......to that
effect'.

Add

All instances of starting/stopping activities are recorded here even though the verbal means used by the teacher may not be entirely consistent with the primary aim, i.e. to start or stop the activity. For example, 'Right - come over here' may verbally appear to be an example of organising people (C1) but on this occasion the teacher has used the statement primarily to stop the activity in progress. It should be categorised as C8.

3.3.2 C5 - Officiates

This category is reserved for the Managerial Function only. The teacher either officiates or teaches. If he is seen to be carrying out the duties of an umpire or referee by word or deed, category C5 is used. Should he have been officiating and then interrupt that Managerial Function, for example, to give information about a Law or Rule, then he is deemed to be teaching which would be categorised as D5.

3.3.3 E3 - Gestures

Further explanation was given regarding the distinction drawn between gesticulates and gestures, and the use of an incomplete demonstration which would be categorised as E3 and not E5. The point was made that there is a difference between a managerial gesture (e.g. pointing) and a teaching gesture (e.g. using the hands to describe a movement). This was felt to be recorded adequately because the Function would appear above the Mode on the coding sheet.

The session concluded at 9.55am

The length of the session: one hour.
One of the original team had withdrawn and another had taken his place. The substitute attended this first practical session. Since the first session another observer had been approached to join the team but he failed to turn up for this session. A member of the original team was unable to join this session due to another engagement. A total of four observers attended (initials BR, BP, LA, TJ)

Session 3: June 3rd 1981

REPORT

1. LIVE CODING OF EDITED TAPES
1.1 Identifying Managerial Acts
The observers were shown the remainder of the edited tape 0/1 with examples of Managerial Acts. These had been coded and could be followed on the prepared coding sheet.

1.2 Identifying Teaching Acts
The observers were shown examples of Teaching Acts from the edited tape 0/1. These had been coded and could be followed on the prepared coding sheet. Only the Acts were identified and not any other dimension of behaviour at this stage. The observers were given the opportunity to try and identify some of the Acts without consulting the coding sheet. Some of the examples were difficult. An encouraging start was made although it highlighted the necessity to listen carefully to the linguistic behaviour of the teacher in
order to note the changes of Act which often occur mid-sentence. It will be very important to adhere closely to the definitions and ground rules in this dimension due to the similarity of some of the categories. In particular, 'Describes Response' must relate a) to a past response, and b) it must be an accurate representation only of that response which does not include any attempt at evaluation or the giving of further information.

2. DISCUSSION

D4 - 'Observes Response' received much attention. The point was made that whilst a teacher may appear to satisfy the definition for this category, he may in reality be neither watching nor listening but mentally planning the next move, for example. Two of the observers made the point repeatedly that observation requires mental activity - one can see without observing, for example.

It is acknowledged that this is a difficult category. No descriptive tool can interpret behaviour and one of the aims of this tool is to describe and record in categories which are mutually exclusive. This necessitates the use of an external frame of reference i.e. one which requires little or no inference on the part of the observer. In practice the observers were saying that for this category one could accurately describe that which is not actually happening. The point is a fair one but it is doubtful if such an instrument can accommodate an adjustment. It may have to be sufficient that it does accurately describe the observable act.

ACTION - To bear this discussion in mind when watching more
tapes and to consider at a later date whether an adjustment could or should be made. It may also be useful to reconsider the definition.
The session concluded at 10.05am
Length of session: one hour.

DIARY
The same four observers attended as had the last session. Two observers have yet to appear. This will cause problems in a group training situation.

Session 4: June 8th 1981

DIARY
This session did not take place as intended. One observer had another engagement. Since the last session it had been decided that the two who had missed sessions two and three should go through 'catch-up' tutorials after Session 5, so they were absent also. As only one of the three observers attended the formal session was cancelled and the opportunity was taken to give the observer present a tutorial on coding practice which was intended for session 6, originally.
At this point in time it is apparent that there is an attendance problem which is going to necessitate a considerable quantity of extra time being devoted to tutorials for those missing the group sessions. It is inevitable due to the staff commitments to work both in and outside of the university. It is important to arrange the
sessions regularly and frequently in order that stability of system use be a realistic goal. Tests for reliability must be completed prior to staff leaving during the vacation. Whilst it is difficult to estimate precisely the extent of irregular attendance, it is important to build into the training period adequate time for 'catching-up' sessions. In this case it is hoped that two sessions per week would leave room for manoeuvre.

Session 4 (repeat): June 9th 1981

REPORT

1. DATES OF THE NEXT SESSIONS

Tutorials
RP - June 10th at 9.00am in P507
LA - June 10th at 2.45pm in P507

Group sessions
June 16th at 9.00am in P507
June 17th at 9.00am in P507

2. LIVE CODING OF EDITED TAPE

2.1 A correction was made to 'Function Subscript' on the coding sheet headed teaching under the timed unit 00-16-45 which should now read 'shares' and not 'Does'.

2.2 Identifying the Teaching Acts from the edited tape 0/1. Observers were encouraged to cover the relevant section of the prepared coding sheet in order that each could attempt to code the Acts on his own. This was extended to include looking at the other dimensions also, with the exception of
Time and Content.

An alteration was made to the Master copy. Unit 00-16-54 included a managerial gesture in between two teaching acts which was deemed to be of significance by the observers and worthy of inclusion. The gesture was used by the teacher to start the activity. This unit is now divided to include a managerial function (A1), Subscript (B1), Act (C8), Mode (E3), Direction (F1), Action (G1). The unit starts at 00-16-56 and ends at 00-16-57.

There was one example of a very neutral evaluative comment which might be referenced under 'Further Information'. The comment was 'Right Michael, we'll give you that one.'

2.3 Identifying the dimensions A, B, C, D, H and I from the lesson tape LH/6/1.

The tape was played straight through to familiarise the observers with the content. Each was given a coding sheet and the relevant page in the Manual was referenced in order that the reported speech only should be followed. This allowed the observers to ignore unit analysis for coding purposes as the reported speech was written out in such a way as to indicate clearly where one unit or Act began and ended.

The tape was played and controlled in such a way as to show approximately three separate functions/acts at any one time. It was then re-run several times. After a preliminary viewing a comment was made that it can sometimes be easier not to look at the monitor but to listen to the linguistic act only. After this had been analysed, then to refer back to the monitor to check on Mode and Direction.
Two of the three observers at this session were attempting to code the section of tape in all of its dimensions. Both were more familiar with the procedure that the other observer. It was felt that more work at coding procedure should have been done before progressing to this practice. The tutorial sessions should redress the balance, it was felt.

Comments included - a tendency to omit E Mode categories; not always putting the Managerial Function on to the sheet when an Act is recorded; a need to learn the Function Subscript definitions.

The session concluded at 10.25am
Length of session: one hour and fifteen minutes.

DIARY

One of the observers was absent for this session.

The situation had been reached where in this group session, all three observers were at different stages. One was quite experienced and was racing ahead wanting the tape to be wound on. A second was inexperienced, struggling to attempt the full coding procedure even though it had not been asked for. It was felt that this observer felt to be under pressure from the experienced observer. The training session was becoming competitive. The third observer fell between the two - occasionally missing items in the Mode dimension but otherwise progressing satisfactorily.

The inexperienced observer was unhappy at the close of this session. It was probable that the tutorial on coding
procedure was coming one session too late at this stage.

Session 5: Tutorials on different days.

REPORT

These tutorials took the place of a group session partly because it was difficult to meet as a group but largely because it was felt to be desirable that coding procedure should be done on an individual basis. One tutorial constituted an introduction to the system from the beginning. For this reason the Report will only outline the major points which were brought up in the sessions as a whole.

1. TIME DIMENSION

If a new unit of analysis begins as the second digit changes, record the later second as the time for the start of the unit.

2. CODING PROCEDURE

Two or more Acts plus the same categories in the other dimensions are coded as ONE i.e. the teacher prescribes to one pupil by speaking and whilst moving and this is immediately followed by the teacher doing exactly the same but to another one pupil (even if the content of the prescription changes) then this is coded as one unit. The teacher behaviour has not changed. However, should one of the categories change, for example, the teacher addresses the next prescription to a group whilst moving, then the act is coded again.
3. FURTHER INFORMATION

If a teacher stops a class and waits for each to be ready until he speaks, count the number of seconds he waits. Should it be over five, do not code this Managerial Act as 'Observes' (which is a Teaching Act) but add the time to the Act CB and include 'observes' in the Mode Dimension. Under five seconds would be added to CB in any case.

Should the waiting be used by the teacher to gaze out of the window, arrange equipment, tie a shoelace, etc, then this should be coded as Non-Interactive behaviour either on-task or off-task according to whether or not the act is relevant to the lesson.

4. GROUND RULES FOR THE UNIT OF ANALYSIS

The last sentence should read, 'in the case of c, record the symbol / in 13 or 14 and leave all other dimension and category boxes clear. Record the time.

DIARY

Observer BR

The videotapes were not used in this brief, half hour tutorial. The required changes were made in the Manual. The coding procedure was revised theoretically.

