

# **TEACHER-BASED ASSESSMENTS**

**A STUDY OF DEVELOPMENT, VALIDITY AND RELIABILITY OF TEACHERS'  
ASSESSMENTS AND ASSOCIATED STRUCTURED ACTIVITIES DEVISED TO  
ASSESS ASPECTS OF THE PRIMARY CURRICULUM FOR THE AGE RANGE  
8-12 YEARS AND THE EVALUATION OF IN-SERVICE PROVISION TO  
FACILITATE SUCH TEACHER-BASED ASSESSMENTS.**

**Anne M. Jasman**

**Submitted for the degree of Ph.D. at Leicester University**

**1987**

**CONTENTS**

<b>INTRODUCTION</b>	<b>1</b>
<b>CHAPTER 1 REVIEW OF THE LITERATURE</b>	
1.1 INTRODUCTION	4
1.2 DISCUSSION OF FACTORS INFLUENCING THE DEVELOPMENT OF PROGRESSIVE PRIMARY SCHOOLING IN BRITAIN	5
1.3 PROGRESSIVE PRIMARY SCHOOLS: MYTH OR REALITY?	13
1.4 THE PRIMARY SCHOOL CURRICULUM: AIMS, CONTENT, PROCESS, ASSESSMENT AND EVALUATION	18
1.5 THE ACCOUNTABILITY MOVEMENT: ITS DEVELOPMENT AND IMPACT ON SCHOOLS AND TEACHERS	31
1.6 TEACHER PROFESSIONAL DEVELOPMENT	43
1.7 SUMMARY	56
<b>CHAPTER 2 ASSESSMENT OF PUPIL LEARNING: A DISCUSSION OF METHODOLOGICAL ISSUES</b>	
2.1 INTRODUCTION	58
2.2 A CRITICAL REVIEW OF EDUCATIONAL ASSESSMENT METHODS	59
2.2.1 Developments in educational assessment	60
2.2.2 Selection of assessment methods	68
2.2.3 Factors influencing teachers' assessments of pupil learning	81
2.3 INVOLVEMENT AND PROFESSIONAL DEVELOPMENT OF TEACHERS IN THE RESEARCH	86
2.4 CONCLUSION	94



## **CHAPTER 3 RESEARCH DESIGN, METHODS OF DATA COLLECTION AND ANALYSIS.**

<b>3.1</b>	<b>SCOPE AND LIMITATIONS OF THE RESEARCH</b>	<b>96</b>
<b>3.2</b>	<b>HYPOTHESES</b>	<b>100</b>
<b>3.3</b>	<b>RESEARCH DESIGN</b>	
3.3.1	The Oracle research programme	102
3.3.2	Selection of schools and classes	103
3.3.3	Selection of pupil sample	107
<b>3.4</b>	<b>METHODS OF DATA COLLECTION</b>	
3.4.1	Observational instruments	107
3.4.2	Assessment procedures: Richmond tests, study skills exercises and teacher-based assessments.	109
3.4.3	Social class of pupils	110
<b>3.5</b>	<b>DEVELOPMENT OF THE TEACHER-BASED ASSESSMENTS</b>	
3.5.1	Formation and functions of teacher discussion groups.	111
3.5.2	The child should be able to listen with concentration and understanding.	116
3.5.3	The child should show some creativity in some fields.	119
3.5.4	The child should be able to acquire information other than by reading.	121
3.5.5	The child should know how to compute in the four arithmetic rules and use mathematical techniques in his/her everyday life.	124
3.5.6	The child should be able to communicate clearly and accurately through speech.	127
3.5.7	Summary of teacher-based assessments: checklists of abilities and structured activities	129
3.5.8	Introduction of materials to schools in the Oracle research programme	130
<b>3.6</b>	<b>VALIDITY AND RELIABILITY OF TEACHER-BASED ASSESSMENTS</b>	<b>133</b>
3.6.1	Validity of the checklists of ability	134
3.6.2	Validity of the structured activities	136
3.6.3	Validity of teachers' assessments	138
3.6.4	Practicality and reliability of teacher-based assessments	139
3.6.5	Reliability and validity checks	140

3.7	PUPIL PERFORMANCE ON THE STRUCTURED ACTIVITIES	142
3.8	METHODS OF ANALYSIS USED IN THE STUDY OF TEACHER-BASED ASSESSMENTS	143
3.8.1	Correlational analysis	144
3.8.2	Factor analysis	144
3.8.3	Analysis of covariance	145
3.8.4	Content analysis	146
3.9	DEVELOPMENT AND EVALUATION OF AN IN-SERVICE EDUCATION PROGRAMME	147
3.9.1	Strategies for the development of in-service provision	149
3.9.2	Research design and methods used for the evaluation of in-service provision	155
3.9.3	Methods of data collection	156
3.9.4	Methods of data analysis	161
3.9.5	Summary	162
3.10	SUMMARY OF RESEARCH DESIGN, METHODS USED IN THE STUDY OF TEACHER-BASED ASSESSMENTS AND THE EVALUATION OF IN-SERVICE PROVISION	164

## **CHAPTER 4: RESULTS OF THE STUDY OF TEACHER-BASED ASSESSMENTS AND DISCUSSION**

4.1	VALIDITY OF CHECKLISTS OF ABILITIES	
4.1.1	Logical review procedures	165
4.1.2	Conclusions	169
4.2	VALIDITY OF STRUCTURED ACTIVITIES	
4.2.1	Logical review procedures	170
4.2.2	Empirical review procedures	173
4.2.3	Conclusions	189
4.3	RELIABILITY OF TEACHER-BASED ASSESSMENTS	190
4.4	TEACHING STYLES AND PUPIL TYPES	193
4.4.1	Teaching styles	194
4.4.2	Pupil types	196
4.5	VALIDITY OF TEACHERS' ASSESSMENTS	197

4.6	<b>FACTORS INFLUENCING THE VALIDITY OF TEACHERS' ASSESSMENTS</b>	201
4.6.1	Discussion of results collected in the first year of the study (September 1976 - July 1977)	203
4.6.2	Discussion of results collected in the second year of the study (September 1977 - July 1978)	209
4.7	<b>SUMMARY OF MAIN FINDINGS</b>	213
<b>CHAPTER 5:</b>	<b>PUPIL PERFORMANCE ON THE STRUCTURED ACTIVITIES</b>	
5.1	<b>THE CHILD SHOULD BE ABLE TO LISTEN WITH CONCENTRATION AND UNDERSTANDING</b>	218
5.2	<b>THE CHILD SHOULD SHOW CREATIVITY IN SOME FIELDS</b>	222
5.3	<b>THE CHILD SHOULD BE ABLE TO ACQUIRE INFORMATION OTHER THAN BY READING</b>	225
5.4	<b>THE CHILD SHOULD KNOW HOW TO COMPUTE IN THE FOUR ARITHMETIC RULES AND APPLY THESE IN HIS/HER EVERYDAY LIFE</b>	228
5.5	<b>SUMMARY OF MAIN FINDINGS</b>	228
<b>CHAPTER 6:</b>	<b>REPORT OF THE IN-SERVICE EVALUATION STUDY</b>	
6.1	<b>INTRODUCTION</b>	233
6.2	<b>PROCESSES IN TEACHERS PROFESSIONAL LEARNING</b>	
6.2.1	Teacher characteristics at the start of the course	234
6.2.2	The process of teachers' professional learning: the course events	237
6.2.3	Linking theory and practice: teachers' experiences of the school placement	248
6.2.4	Processes of teachers' professional learning: discussion of results and some conclusions	259
6.3	<b>TEACHERS' PROFESSIONAL LEARNING: OUTCOMES OF THE COURSE</b>	265
6.3.1	Case studies, school and course files: evaluation of assessed work.	266



6.3.2	Teachers understandings of the concept of assessment: an analysis of line-labelled concept maps	269
6.4	<b>SELF-EVALUATION: AN ACCOUNT OF PROFESSIONAL LEARNING</b>	<b>278</b>
6.5	<b>SELF-EVALUATION: DISCUSSION AND CONCLUSIONS</b>	<b>288</b>
6.6	<b>TEACHERS' EVALUATION OF THE COURSE</b>	<b>291</b>
6.7	<b>SUMMARY AND CONCLUSIONS</b>	<b>291</b>
<b>CHAPTER 7:</b>	<b>CONCLUSIONS AND SIGNIFICANCE OF THE STUDY OF TEACHERS' ASSESSMENTS OF PUPIL PERFORMANCE AND IN-SERVICE EVALUATION</b>	<b>294</b>

## INTRODUCTION

This thesis reports the findings of research carried out as part of the research programme: Observational research and classroom learning evaluation (Oracle), based at Leicester University from 1975 to 1980 and funded by the Social Science Research Council and the subsequent development and evaluation of in-service activities as part of the B.Ed.(Hons) In-service programme offered at Worcester College of Higher Education.

The major objective of the Oracle research was to study the relative effectiveness of primary school teaching. The research programme was developed against a background of criticism of new approaches in primary schools. It emphasised classroom observation as the main research method. Systematic observation of teacher and pupil behaviour would, it was argued, offer teachers and other educators evidence to face these criticisms by providing descriptions of current practice in terms of pupil types and teaching styles derived from the observation of classroom events (Galton *et al.*, 1980; Simon and Galton (eds.), 1980). The programme included studies of transfer to the next stage of education (Galton and Willcocks (eds.), 1983) and the development of new forms of teacher-based assessment, that is, assessment procedures developed by teachers for use by other teachers (Jasman, 1979, 1980, 1981). This research on teacher-based assessments has been the subject of this thesis and was continued on completion of the Oracle programme in order to develop and evaluate in-service training procedures to improve the quality of teacher assessment of pupil performance. The in-service programme was developed as a result of the findings reported here on the validity and reliability of teachers' assessments.

The research described here addresses the following questions relating to the development of new forms of assessment.

- a) Can valid and reliable assessments of the performance of pupils be made in areas of the curriculum which are not amenable to measurement by standardised tests, by the provision of criteria for assessment in terms of observable pupil behaviours and structured classroom activities in which to observe these behaviours?
- b) Do teachers make valid and reliable assessments of pupil performance in areas of the curriculum not amenable to measurement by standardised tests?



- c) If teachers are unable to make valid and reliable assessments of pupils what are the factors contributing to their judgements of pupil performance in such areas?
- d) What is the relationship between pupil performance in those areas of the curriculum not assessed by standardised tests and other pupil characteristics such as age, sex, social class, attainment in the basic skills, pupil type and teaching style?
- e) Does in-service education affect the quality of teachers' assessments by the provision of appropriate information and skills to facilitate assessment of pupil performance in areas other than those tested by standardised achievement tests.

### **Significance of the Study**

The research discussed in this thesis was proposed in response to three major considerations. First, criticisms have been made of teacher effectiveness research and curriculum evaluation for the limited nature of the assessment procedures used to describe pupil performance and progress. Such studies have also failed to consider the aims held to be of importance by teachers as a basis for evaluating teacher effectiveness and the primary school curriculum. If it can be assumed that the aims held by teachers have been an important influence on teaching style and the nature of the curriculum, then these aims should be taken into account when deciding the criteria and assessment procedures to be used in evaluating their effectiveness.

Second, teacher and school accountability for pupil learning, attitudes and behaviour have become increasingly important issues in the 1970's. Concern was originally expressed by government, media and the community as to the quality and standards of primary and secondary education in England and Wales following the period of apparently rapid change and innovation in the 1960's. More recently education cuts and restrictions of funding have resulted in a financial model of accountability gaining acceptability, so that the 1980's have been characterised more with a concern for value for money than maintaining educational quality. Teachers and the education community in general have needed appropriate methods to demonstrate the quality and worth of any changes in the organisation of the curriculum. However, such curricular developments have not always been matched by corresponding developments in assessment, testing and evaluation techniques to monitor the quality and effectiveness of such innovations.

Third, there have been increasing demands on teachers to be involved in the assessment of pupils although the judgements made by teachers have often been mistrusted with respect to their validity and reliability both in the profession and by parents, employers and other agencies outside the school. The need for effective in-service training for teachers in all sectors of the education service has been highlighted by the recent developments in local education authority policy on primary school records; changes in secondary examinations, particularly the introduction of pupil profiling and graded achievement tests. The requirement of the 1980 Education Act for local education authorities to provide for children with special educational needs has also raised the question of whether teachers were equipped to identify such pupils and monitor the quality and effectiveness of the provision made for them.

The purposes of this research are therefore to:-

- a) counter criticisms of previous teacher effectiveness research by developing measures to collect more extensive data on pupil performance to compliment the standardised attainment tests of mathematics, language and reading that were used in the Oracle research programme to investigate the effectiveness of different teaching styles.
- b) provide valid and reliable assessment procedures for teachers to monitor aspects of pupil performance in line with the changed nature of the aims and curriculum of primary schools.
- c) develop in-service training procedures designed to improve the quality of teachers' judgements when this need had been identified.

In Chapters 1 and 2 the relevant literature has been discussed in order to identify and support the various purposes outlined above and the methodological issues and problems associated with the study. The strategy, design and implementation of the research carried out to address the questions detailed above have been considered in Chapter 3. The results of the various phases of the research have been presented in Chapters 4, 5 and 6 before concluding with a discussion of the significance of the findings and the implications of this study of teachers' assessments in Chapter 7.



## CHAPTER 1: REVIEW OF THE LITERATURE

### 1.1 INTRODUCTION

The discussion presented in this chapter focuses on demonstrating the need for the development of new, more appropriate assessment procedures which arose from the emergence of a progressive ideology in primary schools. It is argued here that one of the most important factors in the evolution of progressive practices has been the move away from a pedagogy based upon the class as the basic unit of teaching. As a consequence of this new approach teachers have to be more aware of individual pupil needs. Thus, in order to provide appropriate learning experiences which match a pupil's stage of development, teachers require an efficient and effective way of monitoring and assessing pupil learning. In addition, the need for appropriate assessment procedures in primary classrooms has also been increased by the growing pressures for greater school and teacher accountability.

The individualisation of instruction was, however, only one of the features which characterised progressive primary schools. In writing about such schools as models of institutional innovation, Kogan described the following features as characteristic:

*"The curriculum lacks a teacher imposed structure, since the best teachers know what they want children to know and allow children to become interested in the environment provided by them. The day becomes open with children moving around the classroom and school pursuing their interests. The teacher becomes a counsellor, working with children individually, or in groups or as a whole class as the opportunity and need arise. Primary school buildings allow for flexibility which combines openness with small group spaces and a variety of resource areas. The overt relationship between teacher and child is one of partnership. Children's wishes are respected but they are encouraged to build their own motivation. Individual and small group instruction rather than class instruction is the norm because it respects that children develop at different rates. Parents are part of the process but do not control it. Schools tend to be relatively small even if some of the classes are large."* (Kogan, 1971, p.44)

These characteristics echoed the pronouncements of the Plowden Committee in their report *Children and their primary schools*, where the primary school was described as

*"not merely a teaching shop, it must transmit values and attitudes. It is a community in which children learn to live first and foremost as children and not as future adults. In family life children learn to live with people of all ages. The*

*school sets out deliberately to devise the right environment for children, to allow them to be themselves and to develop in the way and at the pace appropriate to them. It tries to equalise opportunities and to compensate for handicaps. It lays special stress on individual discovery, on first hand experience and on opportunities for creative work. It insists that knowledge does not fall into neatly separate compartments and that work and play are not opposite but complementary. A child brought up in such an atmosphere at all stages of his education has some hope of becoming a balanced and mature adult and of being able to live in, to contribute to and look critically at the society of which he forms a part."* (Plowden, 1967, pp.187-188)

These descriptions of progressivism give some insight into what was considered, in the late 1960's, a model of good primary practice. They stressed the importance of individual and group work based on first hand experience, discovery learning growing from the children's interests and acknowledged that the needs of society would not be met if children were simply educated to do what they were told. This approach had, however, taken several decades to become official orthodoxy. The factors which have resulted in the implementation of new patterns of organisation and curriculum provision based on an individualised approach to instruction are now considered since an understanding of this progressive approach to primary education was important in the development of more appropriate methods of assessment.

The following discussion describes the various influences that have contributed to the evolution of progressive theory and current pedagogic practice in Britain today. Two major themes are examined. First, the contributions of a number of individuals and governmental reports to the development of progressive theory and practice are presented. Second, it is argued that a number of contextual factors such as the demise of the 11+ examination, comprehensive schooling and changes in the organisation of the education system were instrumental in shaping the primary schools of today.

## **1.2 DISCUSSION OF FACTORS INFLUENCING THE DEVELOPMENT OF PROGRESSIVE PRIMARY SCHOOLING IN BRITAIN**

During the last half of the nineteenth century the education of young children in elementary schools was dominated for some thirty years by payment by results following the publication of the Code of Conduct in 1862. The beginnings of a view of elementary schooling which can legitimately be called progressive (Selleck, 1972) can be found in the period following publication of *What is and what might be* (Holmes, 1911). Subsequently, the works of other individual educators and practitioners clearly affected



the development of the primary school.

Montessori (1912) was an important influence. Her methods brought about major changes since they emphasised structured learning, sense training and individualization. The use of apparatus and the attempt to provide for the developmental needs of individual pupils enabled children to work independently in the same room at the same time. The kindergarten movement (Bowen, 1903) based on Froebel's theory and practice, also contributed to the development of individualisation and was institutionalised with the establishment of the Froebel Institute and other colleges for the training of teachers on Froebelian principles. The work of Margaret McMillan and her sister Rachel, which emphasised the improvement of hygienic conditions and the provision of an appropriate environment for young children was influential in bringing new concepts concerning activity and creativity to the infant school particularly, but also affected the work of teachers in the elementary system as a whole (Rusk, 1933; Bradburn, 1976).

Another important influence in the development of modern primary practice resulted from the Dalton Plan which was originated in 1920 by Helen Parkhurst (1922) at the Dalton High School in Massachusetts. The plan embodied ideas from Dewey (1900, 1916) and was based on an assignment system where work was completed by individuals. It was implemented to a much greater extent in British primary schools than in the United States (Kimmins and Rennie, 1932) and was the forerunner of contemporary project and topic-work. It also encompassed the essential features of "*child-centredness*" which underpinned many of the developments in primary education in recent years. Selleck (1972) pointed out that for a teacher who was worried about his or her own everyday practice, as a result of the spread of progressive ideas, the Dalton plan had a particular attraction. It offered a specific methodology, a set of procedures and a system of organisation which would accommodate new methods but which did not entail a total rejection of the old.

A number of individuals set up experimental schools; for example, Macmunn at Tiptree Hall which opened in 1919; Homer Lane, director of Little Commonwealth in Dorset and A.S. Neill of Summerhill. These pioneers did a great deal to publicize the new teaching methods. Similarly, the work of Susan Isaacs at the Malting House Garden School and the publication of her books on the social and intellectual development of young children (Isaacs, 1930, 1933) also contributed to the development of a progressive tradition.



Another important milestone in the development of progressive primary schools was the publication of the Board of Education report, *The Education of the Adolescent* (Hadow, 1926) establishing separate educational provision for pupils aged five to eleven. Although its recommendations were concerned principally with provision for post-primary pupils, an unintended result of separating children at eleven was the opportunity for educational programmes to be developed to meet the individual needs of children in their early years. Reorganisation began after 1928, when the report's recommendations were accepted by the government and by 1938, 48% of children aged seven to eleven were in junior departments (Galton *et al.*, 1980). The Hadow Report of 1931 (p.xvii) focused specifically on the "*primary school*" for the first time. The Committee was required "*to enquire and report as to the course of study suitable for children other than in infant departments up to the age of eleven*" and stressed that the young child was a "*growing organism*" and that his/her curiosity was more likely to be aroused by starting from "*a centre of interest*" than being taught specialized subjects. The report urged that the curriculum should be thought of in terms of "*activity and experience*" rather than as inert pieces of knowledge and that there should be a new orientation of school instruction bringing it into "*closer connection with the natural movement of children's minds*" (Hadow Report, 1931, p.xvii). The committee also recommended an approach that provided for children's interests and for work on individual and group projects. It was argued that full use should be made of the environment and that the "*more closely the design of the primary school approaches that of the open air school the better*" (Hadow Report, 1931, p.xvii). The Hadow (1933) report on infant and nursery schools equally favoured active methods of learning and individualisation. Both reports spoke of the value of individual and group work and suggested that "*the function of the teacher is less that of an expositor than of an advisor and consultant*", (Hadow Report, 1931, p.xxiii).

Given these recommendations regarding individualisation, how far did progressive notions impact on the state primary schools? Selleck (1972) argued that progressivism had become the official orthodoxy by 1939. It was propagated by the training colleges, Board of Education in-service courses and local authority inspectors. This trend was also reflected in some of the new buildings of the period, with the use of more flexible furniture and the provision of larger working spaces in classrooms (Seabourne, 1971).

His Majesty's Inspectorate (Ministry of Education, 1952, prefatory note), however,

indicated that the 1931 Hadow Report's recommendations took a long time to gain ground, especially in junior departments. This lack of change was ascribed to a number of factors; for example, the fact that local authorities concentrated on secondary reorganisation and building; the scholarship system of examination at eleven, which encouraged the view that the primary school's justification lay solely in preparing children for the secondary school and finally as a result of the sheer inertia associated with well established traditions. The size of classes also worked against individualisation; for example, in 1932 over 50% of pupils of primary school age were still in classes of over 40 children (Ross, 1960). It appeared from these comments that the necessary conditions for the implementation of change were lacking.

One important factor in maintaining the status quo can be attributed to the 1931 Hadow Report's recommendations on streaming, based on the psychology of individual differences. This and subsequent Hadow reports placed great emphasis on the classification of children which reinforced the practice of streaming in schools, since it was considered easier to cater for individual interests and needs in an homogenous setting. However, this emphasis on streaming appeared to work against the introduction of individualisation and group work. Teachers continued to use the class as the medium of instruction as they did not have to cope with the range of pupil abilities characteristic of unstreamed classes.

Streaming was also sustained by the continuation of the role of the school as a selection and classification body which identified those pupils who would best benefit from the new secondary provision embodied in the 1944 Education Act. This need for selection was exacerbated by the increased competition for entry to grammar schools and led to the spread of streaming *"with barely credible rapidity through the country"* (Jackson, 1964, p.150). The selection examination was based on tests of the three "R's" and intelligence and led to the re-emergence of the problems associated with payment by results in the 1870's, that is, teaching to the test, even though then as now the three "R's" were viewed

*"not only or primarily as the exact amount of knowledge which could be given, but as an amount of knowledge which could be ascertained thoroughly by examination."* (Sylvester, 1974, p.67)

A basic class teaching approach still predominated, with the emphasis on literacy and numeracy. There was little opportunity for change in curriculum practice or the



development of the distinctive non-selective, non-competitive ethos associated with progressivism, individualisation and co-operative group work. Thus, the theory and practice associated with the progressive tradition were not worked out in any systematic or coherent way until after the Second World War and did not have any noticeable impact on the state primary schools until well into the 1950's. If the continued existence of class instruction, streaming and the need for selection during the 1940's and 1950's tended to reinforce the traditional model of school organisation and its related pedagogy, how then did the changes of the 1960's come about? A number of factors may be advanced although few of these were explicitly concerned with the promotion of change and innovation in the primary sector.

Perhaps the most powerful force in bringing about change in the organisation of primary school classrooms was the move towards comprehensivisation in the late 1950's (Rubinstein and Simon, 1972) which resulted in the demise of the 11+. This stemmed from the social and political movements of that time promoting education for all and equality of opportunity. This view also involved the rejection of the absolute nature of individual differences and an acceptance of the influence of the environment on intelligence and educational attainment. This gave rise to movements against 11+ selection, streaming and towards positive discrimination in favour of certain socially depressed areas (Plowden, 1967) and provision for the lower ability pupil in secondary schools (Newsom Report, 1963). Indeed, it was the recommendations made in the Plowden Report that led to the establishment of Educational Priority Areas (1967, p.151-154). Although this movement was criticised at the time in that "*education could not compensate for society*" (Bernstein, 1970, p.111), the notion of compensatory education and comprehensive schooling did gain acceptability (Newson, 1977).

A climate was thus established for the implementation of comprehensive schooling as a result of this move against selection at eleven. In 1951, twelve comprehensive schools existed largely as a result of the influence of a few powerful local authorities such as the London County Council, West Riding and Coventry. The early 1960's saw the beginning of moves to do away with the selective system and by 1969 the National Union of Teachers had passed a resolution calling for the necessary legislative changes to ensure comprehensive education for all. As the need for selection at eleven became redundant, many local authorities began to abolish or modify the examination. The first example of this was in Anglesey in 1953, when all primary aged pupils transferred to the neighbourhood comprehensive without any form of examination or selective test. The

primary school was, therefore, gradually released from the constraints imposed by the need to achieve 11+ examination success.

Streaming also gradually lost acceptance with teachers. In 1962 the survey carried out by the National Foundation for Educational Research found 85% of primary teachers favoured streaming, 6% had mixed views and 9% were hostile. In 1967 the Plowden committee reported that only 34% of teachers approved of streaming for older pupils and 30% were hostile, the remainder had mixed views (Plowden, 1967, p.287). This pronounced shift in teachers' opinions was particularly noteworthy since educational change has often been remarkably slow to find expression in this way.

A second blow to the 11+ and particularly to the practice of streaming was the report from the committee set up by the British Psychological Society in 1957 to look at the growing criticisms of intelligence tests (Vernon, 1957). They pronounced against the use of such tests to select pupils and deprecated streaming on the grounds that this classification of children determined their educational environment and thus set limits on their attainment in an unacceptable way. The Plowden committee noted that

*"streaming serves as a means of social selection .. more middle class pupils are to be found in upper streams and fewer in lower streams than would be expected from their results in objective tests."* (Plowden, 1967, p.289)

During the 1960's and early 1970's a number of studies on the effects of streaming were carried out (Jackson, 1964; Hargreaves, 1967; King, 1969, 1973; Lacey, 1970) and links between the child's home background, social class and educational experience with achievement became well established (Jackson and Marsden, 1962; Douglas, 1964; Coleman *et al*, 1966; Ford, 1969; Banks and Finlayson, 1973). The Plowden committee also reported this relationship with the statement that Hindley (1962) had shown that

*"even amongst children well below compulsory school age the growth of measured intelligence is associated with socio-economic features."* (Plowden, 1967, p.119)

The move away from selection had another unforeseen consequence in relation to the schooling of children aged five to eleven. This was the introduction of different patterns of organisation in certain local authorities with provision for transfer at ages other than eleven. Mason (1964) then Director of Education in Leicestershire reached the opinion



that it might be possible to avoid the inequalities of selection by introducing a two-tiered secondary system of comprehensive schools. This proposal raised the question of why transfer should be at the age of eleven. Sir Alec Clegg, Chief Education Officer for the West Riding, proposed an alternative, with pupils aged 5-9 in first schools, 9-13 in middle schools and 13-18 in upper schools. Various educational arguments were advanced for middle schools including their capacity to extend the best practices of primary education, their potential for assisting pupils during a critical transitional period of their personal development but as Hargreaves and Tickle noted

*"priority was attached to the claim that without an official sanctioning of middle schools, many areas, contrary to their wishes would simply be unable to provide a viable system of comprehensive education."* (Hargreaves and Tickle, 1980, p.3)

The Plowden committee endorsed a move towards transfer at 12+ and the development of first, middle and upper schools (Plowden, 1967, p.152). However, some local education authorities who subsequently introduced middle schools were not necessarily motivated by educational arguments. Their decision to reorganise appeared to be made in order to maximize the use of available space to meet the needs of an expanding school population resulting from the *"baby boom"* of the post-war years. However, conditions were created for the development of different patterns of organisation in schools. Such major changes in the structure and organisation of the education system clearly had an impact on the nature and provision of primary schooling even though they were not directed towards this end and, as has already been noted, were not always the result of decisions made on the basis of educational principles.

In addition to these changes in patterns of organisation at the local authority level there were also changes in the pattern of building and class organisation within schools which also helped to promote the shift towards progressive primary practices. Open plan schools began to be built in the period following the introduction of government cost and area per place limits in 1950. The central ideas of open planning have been described in various Ministry of Education and Department of Education and Science building bulletins since 1949 (Seaborne and Lowe, 1977, pp.170-171). The new system limited the cost of every school to an amount calculated to a set formula. This was reduced from 200 pounds to 170 pounds in 1950. As building costs rose the average cost per place also rose but generally at a slower rate. This decrease in funds and changes in the amount of circulation space allowed produced different ideas of school design and saw the start of the open plan school. The change of design was justified in terms of educational



principles. Seaborne commented that the role of the teacher had become more varied than was possible to fulfil in a school of more conventional layout.

*"Many primary school teachers had for some years been using highly flexible timetables to enable young children to pursue their individual and group interests rather than be taught in classes of forty and they welcomed the greater informality and the use of more varied resources which were possible in such schools."* (Seaborne, 1971, p.173)

He also commented that sceptics might remark on the convenient way in which educational arguments were found for making economies which were forced on the educational service by the pressure of external events (Seaborne, 1971, p.169). However, the less specialized curriculum and the smaller size of most primary schools did give greater opportunities for trying out new methods from both the architectural and educational view point.

Central and local government officials who were initially responsible for the design of new primary schools were perhaps most influenced by the Hadow Reports and it was at this time that the conditions of implementing the recommendations of these reports became favourable. Other reports such as the Ministry of Education Handbook on Primary Education (Ministry of Education, 1959) have contributed to the ideas of open planning, which received their fullest expression in the Plowden Report (1967).

Ideas took a long time to mature but by the 1960's open plan designs began to transfer to urban areas and primary school building more generally. An example of this type of development was the Eveline Lowe Primary School, which was designed by the D.E.S. Development group in co-operation with the Inner London Education Authority. The educational rationale and building plans have been fully documented in Building Bulletin No.36 and in the Plowden Report (Plowden, 1967, p.400). Although only 10% of primary schools in England and Wales were of open plan design by 1978 (Bennett *et al.*, 1980), it can be argued that the development of such schools contributed to the "quickenning trend" (Plowden, 1967, p.188) towards child-centred education noted in the Plowden Report and more informal methods and practices in primary schools generally.

The preceding discussion has illustrated that the development of progressivism in British primary schools was influenced by a number of contextual factors such as: the introduction of comprehensive schools, with the demise of the 11+ following from this and changes in the patterns of organisation at the local education authority level,

associated with the introduction of different ages of transfer and changes in school building design, which resulted in the development of the open plan school. It has been noted that these factors have not necessarily been designed to promote change in the primary sector but have had that effect. Thus, it appeared that the recommendations embodied in the Hadow and Plowden reports had at last found a suitable climate for implementation by the late 1960's. However, the impact of these recommendations and organisational changes have yet to be examined and the following section will discuss the extent to which the factors identified in the preceeding discussion have influenced practice in the primary school and thus the need for more appropriate methods of assessment.

### **1.3 PROGRESSIVE PRIMARY SCHOOLS: MYTH OR REALITY?**

It was important to have a clear understanding of what characterised progressive schools in considering the central question of how far the recommendations and factors described in the previous section were implemented in schools. The work of Stephens (1974) in her study of practice in both Britain and the United States identified a number of characteristics that can be viewed as basic to the idea of an open or progressive classroom. These matched the recommendations proposed by Plowden Report (1967) and the features identified by Kogan (1971, p.44) described at the beginning of this discussion. They provided a useful checklist for the identification of progressive practice in a primary school, and are given below:

- a) The organisation of the classroom was characterised by most instruction being geared to groups or individuals, with a variety of activities progressing simultaneously.
- b) The pattern of the day usually involved a flexible timetable or an integrated day in an environment rich in materials and resources.
- c) The pupils were allowed freedom of movement and the chance to talk and help each other.
- d) The curriculum was flexible with some integration and stressed attention to the individual child's intellectual, emotional, physical and social needs.



- e) Creative activities were also seen as an important part of the curriculum.
- f) The teacher's role was to encourage open relationships in the classroom, where the teacher became a friend and counsellor seeking to promote self discipline rather than impose rules and regulations.
- g) Teaching methods were designed to provide opportunities for self-directed learning and experimentation, with the pupils pursuing their own interests in an atmosphere of trust, acceptance and respect for their diversity.
- h) There was a minimum of marking and grading for summative purposes. Assessment was viewed as a means of helping identify pupil needs so that learning activities were provided at an appropriate level. Its purpose was not for classifying and selecting individuals as in formal testing.

How far then did the reality of the classroom match this model of progressive primary practice, since it is argued that more appropriate assessment procedures are needed if such changes in practice are observed in primary schools? An extensive literature relating to this area exists from the mid 1960's. Investigations into various aspects of primary education have been reported, particularly the nature of change in primary practice. A number of studies have given descriptions of individual schools and classrooms based on anecdotal material which provided models of what could be done (Blackie, 1967; Brown and Precious, 1970; Kallett, 1970; Jeffrey, 1970; Mitchell, 1970; Sturgess, 1970). Similarly, there were reports and recommendations on the organisation of schools along informal lines, as open-plan schools (Jarman, 1977) or operating the integrated day (Marsh, 1970; Blackie, 1974).

Whitbread (1972) reviewed the spread of progressive or informal schools in the infant and nursery sectors stating that they were relatively common by the 1950's and widespread by the 1960's. Spodek and Walburg (eds) (1975) reported studies in open education, including comparisons between British open, American traditional and open classrooms (Evans, 1975) and an observational portrait of a British infant classroom (Brandt, 1973, 1975).

A number of studies using participant observations and ethnographic methods documented the process and nature of interactions between teachers and pupils (Brown,

1969; Nash, 1973; Sharp and Green, 1975; King, 1978). Other studies used systematic observations to describe the interactions between teachers and pupils (Jackson and Lahaderne, 1967; Resnick, 1972; Boydell, 1974) or the activities and experiences of the pupils (Boydell, 1975a, 1975b). The changing nature of the teacher's role and aims in relation to progressive practice have also been described (Gardner and Cass, 1965; Allender, 1972; Ashton *et al.*, 1975; Ridgeway, 1977; Boydell, 1979) as well as studies on how teachers spent their professional time (Duthie, 1970; Hilsum and Cane, 1971).

Studies which have documented the patterns of organisation in junior school classrooms were of particular interest, since they provided evidence of readily observable changes in primary school practice. Research in this area was described in the work of Moran (1971) and Bealing (1972) who found that *"relatively informal layouts were adopted by the vast majority of teachers"* (Bealing, 1972, p.235). In terms of the time spent in different grouping patterns, Bassey (1978) calculated that a typical pupil in a junior school would spend about nine hours on classwork, ten hours on group work and three hours on self-organised individual assignments. These studies provided evidence of a move away from class organisation and instruction towards individualisation of pupil work and seating arranged in groups. These changes were associated with different conceptions of the teacher's role but it was perhaps assumed that these would result in differences in teaching style.

*"The teacher's major role in a school where the majority of emphasis is on individualized instruction is obviously not to transmit information - the most significant function of the teacher is to make sure that each student always has materials that are particularly appropriate for him at any given time."* (Allender 1972, p.231)

However, the evidence relating to teaching style has suggested that any developments in progressive practice occurred largely at the organisational level, although changes in interaction patterns associated with the extension of the teacher's role might have been expected. According to Bennett's 1975 survey only 8% of fourth year teachers and 9% of third year teachers were working to a Plowden model of primary education. Bennett's findings provided evidence of little general movement towards informality in teaching style in the primary school, at any rate at the upper age level. Only 17% of the teachers in the overall sample were said to be teaching in the informal manner; whilst the majority claimed to use mixed styles.



The more recent government survey (D.E.S., 1978a) of 542 primary schools visited by Her Majesty's Inspectors between 1975 and 1977 showed a trend towards even greater formality with about 75% of teachers in the sample using mainly didactic methods, compared with only 5% teaching by discovery techniques. Bealing's (1972) survey showed that children were organised physically in groups, but they were given individual tasks. The methods which teachers used showed *"evidence of tight teacher control over such matters as where children sit and move"* and it seemed *"highly doubtful that there is much opportunity for children to choose or organise their activities in most classrooms"* (Bealing, 1972, p.236). Thus, the pupils were still directed by the teacher in much the same way as was the case in the elementary school of the early twentieth century.

Many teachers appeared to be the unwilling victims of the rapid expansion of the education system during the 1960's. They found themselves in open plan buildings or reorganised schools which required them to implement the necessary physical reorganisation of their classroom in order to work with groups and individuals. They had to cope with much broader ranges of ability as a result of unstreaming, but they appeared not to bring their practice in line with progressive theory of the time. For example, open plan architecture was not necessarily associated with open or progressive methods nor did the fact that a school was built on traditional lines mean that it used traditional methods. Pluckrose described open plan schools where *"teachers work in tight class units, their desks across the openings instead of doors"* (Pluckrose, 1975, p.54). Adelman and Walker also observed open space designs where

*"uncommitted teachers were wrestling with problems arising from a clash between their own educational assumptions and those imposed upon them by their immediate environment."* (Adelman and Walker, 1974, p.103)

Hurlyn in discussing the incidence of inquiry/discovery learning in open plan schools noted that

*"in some cases the structure of the rooms and the division between the teaching spaces created great difficulties and inconveniences to teaching procedures, mainly because they were designed with different procedures in mind ... that there was no inquiry/discovery work taking place in the open plan schools was not so much because of the design of the schools, but it was because the teachers had other constraints bearing upon them."* (Hurlyn, 1975, p.103)

This appeared to have been in spite of the argument presented by Rintoul and Thorne



(1975) that building trends were best explained in terms of a drive for flexibility and maximum opportunity and that open plan design was intended to help teachers plan their work in accordance with the modern trends in primary education, thus allowing them to experiment and innovate with such practices as the integrated day, team teaching and vertical grouping. Seefeldt reiterated the concerns expressed by Pluckrose (1975), Adelman and Walker (1974) when he noted that

*"when one settles down carefully to observe and scrutinize the teaching-learning process that is taking place in many open plan schools, the observer leaves with a somewhat different and disquieting feeling. Open spaces do not necessarily guarantee freedom in the classroom."* (Seefeldt, 1973, p.356)

It was even less likely that changes in teachers' classroom practice and style would be observed where there were no obvious external factors such as open plan buildings acting to modify practice. However, despite this apparent lack of pupil autonomy or change in teaching style, the pattern of organisation in schools and classrooms had demonstrably altered, with mixed ability being the norm and a shift towards individual assignments for pupils. These moves required teachers to keep more accurate records of pupil work and to be more aware of the individual child's stage of development, as well as particular strengths and weaknesses, in order to promote their learning. Thus the expected role of the teacher changed as described by Wolfson.

*"The most popular interpretation of individualized instruction is that the teacher makes specific recommendations and assignments for each pupil. This interpretation rests on the teacher-as-doctor analogy ... she diagnoses needs, deficiencies or problems and prescribes treatment."* (Wolfson, 1968, p.358)

Such changes in classroom organisation, family grouping and the integrated day encouraged children to study at their own pace so that pupils could be doing different things at different times, so that as Rance argued

*"the problem of how to record what is actually taking place within a class under these conditions faces all teachers who adopt less formal methods of organisation."* (Rance, 1971, p.13)

Thus, such organisational changes in schools and individual classrooms necessitated the development of more appropriate ways of monitoring and assessing pupil activity, even though the recommendations for change in primary practice and the external factors which produced these differences in organisation at the school and classroom level appeared to have had little effect on teaching style. Other factors have also contributed to

the need for the development of more appropriate ways of assessing pupil learning. For example, it will be argued that shifts in the nature of the primary curriculum, through the development of more extended aims of education have had some impact on primary practice. Developments in the primary curriculum during the 1960's and 1970's are now discussed to illustrate those changes in content, processes and evaluation methods which need to be taken into account in the development of more appropriate methods of assessing pupil learning in the primary school.

#### **1.4 THE PRIMARY SCHOOL CURRICULUM: AIMS, CONTENT, PROCESS, ASSESSMENT AND EVALUATION**

Harlen (1976a) argued that changes in the values related to the aims of education were influential in promoting shifts in assessment practices. When it was recognised that the nature of society and employment were changing the traditional aims of primary education and curriculum content were questioned (Plowden, 1967, p.185). Education became concerned with preparing for the new. The progressive curriculum expected pupils to challenge and ask questions, to show independence of thought and action and to pursue their own lines of enquiry. Clearly there was no tradition of measuring these outcomes and different strategies needed to be devised. One way of achieving this was to extend the range of assessment strategies in primary schools. Aims and objectives other than those related to traditional curriculum content and basic skills should be included and thus aspects of children's performance more relevant to independent inquiry and discovery learning would be assessed. However, before considering the need for more appropriate assessment methods, it is necessary to trace the development of the curriculum in primary schools in order to determine what would have to be assessed.

Blyth (1965) identified three main influences on primary curriculum development: the *elementary*, *preparatory* and *developmental* traditions. The elementary emphasised the basic skills, whilst the preparatory tradition viewed the education of young children in terms of preparation for the later stages of education rather than as a stage in its own right. According to Blyth (1965) the third tradition has its roots in the child development movement, where practice was guided by a similar set of beliefs as those of the progressive movement; with an emphasis on the education of the individual, on meeting needs as appropriate to the stage of development and the belief that the curriculum should be centred on the child's own interests and experiences.



In the late 1960's and early 1970's, curriculum developments were based on the introduction of infant patterns of organisation and the extension of the integrated day into the junior classroom (Brown and Precious, 1970; Brearley *et al.*, 1969; Marsh, 1970; Moran, 1971; Blackie, 1974). These recommendations were firmly rooted in the child development tradition identified by Blyth (1965) and were based on the work of Bruner (1966), Piaget and Inhelder (1969) which stressed that each child has different learning needs. It was important, therefore, to cater for these differences by providing an appropriate environment and curriculum which was flexible and adaptable enough to meet these individual needs. It was assumed that the pupils would learn through activity and experience providing there was a flexible timetable which would allow pupils to pursue their own interests and there was a wide variety of resources in the classroom for pupils to choose from.