Observer RP

The coding procedures were gone through using a coding sheet.

The tape LH/6/1 was played, concentrating on coding procedure. The observer found it easier as the tutorial
progressed.
Some difficulties were experienced in differentiating between 'prescribes' and 'describes'. The suggestion was made to think that 'prescribes' is NOT feedback.
A few corrections were made to the Manual ref. LH/6/1 'Complete' section.
100% agreement was achieved on the Time Dimension.
Some difficulty was experienced in identifying 'Officiates'. The comment was put 'How do you recognise officiating if you do not know the game?'
The comment at the end of this two-hour tutorial was that it all seemed so much easier now.

Observer LA
During this one hour tutorial corrections were made to the Manual and there was practice at using the coding sheets, plotting the unit of analysis.
Using the tape LH/6/1 there was complete agreement achieved on using the Teaching and Managerial Acts.
Two corrections were made to the Time Dimension. A ground rule was established to help with the situation which occurs when the digit appears to change as the Function changes and the unit begins. In this case the later second is noted.
The observer noted that he felt more confident after this session.

Observer RT
This long three-hour tutorial was intended to bring the observer up to date as quickly as possible as he had not
attended the previous practical four sessions. As he had read the Manual it was possible to progress quickly through looking at Managerial Acts to including a look at the other teacher behaviour dimensions. No problems were encountered until the Function Subscript was used when the teacher was Managing Equipment. The problem was this: although the teacher gave the command to do X or Y with the equipment, it was the pupils who actually moved the equipment. This should be coded 'Delegates'. The observer commented that it became clearer once he put the subject of the verb in front of the equipment.

In analysing the Teaching Acts from the edited tape 0/1 it was necessary to stress the need to listen carefully to the language of the teacher for clues, e.g. 'I want you to', or 'now this time' usually led into 'Prescribes' (D1). Similarly words such as 'very', 'too', 'good' were indicative of evaluative statements. It was necessary to stress that to use 'Describes' there must be an accurate, unappraised account of the response. Function Subscript 'shares' required careful explanation for 'seeking Information' when a response is waited for and received.
Coding procedure was gone over theoretically and a few sheets of coding prepared before progressing to the tape LH/6/1.

100% agreement was achieved on the Time Dimension. A rule for Mobility was added after the observer experienced some difficulty with this dimension to state that mobility would equal more than one step to achieve a change of location rather than a change of weight which might happen if a teacher took one step one way and another back on the same path.

The observer appeared confident and was picking up the correct cues from the tapes for coding purposes.

Some time was spent over the word 'Good' which a particular teacher used often. According to the system it would be coded 'Evaluates Response' but the observer felt that this teacher said it for something to say and not specifically to evaluate the outcome of the game or the performance of the pupils involved. He would have coded it H1. After a lengthy discussion it was agreed to leave it as 'Evaluates' largely because there was the element of doubt remaining that that the teacher may well have been appraising the score, for example, and so the pupils' performance. The close scores of the games in the tape (4-1, 2-2) may have told the teacher that she had organised her team selection well or that one of the teams in the 2-2 game had done better than she anticipated to achieve a draw. As this was a real possibility it was coded D3. The comment was made that one was persuaded to code things according to the designer's interpretation rather than according to one's own instinct
and perception! Undoubtedly this is true but it is also the strength of a system for although it reflects the designer's thinking, it provides clear Ground Rules which iron out observer instincts which may not agree even with other observers.

This observer also had difficulty knowing when the female teacher was using an official's language in the game situation as opposed to a teacher's. There would appear to be no easy solution to this except much practice and exposure to different games and to look for the clues in the content and context i.e. officiating is very likely to occur in the full game practice rather than in a skill practice. This lengthy tutorial was of particular interest and value. It would seem that this observer reached the same point in his ability to code accurately after three hours instruction as the others after their three group sessions plus a tutorial, there being a difference of approximately one hour. The individual attention coupled with the benefit of hindsight, an improved ability to communicate procedure and an increased awareness of potential difficulties had undoubtedly combined to improve training efficiency and effectiveness. It is difficult to estimate how much this observer has lost through not being party to the other observers' questions, although their significant contributions were put across whenever possible.

All except one of the tutorials was successful if success can be measured by observer response to the session. It was interesting to change the relationship for the
individual tutorial. Both designer and observer sat side by side close to the monitor and playback machine. The designer could see the observer's coding sheet and watch the pencil going down the list of category options before a decision was reached. Observers though out loud without prompting - a factor which the designer found particularly helpful. One observer commented that he felt 'put on the spot' in such a situation - which he found helpful, not obstructive.

One observer, the first to receive a tutorial, has received different treatment. His coding session was done theoretically using the Manual with no reference to the tapes. This was done initially because a) it was an impromptu session when the others did not turn up, b) it had not been carefully planned beforehand, c) time was at a premium and d) the observer was experienced in the use of the technology and other coding procedures and appeared to be quickly satisfied that he knew all there was to know about this procedure. At this stage one could only hope that the observer had not been disadvantaged through this approach.

Session 6: June 16th 1981

REPORT

1. ADDITIONS TO THE MANUAL

Additions to the Manuals were made as outlined on the previous report.
2. LIVE CODING OF EDITED TAPE

Tape LH/6/2 was played through completely in its edited version. Observers were asked to code the tape with the Designer playing the tape back. After one hour's coding the observers were asked to compare their coding sheets with the completed sheet in the Manual.

POINTS RAISED

Is a gesture mobile? NB Not in its own right. Mobility is related to feet for it constitutes a change of location.

Is not the use of an implement e.g. a hockey stick, as an aid to illustrate that as a stick it has a flat side and a round side, an example of 'Uses Aids' and not 'Demonstrates'? This is an important point requiring further consideration. It will be raised in another session.

The Ground Rules for 'Function Subscript' were gone over again at the request of an observer.

A correction was made to the Manual's coding sheet. The very first Managerial Act was changed from C1 to C8.

Observer difficulties in the coding procedure whilst the Designer was manipulating the play-back resulted in the arrangements for the next session being altered.

Tutorials will now take place as follows:-

RT 9.00am June 17th and 2.00pm June 17th
RP 9.30am June 17th
BR 10.00am June 17th and 11.15 June 17th
LA 10.30am June 17th and 3.00pm June 17th.

The session concluded at 10.30am

Length of session: one hour and thirty minutes.
DIARY

The intention was to allow the observers a free rein and code edited tape after having begun to do so in the earlier tutorials. The designer played the whole section through and then played smaller sections several times over. The tape was usually run to show a maximum of two Acts at any one viewing before being re-run. No advice was offered by the designer during this process.

OBSERVATIONS

Although the observers were not prohibited from speaking, scarcely a word was uttered the whole session. Gradually it became evident that the four observers were going at different speeds and that they wanted the tape stopped at different points. Afterwards there was a lengthy discussion on this and two factors were evident:—

- there was slightly too big a jump from the tutorial session to this session;

- Each observer would prefer to operate the tape play-back. A decision was taken to re-organise the next group session immediately into a series of tutorials in which the observers would operate the videotape recorder. This day's material would be gone over in that session. At the end the observers began to compare their codings with the Master copy. It was evident that several mistakes had been made, some in the 'Gestures' category and some in the different Acts which had been selected.

The major concern at this stage is whether or not the
System can be learned and tested in the time available. If sessions become largely individual tutorials, three factors emerge:—
The time scale increases overall;
The designer's commitment to the training programme increases four-fold;
The exchange of ideas between group members falls off.
The time period for training has probably been increased because the designer has attempted to achieve two objectives:—
For observers to learn to use the system, and this they seem to find easier on an individual basis, and to subject the system to review and debate based on the observers' perceptions. This last point can take at least as long to accomplish and needs to be borne in mind when sessions are planned. It is considered that a better system will emerge from a constructively critical atmosphere in which observers are encouraged to express views and the designer is flexible.
This period of 'negotiation' in the training process was proving to be the most informative, useful and stimulating part of the programme so far. It was regrettable that it was appearing to be rushed through the pending necessity to obtain test reliability figures. It needed to have started much earlier in the calendar year although the error rate for the designer would have been even greater than it was already. Analysing these tapes afresh with other observers had brought new features to light. Coding procedures were being analysed in great detail and the Ground Rules would
benefit enormously from this. It would have been beneficial to have gone through a 'dummy run' training an observer before starting the formal training programme with these university staff.

One of the observers who had attended four group sessions but no tutorials withdrew from the team today on the grounds of work pressure.

As no further group sessions were to be held it was decided to discontinue writing formal reports and instead for memos to be made in the designer's diary during the tutorials.

TUTORIALS

June 23rd 1981 : RP

Continued to be unhappy about the distinction between 'Describes' and 'Gives Information'. Practised with extracts from the tape KA/7/1. It takes between 10 and 15 minutes for this observer to relax and follow procedures with the minimum of error rate.