Implicit in these moves towards a child-centred curriculum and integrated day was the notion of enquiry/discovery learning, since as Anthony pointed out

*"the teacher is expected to arrange for the pupils to experience freedom, activity and discovery in learning."* (Anthony, 1982, p.381)

Kemp (1955), Warburton (1964), Gardner (1966), Barker Lunn (1970), Bennett (1976), Aitken *et al.* (1981) all referred to pupils having relative freedom of activity in the progressive classroom even though as Anthony noted

*"discovery methods are not as prominent as such in the researchers' descriptions, but discovery in a broad sense is implied by the idea of learning by (relatively) free activity."* (Anthony, 1982, p.381)

These recommendations were in line with the Plowden committee who endorsed *"the trend towards individual and active learning"* (Plowden, 1967, p.202). In further consideration of the curriculum the Plowden Committee argued that *"children's learning does not fit into subject categories"* (Plowden, 1967, p.203) and advised that class teachers be encouraged to draw up a statement of their aims, relating them to *"guiding principles instead of short term objectives"* (Plowden, 1967, p.187). Thus, the Plowden Report-emphasised the need to provide an appropriate learning environment with maximum opportunity for developing a child-centred curriculum.

A second curriculum trend was in evidence during the 1960's. As the nature of



employment and occupations changed, schools were looked to for a more general education which would produce people who were more adaptable and flexible (Plowden, 1967, p.185). It was argued by Dearden (1968) that this adaptability was a function of having general concepts and principles at one's command and that in illustrating these principles, greater emphasis should be given to mathematics and science within the framework of a general education. These changes in curriculum content, it was argued, would have repercussions on general methodology and encourage moves away from an elementary tradition with its emphasis on basic skills. There was, however, still the need to review the content of the curriculum and to identify the processes, concepts and principles which would be most appropriate for the primary school to contribute towards the preparation of pupils for a changing world.

The Schools Council and the Nuffield Foundation were both instrumental in making the 1970's the era of curriculum development (Stenhouse, 1980a). Some projects offered curriculum materials that responded to existing needs: for example, **Breakthrough to literacy** (Reid, 1975) which applied linguistic theories to primary school work and the **Scope** materials (Taylor and Ingleby, 1975) which met the needs of non-English speaking pupils. Other projects concentrated on particular age ranges such as the **Middle Years of Schooling** project (Badcock *et al.*, 1972; Ross *et al.*, 1975), whilst others reviewed and developed curriculum subject content and methods in areas such as social studies (Schools Council, 1971a), physical education (Schools Council, 1971b), art and craft (Schools Council, 1974), environmental studies (Schools Council, 1973a), mathematical concepts (Schools Council, 1975), science (Schools Council, 1972a) and drama (Stabler, 1978). Similarly, the Nuffield Foundation funded development work in mathematics and science in the primary sector and the Bullock Report (D.E.S., 1975a) stimulated much activity by local education authorities with its recommendations for language across the curriculum.

As an external agency for curriculum development the Schools Council provided a useful resource for teachers, although the degree of their success was questioned in the research study conducted to monitor the impact and take-up of materials (Steadman *et al.*, 1978). However, it was clear from this study that a number of projects were in use; notably **Science 5-13**, **Mathematics in the Primary school**, **S.R.A. Reading Laboratories** and **Breakthrough to literacy**. These were the most popular projects in terms of their take-up by schools being concerned with language, mathematics and science rather than other curriculum areas. Each of these projects also provided guidelines

on the assessment of pupil learning. It could be suggested that teachers responded more favourably to curriculum developments and guidance on what to teach and how to assess it in the areas of mathematics and language since the demise of the 11+ meant there was not longer a test to teach to or to determine the nature of the curriculum.

A third trend in curriculum development emerged during the 1970's. This entailed a shift from external control of curriculum development to the involvement of teachers and schools in the process. Curriculum projects were traditionally designed to have research, development and dissemination phases with curriculum specialists employed to carry out the research. Teachers were not involved until the materials were presented for use in the classroom. This approach gradually lost favour as new theoretical approaches to curriculum development began to emerge. For example, the **Humanities Curriculum Project** (1970), although not designed for use in primary schools, adopted a process model of curriculum development.

Blyth and Clayfield (1982) observed that the limitation of the traditional model of curriculum development lay with the effectiveness of the dissemination and implementation stages. Dissatisfaction with the failure of projects to have much impact on the practice of teachers led to the development of this process model *"in which the emphasis was laid on ways in which teachers and others can effectively develop their own curriculum"* (Blyth and Clayfield, 1982, p.158). Within the Schools Council it tended to supercede the research, design and development model associated with more traditional curriculum planning. As Watkins argued

*"some means must be found of creating for teachers who are not involved in the development phase an experience which will enable them to modify their insights in a way in which 'initiates' have already done. To provide for this, future projects should invest in teacher development alongside curriculum development."* (Watkins, 1972, see Whitehead, 1980, p.21)

This led to the development of such projects as the Schools Council's **Progress in Learning Science** (Harlen et al., 1977a, 1977b) and **Communication Skills in Early Childhood** (Tough, 1976, 1977) projects. These did not provide much in the way of new curricular materials but focused on methods of helping teachers to review the curriculum through the appraisal of children's learning and to devise their own ways of helping pupils develop appropriate skills within the classroom. Both projects were linked to programmes of teacher in-service education. Thus the emphasis of curriculum development was shifting from the identification of what was to be taught to consider



how teachers could develop and operationalise new ideas in their own classrooms.

Stenhouse also argued for curriculum development to be seen as an attack on the separation of theory and practice. He identified the

*"central problem of curriculum study as the gap between our ideas and aspirations and our attempts to operationalise them"* (Stenhouse, 1975, p.3)

and that the only way of closing the gap was by adopting

*"a research and development approach to one's own teaching, whether alone or in a group of co-operating teachers."* (Stenhouse, 1975, p.3)

More recent guidelines (Barnes, 1982; Schools Council, 1983) have, also, stressed the need for the involvement of teachers in the process of curriculum development and offer a number of different starting points for primary curriculum planning. As Barnes stated

*"this approach assumes the importance of colleagues taking joint responsibility for the curricula they are making available to their pupils."* (Barnes, 1982, xv)

Barnes (1982) also developed and extended the description of the primary curriculum by identifying five bases which teachers could use in their planning. These were *content*, *concepts*, *skills*, *problems* and *interest*. Each of these starting points was considered more appropriate for certain areas of the curriculum and teaching methods than others. A *content* curriculum that could be based on an area of experience, a body of knowledge, or a group of phenomena, was suitable for most subjects; for example, science, history, geography, environmental studies, religious education and literature, particularly at the primary stage. A *concept* curriculum was appropriate to all subjects particularly at advanced levels. *Skills-based* planning was most suited to subject areas such as mathematics, literacy and craft which were more easily defined in terms of skills than in any other way. A *problem-solving* approach was seen as a means of enabling pupils to apply skills in more complex and realistic situations or as a means of indirect access to concepts. This approach could be used anywhere, but was most useful in skills-based subjects such as mathematics, design and physical education. Finally, in planning based on *interest* the onus was placed on learners to select issues important to them, the teacher's role was to offer a starting point and a range of possible options and methods. This classification allowed for a number of different methods of teaching, approaches to

learning and descriptions of the nature of the curriculum to be taught. The Schools Council in their publication *Primary practice* (1983) described the curriculum under similar headings to those used by Barnes (1982). These categories for planning were given as subjects, process, the study of problems, areas of knowledge and experience and as seen through a child's eyes.

Thus three trends that have contributed to the development of the primary curriculum have been identified in the preceding discussion. These were the moves towards infant patterns of organisation with an integrated day and more pupil choice; the extension of the curriculum to include a greater number of areas of study, an emphasis on processes and pupil interests and finally a commitment to involve teachers in the process of curriculum development and thus improve the implementation of new curricula. These trends were particularly relevant for junior and middle schools in the late 1960's and 1970's, since it has been generally accepted that progressive practice was more widely in evidence in the nursery and infant sectors by that time (Whitbread, 1972).

As has been suggested previously with regard to changes in patterns of school and classroom organisation such recommendations do not necessarily result in changes in teachers' thinking or observed teaching styles so it is now appropriate to consider the impact of such curriculum developments in primary schools.

Ashton *et al.* addressed this issue in a study of teachers' opinions of the aims of primary education (1975). Teachers' groups were established and through discussion identified an extended range of seventy two aims of primary education (Appendix 1). These aims reflected progressive practice, whilst not neglecting the basic skills (Ashton *et al.*, 1975) identifying knowledge, skills and qualities within the areas of intellectual, aesthetic, personal/emotional, physical, spiritual/religious and social/moral development. This was the first survey which systematically investigated teachers' perceptions of the purposes of primary education. Each teacher rated the seventy two aims as of utmost importance; major importance; important; minor importance; no importance or that it should not be an aim. In 1971 when this survey was carried out a number of aims were considered to be of utmost importance, in that they achieved a mean ranking of at least 4 on a five point scale. These aims included reading, happiness, individuality, positive attitudes to school and good behaviour.

Whilst there was considerable variation between the views of the teachers who



responded to the questionnaire it can be argued that their aims reflected a progressive curriculum perspective overall since these were more highly rated than those relating to basic skills in mathematics and literacy. This discussion has not, however, considered the question of how far the curriculum has been influenced *in practice* by such curriculum research and development. It may well be that the guidelines and recommendations for a progressive primary school curriculum described in the preceeding sections, whilst achieving acceptance in principle by teachers (Ashton *et al.*, 1975), have not been implemented in schools to any great extent.

Research into curriculum practice in the primary school has most often been described in terms of the informal - formal or progressive - traditional dichotomy, on the assumption that curriculum organisation and implementation would reflect one or other of these perspectives. Studies described previously have focused on the primary curriculum in terms of organisational changes; such as grouping and open plan schools, or on aspects of classroom life; such as teacher-pupil interaction, the operation of the integrated day and perceptions of the teacher's role. Other research has focussed on the impact of progressive or traditional schooling on pupil achievement at this time (Daniels, 1961; Gardner, 1966; Wiseman, 1967; Barker Lunn, 1970; Bennett, 1976).

Evidence relating to the primary school curriculum described by Bennett and Jordan (1975) in their study of teaching styles made a distinction between curricula characterised by subject teaching as opposed to integration of subject matter and the degree of pupil choice of activity. The existence of pupil choice or an integrated, thematic approach to the curriculum was associated with a progressive style of teaching. This was shown to be rare, particularly in the upper age groups (Bennett, 1976). A subject-based curriculum with little or no pupil choice which emphasised the basic skills was associated with a traditional style of teaching. This was more commonly found but the majority of teachers tended to use a mixed style employing aspects of both approaches. Further evidence from the study of teaching styles conducted by the Oracle research programme (Galton *et al.*, 1980) also suggested that a progressive primary curriculum as characterised by the Plowden Report (1967) was not much in evidence in the schools studied between 1975 and 1980. Again teachers appeared to use a range of curricular strategies with an emphasis on the basic skills of language and mathematics in much of their teaching.

Another useful indicator of the emphasis placed on traditional and progressive curriculum activities was the amount of time allocated to particular types of activity: for

example, separate subjects versus integrated, thematic work; basic skills work in mathematics and language versus creative activities such as drama, art and music.

It has been shown that a heavy commitment in terms of the time given to the basic skills of mathematics and language was characteristic of primary schools. Bassey (1978) in his survey of 900 primary school teachers reported that the average pupil in a junior school classroom with a school week of twenty seven and a half hours, spent five hours on mathematics, six hours on language, three hours on physical education, one hour on music, two hours on art and craft and four hours on thematic studies (an umbrella term used to include topic and project work as well as separate subjects). The remaining time was spent on assembly (two hours), administration (one hour) and playtimes (two and a half hours). This pattern of organising the curriculum emphasised the basic skills and did not show any significant differences in the range of subjects when compared with practice in primary and elementary schools throughout the twentieth century (Schools Council, 1983, p.26-27). In addition, the questionnaire items relating to integrated studies in the Bassey survey were not responded to in 75% of the cases, suggesting that this was not a common feature of the primary school curriculum.

This evidence would support the conclusion that there has been some change in the nature of the resources and learning materials used in primary schools. Some curriculum development projects have made an impact on the activities of pupils and styles of teaching (Steadman *et al.*, 1978) but basic skills still predominate. Teachers have also incorporated an extended range of aims of education into their thinking about primary schooling, but as with the impact of organisational changes in the primary school classroom, very little change in the curricular emphasis in practice appears to have occurred.

In more general terms there has been much debate as to the true extent of the spread of progressive practice in primary schools in Britain through the 1960's and 1970's. Much of the literature on progressive practice could be argued to relate to the provision for infants not juniors, indeed many of the American reports detailing the changes in British primary schools dealt with younger pupils only (Brandt, 1975; Spodek and Walberg, 1975; Hearn *et al.*, 1973). The more recent survey (D.E.S., 1978a) of primary education conducted by the Her Majesty's Inspectors between 1975 and 1977 has also revealed that far fewer schools were part of what the Plowden Report (Plowden, 1967, p.188) described as the "*quickenning trend*" towards child-centred education than many people



were prepared to believe. However, there were greater opportunities for the development of a progressive primary curriculum during the late 1960's and 1970's. This was clearly reflected in the way teachers thought about the purposes of education as illustrated in the Ashton *et al.* survey of teachers' opinions (1975). These trends towards a more progressive approach to curriculum have considerable implications for the monitoring of standards and the assessment of pupil performance.

Rance (1971) identified a number of consequences of such developments. One result was the breakdown of subject barriers in the school curriculum. Also, changes in the organisation of the classroom meant that it was no longer necessary for every child in every class to be studying the same problem at the same time. It has been noted previously that the demise of the 11+ also helped create the climate which allowed for the development of some measure of progressive practice, in terms of classroom organisation and unstreaming which contributed to the need for more appropriate assessment practices in the primary school. Teachers acquired a far greater degree of autonomy and control by removal of this external constraint, with its associated pressure to achieve maximum success on the 11+ examination. This applied not only to classroom organisation but also to the curriculum. However, with this greater autonomy came a greater responsibility to monitor the extent of pupil learning.

How then were teachers to fulfill this responsibility for the assessment of pupil learning effectively? Given the nature of the curriculum trends described it would appear that schools needed to be very clear about their goals, how these might be expressed and ways in which it was possible to check whether these goals were being achieved. Thus if a shift occurred in the primary school curriculum from basic skills and measurable products to an emphasis on processes and therefore less easily quantifiable outcomes it was likely that this would lead to some confusion regarding the assessment and evaluation of pupil learning. Two different approaches to solving this problem have been suggested in the literature on curriculum evaluation. These approaches are well illustrated in the debates relating to objectives and their role in defining the curriculum and the assessment of pupil learning which are now considered in more detail.

The specification of objectives was central to the development of rational curriculum planning models. Initially, these objectives were pre-specified and described in behavioural terms (Tyler, 1949; Bloom (ed), 1956; Taba, 1962; Mager, 1962). This enabled the evaluation of the effectiveness of a new curricular programme to be made in

relation to the degree of success in achieving these objectives, as measured by *summative* procedures at the end of the period of instruction. Ashton *et al.* argued that

*"if the teacher's aims are to help guide his practice then they should be expressed in behavioural terms; that is to say they should state what the child will actually be able to do when the aim is achieved."* (Ashton *et al.*, 1975, p.15)

This approach to planning by the pre-specification of objectives solely in behavioural terms has been seen by some to be *"most seriously disturbing"* (Blenkin and Kelly, 1981, p.69) since a more goal-orientated, content curriculum was emphasized rather than the process curriculum which was being advocated in the late 1960's and early 1970's.

Dearden (1976) argued that in a progressive primary school teachers should avoid making aims that were prescriptive of content or pupil behaviour by concerning themselves with *"relational"* aims; that is, aims that focused on the child's development of a positive attitude to learning, intrinsic interests, self-expression and autonomy. Eisner (1979) also argued strongly against the use of behavioural objectives alone in curriculum planning stating that

*"goals are not always clear. Purposes are not always precise ... many of our most productive activities take the form of exploration and play. In such activities the task is not one of arriving at a preformed objective but rather to act, often with a sense of abandon, wonder and curiosity, out of such activities rules may be formed and objectives may be created."* (Eisner, 1979, p.100)

More recently, however, the Schools Council recommended that

*"more detailed objectives are needed whatever the approach adopted by the school in describing the curriculum (in order) to plan programmes of work and decide on appropriate ways of assessing their pupil's learning."* (Schools Council, 1983, p.34)

Such activities could only be described using terms such as understanding, insight, appreciation and interest which could not be observed in behavioural terms only inferred from the child's actions. These alternative ways of looking at educational aims and objectives were described as *problem-solving objectives* and *expressive outcomes* and were seen as important adjuncts to behavioural objectives in curriculum planning and evaluation.

Two major strands of development should be taken into account in identifying the



most appropriate way of assessing how far a school has achieved its curricular goals independent of the ways in which the aims and objectives were specified. The first derived from the work done in mental testing and the measurement of achievement, the second from the informed judgements that have been used to evaluate pupil behaviour, teaching methods and materials, as exemplified by the reports prepared by Her Majesty's Inspectorate.

The first strategy enabled the achievement of objectives to be measured in experiments which supposedly controlled non-relevant variables. This was described by Tyler as

*"essentially the process of determining to what extent educational objectives are actually being realized by the program of curriculum and instruction. However, since educational objectives are essentially changes in human beings ... then evaluation is the process for determining the degree to which these changes in behaviour are actually taking place."* (Tyler, 1949, p.105)

This approach to assessment and evaluation was used in the Oxford Primary Science Project (Harlen, 1967). Pupils were tested at the beginning of the trial and on three subsequent occasions during the year. There were two experimental groups and two control groups and the tests were devised to detect behaviours related to the objectives of the experimental programme. The results showed that generally the experimental groups were achieving the objectives but there were differences in their performance between objectives and from one class to another. Harlen noted that

*"it was not possible from the evaluation to learn anything about the conditions which might promote achievement, or how to improve the parts of the material which had been least successful."* (Harlen, 1976a, p.31)

The second strategy was exemplified by the Nuffield Junior Science Project. During 1964 and 1965, some 200 teachers were involved in drafting teaching materials. The teachers were asked to write reports on their work but there were no tests for the children, only general statements about what they might achieve through the work. Crossland justified this approach by saying that

*"the difficulty is that no-one knows what primary children can do and there would be dangers in setting objectives which might become, needlessly, limits."* (Crossland, 1967, p.6)

In the evaluation of the project he noted that

*"the evidence for children's achievements was gained by summarising the opinions of teachers and organisers and from my own limited observations of questioning the children." (Crossland, 1967, p.6)*

There were obvious limitations to both these approaches: one was too restricted to outcomes which could be readily assessed, whilst the other in focusing on processes was more difficult to quantify and subject to observer error. Both these strategies have subsequently been modified and brought closer together. The empirical objectives approach to curriculum evaluation has become less restrictive; looking for unintended as well as intended outcomes. Evidence of processes as well as products has been collected and informal procedures used to monitor these. The informed judgement approach to evaluation has become less unreliable since the development of techniques for classroom observation and interaction analysis.

Such approaches indicated a possible strategy for the development of more appropriate evaluation and assessment procedures for the progressive primary school curriculum. This was exemplified by the Schools Council project Science 5-13 which attempted to blend the behavioural objectives model with a more process-orientated approach in curriculum development and evaluation (Harlen, 1975). The project made three major assumptions. First, that the mode of discovery-based and enquiry learning was the most appropriate to scientific development. Second, that a hierarchical construct of concepts could usefully be used, in this case Piagetian stages of development. Finally, the team considered that these two aims could be realized through an objectives-based approach to curriculum planning.

The team further developed their integrative approach in the evaluation of Science 5-13 with the preparation of checklists to monitor and assess pupil learning in the classroom in the Progress in Learning Science project (Harlen *et al.*, 1977a, 1977b). The project aimed

*"to help teachers gather information about their pupils' abilities, concepts and attitudes and to use this information both for making decisions about classroom experiences and for keeping a record of progress." (Harlen et al., 1977a, p.12)*

This approach was adopted since it was recognised that no appropriate assessment procedures were available that would match the concepts and processes that were an integral part of the Science 5-13 project. Groups of teachers were involved in the development of checklists to match the project's objectives. The substance of this work



was intended to help these teachers' diagnostic abilities so that they would be able to match the child's level of understanding with appropriate experiences that would challenge and develop pupil thinking. In addition, teacher development was further promoted through an in-service course consisting of discussion documents and extensive audio-visual resource materials. In this way the project team developed an approach which took account of the criticisms previously made of both the objectives model and the more subjective methods of assessing pupil learning and evaluating the curriculum.