June 24th 1981 : RP

Continued with the same programme and coding seemed to be speeding up slightly. Went on to look at tape KA/7/2. 100% agreement for Time, Content, Mode, Direction and Function
Subscript. The majority of errors were in the selection of Teaching Acts. He commented on the need for practice. There needs to be included in the definition for 'Describes' and 'Evaluates' reference to the teacher's response. (one hour)

June 24th 1981 : RT

Used tape KA/7/1. There was a tendency to stop the tape too soon instead of letting it run to the end of the Teaching Act prior to putting in all of the categories. Consideration needs to be given to when to place the content code - at the first sign of change which might be Managerial or Teaching. (one hour and thirty minutes)

June 25th 1981 : BR and LA

Both observers had to cancel their sessions today.

June 25th 1981 : RP

Used tape AA/1/1. Two errors on the 'Mobility' dimension which needs checking again. The observer commented that he needs to go over the coding procedures again because he feels to be making errors through not knowing the rules. He finished the session early due to this and asked for the next session to be the next week and not tomorrow to give him time to learn the rules. He seemed annoyed with himself for making these errors. (Half an hour)
June 25th 1981 : RT

Used tape KA/7/2. This observer does not like clues or prompts when coding. His comment was 'Don't clue me in - let me think about it'. This, in line with other systems, does not permit the coding of intentions or acts which might overlap. For example, a teacher finishes giving information to a pupil whilst handing him a ball which is also understood by the child to be his signal to start the activity cannot be adequately coded by this system simultaneously. The child and the teacher may know that to begin is the intention but it is not explicit and cannot be coded.

The observer persists in finding difficult the distinction between 'Prescribes' and 'Gives Information'. Some difficulty was met also in determining the Direction to be coded when the teacher was concerned with the observation of one pupil with others. It was decided to let the tape run and code this Direction if the speech or Act after this observation period indicated the Direction of that observation.

June 26th 1981 : RT

Used tape AA/1/1. The point was clarified that mobility after one act and before another is ignored. The use of edited tapes tends to put the teacher's behaviour out of context. Observers have to code what they see. It had been evident over the weeks that the category
'Gestures' was causing great difficulty. The debate centred around whether or not it was even significant, for examples in the lessons tended not to communicate very much it was thought. If it was not very significant then it might be best to omit it from the system as it was causing difficulties of interpretation. This would be continued to be reviewed.

June 26th 1981 : LA

The point was clarified that behaviour which was neither that of a Manager nor Teacher was not placed in the Function Dimension. It was necessary to revise the coding procedures from the beginning. One of the reasons for this was the time lapse between sessions, but also because this observer has a tendency to look too soon for identification of all the dimensions and categories. Further guidance was given on identification and recording of the Function and Act categories. Practice was given in the recognition of boundaries of an Act within the unit of analysis. Some difficulty initially was experienced in running the tape machine for play-back purposes by the observer. (one hour)

June 29th 1981 : BR

The observer was informed of the revisions which had been made since the last session. It was agreed that should the Direction change during the same teaching Act then the unit of analysis changes and the teaching Act must be tallied
again with the new Direction noted. The remainder of tape KA/7/1 was run through.

In general this observer looked for cues well but attempted to go far too quickly, not always first establishing the boundaries of the Act before going over the Mode. Many errors were made in the Mode Dimension. There is a tendency for the experience of this observer in the work on Teaching Styles to influence the interpretation of a Teaching Act. The comment was made that the observer dislikes coding intensely and this had become increasingly obvious over the last two weeks. There would seem to be no cure for the concentration required for this activity and the potential boredom if it does not appeal to you. (One hour)

June 29th 1981 : RP

Practised tape analysis for most of the session without interruption. A definition and Ground Rule was clarified for Category F4 when a teacher uses one child to present a point to the rest of the class. (one hour)

July 2nd 1981 : BR

Using tape AA/1/1 only six out of fourteen Acts were correct at the first attempt which is very low indeed. Only one complete column (i.e. all dimensions and categories) was entirely correct in the whole session. The observer was totally frustrated. A complete check was done of each unit. Partly because of the extreme frustration felt by this
observer it was interesting that there was a strong tendency to confront rather than debate and negotiate over different points of interpretation. Finished the session by going over the H and I Dimensions.

July 2nd 1981 : RP

Went over the H and I Dimensions.
The observer made the interesting point that early in the training process he had wanted to question a lot but had decided that due to the pressure of time he would simply have to be a 'sponge' in the process and learn to reproduce accurately.
July 3rd 1981 : BR

Continued with the tape AA/1/1. Again only one column was entirely correct. The observer had not grasped that H is an alternative to C/D and so the procedure for using Dimension H was revised.

July 7th 1981 : BR

It was decided today to allow this observer to select the tape to be analysed. Tape LH/6/1 was selected and both the designer and the observer coded it independently. This appeared to be far more successful as an exercise with this observer. The fact that the observer and designer seemed to be more on a par during the coding process rather than in a teacher-learner relationship might have been significant. Changes in the Ground Rules were gone over. Two teaching Acts were wrongly selected by the observer. (One hour).

July 7th 1981 : RP

The changes in the Ground Rules to date were revised. The same section of tape used earlier with the other observer was used, i.e. LH/6/1. The designer and observer agreed on all counts bar two Mobility interpretations.

July 7th 1981 : LA
The changes in the Ground Rules to date were revised. Tape KA/7/2 was used. Some lack of practice was evident initially in machine play plus the identification of Acts but this was soon made up. There was sound reasoning behind Act identification.

July 8th 1981: LA

Continued with tape KA/7/2 with concentration initially on the Teaching Acts. The point was made that Observation direction identification is collective and the Act it accompanies changes when the direction of that same act changes. In other words if a teacher is Giving Information to a Group of pupils and then turns and walks to one pupil and Gives Information to him personally, the Teaching Act is coded again. Concurrent dialogue was assumed to mean concurrent teaching. This must be made clear in the ground Rules. The one or two errors made by this observer were accepted immediately as errors.

At this point in the Observer Training programme it was clear that very little time was left before the observers would leave the university campus for other commitments. Arrangements were made for a final set of tutorials. A test paper would be done without the designer present in the observers' own time. The Ground Rules which were the product of these training sessions were produced by the designer and
a copy each was given to the observers to replace their Training Manuals.

At the end of Session 3 the observers had been given a note about the determination of content validity. The note read:-

As we proceed with the learning of the System and finally the testing of the System, it is also important to consider whether or not the System itself has content validity i.e. is the Teacher Behaviour accurately represented by the Dimensions and Categories of the System?

Questions you may ask yourself might include:-

Would I agree that the two functions of Teacher Behaviour are Managing and Teaching?

Have I witnessed other Functions which would fall outside of those defined?

Are all the principal Acts of Managing and Teaching reported here?

Are those excluded too important to belong to a 'catch-all' category?

Is the main effect of the teacher's comments or non-verbal acts accurately portrayed in the category definitions?

Another vital question to be answered is during classification, are all observer judgements sufficiently low inference? This means is an internal frame of reference used to classify the behaviour leading to high inference or - as is intended - is an external frame used leading to low inference? This can best be judged by asking yourself how
much am I left to infer/guess/impute rather than what can I see and hear and describe without debate?
All observers are asked to bear these points in mind as we proceed with using the System.

All the observers were reminded of the points regarding the establishing of content validity prior to the last tutorial and none reported any difficulty in identifying with the dimensions and categories chosen to describe teacher behaviour in a professional context.

Tutorials

July 9th 1981 : RT

This observer approved of the self-pacing made possible by playing the tape himself, feeling much more confident when he was in charge. He was not interested in going over previous mistakes preferring instead to progress to new material. He feels a need to interpret a teacher's intonation which has to be resisted on occasions for it can lead to high inference. He was much happier with the removal of the category 'Gesticulates' and with the alteration of the definition of Demonstration. This observer tackles coding very deliberately and quite slowly and is largely accurate in his results.

July 9th 1981 : BR
The complaint was made that the Ground Rules had come too late in the exercise to be of any real use. Speed is of the essence with this observer but it is feared that there are many errors as a result. It seems doubtful at this stage that figures on a reliability test would be high enough without further training.

July 9th 1981 : LA

Generally accurate in use of the System particularly when a little more time is taken and the tape is allowed to run over the Act under observation. Appears very keen to want to do well.

July 9th 1981 : RP

This observer was also much happier at operating the tape machine himself and continues to be much happier on a one-to-one-relationship in the tutorial situation, commenting that he now felt freer to ask 'dumb questions'.

During these last tutorials the designer said very little, leaving the observers to their own devices in order to simulate the reliability test situation as closely as possible. A test tape was selected at random from among those which did not have interference or poor quality sound. A section was selected which had not been practised by the observers during the training programme. Each was left instructions
regarding the test procedures and advised to submit the completed coding sheets to the designer as soon after the test as possible. It was not possible to test the observers simultaneously due to their various commitments.

Unfortunately at this late stage one of the observers had to return to the United States unexpectedly and was unable to complete the observer reliability test. Two other observers were not able to attempt the coding until long after their last tutorial and practice. It was quite evident at this late stage that there was not going to be a result from the testing of these observers and that without reliability test results the system as it stood could not be considered available for use. It was decided that BOS 1 had to be developed further along the lines envisaged originally in order that it was available to a wider audience. This should mean that it would be considerably easier to learn and there would be less difficulty in persuading colleagues to use it and test it for reliability. Having taken this decision a final Observer Training Report was written and is presented in Appendix 3.

Preparations were begun for the development of the system which was to be labelled BOS 2.
APPENDIX 3

Observer Training Report

The training of observers is an essential part of instrument design and development for the process includes procedures to validate and test the reliability of the system. For this reason the training was monitored carefully and reported fully. Lessons learned from the exercise are summarised below.

Careful selection of the observers is important. Whilst the level of professional expertise is a consideration, so must be the commitment each can promise and maintain throughout the training period. Failure to keep to the training schedule results in loss of accuracy in system use as points have to be gone over repeatedly. Extension of the programme will also result in failure to keep to original deadlines and this can have an adverse effect on commitment. Those involved in the programme also include technicians required for the setting up of equipment and they need to be aware of the whole programme from the beginning.

The nature of the work needs accurate communication at the recruitment stage. There is little doubt that it can be tedious. Busy people have little time for the challenge of tedium especially when it necessitates spending many hours working voluntarily for others.
The value of the tutorial arrangements far outweighed the group training sessions. After an initial group introduction to the system training, would best be conducted individually. For use of this system, each observer needs to be able to operate the tape deck himself at his own pace.

The training programme would have benefitted from a pilot programme being conducted with one observer beforehand. This would have revealed many of the rather simple errors and established procedures before the training programme began, thereby saving both time and confusion. The final set of Ground Rules would have been produced earlier which would have helped the observers. Additionally it would have been better, had it been possible, to have held the training programme earlier in the academic year to avoid finishing late and losing staff over the long vacation.

Perhaps the most serious criticism of the training programme concerns that to do with the system itself. Although its use provides the observer with a full account of the teacher's behaviour it is at considerable cost to the observer. It takes at least twelve hours to code a one hour lesson. Not only does this require patience and concentration but is also ties the use of equipment for a considerable period of time. One must ask if the time is well spent as well as who is going to be able to spend that amount of time coding lesson tapes. If developed, the system could be used by researchers but would be unsuitable for live coding and probably, therefore, for teachers and
students. In its current form the system requires too long a time to achieve adequate reliability standards and far too great a commitment from the observers.

It was resolved therefore to develop the system in a way which would make it more available to others in the education service. This would mean making it simpler to handle, quicker to learn and able to be used if required for live coding. To some extent this could be achieved without too great a loss of discrimination by collapsing the categories and removing those dimensions considered to be of less significance to the observer.

Details of the development of the system are given in chapter 4. To avoid confusion, the original system was labelled BOS 1 (Bailey Observation System 1) and the development was called BOS 2 (Bailey Observation System 2).
APPENDIX 4  BOS 2 Computer Program

10 N1=0
20 ON ERROR GOTO 1700
30 MODE 7
40 *FX 6,0
50 VDU 2
60 VDU 1,27
70 VDU 1,67
80 VDU 1,70
82 VDU 1,27
83 VDU 1,78
84 VDU 1,12
90 VDU 3
100 REM P.E. PROG by M. DALY 29/11/82
105 REM MODIFIED BY LIZ HEIGHWAY 4/3/85 & AGAIN 11/3/85 TO INCLUDE TOTAL SAMPLES TAKEN AND NUMBERS OF OCCURRENCES IN EACH CATEGORY OF OBSERVATION.
110 PROC INIT: REM INITIALISE
120 PROC INFO:CLS: REM USE OLD OR NEW INFO
130 IF Y$="E" THEN PROC E INFO:CLS: REM ENTER NEW INFO
140 IF Y$="P" THEN PROC G IF FILE: REM GET INFO FROM FILE
150 PROC PRINT INFO?
160 PROC S INFO: REM SAVE INFO?
170 IF 1$="1" THEN PROC S A INFO: REM SAVE INFO
180 PRINT "END OF PROGRAM BYE!": END
190 REM *****************************************************
200 DEF PROC INIT: REM INITIALISE
210 DIM A(500,3): DIM C(19)
220 END PROC
230 REM *****************************************************
240 DEF PROC INFO: REM USE OLD OR NEW INFO?
250 PRINT SPC(12); "P.E. DEPARTMENT"
260 PRINT SPC(11); "TEACHER BEHAVIOR"
270 PRINT "IF IOU WANT TO:"
280 PRINT SPC(11); "PRINT STORE DATA"; SPC(5); "PRESS 'P'"
290 PRINT "; SPC(6); "ENTER NEW DATA"; SPC(8); "PRESS 'E'"
300 Y$=GET$: IF Y$<"P" AND Y$<"E" THEN PRINT CHR$(7): GOTO 300
310 END PROC
320 REM *****************************************************
330 DEF PROC S INFO: REM PRINT INFORMATION?
340 DIM C(19)
350 PRINT TAB(5,2) "IF IOU WANT TO"
360 PRINT "PRINT OUT THE HISTOGRAM PRESS 1"
370 PRINT "PRINT OUT PATTERNS PRESS 2"
380 PRINT "PRINT OUT BOTH PRESS 3"
390 PRINT "NOT PRINT THE RESULTS PRESS 4"
400 Y$=GET$: IF Y$<"1" AND Y$<"2" AND Y$<"3" AND Y$<"4" THEN PRINT CHR$(7): GOTO 400
410 IF Y$="4" THEN 480
420 IF Y$="1" THEN 450
430 PRINT"ENTER TITLE FOR PATTERNS:";INPUT T1$
440 PROCPRINTINFO
450 IF Y$="2" THEN 480
460 PRINT"ENTER TITLE FOR HISTOGRAM:";INPUT T2$
470 PROCPHIS
480 ENDPROC
490 DEFPROCINFO:CLS:REM SAVE INFORMATION ?
500 PRINTTAB(5,2)"DO YOU WANT TO "
510 PRINT"SAVE THE RESULTS (Y/N)"
520 Y$=GET$:IF Y$<>'Y'ANDY$<>'N' THEN 520
530 ENDPROC
540 REM ****************************
550 DEF PROCENINFO:REM ENTER NEW INFO
560 PROCSDISPLAY:REM SET UP DISPLAY
570 AB=1
580 PRINTTAB(2,23)"NO. ";AB;:INPUT A(AB,1),A(AB,2),A(AB,3)
590 IF A(AB,1)=0 THEN N1=AB-1:GOTO 710
600 IF A(AB,1)=-1 THEN AB=AB-2:GOTO 690
610 IF A(AB,1)>1 OR A(AB,1)<0 THEN AB=AB-1:GOTO 690
620 IF A(AB,2)>1 OR A(AB,2)<0 THEN AB=AB-1:GOTO 690
630 IF A(AB,3)>1 OR A(AB,3)<0 THEN AB=AB-1:GOTO 690
640 C(A(AB,1))=C(A(AB,1))+1
650 C(A(AB,2))=C(A(AB,2))+1
660 C(A(AB,3))=C(A(AB,3))+1
670 PRINTTAB(2,22)"
680 PRINTTAB(2,22)"NO. ";AB;"A(AB,1);A(AB,2);A(AB,3)
690 PRINTTAB(2,23)"
700 AB=AB+1:GOTO 580
710 ENDPROC
720 REM ****************************
730 DEF PROCSDISPLAY:REM SET UP DISPLAY
740 PRINT"";
750 PRINT"A. MANAGES";SPC(5);"1. PEOPLE"
760 PRINTSPC(15);"2. EQUIPMENT"
770 PRINTSPC(15);"3. TASK"
780 PRINTSPC(15);"4. POLICY"
790 PRINT"B. TEACHES";SPC(5);"5. PRESCRIBES RESPONSE"
800 PRINTSPC(15);"6. GIVES FEEDBACK"
810 PRINTSPC(15);"7. OBSERVES"
820 PRINTSPC(15);"8. GIVES INFORMATION"
830 PRINTSPC(15);"9. SEEKS INFORMATION"
840 PRINT"C. ";SPC(12);"10. OTHER INTERACTION"
850 PRINTSPC(15);"11. NON INTERACTION"
860 PRINTSPC(15);"12. INADEQUATE RECORD"
870 PRINT"D. DIRECTION";SPC(2);"13. ONE"
880 PRINTSPC(15);"14. GROUP"
890 PRINTSPC(14);"15. CLASS"
900 PRINT"E. MODE";SPC(7);"16. AUDIBLE"
910 PRINTSPC(14);"17. SILENT"
920 PRINTSPC(14);"18. PHYSICAL"
930 PRINTSPC(14);"19. AUDIBLE PHYSICAL"
940 PRINT"(0,0,0 TO END. -1,-1,-1 TO AMEND"
950 ENDPROC
960 REM ****************************
970 DEF PROC GETFILE:CLS:REM GET INFO FROM FILE
980 PRINTTAB(4,2)"MAKE SURE THE DISC CONTAINING"
990 PRINT" THE FILE YOU WANT"
1000 PRINT" IS IN THE DISC DRIVE"
1010 INPUT TAB(1,15);"ENTER NAME OF FILE:"F$"
:GOTO 1010
1020 IF LEN(F$)>7 THEN PRINTCHR$(7);TAB(1,15);" :GOTO 1010
1030 X=OPENUP(F$);*OPT1,2
1040 IF X=O THEN PRINTTAB(5,15);"FILE NOT FOUND";CHR$(7):S=INKEY(200):PRINTTAB(5,15);"
:GOTO 1010
1050 AB=1
1060 REPEAT: INPUT$X,A(AB,1),A(AB,2),A(AB,3)
1070 PRINTA(AB,1),A(AB,2),A(AB,3)
1080 C(A(AB,1))=C(A(AB,1))+1
1090 C(A(AB,2))=C(A(AB,2))+1
1100 C(A(AB,3))=C(A(AB,3))+1
1110 AB=AB+1
1120 UNTIL EOF$X
1130 N1=AB-1
1140 CLOSE$X:S=INKEY(200):ENDPROC
1150 REM ****************************
1160 DEF PROC PRINTINFO:REM PRINT INFORMATION
1170 VDU2:PRINT SPC(30);"P.E. DEPARTMENT TEACHER BEHAVIOUR"
1180 PRINT$S
1190 PRINT:PRINT
1200 PRINT" 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15";
1210 PRINT" 16 17 18 19"
1220 FOR AB=1 TO N1
1230 IF A(AB,1)>A(AB,2) THEN C=A(AB,1):A(AB,1)=A(AB,2):A(AB,2)=C
1240 IF A(AB,2)>A(AB,3) THEN C=A(AB,2):A(AB,2)=A(AB,3):A(AB,3)=C:GOTO 1230
1250 IF A(AB,2)=0 THEN 1280
1260 PRINTSPC((A(AB,1)*4)-1);"*";SPC(((A(AB,2)-A(AB,1))*4)-1);"*"
1270 PRINTSPC(((A(AB,3)-A(AB,2))*4)-1);"*"
1280 NEXT
1285 PRINT"""NUMBER OF SAMPLES TAKEN : ";N1
APPENDIX 4  Bos 2 Computer Program