As suggested previously, the development of the primary curriculum was characterised by a tension between approaches to education that began from a concern with the end result and those which focussed on the child and processes of learning. This was reflected in the debate on the nature of aims and objectives; pre-specified objectives being seen as restrictive and favouring products through a content-based, basic skills curriculum whilst relational aims, problem-solving objectives and expressive outcomes have been seen as enabling a process-orientated, child-centred curriculum to develop. This conflict between processes and products has been illustrated in the developments of methods used to evaluate the curriculum (Tawney (ed.), 1976; Hamilton, 1976).

This debate was not, however, apparent to the more general public. By the mid 1970's, it appeared to the outsider that there had been a dramatic change in the pattern of primary education at an organisational level; with vertical and family grouping, mixed ability, the building of open-plan schools and reorganisation to allow for different ages of transfer. In terms of the curriculum, there was no obvious indication that pupils were learning the basics. Table charts with stars indicating the progress of individual children were not longer to be seen. With less emphasis on phonic methods of teaching reading children no longer came home with their "tin of words". The curriculum was integrated with few recognisable subjects on the time-table and unfamiliar activities like new mathematics, project work and creative writing appeared. The greater emphasis on creative writing meant that the children's work was often displayed on the walls of the classroom irrespective on the correctness of the grammar.

Not surprisingly, parents found it difficult to understand the complex theories behind these changes and the public at large readily jumped to the conclusion that the primary school no longer encouraged children to learn their tables, do simple arithmetic and learn to read. This view was reflected, some would say encouraged, in the media with even the influential Times labelling primary teachers of the day as the "*wild men of the classroom*"

(Times editorial, October 13, 1976). Thus, the next question to be addressed in this discussion relates to the factors which contributed to this change in perception and the impact of such pressures on teachers and assessment methods, in particular those attempting to develop an approach approximating to the Plowden model of a progressive primary school.

## **1.5 THE ACCOUNTABILITY MOVEMENT: ITS DEVELOPMENT AND IMPACT ON SCHOOLS AND TEACHERS**

It would be oversimplistic to attribute the change in the public image of education solely to the failure of schools to communicate effectively with their clientele and to the sensationalism of the media. At a time when the economy of the country was in decline, other factors concerned with the economics and the structure of the education service also contributed to an unease about modern primary practice. The Black papers (Cox and Dyson, 1969a, 1969b, 1970, 1975, 1977) argued with conviction that there was a decline in the quality of education and standards; that the non-competitive ethos of the progressive school undermined the nation's ability to compete in world markets; that the non-selective comprehensive school worked against clever working class pupils receiving an academic education and that all normal children should be able to read by the age of seven. Another aspect of the debate centred on the view that the service did not provide value for money since it was suggested that standards fell as expenditure on education increased. The Society of Education Officers identified a number of questions which highlighted this concern.

*"Why after 30 years of a sustained high level of investment in education are standards not higher, opportunities more equal? Why are so many of our children (20%) still grossly backward? In backing education heavily to win for society has the nation been backing the wrong horse?" (Society of Education Officers, 1975, p.103)*

Doll identified the declining school population in the United States, with the subsequent decrease in school staff and reduction in the total expenditure on education as the context where

*"objectives of teaching and learning are receiving new attention and performance criteria and other standards of accountability are being sought and applied." (Doll, 1974, p.12)*



He concluded that

*"schools are now caught in simultaneously operating processes of "loosening" their programs in the open school during a revival of progressivism and of "tightening" their programs in response to national assessment and state-wide testing." (Doll, 1974, p.12)*

Both these processes identified by Doll of "loosening" and "tightening" were equally relevant to British primary education in the mid 1970's.

The political and economic climate was changing. The co-operative, partnership view of education during the growth economy of the 1960's which was characterised by Ginsberg *et al.* as the *"vehicle of social reform and investment in education (which) was seen as essential to economic success"* (Ginsberg *et al.*, 1977, p.6) was no longer acceptable. The climate of the 1970's was one where the demands of interest groups to be more involved in the control of the curriculum and the requirement for greater school and teacher accountability became politically quite acceptable in a cash restricted economy (Kogan, 1975; Fowler, 1981). Education was the scapegoat, apparently it had failed to meet the needs of the economy, particularly the wealth-creating sectors of industry, since the required skilled labour force and entrepreneurial spirit of enterprise were not in evidence.

MacDonald argued, however, that the accountability movement was not primarily a response to the financial climate but more of an administrative pressure and noted that

*"managerial concern for accountability was an effect of educational and local government expansion in the 1960's, a consequence of the need for more sophisticated thinking about increasingly complex administration." (MacDonald, 1978, p.28)*

Education officers supported this need for greater accountability so that output measurement would provide information to assist decision-making in a more effective service and cautioned that

*"single and simple measures will seldom be appropriate: they will need to be correlated with other measures often relating social and economic factors." (Society of Education Officers, 1975, p.113)*

Such views and concerns led to mounting pressures on schools which gathered momentum during the late 1970's. Demands for increased teacher accountability, for the

monitoring of standards and for greater public control of the curriculum have all had their effect on the work of primary schools. It has been argued (Blenkin and Kelly, 1981) that it was the sequence of events at the William Tyndale School (Auld, 1976) that triggered off this process. Whatever the initial catalyst, it was further fuelled by the Great Debate following on from Callaghan's Ruskin College speech (1976), the media attention afforded to the research on primary schools conducted by Bennett (1976) and the recommendations of the Taylor Report (D.E.S., 1977b) for more parental involvement in schools.

Much of the criticism of primary schooling which found expression in the demand for more accountability within the educational service centred on the impact of progressive methods on the quality of pupil learning. There were concerns expressed, such as in the Black Papers (Cox and Dyson, 1969a, 1969b, 1970, 1975, 1977), that standards of attainment in the basic skills were falling and the assumption was made that this was the result of progressive practices. This issue of falling standards became the subject of a political and economic debate as well as a question for educational research to address. The literature on the relative merits of progressive or traditional methods is now fairly extensive, but no firm conclusions have been drawn in favour of one or other pedagogy.

Earlier research focused on the differences between schools using such progressive and traditional methods. Daniels compared two streamed and two unstreamed junior schools and concluded that, as well as significantly increasing the mean scores of pupils in reading, English tests and arithmetical attainment there

*"appears to be fairly definite evidence that the policy of non-streaming, as compared with streaming, significantly increases the average I.Q. of children in the junior school by about three points." (Daniels, 1961, p.127)*

Lovell suggested that

*"overall there is no evidence whatever of any deterioration in reading standards in informal junior schools. Although there is no evidence that these schools bring superior standards in reading they may well benefit children in other ways." (Lovell, 1963, p.76)*

The work of Gardner (1966) was carried out in response to a growing number of experimental schools in the junior sector. The study compared the results of pupils on a range of tests and assessments at the ages of seven and ten in progressive schools to those of pupils in more traditional settings identified as controls. It was found that on the



whole experimental schools had more instances of surpassing their controls except in the case of arithmetic, especially in mechanical arithmetic, where the controls were slightly ahead. There were five tests in which no control school was superior. These were listening and remembering, general information, the English paper 2, interests and tests of moral conduct. The experimental schools were found to be superior overall in listening and remembering; neatness, care and skill; ingenuity; free drawing and painting; English; composition and interest.

Evidence of the impact of progressive methods on the performance of pupils indicated that schools adopting progressive methods had fewer "*backward children*" (Warburton, 1964) and more "*bright*" children (Wiseman, 1967). This effect was noted to be more marked with reading than with arithmetic. Wiseman also noted that there was also

*"an outstanding improvement in the energies and enthusiasms of the children and their attitudes towards school and learning."* (Wiseman, 1967, p.110)

These progressive methods were also associated with high scores on tests of divergent thinking (Haddon and Lytton, 1968). The Plowden Committee reported results of a cross-sectional study of attainment in matched streamed and unstreamed schools conducted by the National Foundation of Education Research in 1963 and noted that

*"at seven and at ten children in streamed schools did somewhat better than those in unstreamed schools ... By ten the lead of children in streamed schools had been reduced in all tests and there was no significant difference in reading."* (Plowden, 1967, Vol.II, p.291)

Other results (Mc V. Hunt, 1961; Gardner, 1966) also suggested that children taught by informal methods made a slower beginning but caught up later.

In 1970, Barker Lunn conducted a longitudinal study of children in a sample of streamed and non-streamed junior schools. Evidence was reported based on pupil achievement in English tests, problem and conceptual arithmetic tests collected for two types of teachers; one group classed as progressive (unstreamed), the other group as traditional (streamed). Comparisons showed that while progressive teachers' pupils did fractionally better in the English tests, the traditional teachers' pupils did decidedly better on the two arithmetic tests. Acland (1973) later urged caution in interpreting the outcomes of such surveys and argued that the size of the impact of streaming was not big enough to justify the replacement or retention of streaming in terms of its effects on academic

performance. In the United States similar research was carried out by Silberman (1970) which also suggested that informal schools did not achieve greater academic standards.

Bell *et al.* (1976) found that formal groups were more proficient in reading and in mathematics but a derived score estimating "*academic success*" was higher for the informal groups. In a later study Bell *et al.* (1977) compared 40 children randomly selected from traditional classrooms with 40 children from open space or informal classrooms and found that the former were superior in mathematical problem solving and reading but less willing to accept responsibility for "*negative intellectual events*".

By this time concerns were also being expressed about the less able and the gifted child in the progressive primary classroom (Ogilvie, 1973; Tempest, 1974). Evidence regarding the impact of changes in practice on these groups suggested that removing streaming led to a reduction in the spread of attainment between the most and least able pupils (Daniels, 1961; Bennett, 1976). It has not been demonstrated if this decrease was due to less able pupils doing better and/or the more able pupils doing less well. However, with progressive methods the average and below average pupil revealed more positive attitudes to teachers (Barker Lunn, 1970).

By the middle 1970's, there were far fewer streamed classes and attention focused on describing teaching methods in more specific terms. The crude dichotomy between streamed and unstreamed, progressive and traditional schools was questioned and research initiated to explore more fully the range of primary practice and the impact of this on pupil achievement. Researchers focused on the description of teaching style. Bennett (1976) identified three general teaching styles: *informal*, *mixed* and *formal*. These were derived from amalgamating twelve teaching styles which he identified from the analysis of questionnaire data relating to such variables as classroom organisation, pupil choice, curriculum and assessment methods. In comparing informal and formal teaching styles, Bennett came to the general conclusion that pupils taught by formal methods made more academic progress than those taught by informal methods. However, a subsequent re-analysis of the data (Aitken, Bennett and Hesketh, 1981) suggested that no significant differences in outcome were attached to the traditional - progressive dichotomy.

There has been a continuing debate in recent literature (Anthony, 1979; Gray and Satterley, 1981; Anthony, 1982) as to the merits of the apparently conflicting results of the earlier studies carried out by Kemp (1955), Warburton (1964), Gardner (1966) which



favoured progressive schools, when compared with those of Barker Lunn (1970), Bennett (1976), Gray and Satterley (1979) and Aitken *et al.* (1981) which appeared to favour traditional teaching. Anthony argued that "*pupil interest appears to be a condition for progressive teaching to work well*" (Anthony, 1982, p.381) but the descriptions of progressive teaching offered by Barker Lunn, Bennett, Gray and Satterley did not take this factor into account.

This debate has been further extended in teacher effectiveness research where the adequacy of measures used to assess pupil achievement have been criticised (Gardner, 1966; Elliott, 1978c). Even Bennett (1976) argued that teachers aimed to engender different outcomes in their pupils. Formal teachers have tended to lay greater emphasis on promoting a high level of academic attainment and the acquisition of skills in number and reading, whereas informal teachers have valued social and emotional development, stressing the importance of self-expression, creativity, enjoyment of school and the development of independence through the facilitation of enquiry and discovery learning. If these were teachers' aims then it must be argued that they have some impact on the quality of pupil outcomes. The effectiveness of a teacher or a primary school should not, therefore, be measured by reference to the attainment of pupils in a narrow range of cognitive skills and pupil progress should not be defined solely in terms of reading, mathematics and English achievement.

In the context of the accountability movement Barth (1975) also argued that it was quite illogical for a school to be censured by an outside agency for not producing outcome "A" when its whole philosophy and practice were directed to outcome "B". This has raised the question of how schools which aimed to produce such alternative outcomes could demonstrate their effectiveness and account to society for their actions, and has provided a further justification for developing new assessment procedures within this research.

However, in the climate of accountability in the late 1970's, the focus of assessing pupil performance was shifting back to the basics. Both parents and politicians readily evaluated the work of the primary school by reference to the scores pupils attained on tests of reading and mathematical competence; local authorities also yielded to the pressure and made the same type of narrow evaluation of primary schools. At the national level, the Assessment of Performance Unit, although originally established in 1975 to monitor standards in science, mathematics, language, aesthetic, physical, personal and

social development, has focused on monitoring science, mathematics and language.

This focus on assessing basic skills achievement using some form of standardised test was, however, in direct conflict with the ethos of primary schools at that time, since a climate of anti-testing had arisen within the primary school with the advent of progressivism which extended to all types of grading and assessment. This was partly because the reliability and validity of the 11+ as a predictor of future success had been questioned (Vernon, 1957) and partly because group standardised testing of the three "R's" had not provided the necessary information on individual children's development in line with progressive ideology. Alternative methods which involved the classification of children were also suspect. Blyth noted that

*"tests and record cards alike are carefully designed and the former are creditably free from bias, but both are of limited value for prognosis. Now, as more and more local education authorities are abandoning the eleven plus examination, various series of standardised tests are increasingly used as a principle basis for secondary school selection or for initial assessment within the comprehensive school ... where objections are raised to this type of classification of children ... they embody a protest against over-emphasis on the apparently measurable and predictable aspects of the educative process in general."* (Blyth, 1965, p.89)

The Plowden Report concluded that *"it is not possible to describe a standard of attainment that should be reached by all or most of the children"* (Plowden, 1967, p.21), although they recommended that recurring surveys of reading and mathematics attainment should continue. Rance also argued in line with the climate of the time that

*"too efficient and systematic a recording of a child's school life can prove self defeating, especially when based entirely on written or verbal attainment tests, for these can be disastrous for the growth of the primary school child if they are subsequently used as a basis for streaming at the secondary level."* (Rance, 1971, p.13)

In describing the practice of schools in the late 1960's, Dean identified the move away from marks and grades as related to

*"the feeling that each individual is unique and is not really comparable with other individuals; the feeling that if something is worth learning and is right for the stage the child is at, then it should need no artificial boost, but should provide satisfaction in the doing."* (Dean, 1972, p.35)

Pluckrose was even more forceful in condemning the use of marks, grades and stars since he argued that



*"these forms of tabulation cannot show growth of personality or developing social confidence; they cannot indicate group leadership or an emerging sense of responsibility for younger members of the community." (Pluckrose, 1975, p.51)*

He suggested that teachers in open classrooms should be prepared to record progress in both the social and academic fields, but offered little practical guidance on how this might be achieved.

There was some evidence to suggest that the use of marks and grades largely disappeared from primary schools during this time. Bassey (1978) found that the majority of junior school teachers rarely or never used stars, grades, house points or a numerical mark on an individual child's work, preferring to convey approval through the use of ticks and/or written comments. This trend may well have occurred as a result of the climate of anti-testing that arose with the advent of progressivism and criticisms of the 11+. Blenkin and Kelly (1981) suggested that the increased demand for cumulative records of a pupil's progress at school was a result of curriculum development that involved changes in objectives, organisation of subject content and teaching method as well as providing some form of account of school or teacher effectiveness. They also recommended that such records should include statements of achievement in the cognitive field and also contain qualitative judgements about children's aesthetic, emotional or moral development. The solution to keeping track of pupils in the primary school was seen in terms of extensive school records, supplemented by teachers' comments and pupils' classwork with fewer tests and grades.

Some local education authorities made the effort to co-ordinate and standardize the record of a child's progress through their schools; for example, Stoke on Trent and Surrey. These records usually included aspects of the child's health and physical development, personality, academic attainment, social background, emotional and psychological problems but recommendations for assessing pupils' learning were still seen in terms of standardised attainment tests administered once a year. Teachers, however, still tended to record pupil progress in terms of anecdotal comments and class work. The collection of examples of a child's work: displays, class folders, models, was seen as an ideal way of keeping a record since it saved time, clearly indicated the type of teaching employed and the standard reached by the class or group within the school (Rance, 1971). These records were usually based on anecdotal comments where the teacher selected the information to be recorded; test results, particularly in reading

attainment, were also kept. The responsibility for deciding what was to be recorded was seen to lie with the school or the classroom teacher.

Dean (1972) also noted that diagnostic materials were necessary to help teachers assess such abilities as study skills and concept development and provide a reinforcement or check to personal judgement. This suggested that teachers were expected to rely on their own judgements for the assessment of the progressive primary curriculum and pupil learning, particularly the study skills needed for independent enquiry since alternative assessment procedures were unavailable. This provided a further justification for the development of more appropriate methods of assessment for the primary school curriculum, since teachers were faced with the need to justify their practice in this climate of anti-testing as the accountability movement gained ground.

However, the teaching profession's response to this pressure for accountability was not initially to improve the quality of their judgements nor to develop a more rigorous approach to the assessment of pupil learning but to produce increasingly more complex record keeping systems. The quality of such anecdotal primary school records was questioned since they did not provide the quick and easy guide to pupil performance that had previously been obtained from the use of test results, marks and grades. In addition, emphasis was increasingly being placed on the quality of information provided to secondary schools in the absence of the 11+ examination.

Lynch discussed this problem in relation to middle schools and noted that often inadequate records transferred to upper schools, creating difficulties not just on

*"the curricular ground to be covered but also the style of subjects and teaching, traditional or modern"* (Lynch, 1980, p.110).

The Green Paper Education in Schools (D.E.S., 1977a) also drew attention to the need for more accurate records and the Director of the Assessment and Performance Unit noted that

*"this undoubtedly entails for the 9-13 middle schools a serious obligation to give the greatest possible care to assessment and record keeping."* (Marjoram, 1978, p.35)

In the late 1970's teachers were expected to provide more publicly accessible information on pupil performance. There was a re-emphasis on accountability in the



tradition of payment by results and that achieved through the 11+ examination system. The good school or teacher in the public mind was still the school or teacher with the best results. The comprehensive system and the removal of the 11+ resulted in relatively weak forms of accountability since it became the responsibility of schools and teachers to monitor standards and the progress of pupils. The 1970's had marked a swing towards external agencies, such as the Assessment of Performance Unit, who were charged with performing the task of monitoring standards and an increase of blanket testing by local authorities using standardised tests (Galton, 1978a). Elliott (1980a) drew attention to the link between this emphasis on monitoring procedures and the call for greater school and teacher accountability. The Green Paper Education in Schools (D.E.S., 1977a) also argued that the

*"growing recognition of the needs of schools to demonstrate their accountability to the society which they serve is seen as requiring a coherent and soundly based means of assessment for the education system as a whole, for schools and for pupils." (D.E.S., 1977a, p.61)*

However, teachers' assessments and school records, which it could be argued provided valuable information to schools and class teachers, did not appear to meet the requirements of the local education authorities and central government for monitoring standards and the responsibility for deciding what should be recorded shifted away from teachers.

Recommendations were issued from central government to local education authorities which suggested that monitoring systems should be regularised. The Green Paper, Education in Schools (D.E.S., 1977a), although it rejected league tables of school performance, did recommend that local authorities should *"try to achieve a greater degree of uniformity in their approach to schools"* (paragraph 3.7) and that there should be *"consistency within local education authorities and wherever possible between authorities"* (paragraph 9) with regard to monitoring standards. Most local education authorities now have standardised record cards throughout their schools in response to the demands for greater accountability. For example, Rochdale (1978) has a book which has to be completed for each pupil every school year in respect of detailed lists of concepts, skills and qualities.

More recently critics of these moves towards externally imposed accountability systems have examined the impact of such procedures on schools. Alexander (1984) noted that one outcome of this shift of emphasis to external accountability was the

internalisation of societal pressures and perceptions. These were then operationalised in the school curriculum which resulted in child-centred and integrationist arguments being overridden.