> 1290 PRINT:PRINT"------------------------------------------
1300 PRINT"------------------------------------------"
1310 VDU 3:ENDPROC
1320 DEF PROC PHIS:REM PRINT HISTOGRAM
1330 VDU2
1340 PRINT:PRINTT2$:PRINT
1350 PRINT:PRINT"HISTOGRAM"
1360 FOR X=100 TO 5 STEP -5
1370 IF X=100 THEN PRINT;X;" I";SPC(1);GOTO 1400
1380 IF X=5 THEN PRINT;X;" I";SPC(1);GOTO 1400
1390 PRINT;X;" I";SPC(1);
1400 FOR I=1 TO 19
1410 C6=C(I)/N1*100
1420 IF C6>X THEN PRINT"* ";:GOTO 1440
1430 PRINT"
1440 NEXT
1450 NEXT
1460 PRINT"O -------------------------------------------"
1470 PRINT"----------------------------------"
1480 PRINT"2 3 4 5 6 7 8 9 10 11 12 13 14 15"
1490 PRINT"16 17 18 19"
1500 PRINT:PRINT"SUMMARY OF HISTOGRAM"
1510 REM ********************************
1520 FOR I=1 TO 19
1530 PRINT I;SPC(10);C(I)/N1*100;:PRINTTAB(32)C(I);" OCCURRENCE"
1540 NEXT:PRINT' '''NUMBER OF SAMPLES TAKEN: ";N1'":VDU3:END PROC
1560 DEF PROC SAINFO:REM SAVE INFO ON FILE
1570 CLS
1580 PRINT:PRINT"FILE WILL BE SAVED ON DISC"
1590 PRINT"IN DISC DRIVE"
1600 INPUT TAB(14,4)"ENTER NAME OF FILE:"F$
1610 IF LEN(F$)>7 THEN PRINTCHR$(7);TAB(14,4);"
1620 X=OPENOUT(F$)
1630 FOR AB=1 TO N1
1640 PRINTX,A(AB,1)
1650 PRINTX,A(AB,2)
1660 PRINTX,A(AB,3)
1670 NEXT
1680 CLOSEX:ENDPROC
1690 REM ********************************
1700 CLOSEX:IF ERR=15 THEN PRINT "FILE MUST BE DATA FILE";
1710 PRINT ERR "AT LINE "ERL
1720 STOP
APPENDIX 5

Computer Print-out

SPM/1

<table>
<thead>
<tr>
<th>P</th>
<th>HISTOGRAM</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 I</td>
<td></td>
</tr>
<tr>
<td>95 I</td>
<td></td>
</tr>
<tr>
<td>90 I</td>
<td></td>
</tr>
<tr>
<td>85 I</td>
<td></td>
</tr>
<tr>
<td>80 I</td>
<td></td>
</tr>
<tr>
<td>75 I</td>
<td></td>
</tr>
<tr>
<td>70 I</td>
<td></td>
</tr>
<tr>
<td>65 I</td>
<td></td>
</tr>
<tr>
<td>60 I</td>
<td></td>
</tr>
<tr>
<td>55 I</td>
<td></td>
</tr>
<tr>
<td>50 I</td>
<td></td>
</tr>
<tr>
<td>45 I</td>
<td></td>
</tr>
<tr>
<td>40 I</td>
<td></td>
</tr>
<tr>
<td>35 I</td>
<td></td>
</tr>
<tr>
<td>30 I</td>
<td></td>
</tr>
<tr>
<td>25 I</td>
<td></td>
</tr>
<tr>
<td>20 I</td>
<td></td>
</tr>
<tr>
<td>15 I</td>
<td></td>
</tr>
<tr>
<td>10 I</td>
<td></td>
</tr>
<tr>
<td>5 I</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
<table>
<thead>
<tr>
<th>CATEGORY</th>
<th>%AGE</th>
<th>OCCURRENCES</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>17.2588833</td>
<td>34 OCCURRENCES</td>
</tr>
<tr>
<td>2</td>
<td>2.03045685</td>
<td>4 OCCURRENCES</td>
</tr>
<tr>
<td>3</td>
<td>7.6142132</td>
<td>15 OCCURRENCES</td>
</tr>
<tr>
<td>4</td>
<td>0.507614213</td>
<td>1 OCCURRENCES</td>
</tr>
<tr>
<td>5</td>
<td>27.4111675</td>
<td>54 OCCURRENCES</td>
</tr>
<tr>
<td>6</td>
<td>6.59898477</td>
<td>13 OCCURRENCES</td>
</tr>
<tr>
<td>7</td>
<td>22.3350254</td>
<td>44 OCCURRENCES</td>
</tr>
<tr>
<td>8</td>
<td>10.1522843</td>
<td>20 OCCURRENCES</td>
</tr>
<tr>
<td>9</td>
<td>2.03045685</td>
<td>4 OCCURRENCES</td>
</tr>
<tr>
<td>10</td>
<td>2.53807107</td>
<td>5 OCCURRENCES</td>
</tr>
<tr>
<td>11</td>
<td>1.52284264</td>
<td>3 OCCURRENCES</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>0 OCCURRENCES</td>
</tr>
<tr>
<td>13</td>
<td>15.7360406</td>
<td>31 OCCURRENCES</td>
</tr>
<tr>
<td>14</td>
<td>36.5482233</td>
<td>72 OCCURRENCES</td>
</tr>
<tr>
<td>15</td>
<td>46.1928934</td>
<td>91 OCCURRENCES</td>
</tr>
<tr>
<td>16</td>
<td>55.8375635</td>
<td>110 OCCURRENCES</td>
</tr>
<tr>
<td>17</td>
<td>23.857868</td>
<td>47 OCCURRENCES</td>
</tr>
<tr>
<td>18</td>
<td>0</td>
<td>0 OCCURRENCES</td>
</tr>
<tr>
<td>19</td>
<td>18.7817259</td>
<td>37 OCCURRENCES</td>
</tr>
</tbody>
</table>

NUMBER OF SAMPLES TAKEN: 197
Categories:

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>14</td>
<td>16</td>
</tr>
<tr>
<td>1</td>
<td>15</td>
<td>17</td>
</tr>
<tr>
<td>1</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>5</td>
<td>14</td>
<td>16</td>
</tr>
<tr>
<td>5</td>
<td>14</td>
<td>16</td>
</tr>
<tr>
<td>3</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>7</td>
<td>14</td>
<td>17</td>
</tr>
<tr>
<td>7</td>
<td>14</td>
<td>17</td>
</tr>
<tr>
<td>7</td>
<td>14</td>
<td>17</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>14</td>
<td>17</td>
</tr>
<tr>
<td>6</td>
<td>13</td>
<td>16</td>
</tr>
<tr>
<td>7</td>
<td>14</td>
<td>17</td>
</tr>
<tr>
<td>7</td>
<td>15</td>
<td>17</td>
</tr>
<tr>
<td>7</td>
<td>14</td>
<td>17</td>
</tr>
<tr>
<td>7</td>
<td>14</td>
<td>17</td>
</tr>
<tr>
<td>8</td>
<td>13</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>2</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>5</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>5</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>7</td>
<td>14</td>
<td>17</td>
</tr>
<tr>
<td>8</td>
<td>13</td>
<td>19</td>
</tr>
<tr>
<td>7</td>
<td>14</td>
<td>17</td>
</tr>
<tr>
<td>3</td>
<td>14</td>
<td>16</td>
</tr>
<tr>
<td>7</td>
<td>14</td>
<td>17</td>
</tr>
<tr>
<td>7</td>
<td>14</td>
<td>17</td>
</tr>
<tr>
<td>7</td>
<td>14</td>
<td>17</td>
</tr>
<tr>
<td>7</td>
<td>14</td>
<td>17</td>
</tr>
<tr>
<td>8</td>
<td>15</td>
<td>19</td>
</tr>
<tr>
<td>8</td>
<td>15</td>
<td>19</td>
</tr>
<tr>
<td>5</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>5</td>
<td>13</td>
<td>19</td>
</tr>
<tr>
<td>1</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>5</td>
<td>15</td>
<td>19</td>
</tr>
<tr>
<td>5</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>16</td>
</tr>
<tr>
<td>---</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>4</td>
<td>13</td>
<td>16</td>
</tr>
<tr>
<td>5</td>
<td>15</td>
<td>19</td>
</tr>
<tr>
<td>1</td>
<td>14</td>
<td>16</td>
</tr>
<tr>
<td>6</td>
<td>13</td>
<td>16</td>
</tr>
<tr>
<td>5</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>7</td>
<td>14</td>
<td>17</td>
</tr>
<tr>
<td>7</td>
<td>14</td>
<td>17</td>
</tr>
<tr>
<td>6</td>
<td>13</td>
<td>19</td>
</tr>
<tr>
<td>8</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>7</td>
<td>14</td>
<td>17</td>
</tr>
<tr>
<td>6</td>
<td>13</td>
<td>16</td>
</tr>
<tr>
<td>3</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>8</td>
<td>15</td>
<td>19</td>
</tr>
<tr>
<td>5</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>8</td>
<td>13</td>
<td>19</td>
</tr>
<tr>
<td>1</td>
<td>13</td>
<td>16</td>
</tr>
<tr>
<td>8</td>
<td>15</td>
<td>19</td>
</tr>
<tr>
<td>6</td>
<td>15</td>
<td>16</td>
</tr>
</tbody>
</table>
8   14   19
7   14   17
9   14   16
9   14   16
5   14   16
5   13   19
6   13   16
6   13   16
1   15   16
5   15   16
1   13   16
8   15   16
5   15   16
5   14   16
10  13   16
5   15   19
5   15   19
5   15   16
<table>
<thead>
<tr>
<th></th>
<th>5</th>
<th>15</th>
<th>19</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>13</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>15</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>15</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>15</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>15</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>15</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>15</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>15</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>15</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>14</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>15</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>13</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>14</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>14</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>15</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>14</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>14</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>----</td>
<td>----</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>15</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>15</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>15</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>13</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>15</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>15</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>15</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>14</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>13</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>15</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>14</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>14</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>14</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>14</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>13</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>15</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>14</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>----</td>
<td>----</td>
<td>---</td>
</tr>
<tr>
<td>7</td>
<td>14</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>14</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>15</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>14</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>15</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>15</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>15</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>14</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>14</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>13</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>14</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>15</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>14</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>15</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>14</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>15</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>15</td>
<td>16</td>
<td></td>
</tr>
</tbody>
</table>
## APPENDIX 6

**TEST TAPE PL2/1 (1)**

<table>
<thead>
<tr>
<th>OB1</th>
<th>OB2</th>
<th>OB1</th>
<th>OB2</th>
<th>OB1</th>
<th>OB2</th>
<th>OB1</th>
<th>OB2</th>
<th>OB1</th>
<th>OB2</th>
<th>OB1</th>
<th>OB2</th>
</tr>
</thead>
<tbody>
<tr>
<td>T/M</td>
<td>T/M</td>
<td>D</td>
<td>D</td>
<td>M</td>
<td>M</td>
<td>T/M</td>
<td>T/M</td>
<td>D</td>
<td>D</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>16</td>
<td>1</td>
<td>1</td>
<td>15</td>
<td>15</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>13</td>
<td>13</td>
<td>17</td>
<td>17</td>
<td>5</td>
<td>5</td>
<td>15</td>
<td>15</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>13</td>
<td>13</td>
<td>18</td>
<td>18</td>
<td>5</td>
<td>5</td>
<td>15</td>
<td>15</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>13</td>
<td>13</td>
<td>18</td>
<td>18</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>15</td>
<td>15</td>
<td>17</td>
<td>17</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>15</td>
<td>15</td>
<td>16</td>
<td>16</td>
<td>1</td>
<td>1</td>
<td>14</td>
<td>14</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>15</td>
<td>15</td>
<td>16</td>
<td>16</td>
<td>1</td>
<td>1</td>
<td>14</td>
<td>14</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>15</td>
<td>15</td>
<td>19</td>
<td>19</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>15</td>
<td>15</td>
<td>19</td>
<td>19</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>15</td>
<td>15</td>
<td>19</td>
<td>19</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>15</td>
<td>15</td>
<td>19</td>
<td>19</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>15</td>
<td>15</td>
<td>19</td>
<td>19</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>16</td>
<td>2</td>
<td>2</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>16</td>
<td>2</td>
<td>2</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>15</td>
<td>15</td>
<td>16</td>
<td>16</td>
<td>5</td>
<td>5</td>
<td>15</td>
<td>15</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>15</td>
<td>15</td>
<td>16</td>
<td>16</td>
<td>8</td>
<td>8</td>
<td>15</td>
<td>15</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>16</td>
<td>5</td>
<td>5</td>
<td>15</td>
<td>15</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>16</td>
<td>5</td>
<td>5</td>
<td>15</td>
<td>15</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>16</td>
<td>1</td>
<td>1</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>16</td>
<td>6</td>
<td>6</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>16</td>
<td>6</td>
<td>6</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>16</td>
<td>6</td>
<td>6</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>16</td>
<td>5</td>
<td>5</td>
<td>15</td>
<td>15</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>16</td>
<td>5</td>
<td>5</td>
<td>15</td>
<td>15</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>15</td>
<td>15</td>
<td>16</td>
<td>16</td>
<td>6</td>
<td>6</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>15</td>
<td>15</td>
<td>16</td>
<td>16</td>
<td>5</td>
<td>5</td>
<td>15</td>
<td>15</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>16</td>
<td>5</td>
<td>5</td>
<td>15</td>
<td>15</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>5</td>
<td>8</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>16</td>
<td>5</td>
<td>5</td>
<td>13</td>
<td>13</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>OB1</td>
<td>OB2</td>
<td>OB1</td>
<td>OB2</td>
<td>OB1</td>
<td>OB2</td>
<td>OB1</td>
<td>OB2</td>
<td>OB1</td>
<td>OB2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T/M</td>
<td>T/M</td>
<td>T/M</td>
<td>T/M</td>
<td>T/M</td>
<td>T/M</td>
<td>T/M</td>
<td>T/M</td>
<td>T/M</td>
<td>T/M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>15</td>
<td>15</td>
<td>17</td>
<td>17</td>
<td>10</td>
<td>10</td>
<td>13</td>
<td>13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>16</td>
<td>10</td>
<td>10</td>
<td>13</td>
<td>13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>16</td>
<td>6</td>
<td>6</td>
<td>15</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>16</td>
<td>5</td>
<td>5</td>
<td>15</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>16</td>
<td>1</td>
<td>1</td>
<td>15</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>15</td>
<td>15</td>
<td>17</td>
<td>17</td>
<td>1</td>
<td>1</td>
<td>14</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>15</td>
<td>15</td>
<td>16</td>
<td>16</td>
<td>1</td>
<td>1</td>
<td>14</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td>17</td>
<td>17</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>15</td>
<td>15</td>
<td>16</td>
<td>16</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td>17</td>
<td>17</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>14</td>
<td>14</td>
<td>17</td>
<td>17</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>14</td>
<td>13</td>
<td>16</td>
<td>16</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>15</td>
<td>15</td>
<td>16</td>
<td>16</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>16</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>16</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td>17</td>
<td>17</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>14</td>
<td>13</td>
<td>16</td>
<td>16</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>14</td>
<td>14</td>
<td>17</td>
<td>17</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>14</td>
<td>13</td>
<td>17</td>
<td>17</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td>17</td>
<td>17</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td>17</td>
<td>17</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>16</td>
<td>5</td>
<td>5</td>
<td>13</td>
<td>13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>16</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>14</td>
<td>13</td>
<td>16</td>
<td>16</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td>17</td>
<td>17</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>14</td>
<td>13</td>
<td>16</td>
<td>16</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>16</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>16</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>16</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OB1</td>
<td>OB2</td>
<td>OB1</td>
<td>OB2</td>
<td>OB1</td>
<td>OB2</td>
<td>OB1</td>
<td>OB2</td>
<td>OB1</td>
<td>OB2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T/M</td>
<td>T/M</td>
<td>D</td>
<td>D</td>
<td>M</td>
<td>M</td>
<td>T/M</td>
<td>T/M</td>
<td>D</td>
<td>D</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>14</td>
<td>13</td>
<td>16</td>
<td>16</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>15</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>15</td>
<td>15</td>
<td>17</td>
<td>17</td>
<td>6</td>
<td>6</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>14</td>
<td>13</td>
<td>16</td>
<td>16</td>
<td>9</td>
<td>9</td>
<td>14</td>
<td>14</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>17</td>
<td>17</td>
<td>3</td>
<td>3</td>
<td>14</td>
<td>14</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>17</td>
<td>17</td>
<td>10</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>17</td>
<td>17</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td>17</td>
<td>17</td>
<td>16</td>
<td>17</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>17</td>
<td>17</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td>17</td>
<td>17</td>
<td>16</td>
<td>17</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>17</td>
<td>17</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td>17</td>
<td>17</td>
<td>16</td>
<td>17</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>14</td>
<td>14</td>
<td>16</td>
<td>16</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>16</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>15</td>
<td>15</td>
<td>17</td>
<td>17</td>
<td>5</td>
<td>5</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>14</td>
<td>14</td>
<td>16</td>
<td>16</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>15</td>
<td>15</td>
<td>17</td>
<td>17</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td>17</td>
<td>17</td>
<td>9</td>
<td>9</td>
<td>14</td>
<td>14</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td>17</td>
<td>17</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td>17</td>
<td>17</td>
<td>6</td>
<td>6</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>16</td>
<td>1</td>
<td>1</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td>17</td>
<td>17</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>14</td>
<td>13</td>
<td>16</td>
<td>16</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>16</td>
<td>1</td>
<td>1</td>
<td>14</td>
<td>14</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>14</td>
<td>14</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td>17</td>
<td>17</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>14</td>
<td>14</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td>17</td>
<td>17</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td>2</td>
<td>2</td>
<td>15</td>
<td>15</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
<td>1</td>
<td>1</td>
<td>15</td>
<td>15</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
</tr>
</tbody>
</table>
APPENDIX 7  Computer Program used to scale down samples