*"Curriculum areas concerned become the object of research ... they become more systematised ... the less, therefore, it is possible to sustain the child-centred argument that the area must be defined in terms of how the child perceives the world ... Codification and systematisation widens the gap between the basics and the rest of the curriculum."* (Alexander, 1984, p.68)

Blenkin and Kelly have also argued that

*"to view and evaluate the curriculum of the primary school by reference to goals or purposes external to it, no matter how those goals or purposes are defined is to put at risk the essence of that curriculum, which is its focus on the processes of education. If education is concerned with intrinsic value, then it cannot be defined by reference to any external goal."* (Blenkin and Kelly, 1983, p.21)

However, within the educational system goals have increasingly been set by the local education authorities through monitoring by means of standardised tests and by central government through recent publications on the curriculum (D.E.S., 1983b). Alexander (1984) also noted that this process was much in evidence in local education authority record cards which became consistently more elaborate as a response to the increasing accountability pressures from central government. For example, the Inner London Education Authority's primary record (I.L.E.A., 1979) illustrated the kind of detail and emphasis expected in such records. It has ten curriculum headings as follows: oral language, reading development, written language, number, measures, geometry, sets, creative abilities, handwriting, topics/projects. A wealth of detail was invited for language and mathematics with pre-specified skills and concepts; but only broad headings provided for the rest of the curriculum, where the teacher was not invited to specify what if anything the pupil has learnt. This format stressed the importance of the basics and Alexander argued that

*"if there is not some codification of the skills, attitudes and concepts concerned in the curriculum areas such as the humanities, art or drama then there can be no strong case or justification for twenty minutes of the week let alone three hours."* (Alexander, 1984, p.69)

The most likely outcome of such pressure is, therefore, a decreasing emphasis within the curriculum on those activities which have characterised a more progressive school. Also, teachers increasingly lose control over curriculum content and processes as they



become more constrained by the demands of external monitoring of their performance and pupil learning. This was in direct contrast to the control afforded to teachers in the 1960's and the early 1970's where the curriculum and assessment of pupils was seen to be the professional responsibility of the teacher.

The societal aspects of such externally imposed systems of accountability have been the subject of much debate in the literature (Rawls, 1971; House, 1973a, 1973b; Sockett, 1976; Elliott, 1976, 1978a, 1979a, 1980a). It was argued that the model was anti-educational in that excellences or education goals could not be defined in pre-specified behavioural terms. The model also assumed that improvement in educational quality was synonymous with raising the level of test scores. Sockett (1980) argued that the model was also politically unacceptable in that it was anti-democratic in reconciling the private wants or needs with the public interest, since it conceived the educational enterprise as subject only to the criteria of public benefit. The teacher who worked in a democratic context was expected to develop a child's wants and expectations within a moral framework; to promote an understanding of what was publicly desirable, what it was to consider matters through the eye of the public interest and to promote understanding and skill in weighing up such conflicting claims morally and prudentially. Sockett argued that no notion of public benefit could permit the development of these three areas of understanding.

The debate on accountability was also centred on the form of accountability that was appropriate to education. It has been argued that if teaching was considered a profession then it was more appropriate for teachers and others within the service to be answerable to their peers and accept a moral responsibility for their actions, rather than have to account for results as was common in the commercial and industrial context. Sockett (1980) addressed three questions relating to this issue: was the school or teacher accountable, to whom and for what should the school or the teacher be accountable?

He argued that it was perhaps more appropriate for teachers rather than the school to be held accountable since it was the accountability of teachers as a group that would be examined. A more complex issue was raised as to whom teachers should be held accountable since there were differences in the views of teachers and other interested parties. The third question also raised a number of difficulties, most importantly, the variety of educational purposes and the lack of clear priorities for teachers. Should they go for outcomes in terms of examination success or should they pursue more general

educational aims? Should they work towards providing for the individual's needs or society's needs? Was it their professional practices, the outcomes resulting from them or both aspects for which teachers should be held accountable?

These moves towards accountability and teacher evaluation by outside agencies have emphasised basic skills attainment and global statements regarding the quality of a school or a teacher's classroom performance, although some local authorities, such as Salford, moved away from such an overt accountability approach. This authority introduced a detailed primary school booklet that covered all facets of primary school life. School assessment was seen as a contribution to staff development, curriculum improvement and organisational evaluation (Salford, 1977, 1982). Thus an alternative conception of accountability began to emerge which was based on the professional development and in-service training of teachers. This approach stressed the need to develop teaching skills, assist the dissemination of new curricula and improve teacher effectiveness at the level of the classroom.

Teachers were, therefore, confronted with two approaches to accountability. The first relied on external agencies determining criteria and methods of assessment which predominated in the United States. The second approach stressed the internal monitoring of the quality of education through school and staff development, which has been the more common response in Britain. This second approach has, however, had to address the question of how, without the necessary expertise, schools and individual teachers can respond to the requirement for more effective internally-controlled monitoring procedures which satisfy both the demands of external agencies and interested parties for accountability without compromising the breadth of educational aims and changed priorities for primary schooling which have been established over the last seventy years. An attempt has been made to achieve this through a variety of approaches to the professional development of teachers.

## **1.6 TEACHER PROFESSIONAL DEVELOPMENT**

The advocacy of teaching as a profession has been a relatively recent phenomenon. An important factor in the development of this view was the James Report on Teacher Education and Training (D.E.S., 1972). The report advocated that teachers should all be graduates, indicating a level of training commensurate with professional status. The



report also recommended that teachers should be entitled to release for in-service training for periods equivalent to one term in every seven years, supporting the importance of opportunities for the professional development of teachers. This move was further endorsed in the Green Paper, *Education: A Framework for Expansion* (D.E.S., 1975b) and *Making INSET Work* (D.E.S., 1978b).

A number of principles regarding the direction of professional development emerged from the report requested in 1974 by the Induction and In-Service Training Sub-Committee of the Advisory Committee on the Supply and Training of Teachers on in-service education and training of teachers (A.C.S.T.T., 1974). The most important of these was that teachers should identify in-service needs within the context of the school, re-iterating the recommendations of the James Report that

*"In-service training should begin in schools. It is here that learning and teaching takes place, curricula and techniques are developed and needs and deficiencies revealed."* (D.E.S., 1972, paragraph 2.2)

These recommendations when coupled with moves towards greater accountability of the teaching profession during the 1970's led to an increased emphasis on school-focused and school-based work centred on the improvement of schools, the curriculum, the quality of teaching and pupil learning. The Oracle research programme was planned in this climate of accountability where responses based on teacher self-evaluation and involvement in research and curriculum development were being advocated. Thus, this research was committed to the involvement of teachers in developing and implementing the various assessments of pupil performance which are discussed in later chapters. A number of alternative strategies for teacher professional development have been proposed which are now considered in more detail, since a critical appraisal of developments in this field was central to the ways in which teachers became involved within this research study.

In Britain, the moves to counter externally imposed accountability systems have been closely linked to the notion of the *"extended professional"* (Stenhouse, 1975, p.143) and ways of implementing and facilitating the professional development of teachers through self-evaluation (Elliott, 1978a, 1980b). This movement has been based on the assumption that teachers determine the quality of education through their activities in schools and classrooms. A number of ways have been developed to enhance professional development. First, the involvement of teachers in curriculum development and

dissemination provided the initial thrust towards the idea of an *"extended professional"*. Second, where teachers could not be directly involved in curriculum research this need was to be met by the provision of more appropriate *"school-focused"* (Eraut, 1972) and *"school-based"* (Elliott, 1977, 1979a, 1980b) professional development for teachers. Finally, teachers were encouraged to become part of the educational research community through participation in *"action research"* (Elliott, 1978b, 1978c, 1979b; Nixon, 1981), *"collaborative research"* (Black 1980) and *"self-evaluation"* by monitoring their own classroom performance (Elliott, 1978a, 1983).

Such a wide range of developments indicated that a quite different approach to teacher accountability emerged in Britain when compared with the United States. It will be argued here that this was closely linked to the tradition of teacher autonomy in the classroom and an increasing involvement in curriculum development and dissemination which set the stage for the acceptance of self-evaluation and professional development as an appropriate response to the pressures for accountability.

Early curriculum developments in Britain were based on the *research, development, diffusion model* of planned organisational change proposed by Havelock (1971). The main problem encountered with this type of curriculum development was the actual implementation of change once materials were introduced into schools. Ruddock and Kelly (1970) conducted research into the strategies and experiences of disseminators in order to establish necessary conditions for successful curriculum dissemination. They concluded that lack of success could be attributed to two main factors. First, teacher re-education was seldom achieved as little emphasis was placed on this type of follow-up activity and second, the importance of motivating teachers was not widely appreciated. This highlighted the need to provide for the in-service training and professional development of teachers to cope with any curricular changes which were advocated. This need was also supported by the Schools Council publication *Dissemination and In-service training* (1972b) which considered the problems of implementing curriculum innovation and supported an increased involvement of teachers in the process of curriculum development.

One example of curriculum development and dissemination which addressed this problem and developed ways of helping teachers change their practice was the Humanities Curriculum Project (1970; Stenhouse, 1971a, 1971b). This research concentrated on the technical problems of operating a discussion-based form of teaching



in which the group, including both teachers and pupils, critically examined evidence whilst considering controversial issues. Both teachers and pupils were seen as learners engaged in developing a shared understanding. The discussion was chaired by the teacher who was expected to take a neutral approach. This role was obviously an unfamiliar one which required a different style of teaching and interaction with pupils. As a result, the role definition of neutral chairperson became the source of controversy for teachers involved in the project.

The project had, therefore, to address the question of the professional development of teachers so that they could undertake this role and operate effectively in discussion. Researchers worked in collaboration with teachers to develop a procedural discipline for discussions of this kind. The research team also advocated a more general involvement of teachers in the research process as a basis for improving the quality of teaching. Stenhouse argued strongly for this approach of "*teacher as researcher*" (Stenhouse, 1975, pp.143-165) as a means of promoting professional development where teachers took control of educational research in their own schools and classrooms. There was evidence that some success in changing teachers' practice was achieved (MacDonald, 1971) despite the difficulties encountered as a result of the background assumptions of both schools and, more particularly, those of the teachers involved.

Alternative strategies to professional development have been based on the notion of *problem-solving* or *client-centred* approaches which were prompted by the intervention of an outsider or someone from within the institution. This approach involved an initial evaluation of the system followed by the introduction of changes appropriate to the problems revealed (Hoyle, 1973). Elliott and Adelman (1975) and Eraut (1972) both proposed client-centred models in which an external agent acted as consultant. Professional development was viewed as a collaborative venture where teachers and significant others negotiated processes and outcomes. However, the teachers were not necessarily the sole researchers, but viewed as partners in the research process which was designed to promote teacher professional development.

The most explicit response to the demands for teacher accountability was based on the idea of self-evaluation for professional development as exemplified by the work of Elliott arising from the Ford Teaching Project. Elliott outlined a theory of method for self-evaluation derived mainly from this project (Elliott and Adelman, 1975) which was an attempt to help 40 teachers in twelve East Anglian schools to develop their competence

at self-monitoring in the classroom. Teachers, he argued, became aware of the consequences of their actions and the extent to which they could be held responsible for them by reflecting on their own practice. This process has three stages: first, the teacher has to "*ask a question*" to identify those actions that have been performed directly at will; second, the consequence of those actions has to be determined and finally teachers have to assess the extent to which they have been responsible for what happened within the classroom. Elliott argued that implicit in the notion of the self-monitoring teacher was the proposition that

*"a person becomes accountable to others whenever he is capable of rendering an explanatory account of his conduct to others."* (Elliott, 1978a, p.51).

This process of self-monitoring was described as a response to societal pressures but it also served as a means of professional development for teachers by emphasising the importance of classroom activities and the need to reflect on the quality of their own classroom performance.

Elliott (1979b, 1980b) has since developed his thinking and practice to support action research for teachers through the Classroom Action Research Network, based at the Cambridge Institute of Education. He has also made distinctions between four kinds of self-evaluation and discussed their implications for the professional development of teachers (Elliott, 1983). The first was the notion of *unreflective self-evaluation* based on *tacit practical knowledge*. He argued that the only type of professional development this could foster was "*development towards competence*" or "*the skilful utilization of existing stocks of tacit professional knowledge*" (Elliott, 1983, p.229).

The second type of self-evaluation was described as *practical deliberation*. This concept was based on the idea that the outcome of deliberation was a decision or choice about the best means of achieving a certain end in a particular concrete situation. This involved reflecting on both means and ends jointly and was grounded in a tacit knowledge of tradition. Elliott argued that this enabled autonomous professionals not only to reproduce and maintain traditional practice but also to develop their practice in ways that were appropriate to changing social conditions thus engaging in a process of self-development.

*"The intent is to enlighten practitioners about the way their activities are subject to hidden constraints, the hidden motives for performing them, and their intended outcomes."* (Elliott, 1983, p.245)



The third type of self-evaluation postulated by Elliott (1983) was based on an explicit knowledge of technical rules of the type frequently applied at the local education authority level when advocating the use of checklists for staff appraisal and school self-evaluation. This was seen as a way of improving competency but not necessarily contributing to professional development, since the checklists served as a reminder of what ought to be done or the "*prescribed rules*" and thus simply enabled teachers to rapidly check their performance against them.

Elliott (1983) also argued for a fourth type of self-evaluation or *consciousness-raising* which involved not only practical deliberation on the part of individuals but also collective political action by the practitioner group as a whole to modify or remove constraining structures. Within the education system this would thus involve teachers in bringing about change in institutional structures that would enable them to develop themselves professionally through deliberation and discussion with each other. Elliott also argued that this process could also be enhanced by the use of external consultants to validate the self-evaluations of practitioners.

The work of Argyris and Schon (1976) on professional development has some bearing on these notions of self-evaluation proposed by Elliott (1983). The basic premise of this theory of professional learning has been that theories of action determine practice and these theories can be attributed to explain or predict a person's behaviour. Within any *theory of action*, two components were distinguished: *espoused theories* and *theories-in-use*. Espoused theories were used to justify or describe behaviour, theories-in-use reflected what a person actually did or how they operationalised their espoused theories. Argyris and Schon maintained that professional practice was made up of a number of inter-related theories of action that specified

*"for the situations of the practice the actions that would, under the relevant assumptions yield the intended consequences."* (Argyris and Schon, 1976, p.6)

Thus, professional development would occur when opportunities were created for discrepancies between theories-in-use and espoused theories to be made explicit and for new theories of action to be acquired so that theories-in-use became more compatible with espoused theories. This process was similar to the notion of self-evaluation advocated by Elliott and indicated a basis for exploring how teachers could be encouraged to collect valid and reliable evidence which promoted reflection on classroom practice and led to

self-evaluation and professional development.

These moves towards encouraging teachers to enhance their own professional development through the study of classrooms and particularly their own teaching, have been linked to changes in educational research methodology. This has taken the form of moving from the experimental and quasi-experimental scientific paradigms (Campbell and Stanley, 1963) to an ethnographic case study approach (Parlett and Hamilton, 1972; Hamilton, 1976, 1980; Stubbs and Delamont, 1976; McAleese and Hamilton, 1978; Cohen and Manion, 1980; Simons, 1980; Stenhouse, 1980b). This shift has been evident in both sociological (Burgess, 1985; Hammersley and Atkinson, 1983) and psychological research methods (Rowan, 1976; Reason and Rowan, 1980; Amabile and Stubbs, 1982; Bennett *et al.*, 1984) applied to education.

This has resulted in increasing numbers of studies of cases performed by teachers within their own classrooms and schools (Enright, 1979; Armstrong, 1980; Gower, 1981; Jackson, 1981; Eyers and Richmond, 1982; Rowlands, 1982) where teachers' practical knowledge (Elbaz, 1982) was seen as an important contribution to an understanding of schools, teaching and learning. Ultimately, it was argued teachers would have the ability to change their educational practice since they were directly involved in the generation of educational theory. In particular, action research was seen as a means whereby teachers could research in their own classrooms as a means of professional development (McNiff, 1984) and contribute to the development of educational theory (Stenhouse, 1975; Elliott, 1978b, 1978c, 1980b; Whitehead, 1980; Grundy and Kemmis, 1981; Grundy, 1981; Torbett, 1981). The ways in which teachers have become involved in such research and the implications of this movement for effectively linking theory and practice are now considered.

Traditionally the role of the teacher in research has been as the "*subject*" of research. They have been asked to allow researchers into their classrooms to observe events and monitor interactions. They have cooperated with researchers; completing questionnaires and giving interviews which have been subsequently analysed and interpreted by the researcher. Teachers have not contributed actively to the formulation of the research questions, the decisions related to data collection or the interpretation of results. However, they have been expected to act on the findings of research projects by modifying their behaviour in the light of the results and also to adopt innovations and change the pattern of the curriculum and their teaching styles as a result of new



developments.

This has, in the past, produced a feeling of separation between theory, as propounded by the curriculum developer and researcher, and practice, as experienced by the teacher in the classroom (Stenhouse, 1975, 1980a; Torbett, 1981). In order to reduce this separation of theory and practice, teachers have been increasingly involved in educational research, either as researchers in their own right using case study and action research methods, or in collaboration with researchers or consultants in the study of particular questions of relevance to professional practice. The principles underlying action research and collaborative research are now considered in more detail since they were important to the conceptualisation of professional development, in particular the establishment of links between theory and practice which are considered in this research.

Halsey (1972) defined action research as "*a small-scale intervention in the functioning of the real world and a close examination of the effects of such intervention*". Cohen and Manion (1980) further refined this definition to describe it as *participatory, situational* - in that it was concerned with diagnosing a problem in a specific context; it was usually *collaborative* and *self-evaluative* since modifications were continuously evaluated within the on-going situation with the ultimate aim of improving practice.

Grundy (1981) and Grundy and Kemmis (1981) have argued that three types of action research can be identified; the *technical, practical* and *emancipatory*. Technical action research was described as initiated by a person or group, possessing authority or expertise, who wished others to achieve certain ends they had conceived. Practical action research was of the type advocated by Elliott and Stenhouse and closely related to the notion of practical deliberation (Elliott, 1983). Emancipatory action research (Habermas, 1971a, 1971b, 1974) was seen as complementary to this but going beyond it, in that it required the collective action of a group within an institution to counter the restrictions imposed by organisational structures. This type of action research was also characterised by the generation of critical theorems which identified distorted practices and self understandings, and explained the mechanisms that caused them.

Collaborative research was an alternative strategy advocated by some researchers to promote teacher professional development (Black, 1980; Hunter, 1980; Jasman and Ashby, 1980).

Black (1980), described the Diagnostic Assessment Project of the Scottish Council for Research in Education as a collaborative research project in that

*"the particular activities to be carried out were to be the subject of negotiation between teacher and researcher collaborating as equal professionals, the former as subject expert and practitioner and the latter as "enabler", "agent provocateur" and "assessment expert"."* (Black, 1980, p.3)

Although the development of collaborative research has been closely linked to *"teachers as researchers"* and action research (Smetheram, 1980), an essential difference has emerged between the roles of the teacher and researcher in these two models. In collaborative research these were distinctive, each participant offering a particular expertise rather than the teacher taking on the role of the researcher, whilst *"teacher as researcher"* implied teachers becoming researchers, usually employing action research and case study methodologies.

Thus the quality and degree of involvement of teachers in determining the agenda and methodology for the research has varied depending on the emphasis and focus of the particular problems addressed. For example, Stenhouse and Elliott have advocated the use of action research as a means of addressing the study of classrooms. Teacher involvement has been argued for by Stenhouse (1975) who emphasised *"teachers as researchers"* and the use of case study approaches when studying the curriculum and developing educational theory (Stenhouse, 1980b). Whilst, Elliott (1978a) has developed ideas on teacher self-evaluation as a result of the Ford Teaching Project (Elliott and Adelman, 1976) and pressures for increased school and teacher accountability.

Other directions designed to promote professional development have emerged from these original strategies. Ways in which outside consultants can support school-based curriculum development utilising a client-centred model of professional development have been considered (Miskimmin, 1980; Verrier, 1981). As a development of the *"teacher as researcher"* movement ways in which teachers can assist each other and undertake their own classroom research have been proposed (Ebutt and Partington, 1982; Ruddock, 1982). Finally the advocacy of increased involvement for teachers in their own professional development has led to consideration of how in-service providing institutions can contribute to school-focused in-service (Golby *et al.*, 1981, Chambers and Powney, 1982) and the linking of in-service education with initial teacher training (Ashton *et al.*, 1982).



Such recommendations for teacher involvement in research or professional development have not, however, addressed the issue of how teachers identify the need for "confrontation" of what they actually do in the first place, their ability to carry out effectively the study of their own teaching or engage in the generation of educational theory. They also appeared to ignore the constraints which operated on the teacher in the classroom often working against any change in practice, as Calderhead (1984) suggested

*"It has sometimes been argued that since teachers' and pupils' behaviour is shaped by various constraints, both teachers and pupils find themselves thrust into fulfilling particular roles within the classroom."* (Calderhead, 1984, p. 104)

It has also been suggested that teachers in general have not actively sought to change classroom practices, since such activities were threatening and undermined their sense of security. Thus coping strategies (Calderhead, 1984) were often developed by teachers. These were characterised as routinised tactics in the classroom, grounded in a survival approach rather than a reflective professional view of teaching and learning. Woods (1979) has suggested a number of factors which may operate against the recognition of the need to examine practice by teachers. He argued that it was unlikely that this would happen without intervention since teachers were often unaware of the constraints operating on them, or did not wish to admit to adopting certain coping strategies for reasons of self-esteem or professional image. Woods suggested that in the face of quite powerful constraints teachers felt a strong personal need to maintain a professional "front".

In addition, Woods (1979) argued that an ethnographic approach using participant observation was most useful in uncovering some of these hidden aspects of teaching since the effects of such constraints on classroom practice were not in themselves easily observed. This evidence suggested that professional development for the majority of teachers was not an activity which they would enter into lightly and if imposed would result in considerable resistance to change.

A number of different approaches have been advocated to support teacher research and professional development in order to facilitate this confrontation of issues, the recognition of the need to change practice and the development of the necessary skills to achieve these ends. The facilitation of such teacher involvement has also been closely linked to the climate created between participants and ways of promoting reflection and analysis of the

reality of professional life, whether focusing on the pupils, teachers or the organisation as a whole. Various roles have been suggested for the participants. Both Elliott (1977) and Eraut (1972) have proposed client-centred models of professional development where teachers and significant others were involved. They also cautioned on intervening prescriptively in the client's life. As Day noted

*"the interventionist aims to seek questions which are perceived by the client as relevant to his needs, to investigate answers to these questions collaboratively and to place the onus of action on the client himself." (Day, 1981, p.7)*

Day also drew attention to the importance of the affective area in professional development and in-service provision. He argued that in work concerned with questioning the teacher's self-image the affective relationship was of prime importance. Hoyle (1973) suggested that many programmes did not take account of concerns such as anxiety, status and identity. Thus the quality of the intervention has been of critical importance in promoting change and professional development.

Havelock (1971) has suggested three roles that change agents might successfully adopt *catalyst, solution giver* and *process helper*. He considered the third, that of process helper as the most important. As Day concluded

*"Success in the work with teachers seemed to have been achieved partly because the emphasis was placed on a collaborative process in which key elements were trust, openness and initial support, and partly because the methods used and data collected had been directly related to the teacher's action in the classroom." (Day, 1981, p.16)*

Day (1981, 1985) suggested that work with teachers should be based on the following assumptions:

- 1. Teachers have the capacity to be self-critical and should be offered the means by which they can begin to judge their own level of effectiveness and set new levels to be attained.*
- 2. Motivation for learning must arise out of the identification of a problem by the client, because long-term change is most likely to occur when learning is self-motivated and when knowledge is self-appropriated and self-actualised (Rogers, 1967).*
- 3. The interventionist must help the teacher in processes of internalisation rather than compliance or identification (Kelman, 1961). (Day, 1981, p.7)*

Such strategies explicitly required teachers to confront the reality of their practice in



the classroom or school in order to determine an agenda for improvement. In this way the responsibility for rendering an account of their actions is placed in the hands of teachers not external agencies. Argyris and Schon (1976) placed great emphasis on the conditions that would facilitate this process of professional development in constructing new theories-in-use. In particular, they should

*"produce data that help the individual to learn; help individuals gain insight into the conditions under which their defenses as well as their theories-in-use inhibit and facilitate their growth and growth of others; provide information from which individuals can design programmes for self-improvement, gain help from others and evaluate their progress; and help individuals learn how to discover their own theories-in-use and generate new ones - that is, learn to generate directly observable data, infer theories-in-use, alter theories-in-use and test new theories of action."* (Argyris and Schon, 1976, pp.38-39)

Day proposed a number of principles for enabling learning about theories in action which included: first, that effective learning occurred in response to confrontation of problems by the learner; second, that decisions about teaching stemmed from reflection on the effects of previous actions (Sutcliffe and Whitfield, 1976; Calderhead, 1984); third, that effective confrontation of problems required the maximizing of valid information and finally, that video recordings of lessons provided direct evidence of theories-in-use, and were, therefore, especially effective in promoting confrontation. Similar principles were used in the Ford Teaching Project where teachers needed particular help in reflecting on their classroom practice and were encouraged to check their understandings of events and actions with how these were perceived and received by the pupils and by an outside observer; as well as recording events by photographs, slides, audio-tape and video-tape.

Teachers have required assistance in undertaking this kind of research and professional development since evidence from a number of studies using the process of "triangulation" or the collection of data from a number of sources to illuminate events, also illustrated how often the perceptions of participants differed. This would suggest that teachers have not acquired the necessary skills of collecting valid information about their practice, which would enable them to reflect on this in such a way as to provide a basis for professional development.

Further research has indicated that teachers were often unaware of discrepancies between their theories-in-use and their espoused theories (Keddie (ed), 1973; Day, 1981). Studies of teachers' perceptions of pupils, which are discussed in more detail in

the following chapter, have also clearly demonstrated that a common repertoire of stereotyped roles have been assigned to pupils which has given rise to concern for the impact of such expectations on the performance of pupils. Given that one of the purposes of this research was to develop more appropriate assessment procedures for use by teachers in primary schools as a means of countering externally imposed accountability systems, this mismatch between teachers' perceptions of events and the realities of the classroom as recorded or observed by others gave rise to some concern.

Thus, teachers have often failed to recognise the need for change and require some form of intervention in order to examine their practice. The nature of this intervention has varied with researchers but has always involved some form of data collection which presents the teacher with evidence of their own practice or that of their colleagues. In the case of this research, the assumption was made that the assessment of pupil learning in the classroom context would provide teachers with a suitable method of confronting the realities of the classroom and their own practice as well as provide the basis for the development of more appropriate assessment methods for the extended aims of education under consideration.

In addition, if the validity and reliability of teachers' assessments of pupils was questionable, then these strategies would go some way towards minimising these effects within the research programme. Thus within the Oracle research programme there was a commitment not only to view the teachers as subjects of research by monitoring their classroom activities in studying teaching styles but also to involve them actively in the development of the teacher-based assessment procedures appropriate to the extended aims of primary education, as identified by Ashton *et al.* (1975). This was based on the assumption that if teachers were involved in the development, piloting and evaluation of such procedures they would be more accessible to other teachers involved in the study who would be required to complete these assessments but were not party to the process of development. These are considered in more detail in the following chapter where methodological concerns related to the assessment of pupil learning are discussed. Thus, although the involvement of teachers in research and in setting the agenda for their own professional development was seen as highly desirable, there were considerable problems associated with this approach.



## 1.7 SUMMARY

This review of the literature has attempted to trace the development of new patterns of organisation and curricular provision which ultimately resulted in the emergence of a progressive ideology in primary schools based on the idea of an individualised approach to pupil learning. A number of contextual factors have been considered which, it has been argued, were influential in this development including the recommendations of prominent educationalists, government reports, the impact of secondary school reorganisation, the removal of the 11+ examination and changes to the design of primary school buildings.

The nature and extent of progressive primary practice has been considered in relation to the changes in patterns of classroom organisation, teaching style, and the curriculum as well as the relative effectiveness of this pedagogical position in terms of pupil attainment. This discussion suggested that there was a need to provide appropriate assessment of the wide range of aims considered important by primary school teachers. The impact of an objectives approach to curriculum planning and evaluation has also been discussed in order to demonstrate the need for more effective means of assessing pupil learning. This need has also been highlighted by the moves towards increased school and teacher accountability during the late 1970's.

As a response to the growing demands for accountability a more extended role for the teacher as curriculum specialist, researcher and evaluator emerged, with an increased emphasis on teacher self-evaluation and professional development. In particular, the notion of action research and collaborative research were seen as ways of enhancing the professional development of teachers, bringing theory and practice closer together and facilitating the dissemination of educational research and curriculum developments. Thus the forces contributing to the accountability movement have been discussed and the responses of research and schools considered with regard to school assessment practices.

It has been noted throughout this review of the literature that concerns have been expressed regarding the methods of educational assessment used in curriculum evaluation, teacher effectiveness research and the monitoring of educational standards. Any consideration of the forms of assessment to be used in the Oracle research programme had to take account of these concerns in order to provide an adequate system of assessing primary practice whether this was progressive or traditional in its ideology. In addition, the involvement of teachers in such research has been advocated. A

consideration of different forms of educational assessment, their appropriateness for the purposes of the Oracle research and the issue of teacher involvement in this research are the subjects of discussion in the following chapter.



## **CHAPTER 2: ASSESSMENT OF PUPIL LEARNING: A DISCUSSION OF METHODOLOGICAL ISSUES**

### **2.1 INTRODUCTION**

It has been noted previously that this study of teachers' assessments of pupils was initially conducted within a five year research programme (Oracle). The intentions of the programme as a whole were: first, to provide a more extensive description of pupil behaviour, progress and performance and teacher behaviour in the classroom than had been offered by previous research and second, to assess pupil progress and performance to provide information on the effectiveness or otherwise, of particular teaching styles (Galton and Simon (eds), 1980). Thus, the primary purpose for the assessments used in the research programme was to describe pupil attainment. However, in the preceding chapter an additional set of purposes have been argued for. First, such procedures should also aid teachers in accounting for the time pupils spend in school and the quality of their learning in relation to the more extended aims of teachers resulting from the impact of a more process orientated curriculum and progressive pedagogy. Second, the development of a teacher's ability to assess and monitor pupil learning would provide an appropriate focus for professional development.

Three important factors which contributed to the need for more appropriate assessment methods in schools were considered. The first factor described arose from the need to keep track of the complex classroom organisation which resulted from moves towards unstreamed, more open and progressive primary classrooms. As a consequence of this development, teachers had to meet individual pupil needs and ensure a more effective match between learning activities and stages of development. A shift in curriculum emphasis from products to processes was the second factor identified. It became apparent that information provided by more formal assessment methods about pupil achievement and performance was inadequate for describing pupil learning in a process curriculum, particularly at transition points within the educational system. The third factor which prompted the development of new methods of assessment was the need to provide teachers with a way of coping with the moves towards external accountability systems which used relatively crude measures of pupil attainment in the monitoring of standards. This threat also increased the need for more teacher responsibility for the assessment of pupil learning within the school and therefore more appropriate provision for in-service education and teacher professional development in this area.

Consideration is now given to the developments that resulted from these factors since these have informed the direction and methodology used within this study. First, a critical review of educational assessment methods is presented. Second, the various forms of assessment, both formal and informal, which were available are examined in order to judge their suitability for the purposes of the Oracle research. This review particularly focuses on the availability and appropriateness of measures for the assessment of the aims of primary education as formulated in the Ashton *et al.* (1975) survey of teachers' opinions. Methodological issues associated with the validity and reliability of formal and informal assessments are also discussed before considering the ways of improving validity and reliability through a variety of strategies including the involvement of teachers in the development and implementation of these procedures.

## 2.2 A CRITICAL REVIEW OF EDUCATIONAL ASSESSMENT METHODS

The term educational assessment has been defined as all the

*"processes and products which describe the nature and extent of children's learning, its degree of correspondence with the aims and objectives of teaching and its relationship with the environments which are designed to facilitate learning."*  
(Satterley, 1981, pp.3-4).

Thorndike and Hagen (1969) in discussing psychological and educational measurement and evaluation in general, differentiated methods used to appraise and describe a person into two main categories; first, those which depended on setting up special test situations which were described as *"formal assessments"* and second, those methods which involved an evaluation based on a naturally occurring situation or *"informal assessments"*.

Formal assessments were characterised as occurring at a specified time and place, consisting of a set of tasks uniform for each person tested, and were perceived as test situations by the people being appraised. Informal assessment was characterised as extending over an indefinite period, based on situations that varied from person to person and which were not perceived as a test by the person being appraised. However, there were some situations where this distinction became blurred, for example, when an observer noted the specific errors a child made when s/he read aloud. Alternatively, it



was possible to set up a situation where the conditions were contrived, and thus constituted a test, but the behaviour of the individual was observed and judged by an outsider.

These two major groupings of assessment methods have been further categorised. For example, most texts have distinguished between norm referenced and criterion referenced measurement (Thorndike and Hagen, 1969; Satterley, 1981; Popham, 1981). Harlen (1978) also identified pupil referenced assessments. More recent categorisations have been that of diagnostic assessment (Black and Dockrell, 1981) and informal teacher-based methods of assessment utilising checklists, criteria for monitoring pupil progress and structured observations of pupil behaviour made by the teacher in the classroom (Tough, 1976, 1977, 1979; Harlen *et al.*, 1977a, 1977b). A critical review of such methods of assessment is now presented since an understanding of the key issues and problems of various approaches to educational measurement was important in the development of appropriate assessment procedures in this study.

### 2.2.1 Developments in educational assessment.

Educational "*testing*" as an exemplar of formal assessment has been defined by Kitwood as

*"the indirect and relative measurement of certain attributes, competencies or achievements of persons, using standardized instruments of assessment."*  
(Kitwood, 1977, p.1)

This type of testing was first envisaged as a major feature of education almost 100 years ago. It arose from the concerns of early experimental psychology to measure the nature of individual differences and the development of statistical techniques for carrying out such studies. In Britain, this movement was pioneered by Sir Francis Galton in his studies on heredity which arose largely as a result of the need to consider the implications of Darwin's theory of evolution and resulted in his work on correlation and regression. This was later developed under the leadership of Pearson. In the United States, Catell began work on individual differences in sensory and motor performances and their relationship with academic performance. This work was further developed by Thorndike and the Teachers College of Columbia University, where he was based, became the centre for the spread and development of standardised educational tests. At the turn of the century Rice conducted a series of standardised testing programmes in order to establish

the norms for different grade levels from the average scores of pupils. In 1897, a spelling test was given to some 33,000 pupils; between 1902 and 1903 language and arithmetic examinations were administered to some 14,000 pupils (Popham, 1981).

The period between 1900 and 1915 was one of initial development and saw the emergence of the first Binet intelligence scales in 1905 and the subsequent development of the Stanford-Binet Intelligence Tests by Terman in 1910. Standardised achievement tests in different subjects began to appear as exemplified by Stone's arithmetic tests, Buckingham's spelling tests, and Trabue's language tests. Thorndike's first handwriting scale was developed and Otis and others initiated work on group tests of intelligence. The next fifteen years saw a rapid increase in the number and types of tests. Standardised tests were developed for all school skills and content areas of the curriculum. Group intelligence tests were also produced in great numbers. Personality questionnaires and inventories also came into being.

Thorndike and Hagen (1969) identified the period from 1930 to 1945 as a time of critical appraisal when questions were asked about the validity of objective tests, their use in the classification and selection of individuals and their role in the heredity-environment debate. Criticisms were made on the grounds that such tests were limited in scope and they emphasised restricted and traditional objectives. Also, the whole philosophy of the use of numbers and quantification to express psychological properties was questioned. Thorndike and Hagen noted that it was a period when the centre of attention shifted from

*"measuring a limited range of academic skills to "evaluating" achievement of the whole range of educational objectives. It was a period in which the holistic, global, projective methods of personality appraisal came to the fore." (Thorndike and Hagen, 1969, p.6)*

The next fifteen years saw the development of large scale testing programmes, particularly in the United States, such as those administered by the College Entrance Examination Board. In Britain, as was noted previously, the immediate post-war period was characterised by the use of the 11+ for selection with a subsequent move away from such tests. Traditionally norm referenced tests were used to measure pupil achievement in such contexts. These tests tended to focus on differences between individuals and groups, making use of some norm to enable comparisons to be made. Thus a norm referenced standardised test was one

*"which has been given to a large number of children in controlled conditions and*



*from the results "norms" have been established for different groups of children, usually age groups. The result of giving the test to any child can therefore be compared with the average for a particular group.*" (Harlen, 1978, p.12)

Popham defined this type of test as being *"used to ascertain an individual's status with respect to the performance of other individuals on that test"* (Popham, 1981, p.26) whilst Satterley defined it as *"a test explicitly designed to be used to make comparisons between individuals with norms for interpretation "* (Satterley, 1981, p.354). These tests were most often used as a means of selection and as a basis for prediction of future performance as in the case of the 11+ examination. Such tests have also been used as the means of monitoring standards in schools; for example, the use of various N.F.E.R. tests by local education authorities to test basic skills in primary schools.

Alongside the increasing use of norm referenced tests in schools has been the development of criterion referenced tests. These have been designed to compare individuals to *"a defined behavioural domain"* (Popham, 1981, p.27) rather than with one another so that information was provided about *"a pupil's performance in relation to a specified level of ability or skill or knowledge which is the criterion"* (Harlen, 1978, p.13). Items on criterion referenced tests were devised so that the score could be expressed in terms of the degree of mastery of that particular set of skills.

Thus, the distinction most often made between norm and criterion referenced assessment was the basis used for interpretation. It has been suggested (Millman, 1974; Hambleton *et al.*, 1978) that norm and criterion referenced interpretations could be made from most tests, although the quality of such interpretations varied. A variety of terms have been used to describe such criterion referenced tests: for example, domain referenced, objectives referenced, competency and mastery. The use of these different terms for criterion referenced assessment appeared to be related to the degree of domain definition and instrument specification (Black and Dockrell, 1984). Criterion referenced assessment has thus taken a variety of directions. Nitko (1980) described it as

*"a many-faceted concept with a multitude of specific instances that are quantitatively quite different from each other but have characteristics in common so that the same term can apply."* (Nitko, 1980, p.461)

The initial development of criterion referenced tests arose from the need to assess programmed learning and subsequently the effects of curriculum innovation. Early curriculum research focused on measuring the attainment of objectives on norm

referenced standardised tests. Much criticism was leveled at this approach since the tests were not designed to measure the particular objectives that were built into many curriculum development projects (Popham and Husek, 1969; Steadman, 1976).

During the 1970's, such tests were designed which measured the specific content, concepts and skills associated with a particular curriculum project. These were then scored in relation to achievement of defined criteria or objectives and not against the norm for particular age groups. In this way the curriculum development was evaluated with respect to the achievement of the aims and objectives of that programme rather than some notional standard of attainment represented in the results of a norm referenced test. The development of criterion referenced assessment has more recently been in the assessment of minimal competencies for certification and professional examinations (Popham and Lindheim, 1980).

By the late 1960's, there was a definite climate against testing, whether norm or criterion referenced, particularly in primary schools but the last fifteen years has seen the re-emergence of the testing movement to a position of strength with an emphasis on accountability both in the United States and in Britain. As the demand for educational accountability grew during the 1970's the need to develop appropriate methods of educational assessment for the monitoring of standards also gained considerable momentum. This need highlighted two major problems which have always confronted test developers. As Goldstein noted

*"The first arises from the varied aims and methods of teaching, each of which can make out a case for using its own particular test instruments designed to assess its own set of objectives. ... The second problem, which really stems from the first, is the relative impermanence of all educational tests."* (Goldstein, 1979, p. 211)

These problems were particularly associated with the use of both norm referenced and criterion referenced tests for the monitoring of standards. For example, how could tests be designed to assess the different educational aims and priorities of individual local education authorities, schools and teachers? If criterion referenced assessment was used to monitor standards, how could the results on different tests designed to assess a particular curriculum be compared? If norm referenced test results were used to monitor performance over time then some way of comparing the attainment of different populations taking the same test at a particular time or the same population taking tests over time at different levels of difficulty had to be found. This was necessary if



comparisons were to be made between standards achieved. The basic problem was outlined by Choppin (1969) as follows

*"If results on an item are to be interpreted, then one needs to know how difficult the item is and to what extent it discriminates between people of different ability. Unfortunately, the conventional measures of item difficulty and discrimination are 'sample-bound', which is to say, they are extremely dependent on the nature of the sample of people who provide the data." (Choppin, 1969, p.237)*

A significant development in test design, that of the Rasch model, emerged which went some way to countering the problems of interpretation found with existing methods of assessment used for monitoring standards. The Rasch model (Rasch, 1960) was based on latent trait theory. This theory has underpinned much psychological and educational measurement research where it has been assumed that there are one or more underlying and unobservable factors or traits which characterised an individual and determined his or her observed responses. Thus a set of indicators such as total test scores or, in the case of the Rasch model, responses to individual items, was related to a set of such unobservable factors or traits.

The essential characteristic of this probabilistic model was that the item statistics were "person-free" and the person measurements were "item-free" (Dobby and Duckworth, 1979). Thus, the use of the Rasch model to describe test items in terms of their different levels of difficulty enabled comparisons to be made over time for different groups of students and different tests. As a result of this development it was argued that if the Rasch model provided a reasonably reliable and valid measure of an item's difficulty, it could be used in the monitoring of standards. This could then be achieved by assessing different samples of the population at different times and comparisons could legitimately be made between these results to evaluate any improvement or decline in standards (Satterley, 1981).

This was of particular significance in the work of the Assessment of Performance Unit which was set up to monitor standards in primary and secondary schools during 1975. The original brief of the Assessment of Performance Unit was to promote the development of methods of assessing and monitoring the achievement of children at school, and to seek to identify the incidence of under-achievement. The following tasks were laid upon it:

1. *To identify and appraise existing instruments and methods for assessment*

*which may be relevant for these purposes.*

2. *To sponsor the creation of new instruments and techniques for assessment, having due regard to statistical and sampling methods.*
3. *To promote the conduct of assessments in cooperation with local education authorities and teachers.*
4. *To identify significant differences of achievement related to the circumstances in which children learn, including the incidence of under-achievement, and to make findings available to all those concerned with resource allocation within the Department, local education authorities and schools. (D.E.S., 1974)*

The Assessment of Performance Unit was expected to monitor standards across the curriculum, and groups were established to design testing procedures in the areas of mathematics, language, science and a first foreign language. The procedures designed within the National Foundation for Educational Research for the assessment of language and mathematics were specifically based on the Rasch model. Considerable work was undertaken to develop large item banks where each item was independently described in terms of its level of difficulty. Thus it was anticipated that

*"tests can be tailor made to suit the exact requirements of the user (and) to help ensure their suitability for the user." (Dobby and Duckworth, 1979, p. 49)*

A variety of users were envisaged, for example, teachers could draw on specific items which were appropriate for the curriculum they had taught, local education authorities and central government could use such items in the construction of tests to monitor standards as exemplified by the Assessment of Performance Unit. The Rasch model was the basis of these developments since without this method of describing individual items it would have been impossible to make comparisons between the variety of tests which could be produced by the selection of different items.