5 *FX,60
10 INPUT "INPUT MIN. SIZE OF SAMPLE "M
20 INPUT "INPUT WHICH TEACHER "WT$
30 INPUT "INPUT NO. OF CATEGORIES "NC
40 DIM OC(NC)
50 INPUT "INPUT SAMPLE SIZE "SS
60 FOR I=1 TO NC:PRINT "CATEGORY";I
70 INPUT "NO. OF OCCURRENCES "OC(I)
80 OC(I)=OC(I)*M/SS:NEXT I
84 VDU 2
90 PRINT "CATEGORY TOTALS TO SMALLEST SAMPLE"
100 PRINT:PRINT "TEACHER";WT$
110 FOR I=1 TO NC:PRINT "CATEGORY";I;:PRINT ""
";OC(I);"OCCURRENCES":NEXT I
125 END
## APPENDIX 8

### TABLE 10

#### MALE CATEGORIES

**TOTAL SAMPLES = 6 x 122 = 732**

**OVERALL; M'MENT; TEACHING**

<table>
<thead>
<tr>
<th>CAT</th>
<th>SUM X</th>
<th>X</th>
<th>On-1</th>
<th>%</th>
<th>SUM X</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>87.73</td>
<td>14.62</td>
<td>5.02</td>
<td>11.98</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>63.45</td>
<td>10.58</td>
<td>8.69</td>
<td>8.67</td>
<td>191.04</td>
<td>26.1</td>
</tr>
<tr>
<td>3</td>
<td>34.38</td>
<td>5.73</td>
<td>3.06</td>
<td>4.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>5.48</td>
<td>0.91</td>
<td>0.65</td>
<td>0.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>179.7</td>
<td>29.95</td>
<td>5.77</td>
<td>24.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>57.65</td>
<td>9.61</td>
<td>2.98</td>
<td>7.88</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>111.74</td>
<td>18.62</td>
<td>4.64</td>
<td>15.26</td>
<td>489.74</td>
<td>66.9</td>
</tr>
<tr>
<td>8</td>
<td>101.3</td>
<td>16.88</td>
<td>6.07</td>
<td>13.84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>39.35</td>
<td>6.56</td>
<td>3.31</td>
<td>5.38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>18.07</td>
<td>3.01</td>
<td>2.37</td>
<td>2.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>16.81</td>
<td>2.8</td>
<td>1.93</td>
<td>2.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>16.01</td>
<td>2.67</td>
<td>3.25</td>
<td>2.19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>166.59</td>
<td>27.77</td>
<td>13.22</td>
<td>22.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>240.1</td>
<td>40.02</td>
<td>18.82</td>
<td>32.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>292.49</td>
<td>48.75</td>
<td>20</td>
<td>39.96</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>493.28</td>
<td>82.21</td>
<td>9.19</td>
<td>67.39</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>131.29</td>
<td>21.88</td>
<td>4.8</td>
<td>17.94</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>3.94</td>
<td>0.66</td>
<td>0.68</td>
<td>0.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>87.58</td>
<td>14.6</td>
<td>5.17</td>
<td>11.97</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### TABLE 11

**FEMALE CATEGORIES**

**TOTAL SAMPLES = 5 x 122 = 610**

**OVERALL; M'MENT; TEACHING**

<table>
<thead>
<tr>
<th>CAT</th>
<th>SUM X</th>
<th>X</th>
<th>On-1</th>
<th>%</th>
<th>SUM X</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>92.78</td>
<td>18.56</td>
<td>3.13</td>
<td>15.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>37.79</td>
<td>7.56</td>
<td>6.39</td>
<td>6.2</td>
<td>167.86</td>
<td>27.52</td>
</tr>
<tr>
<td>3</td>
<td>30.55</td>
<td>6.11</td>
<td>3.08</td>
<td>5.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>6.74</td>
<td>1.35</td>
<td>0.94</td>
<td>1.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>145.68</td>
<td>29.14</td>
<td>3.97</td>
<td>23.88</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>53.71</td>
<td>10.74</td>
<td>2.84</td>
<td>8.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>100.76</td>
<td>20.15</td>
<td>16.97</td>
<td>16.52</td>
<td>410.9</td>
<td>67.36</td>
</tr>
<tr>
<td>8</td>
<td>67.14</td>
<td>13.43</td>
<td>5.56</td>
<td>11.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>43.61</td>
<td>8.72</td>
<td>5.99</td>
<td>7.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>19.93</td>
<td>3.99</td>
<td>1.76</td>
<td>3.27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>5.29</td>
<td>1.06</td>
<td>0.54</td>
<td>0.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>6.01</td>
<td>1.2</td>
<td>1.7</td>
<td>0.98</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>158.08</td>
<td>31.62</td>
<td>8.24</td>
<td>25.91</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>199.75</td>
<td>39.95</td>
<td>9.61</td>
<td>32.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>241.17</td>
<td>48.23</td>
<td>12.77</td>
<td>39.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>416.79</td>
<td>83.36</td>
<td>18.25</td>
<td>68.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>119.29</td>
<td>23.86</td>
<td>17.03</td>
<td>19.56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>3.48</td>
<td>0.7</td>
<td>1.1</td>
<td>0.57</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>58.56</td>
<td>11.71</td>
<td>6.36</td>
<td>9.6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX 8

---

TABLE 12
---

MALE CATEGORIES
---

TOTAL SAMPLES = 4 x 122 = 488
---

GYMNASTICS
---

<table>
<thead>
<tr>
<th>CAT</th>
<th>SUM X</th>
<th>X</th>
<th>On-1</th>
<th>%</th>
<th>SUM X</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>47.65</td>
<td>11.91</td>
<td>5.93</td>
<td>9.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>86.43</td>
<td>21.61</td>
<td>16.27</td>
<td>17.71</td>
<td>155.1</td>
<td>31.78</td>
</tr>
<tr>
<td>3</td>
<td>17.12</td>
<td>4.28</td>
<td>3.29</td>
<td>3.51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>3.9</td>
<td>0.98</td>
<td>0.68</td>
<td>0.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>112.56</td>
<td>28.14</td>
<td>10.05</td>
<td>23.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>34.29</td>
<td>8.57</td>
<td>4.88</td>
<td>7.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>77.39</td>
<td>19.35</td>
<td>4.12</td>
<td>15.86</td>
<td>305.84</td>
<td>62.67</td>
</tr>
<tr>
<td>8</td>
<td>58.32</td>
<td>14.58</td>
<td>6.38</td>
<td>11.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>23.28</td>
<td>5.82</td>
<td>2.77</td>
<td>4.77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>5.66</td>
<td>1.41</td>
<td>1.66</td>
<td>1.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>12.36</td>
<td>3.09</td>
<td>4.33</td>
<td>2.53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>8.49</td>
<td>2.12</td>
<td>2.99</td>
<td>1.74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>98.32</td>
<td>24.58</td>
<td>14.25</td>
<td>20.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>122.64</td>
<td>40.88</td>
<td>9.74</td>
<td>25.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>210.43</td>
<td>52.61</td>
<td>19.44</td>
<td>43.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>318.5</td>
<td>79.63</td>
<td>9.73</td>
<td>65.27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>93.52</td>
<td>23.38</td>
<td>3.09</td>
<td>19.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>5.03</td>
<td>1.26</td>
<td>1.46</td>
<td>1.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>51.43</td>
<td>12.86</td>
<td>4.51</td>
<td>10.54</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### TABLE 13