However, the use of the Rasch model in describing items in this way was subject to criticism (Goldstein and Blinkhorn, 1977, 1982; Goldstein, 1979, 1980). Goldstein (1979) argued that, although the model might have an important role to play in attempts to isolate unidimensional traits, and thus develop latent trait theory, it did not necessarily fit the requirements of educational testing. He particularly criticised the concept of item banking in the following way:

*"If we suppose that each of the items in the bank has a prescribed difficulty value, then it is strictly meaningless within the context of the Rasch model to speak of one item as being more applicable to one point in time rather than another. The only*



*meaning which can be attached to such a statement must be in terms of difficulty values. For example, suppose we have two items, one of which is more applicable in 1975 than 1980 and the other which is more applicable in 1980 than in 1975. Then these two items will have different relative difficulties in the two years and indeed their relative difficulty might become reversed between 1975 and 1980. Hence by definition, they cannot belong to a single common Rasch scale extending over the five year period 1975-1980.... Thus an item bank which is designed so that out-of-date items can be replaced is a strictly non-Raschian concept. Similar logic applies to the so called 'tailored testing' procedures .. where it is claimed that items can be selected from an item bank to suit different criteria." (Goldstein, 1979, p. 217)*

The momentum to the development of large item banks as envisaged in the 1970's has, however, decreased as the debate on the appropriateness of the Rasch model to educational testing continued. In addition, the availability of funds for such large scale testing programmes was restricted during the 1980's. It might also be argued that the school-based curriculum movement, coupled with increasing concern for the involvement of teachers in pupil assessment (Harlen, 1978) also contributed to the demise of the Rasch model in educational testing.

Black and Dockrell (1984) argued that developments in formal assessment, in particular criterion referenced testing, were of little relevance to practising teachers since they were inappropriate for classroom assessment purposes. The purposes of educational assessment which were most ably served by the use of norm and criterion referenced tests were those of prediction, selection, curriculum evaluation and monitoring of standards, thus providing statements of pupil attainment for external consumption. These purposes although fulfilling part of the function of educational assessment have not necessarily met the day to day needs of teachers for the assessment of pupil learning so that they might match pupils levels of development more accurately, evaluate their own teaching strategies or provide feedback to pupils on achievement of certain aims and objectives. Among the reasons Black and Dockrell (1984) suggested for the lack of acceptability of criterion referenced tests, in particular, were:

- 1. that the emphasis in the existing literature on precisely defined behavioural objectives (Monjan and Glassner, 1979) was not always compatible with existing pedagogic practice;*
- 2. that the emphasis on complex psychometric procedures for item review was not always of immediate relevance to the needs of diagnostic assessment in schools;*
- 3. that we were focusing on domains of a smaller scale than seemed to be typical of minimum competency tests;*

4. *that if technical suggestions being put forward in the American literature were to be accepted as necessary for criterion-referenced test design and construction, then diagnostic tests would have to be provided by central agencies, and teachers not encouraged to construct them.*

(Black and Dockrell, 1984, p.38)

Black and Dockrell argued that a more appropriate method of interpretation for teachers was that of diagnosis. Diagnostic assessment was described as *"a form of assessment designed primarily to help pupils learn and teachers to teach"* (Black and Dockrell, 1984, p.12). The purposes of such diagnostic assessments have been related to questions of the teacher's success in teaching a class, with the identification of and the reasons for pupils failing. Therefore, a diagnostic test would be one that was designed to provide information that would help pupils learn and teachers teach. In format such tests resembled criterion referenced tests but the interpretation of the results was seen as aiding teacher assessment through the description of pupil learning rather than providing a statement of performance in the form of a grade or mastery level.

An alternative basis for interpretation was proposed by Harlen (1978), called pupil referenced assessment, where interpretations were not made by comparing performance with either a norm for the particular age group nor to a pre-determined criterion. These measures were described as an approach to assessment which enabled comparisons to be made between the pupil's present performance and previous levels of performance (Harlen, 1978). It was suggested that this approach was commonly used by teachers in monitoring the progress of pupils in class. This type of comparison was seen as less invidious than comparisons with norms and criterion levels since it valued the individual and judgements were made in the context of that individual alone. Although norm and criterion referenced tests could be used for this purpose, the essential difference between these forms of assessment and pupil referenced assessment lay in the method of interpreting the data obtained from using such methods.

These differences in the basis for interpretation indicated that *"different educational purposes required differing educational tests and differing uses of those tests"* (Popham, 1981, p.10). Such developments in formal methods of assessment suggested that the key issue for consideration in the selection of the assessment procedures in this research lay in the way in which the evidence derived from such assessments was to be interpreted and for what purpose. The following discussion addresses in more detail the criteria that were



applied in the selection of assessment methods within the Oracle research programme, given that one of the purposes was to measure progress and performance of different pupil types taught by different teaching styles.

### 2.2.2 Selection of assessment methods

The intention of the Oracle programme, in addition to providing descriptions of primary practice, was to offer valid and reliable assessment procedures to monitor aspects of pupil learning in line with the extended nature of primary schools teachers' aims and curriculum thus providing a more accurate picture of the effectiveness of different teaching styles. The assessment of pupil learning had, therefore, to take account of the developments in primary curriculum and pedagogy that have been described in Chapter 1. For this reason the description of aims of primary education found in the Ashton *et al.* (1975) study of teachers' opinions (see Appendix 1) provided a useful starting point for considering what aspects of the primary school curriculum should be assessed within the Oracle programme. It has been assumed that teachers, when they articulate their aims in this way, reflect the nature of their practice and, therefore, these aims should represent more closely their actions in the classroom and provide a more rational basis for judging effectiveness of teaching style.

The particular characteristics of the assessment procedures required in order to investigate the relationship of pupil achievement to different teaching styles are now considered with regard to providing data on the progress and performance of pupils on the curriculum aims identified as of major or utmost importance in the Ashton *et al.* (1975) study. Popham (1981) suggested a number of criteria which could be used in judging the suitability of assessment procedures. These were the quality of the descriptions of measured behaviours, the nature of the comparative data, the number of items per measured behaviour, the scope of the measurement, validity, reliability, practicality with respect to the ease of administration, scoring and cost. Although these criteria were originally designed to review the appropriateness of norm and criterion referenced tests in the evaluation of large scale curriculum development projects they also provide a useful basis for the evaluation of the appropriateness of both "*formal*" and "*informal*" assessment procedures within the Oracle research programme where a variety of assessment procedures were employed.

The first criterion as described by Popham (1981) related to how well a curriculum

developer or evaluator specified the domains that were to be tested prior to construction of the test. This criterion was of particular importance in the selection of appropriate tests for curriculum evaluation since much of the criticism leveled at norm referenced tests related to the lack of adequate description of the pupil learning that would be tested and the likely mismatch between the domains tested and those taught within the curriculum. In addition, this criterion was relevant to the selection of assessment procedures for investigating the effectiveness of different teaching styles since traditional forms of norm referenced testing have not assessed the progressive pedagogy and extended aims of primary teachers reflected in the Ashton *et al.* survey (1975).

The quality of the comparative data was also important since the analysis of test scores was to provide information on the sample population for the analysis of pupil progress and performance. Norm referenced tests have usually been used when it has been necessary to differentiate between people; for example, when sampling and matching students, classes or schools prior to study and when the test results were used to compare performance with norms for a national or local group (Morris and Fitz-Gibbon, 1978). It was, therefore, appropriate to use this type of test for certain purposes in the study of pupil progress and performance; for example, in the selection of a sample of pupils to represent the range of ability levels in each class and to investigate the progress of pupils on the basic skills since such tests have been used in previous research, enabling comparisons to be made where appropriate.

However, this type of test has a particular weakness in relation to the study of pupil progress. Such tests have been characterised by the scores following a normal distribution curve and, therefore, have not discriminated effectively within the central range of scores, so that differences in pupil performance resulting from teaching style differences have been difficult to detect. The tests used in the Oracle research were required to be maximally sensitive to changes in pupil achievement. Thus, criterion referenced tests were considered preferable since these have been designed to assess the level of pupil achievement with respect to mastery of pre-determined objectives. If the criterion referenced tests chosen provided an adequate description of the qualities to be assessed and matched the aims and objectives of primary teachers then they would provide the most appropriate method of assessment.

The number of items per measured behaviour was considered important in criterion referenced testing since each objective should have a number of similar items to test a



particular skill or concept. Whereas norm referenced tests have been designed to provide a large number of different items to discriminate effectively between pupils. Ideally any test chosen to describe pupil progress and performance in relation to primary school practice and teaching style should attempt to fulfil both these functions, since discrimination between pupils and between performance in various areas of the primary school curriculum was considered important in the study of the effectiveness of different teaching styles. This criterion is considered in more detail later in this chapter where specific tests are reviewed.

The scope of a norm or criterion referenced measurement has described how well the constructed test matched the pre-determined objectives which it has been designed to measure. This criterion has been of great importance when designing a new test. It was also relevant to this research; first, in respect of how well the measures matched the various aims of primary school teachers, and second, in devising a shortened version of the standardised attainment test finally selected (see Appendix 4).

The validity of any assessment method must be demonstrated since the data derived from them should accurately describe the areas of achievement which they purport to measure. Thus validity has been demonstrated by a number of different means relating to the purpose of the assessment. Thorndike and Hagen (1969) referred to *content*, *construct* and *criterion-related* validity. Guildford (1973) has described *predictive* and *factorial* validity. Popham (1981) identified three categories of validity endorsed in the Standards for Educational and Psychological Tests (American Psychological Association, 1974): *content*, *criterion-related* (*predictive* and *concurrent*) and *construct* validity. The 1974 Standards offered the following definition of content validity, which has most often been applied to psychological tests of skills or knowledge:

*"To demonstrate the content validity of a set of test scores, one must show that the behaviours demonstrated in testing constitute a representative sample of behaviours to be exhibited in a desired performance domain. Definition of the performance domain, the user's objectives, and the method of sampling are critical to the claims of content validity."* (American Psychological Association, 1974, p.98)

The demonstration of content validity has, therefore, related to how adequately the items or content of an assessment samples the area under consideration. Difficulties associated with the content validity of norm referenced tests in teacher effectiveness research have already been noted. Criterion referenced assessments have, however, been designed to overcome this problem since they have not compared pupils with a norm for

the age group but to particular objectives and, therefore, have been designed specifically to test particular performance domains. As such they have tended to replace the traditional norm referenced test where an appropriate test has been available to assess the particular aspects of pupil learning under investigation.

Criterion-related validity has been defined as an attempt to correlate performance with an independent or external criterion (Popham, 1978). A distinction was made between concurrent and predictive validity in terms of the time between administrations. In the case of concurrent validity there was no intervening time between the administration of the measure and the external criterion, whilst for predictive validity an interval of time has occurred between test administration and the collection of external criterion data.

Three steps have been identified in the demonstration of construct validity; first, the identification of hypothetical constructs presumed to account for test performance; second, the derivation from the theory underlying the construct of one or more hypotheses regarding test performance and third, the empirical testing of these constructs. The construct validity of assessment procedures has particular relevance to evaluating methods used to measure particular attributes such as pupil personality and attitudes. These were, however, relevant to the assessment of pupil progress and performance in relation to the aims of teachers which were the subject of this research since it has been assumed that these aims could be defined in behavioural terms and that the resulting criteria for assessment would reflect underlying pupil abilities or traits.

In addition to being valid, any assessment procedure must measure as accurately as possible since measurement errors may mask differences in the performance of different groups of pupils. The reliability of an assessment procedure has, therefore, provided an estimate of how much confidence can be placed in its precision. Popham (1981, p.128) has suggested four ways of establishing the reliability of norm and criterion referenced tests. These were *stability*, *equivalence*, *equivalence and stability*, and *internal consistency*.

Stability estimates of reliability have been based on the consistency of a test's measurement over time. Typically the measurement of test stability has been made by giving a test to a group of pupils on two occasions usually separated by several weeks. When the correlation between these two sets of scores; that is, the stability or test-retest coefficient of the test has been high then it has been assumed that the test was stable. A



second form of reliability or equivalence has relied on the use of two or more forms of the same test, in the expectation that these forms yield consistent data on the same pupils. This has been of particular importance where two forms have been used interchangeably. The third form of reliability has been associated with stability and equivalence where it has been necessary to administer different forms of a test after a time delay. The fourth type of reliability or internal consistency has been related to the consistency of a test's internal elements or items. This has typically been calculated using the split-half technique where the test has been divided into two equal parts and the scores achieved on each half correlated. The resulting correlation coefficient was considered an estimate of the degree to which the two halves of the test performed their functions consistently. Reliability issues are again considered in more detail later in this chapter where specific assessment procedures are reviewed. The last criterion proposed by Popham (1981) related to the practicalities of administration, cost and scoring which were also considered in the selection of assessment procedures in this research.

Thus taking account of the considerations listed above the following conclusions regarding the appropriateness of using particular assessment procedures in this research were reached. Norm referenced tests were considered suitable for the initial description of pupil performance in order to select the stratified random sample of pupils within the Oracle research programme. Criterion referenced tests were appropriate for assessing the aims of primary school teachers if it were possible to match these aims and the pre-determined objectives of such tests. In addition, criterion referenced tests would provide a more appropriate means of reflecting progress in these areas.

However, such tests were required to provide evidence of pupil performance in areas of the curriculum that reflected the extended aims of teachers as described in the Ashton *et al.* survey (1975). The assumption was made that these aims, because they were formulated by teachers, would provide an accurate reflection of current primary school practice at the inception of this research. Given that these seventy two aims of primary education provided an adequate description of the curriculum then some selection from the seventy two aims listed within the survey was necessary since it was outside the scope of this investigation to consider all the areas which included aspects of intellectual, moral, social, aesthetic, emotional/personal, physical and spiritual/religious development.

This selection was made with regard to three criteria, which were determined by the nature of the overall research programme. First, those aims which could be assessed by

established means such as the Richmond Tests of basic skills or the "*tests*" of study skills were excluded. Second, those aims unrelated to intellectual development were not considered since the main focus of the Oracle research programme was to investigate the effectiveness of different teaching styles in terms of pupil achievement. Third, only those aims which were judged by teachers in the original survey to be of "*utmost*" or "*major*" importance were considered. These aims were related to the factors of basic skills, intellectual autonomy and competence identified in the Ashton *et al.* study (1975, p.66). These have been listed and their rank order given in Table 2.1 overleaf.

A number of tests were available for selection details of these can be found in the wide range of references now available (Buros, 1972 and 1974; Gillham and Hesse, 1970, 1973, 1976; Pumfrey, 1976; Sumner and Robertson, 1977; Vincent, *et al.* 1977; Bentley and Marsden, 1983; Harlen, 1983; Raban, 1983; Levy and Goldstein, 1984). Of the norm and criterion referenced standardised tests suitable for the assessment of pupil performance in some of the areas of the curriculum described by Ashton *et al.* (1975) survey of the aims of primary school teachers were a range of measures developed by the National Foundation for Educational Research, the Bristol Achievement Guides (Brimer, 1969), Moray House tests and the Richmond tests (France and Fraser, 1975).

The tests chosen for this research were constructed from the range of measures known as the Richmond tests (France and Fraser, 1975) based on an earlier version of one of the more popular American test batteries, the IOWA test of basic skills. The Richmond version attempted to move away from the norm referenced approach of the American test. It approximated a criterion referenced test as each part was divided into a series of sub-tests measuring component skills or objectives. For this reason it has been described as a *domain* or *objective referenced* test by Galton and Simon (eds) (1980) since the level of performance necessary for mastery of each skill was not specified. The specific characteristics of the tests used are presented in Appendix 4. The use of such modified norm referenced measures met some of the criticisms of using standardised tests in monitoring pupil performance. They not only provided for adequate discrimination between individual pupils for sampling but also gave useful information on the performance of pupils in relation to sub-sets of skills in each of the areas tested thus providing a suitable instrument for this research. As with most standardised tests of attainment for primary aged children the Richmond tests were designed to measure aspects of the basic skills of language and mathematics. Thus aims such as 26, 12, 55, 8, 5 and 71 (see Table 2.1) were assumed to be covered by these tests.



**Table 2.1<sup>1</sup>: Intellectual aims ranked as of "major" or "utmost" importance**

Rank	Aim	The child should:
2	26	be able to read with understanding material appropriate to his age group and interests.
4	12	be able to read fluently and accurately at a minimum reading age of eleven.
10	21	be able to convey meaning clearly and accurately through speech for a variety of purposes; for example, description, explanation, narration.
12	3	know how to acquire information other than by reading; for example, by asking questions, experimenting, from watching television.
13	50	be able to listen with concentration and understanding.
15	4	know how to use mathematical techniques in his everyday life; for instance, estimating distance, classifying objects, using money.
15	55	have a wide vocabulary.
20	29	know how to compute in the 4 arithmetic rules using his knowledge of; for instance: number, multiplication tables and different units of measurement.
23	9	be developing the ability to make reasoned judgements and choices based on the interpretation and evaluation of relevant information.
24	65	have a questioning attitude towards his environment.
25	8	be able to write legibly and know how to present his work attractively.
30	60	be developing his inventiveness and creativity in some fields; for example, painting, music, mechanical things, poetry, movement.
32	57	be developing the ability to plan independent work and organize his own time.
33	5	know the correct spelling of a basic general vocabulary.
34	71	know how to think and solve problems mathematically using the appropriate basic concepts of, for example, the number system and place value, shape, spatial relationships, sets, symmetry, and the appropriate language.

---

<sup>1</sup> Taken from: Ashton *et al.*, (1975) *The aims of primary education: a study of teachers' opinions*. London: Macmillan Education.

However, the use of Richmond tests alone was rejected since no suitable descriptions of behaviour appropriate to some of the remaining aims were available. Since these instruments did not assess relevant aspects of the primary school curriculum as expressed by the aims of the teachers, it was necessary to decide what alternative strategies to use to assess these aims; for example criterion referenced tests, or to develop diagnostic, pupil referenced or informal measures such as teachers' assessments, based on checklists and the observation of pupils engaged in a variety of classroom learning activities.

Although diagnostic tests and assessment procedures have gained in acceptance and use in primary schools particularly for screening and classroom teaching purposes they have also been developed largely in the areas of language and mathematics attainment and for monitoring the social and behavioural development of children (Neale, 1958; Schonell and Schonell, 1963; Stott and Marsden, 1970; Daniels and Diack, 1970; Gillham and Hesse, 1976; Stott, 1978). Thus, such tests as were available did not meet the requirements of assessing the aims given in Table 2.1 which covered different aspects of the primary curriculum.

Diagnostic tests have not generally been used in teacher effectiveness research since they have been specifically designed to inform the teacher of pupil attainment for classroom purposes rather than large scale educational research. However, an assessment procedure, described as a series of "*tests*" of study skills (Galton and Simon, (eds), 1980, p.17) was developed in the Oracle research to describe pupil performance related to the skills necessary to pursue independent study and reflected the aims relating to this in the Ashton *et al.* (1975) study (see Table 2.1).

The development of these tests has been described in detail in *Progress and Performance in the Primary Classroom* (Galton and Simon (eds), 1980, pp.101-113). These measures were developed particularly to assess aspects of topic and project work in the primary curriculum which have not generally been recorded in previous research into the achievement of pupils of this age range. The tests made use of a number of classroom-like activities which were subsequently assessed using a scoring scheme devised by teachers. They resembled diagnostic tests in that they were designed to assess specific skills related to aspects of pupil learning in the classroom context.

Pupil referenced assessments have rarely been used as product measures in teacher effectiveness research because of the methodological and practical problems involved



when comparing assessments made by different teachers of different pupils using different criteria. However, the ideas behind pupil referenced assessment provided an alternative basis for the development of the additional assessment methods where teachers were asked to assess their pupils using new forms of assessment which are described in the following sections.

Despite the development of additional diagnostic tests, it was still unlikely that any new test would meet the need for assessing such aims as *"the child should be able to convey meaning clearly and accurately through speech"* so other informal assessment procedures were also considered. The use of teachers' assessments based on their knowledge of the pupils and observations they could make in the classroom appeared to be the only appropriate and practical method of obtaining data on pupil attainment in the remaining areas identified in the Ashton *et al.* survey (1975). Thus, a third type of assessment called *teacher-based assessment* was also included in the Oracle research. These assessments comprised checklists of abilities linked to each aim and were used by teachers to rate pupil performance. These checklists also formed the basis for a number of structured activities which served as moderation instruments within the research.