**MALE CATEGORIES**

TOTAL SAMPLES = 5 x 122 = 610

GAMES LESSONS

<table>
<thead>
<tr>
<th>CAT</th>
<th>SUM X</th>
<th>X</th>
<th>On-1</th>
<th>%</th>
<th>SUM X</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>72.38</td>
<td>14.48</td>
<td>5.49</td>
<td>11.87</td>
<td>123.97</td>
<td>20.32</td>
</tr>
<tr>
<td>2</td>
<td>21.92</td>
<td>4.38</td>
<td>0.82</td>
<td>3.59</td>
<td>132.54</td>
<td>21.36</td>
</tr>
<tr>
<td>3</td>
<td>26.84</td>
<td>5.37</td>
<td>2.91</td>
<td>4.82</td>
<td>128.76</td>
<td>21.50</td>
</tr>
<tr>
<td>4</td>
<td>2.83</td>
<td>0.57</td>
<td>0.86</td>
<td>0.46</td>
<td>2.87</td>
<td>0.47</td>
</tr>
<tr>
<td>5</td>
<td>148.72</td>
<td>29.74</td>
<td>2.53</td>
<td>24.38</td>
<td>149.39</td>
<td>24.31</td>
</tr>
<tr>
<td>6</td>
<td>47.45</td>
<td>9.49</td>
<td>3.72</td>
<td>7.78</td>
<td>47.42</td>
<td>7.77</td>
</tr>
<tr>
<td>7</td>
<td>87.68</td>
<td>17.54</td>
<td>5.03</td>
<td>14.37</td>
<td>429.34</td>
<td>70.38</td>
</tr>
<tr>
<td>8</td>
<td>103.15</td>
<td>20.63</td>
<td>6.64</td>
<td>16.91</td>
<td>103.79</td>
<td>16.92</td>
</tr>
<tr>
<td>9</td>
<td>42.34</td>
<td>8.47</td>
<td>6.09</td>
<td>6.94</td>
<td>42.42</td>
<td>6.93</td>
</tr>
<tr>
<td>10</td>
<td>17.39</td>
<td>3.48</td>
<td>2.21</td>
<td>2.85</td>
<td>17.43</td>
<td>2.85</td>
</tr>
<tr>
<td>11</td>
<td>14.32</td>
<td>2.86</td>
<td>2.38</td>
<td>2.35</td>
<td>14.37</td>
<td>2.36</td>
</tr>
<tr>
<td>12</td>
<td>24.97</td>
<td>4.99</td>
<td>6.02</td>
<td>4.09</td>
<td>25.06</td>
<td>4.09</td>
</tr>
<tr>
<td>13</td>
<td>136.36</td>
<td>27.27</td>
<td>13.74</td>
<td>22.35</td>
<td>136.37</td>
<td>22.36</td>
</tr>
<tr>
<td>14</td>
<td>186.51</td>
<td>37.3</td>
<td>22.56</td>
<td>30.58</td>
<td>186.83</td>
<td>30.59</td>
</tr>
<tr>
<td>15</td>
<td>247.13</td>
<td>49.43</td>
<td>22.18</td>
<td>40.51</td>
<td>247.42</td>
<td>40.53</td>
</tr>
<tr>
<td>16</td>
<td>370.8</td>
<td>74.16</td>
<td>19.78</td>
<td>60.79</td>
<td>371.02</td>
<td>60.80</td>
</tr>
<tr>
<td>17</td>
<td>102.32</td>
<td>20.46</td>
<td>5.36</td>
<td>16.77</td>
<td>102.77</td>
<td>16.79</td>
</tr>
<tr>
<td>18</td>
<td>2.21</td>
<td>0.44</td>
<td>0.41</td>
<td>0.66</td>
<td>2.22</td>
<td>0.66</td>
</tr>
<tr>
<td>19</td>
<td>93.27</td>
<td>18.65</td>
<td>11.17</td>
<td>15.29</td>
<td>93.45</td>
<td>15.30</td>
</tr>
</tbody>
</table>
APPENDIX 8

TABLE 14

FEMALE CATEGORIES

TOTAL SAMPLES = 4 x 122 = 488

GAMES LESSONS

<table>
<thead>
<tr>
<th>CAT</th>
<th>SUM X</th>
<th>X</th>
<th>On-1</th>
<th>%</th>
<th>SUM X</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>86.41</td>
<td>21.6</td>
<td>7.52</td>
<td>17.71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>14.44</td>
<td>3.61</td>
<td>2.69</td>
<td>2.96</td>
<td>132.62</td>
<td>27.18</td>
</tr>
<tr>
<td>3</td>
<td>26.83</td>
<td>6.71</td>
<td>3.24</td>
<td>5.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>4.94</td>
<td>1.23</td>
<td>0.41</td>
<td>1.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>113.9</td>
<td>28.48</td>
<td>4.71</td>
<td>23.34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>47.17</td>
<td>11.79</td>
<td>2.6</td>
<td>9.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>74.54</td>
<td>18.64</td>
<td>19.44</td>
<td>15.27</td>
<td>327.4</td>
<td>67.09</td>
</tr>
<tr>
<td>8</td>
<td>53.58</td>
<td>13.4</td>
<td>5.94</td>
<td>10.98</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>38.21</td>
<td>9.55</td>
<td>5.31</td>
<td>7.83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>16.22</td>
<td>4.05</td>
<td>2.14</td>
<td>3.32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>3.2</td>
<td>0.8</td>
<td>0.37</td>
<td>0.66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>8.56</td>
<td>2.14</td>
<td>2.62</td>
<td>1.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>137.34</td>
<td>34.33</td>
<td>7.65</td>
<td>28.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>161.68</td>
<td>40.42</td>
<td>9.9</td>
<td>33.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>177.22</td>
<td>44.31</td>
<td>11.46</td>
<td>36.32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>352.78</td>
<td>88.2</td>
<td>19.5</td>
<td>72.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>87.51</td>
<td>21.88</td>
<td>20.21</td>
<td>17.93</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>3.28</td>
<td>0.82</td>
<td>1.23</td>
<td>0.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>31.78</td>
<td>7.95</td>
<td>2.14</td>
<td>6.51</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX 8

TABLE 15

FEMALE CATEGORIES

TOTAL SAMPLES = 2 x 122 = 244

GYMNASTICS

<table>
<thead>
<tr>
<th>CAT</th>
<th>SUM X</th>
<th>X</th>
<th>On-1</th>
<th>%</th>
<th>SUM X</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>28.34</td>
<td>14.17</td>
<td>4.5</td>
<td>11.62</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>46.91</td>
<td>23.46</td>
<td>4.91</td>
<td>19.23</td>
<td>85.91</td>
<td>35.21</td>
</tr>
<tr>
<td>3</td>
<td>5.35</td>
<td>2.68</td>
<td>0.68</td>
<td>2.19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>5.31</td>
<td>2.65</td>
<td>2.2</td>
<td>0.41</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>49.42</td>
<td>24.71</td>
<td>0.75</td>
<td>20.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>17.25</td>
<td>8.63</td>
<td>1.79</td>
<td>7.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>18.71</td>
<td>9.36</td>
<td>3.13</td>
<td>7.67</td>
<td>146.31</td>
<td>59.96</td>
</tr>
<tr>
<td>8</td>
<td>35.5</td>
<td>16.75</td>
<td>5.84</td>
<td>13.73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>27.43</td>
<td>13.72</td>
<td>0.81</td>
<td>11.24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>6.36</td>
<td>3.18</td>
<td>2.94</td>
<td>2.61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>1.62</td>
<td>0.81</td>
<td>0.41</td>
<td>0.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>3.78</td>
<td>1.89</td>
<td>0.44</td>
<td>1.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>63.12</td>
<td>31.56</td>
<td>9.66</td>
<td>25.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>56.85</td>
<td>28.43</td>
<td>3.67</td>
<td>23.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>119.18</td>
<td>59.59</td>
<td>12.88</td>
<td>48.84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>185.69</td>
<td>92.84</td>
<td>8.59</td>
<td>76.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>32.48</td>
<td>16.24</td>
<td>9.75</td>
<td>13.31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>1.08</td>
<td>0.54</td>
<td>2E-2</td>
<td>0.44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>19.91</td>
<td>9.95</td>
<td>0.69</td>
<td>8.16</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>