These teacher-based assessments are the subject of study in this dissertation and have been designed to match the extended range of teachers' aims identified in the Ashton *et al.* study of primary school teachers' opinions completed in 1975 and compliment the other assessment procedures used within the Oracle research programme. However, research which has the intention of supplementing existing procedures by the use of such informal assessments, whether these were to provide measures for accountability, to evaluate the effectiveness of different teaching styles or the curriculum, must take account of the particular problems posed by the use of procedures involving teachers' assessments since in these more *"informal"* procedures, teachers' impressions provided the basis for formulating a judgement of the quality of educational achievement. The use of such procedures has been strongly criticised on the grounds of bias of the assessor and the general unreliability of such procedures. Thorndike and Hagen noted that

*"evaluations based on these observations have serious limitations. We are likely to find little uniformity from person to person in either the situations observed or the standards of judgement of the observers."* (Thorndike and Hagen, 1969, p.20)

The reliability of teacher's assessments of their pupils was investigated as early as 1945 by Burt when evidence relating to 11+ selection suggested that more reliance was

placed on teachers' assessments of pupils' aptitudes in preference to their performance on the tests and examinations. Burt suggested that

*"where teachers are constantly thinking during the everyday work of the classroom (e.g. industry or attainments in the three R's), the assessments of competent teachers appear to be quite "reliable" - indeed more "reliable" than those of standardised tests, though in the case of attainments there is little to choose."* (Burt, 1945, p.85)

Thorndike and Hagan (1969) identified two particular sets of circumstances where these effects have been observed; first, where the rater has been unwilling to take the time and effort required to complete the appraisal effectively and second, where the rater has identified with the individual being assessed and thus been unprepared to make an honest judgement which might have affected that individual in some way. This has been called the *"generosity error"*.

Even when a group of raters have shown a willingness and motivation to provide valid judgements then a number of other factors have operated to limit the validity and reliability of such judgements. These factors have included: lack of opportunity to observe, the covertness of the attribute, the ambiguity of the quality to be observed, lack of uniform standards for reference and specific rater biases and idiosyncracies. Such factors have produced a number of distortions which have consistently worked to reduce the validity and reliability of any rating schedule. These errors have been described by Thorndike and Hagan (1969) as the *"halo effect"*. This characteristic of ratings has resulted from the tendency to assess in terms of an overall impression without differentiating specific aspects of performance and for a total reaction to a person to colour the judgement of each specific criterion or category to be assessed.

The validity of teachers' assessments has usually been taken for granted since such measures have relied on the ability of teachers to know what they have wanted to assess. However, this assumption has been questioned in studies where the use of criteria derived from Bloom's Taxonomy of Educational objectives (1956) in the Certificate of Secondary Education, Mode 3 was investigated (Poole, 1972; Fairbrother, 1975). Teachers did not use the categories as intended since as Wood and Naphtali (1975) noted

*"at best teachers may manage to fit their own differentiated schema of attributes into the categories they are given; at worst, these categories are liable to be greeted with blank incomprehension and quickly discarded in favour of a seat of the pants approach."* (Wood and Naphtali, 1975, p.151)



Wood and Naphtali (1975) investigated the attributes of achievement which teachers actually assessed as opposed to those they were supposed to assess. They identified the following constructs as being used by teachers to assess achievements and differentiate between pupils:

- a) The involvement of pupils in the learning situation
- b) The ability a pupil has in a subject.
- c) The overall ability of the pupil.
- d) The behaviour of the pupil.
- e) The quality and tidiness of the work presented.
- f) The interest displayed by the pupil in the subject.

Becher, Eraut and Knight (1981) in their study of accountability in primary schools noted a similar problem with establishing the validity of teachers' judgements when using checklists provided by local education authorities. For example, Avon expected teachers to use the Bullock Report categories (D.E.S., 1975a) in assessing pupils' writing in terms of poetic, transactional and expressive qualities and also to use the Tough's (1976, 1977) categories for the appraisal of children's use of language. They concluded that it was often tacitly assumed that teachers knew what these categories meant.

The demonstration of the validity and reliability of such informal assessment procedures has more recently become an issue in the literature as a result of the problems experienced in the development of diagnostic assessment by teachers for different curriculum subject areas (Black and Dockrell, 1984). These were similar to those associated with demonstrating the validity and reliability of teacher-based assessments. First, the validity of the criteria or checklists themselves had to be established. Second, the validity and reliability of the diagnostic tests or structured activities which were designed to moderate the teachers' assessments had to be demonstrated. Finally, the validity and reliability of the teachers' assessments which it was intended to use to describe pupil attainment needed to be investigated.

Black and Dockrell (1984) argued that the concepts of validity and reliability which were applied to norm and criterion referenced standardised tests were not used by teachers and did not "*actually work*" (Black and Dockrell, 1984, p.66) when applied to the types of diagnostic criterion referenced assessments developed by teachers. They argued that much of the work on criterion referenced assessment which emanated from

the United States was not appropriate for the review of criterion referenced diagnostic assessments such as those developed in the Scottish Council for Research in Education project. They proposed an alternative strategy for considering validity and reliability. Their purpose was to build up evidence relating to certain questions about various aspects of the instruments. This involved

*"raising questions (issues) relating to a number of foci. These comprised components of the instrument (items), the whole instrument (tests) and decisions made as a consequence of using the instruments (scores)."*

*"Evaluation of the instruments therefore involved gathering data on an issue relating to a particular focus using one or more procedure. The procedures used were either of a form which involved studying the focus without having applied the instrument (logical review) or studying data which had been gathered from application of the instrument (empirical review)." (Black and Dockrell, 1984, p.66)*

Although this research was not conducted at the time of this study similar strategies were evolved to evaluate the validity and reliability of the teacher-based assessments which centred on demonstrating how well the checklist criteria and structured activities matched the aim under consideration and whether these could be applied in practice.

Given these difficulties such assessments have tended to be used only when other alternative methods have not been available to measure the particular attributes under consideration. Thus the decision to utilize teacher-based assessments in the form of checklists, structured activities and observations meant that ways of improving the validity and reliability of such methods had to be considered if they were to be used as a method of recording pupil performance within the research. The following discussion relates to ways which have been suggested in the literature which would aid the development of valid and reliable procedures utilising teachers' assessments to match the aims of primary education identified in the Ashton *et al.* survey (1975).

Suggestions for improving the validity and reliability of teachers' assessments have focused on refining the definitions of the qualities that have to be rated by the use of checklists, criteria and profiles; the development of more systematic techniques for observation and appraisal and reducing their bias by provision of clear instructions and relatively simple instruments for assessment of pupil learning. A number of specific recommendations have been made for improving the validity of teachers' assessments. Wood and Naphthali (1975) suggested that one improvement would be to develop



*"behaviourally-anchored scales based on a consensus view of what most teachers considered achievement to consist of." (Wood and Naphali, 1975, p.160)*

Such scales were characterised by each point on the scale having a more or less explicit description of the behaviour or performance which qualified for the mark or grade in question. They further argued that the scales should be based on what most teachers think, and that

*"more strenuous attempts should be made to provide descriptions of attainment which will ring bells in teachers' minds and so facilitate assessment and lead to greater fairness." (Wood and Naphali, 1975, p.160)*

This strategy was adopted within this research. In addition to the problems associated with teachers' misunderstandings or misinterpretations of the actual criteria or categories used, another set of factors have influenced the validity and reliability of ratings relating to the way in which such assessments have been carried out. The following considerations were taken into account in this study in the development of the teacher-based assessments in order to minimise the effects of errors associated with the use of ratings in general.

- a) As far as possible assessments were based on a review of actual pupil performance and observable behaviour.
- b) The activities observed were as objective and as distinct as possible.
- c) The teachers were warned about the halo effect.
- d) The assessments were expressed in a simple numerical scale.
- e) Marks were thought of as relative to a predetermined description of performance level.
- f) Teachers were asked to rate their class on a limited range of qualities.

The validity and reliability of teachers' ratings and their use in this research has, therefore, been problematic. As far as possible, the factors described above have been taken into account in the development, implementation in schools and in the way in which the results derived from the teacher-based assessments have been interpreted. However, teachers have always been required to make such judgements as part of their professional responsibility. They have been expected to use particular pupil attributes as a basis of judging their achievement. The validity and reliability of these judgements have been shown to be subject to a variety of influences, in addition to those already described,

which bias their judgements of pupil attainment in particular ways. Such pupil and teacher characteristics merit further consideration since it has been argued that

*".. we frequently make judgements without being aware of it; we make them in a haphazard manner without giving due regard to our sources of evidence which can easily be unreliable, the evidence itself can be poorly selected and misinterpreted."*  
(Downey, 1977, p.73)

Much research has been carried out to investigate the impact of a variety of pupil and teacher characteristics on teachers' judgements. This research is now reviewed since such evidence needed to be taken into account in the research design so that any influence of these factors could be controlled for in the analysis and interpretation of the results. Thus the following section reviews studies investigating factors which have been shown to contribute to teachers' assessments of pupil learning.

### **2.2.3 Factors influencing teachers' assessments of pupil learning.**

A large number of studies have been carried out into teachers' perceptions of pupils. These can be broadly grouped into those with a psychological perspective and those with a sociological stance. The former type of study has focused on the degree to which certain constructs and biases have influenced teachers' perceptions of their pupils, as well as the relationship of these perceptions to the performance of pupils and teacher-pupil interactions. A number of factors influencing teacher perception have been investigated including pupil characteristics such as: personality, attitudes, race, sex, age, classroom behaviour, ability and social class; as well as teacher characteristics such as: age, sex, teaching style and experience. Details of these studies are now discussed as this evidence has direct relevance to the development of teacher-based assessments such as those which, it is argued, are appropriate for assessing the extended aims of primary teachers.

Herbert's study (1974) investigated how many factors might usefully be identified from teachers' ratings of classroom behaviour. Five clusters of pupil personality characteristics were identified for further study, these included; *personality problems or introverted neuroticism*, conduct problems or *extraverted neuroticism*, *competence* and *social extraversion* towards adults and/or children. The results of this study indicated that the teachers' notions of stability in children were closely related to their general approval or disapproval of pupils. When given a large list of behaviours to consider, teachers did not just use a single concept of approval or disapproval when rating, but used a



conceptual scheme which differentiated between "*badness*" in the sense of low competence (poor motivation, lack of persistence), "*badness*" in the sense of naughtiness, aggression, defiance or playing up and "*badness*" in the sense of being an introverted neurotic.

Evidence that teachers rate the class behaviour of boys as being less appropriate than that of girls has been found in a number of studies (Davidson and Lang, 1960; Douglas, 1964; Kellmer-Pringle *et al.*, 1966; Davie *et al.*, 1972; Brandis and Bernstein, 1974; Ingleby and Cooper, 1974; Stevenson *et al.*, 1976). Douglas (1964) also found that primary school teachers assessed girls as having a more serious attitude to their work and as being more satisfactory pupils than the boys, who were often reported as restless and inattentive in class. Hartley (1978) also found the expected trend in that boys were rated at the more "*unfavourable*" extreme of each scale, although the opposite occurred in some reception classes. The five scales used related to whether the pupil was able/unable to concentrate, immature/mature, tidy/untidy, gentle/rough, noisy/quiet. Hartley (1979) noted that these factors were associated with the concept of the "*successful*" first school pupil. It can be argued, however, that these ratings may in fact reflect the reality of the situation rather than teacher bias. Without evidence of pupil behaviour this cannot be determined.

Taylor (1976) studied the perceptions of 24 male and 24 female primary teachers, from 18 urban and rural schools, using Kelly's repertory grid (1955). Taylor had two objectives: first, to map out the attributes which teachers used to explain and predict the activities and performance of the children in their classrooms and second, to find out whether teacher characteristics of sex and formality contributed in any way to a teacher's perceptual framework. The degree of formality of the teachers was measured using the Minnesota Teacher attitude inventory. Teachers were asked to consider the important ways that primary children were alike. The most frequently produced constructs in this study included the following: the pupils were well behaved, good readers, good at maths, intelligent, good home background. These accounted for 37% of all the characteristics used by teachers. Further analysis showed that academic achievement was the most dominant pupil characteristic in the teachers' perceptions with 44% of the constructs falling in this category. Male teachers made more use of academic constructs since the proportion was significantly higher than in the case of female teachers.

Nash (1976) addressed the central question of whether a teacher's attitudes influenced

the performance of pupils in school. This study was carried out with 8 junior school teachers with some 236 pupils. Core constructs identified by the use of Kelly's (1955) repertory grid technique were: hard-working - lazy, mature - immature, well-behaved - poorly behaved. These teachers seemed to perceive their pupils primarily in terms of their work habits, maturity and classroom behaviour. Pupils were judged not in terms of their academic ability but more in relation to their personality attributes which were regarded as important to good progress at school. The differences in emphasis on academic and personality traits seen in the Nash (1976) and Taylor (1976) studies may have resulted from the fact that Nash worked in a progressive environment stressing personal characteristics whilst Taylor worked in a more traditional context which would be more likely to stress academic performance. Such differences in teaching style and their influence on teacher perceptions are further investigated in this study.

Worral and Cowan (1983) investigated the role of situational biases in teacher impressions. They found that even allowing for possible correction of the first impression there was still a residual element in the teachers' judgements related to their perceptions of bad or good pupils. The view has been expressed that such impressions can be crucial determinants of later stable impressions (Good and Brophy, 1974; Rogers, 1982). An opposite view has suggested that in the real-life classroom, where a teacher has day to day interactions with his or her children, opportunities to develop a more veridical picture of the child and correct biased impressions would have occurred.

Teachers' estimates of intelligence for children about whom they have no other information have been shown to be influenced by the child's sex and social class (Cooper *et al.*, 1975) as well as ostensibly unrelated factors such as physical attractiveness (Clifford and Walster, 1973) or even name (Harari and McDavid, 1973). Goodacre (1967) showed that teachers of infants when making a subjective judgement of children's reading ability considered children they thought came from middle and upper working class backgrounds to have a higher reading ability than those they judged to come from lower class homes. When these children were tested on the N.F.E.R. National Survey 7+ Reading Attainment Test this marked difference was not so apparent. The teachers seemed to be reflecting the standards they applied to different social backgrounds; middle class children were expected to perform better than working class children and so they did in the eyes of the teachers even if the measured attainment did not reflect such marked differences.



Morrison and McIntyre (1969) noted that several studies of teachers' ratings of the personality traits and attainment of pupils have shown that teachers in British schools tend generally to discriminate three clusters of characteristics: first, pupil attainment, with attention centred on general ability, English and arithmetic; second, general classroom behaviour and attitude towards teachers and finally, an assessment of social traits which relate particularly to the teacher's perception of how well a pupil has "*got on*" with other pupils (Hallworth, 1962; McIntyre *et al.*, 1966).

McIntyre *et al.* (1966) investigated the variation in teachers' ratings on the basis of teacher characteristics. The assessment pattern was shown to be modified in relation to particular environments; for example, primary or secondary school, middle or working class home backgrounds; age, sex and educational attitudes of teachers. Male and female teachers differed in the extent to which their ratings of attainment were independent of ratings of classroom behaviour. Teachers tended to make a more uniform generalised assessment of a girl than they did of a boy; younger teachers showed more concern with classroom behaviour and older teachers with attainment; the tough-minded teacher stressed quietness, whilst the particular interests of some teachers were reflected in their use of a single trait category, as in the case of male teachers rating the games ability of boys.

The second perspective on the study of teachers' judgements of pupils has drawn on a sociological analysis of the process of *typification* or *stereotyping* of pupils. Hargreaves (1977) described three models that have been used to account for the process of pupil typification. The *ideal matching* model was associated with the early work of Becker (1952) who studied social class variations in teacher-pupil relationships. Interview data drawn from discussions with 60 teachers in Chicago schools produced a picture of the ideal pupil. He argued that teachers, therefore, have an image of the ideal client against which actual clients were matched and evaluated as either good or bad. In the classroom, the good pupils were those who approximated closely to the teacher's ideal image, so presenting him/her with few problems. A good pupil co-operated with the teacher, enabling him/her to do his/her job effectively and with the minimum of conflict or stress. This type of pupil, according to the teacher's own descriptions, tended to come from a middle or upper middle class home. A bad pupil deviated from the ideal by not displaying ability and successful learning, in not conforming to classroom rules and regulations and offending the teacher's moral standards. Perceived social class was identified as one of the factors by which teachers tended to classify their pupils. This form of stereotyping of

pupils has also been shown in later studies (Jackson, 1964; Lacey, 1970; McIntyre *et al.*, 1966).

The *characteristics* model was based on early studies of pupil perception and expectations (Hallworth, 1962; Taylor, 1962; Musgrove and Taylor, 1969). In this model teachers and pupils typify each other by bringing together certain groups of characteristics. Whilst, the *dynamic interactionist* model as expounded by Gannoway (1976) has taken account of changes in perception and typification over time with regard to pupil typification of teachers. Hargreaves *et al.* (1975) identified four stages to this kind of typification. First, the teacher hypothesised that the pupil was a certain kind of person known as the *speculation* phase. This early hypothesis was then put to the test and extended during *elaboration*. *Stabilization* occurred when the teacher developed a highly complex but stable typification which was then resistant to change. Finally, *transformation of type* occurred if the teacher changed this stable typification. This last stage had, however, only been posited and had not been developed from the study.

Other studies have investigated the characteristics of teachers that have influenced their perceptions of pupils. For example, King (1978) investigated the meanings given by three teachers to their work in infant classrooms. He found that the processes of typification were inseparable from those of assessing the work and behaviour of individual children. He concluded that

*"typification, assessment, teaching, learning, controlling and being controlled were all aspects of the same flow of action and interaction in the classroom,"* (King, 1978, p.58)

Two important points were identified from the study. First, teachers drew mainly on their own experience of the child in the classroom situation. Second, three aspects of the child's classroom behaviour were noted and built into the pupil typification. These were; compliance with the teacher's rules of behaviour, relationships with other children and learning progress in terms of development not intelligence. King argued that the consequences of such typifications were real for both teachers and pupils in that they affected the classroom arrangements, the ways teachers dealt with an individual pupil's work and the methods used to control pupil behaviour, in other words, the style of teaching used in the classroom. Thus, it can be argued that the teacher's style would also affect the way in which teachers have formed their judgements of pupils and this factor has also been investigated in this study.



The relative merits of each of the three models of typification described were not of central concern to this study. However, both psychological and sociological research have demonstrated that a variety of different factors have contributed to and influenced the way teachers have perceived and made judgements of pupils' abilities and behaviour. A number of studies have indicated that characteristics of pupils such as social class, ability, sex, age and classroom behaviour have been used as a basis for judging pupils' ability and achievement thus contributing to the inaccuracy and unreliability of teachers' assessments. All these factors have been the subject of research in this study so that their effects could be accounted for in the interpretation of the data on pupil performance derived from the teacher-based assessments. Other pupil characteristics that have been shown to influence the way teachers have judged pupils were beyond the scope of this study. Factors which have not been taken into account but have been investigated by other researchers include: the misrepresentation of the ability of pupils (Rosenthal and Jacobsen, 1968); race (Coard, 1971; Cooper *et al.*, 1975) and staff room reputation of pupils (Hargreaves, 1972).

In conclusion, this section has attempted to describe developments in educational measurement and the rationale for selecting particular assessment methods. In addition, the problems and solutions associated with using teacher-based assessments have been considered. Despite the inherent difficulties described it was considered important not to undervalue the contributions which teachers could make to the assessment of pupil learning in order to provide evidence related to the aims of education not assessed by the Richmond tests and tests of study skills previously described.

### **2.3 INVOLVEMENT AND PROFESSIONAL DEVELOPMENT OF TEACHERS IN THE RESEARCH.**

It has been suggested in the previous chapter that it was considered desirable to use the teachers' assessments of their pupils for a variety of reasons. It has been argued that they would reflect teachers' aims for an extended primary curriculum more accurately, since other types of assessment have not made use of a teacher's knowledge of pupils' everyday performance in class. Indeed, it was argued by Alexander (1984) that although formal procedures were more accessible to public scrutiny than informal measures, they did not necessarily have any stronger claim to objectivity. He advocated that teachers'

assessments should be used since they provided a more valid and reliable basis for classroom diagnosis and action. Her Majesty's Inspectorate have also argued that formal procedures have tended to neglect the diagnostic dimension of assessment and

*"concentrated overmuch on generalised and summative judgements of attainment"*  
(D.E.S. 1983a, paragraph 614).

It was, therefore, hoped that by providing alternative methods of assessment making use of their knowledge of pupils, teachers would be able to counter the criticisms of that time by demonstrating those aspects of pupil learning that were not assessed by means of tests of the basic skills. It can also be argued that any development in the ability of the teacher to accurately diagnose pupil needs, provide appropriate instructional materials and assess their performance validly and reliably might lead to improvements in the quality of teaching and learning in the classroom and of educational standards in general. Thus, the use of teachers' assessments was seen as desirable since this would provide for teacher involvement within the Oracle research and a focus for professional development at a later stage of this study. A brief summary of the issues and directions taken in this regard is now presented.

The forces leading to a greater involvement of teachers in research into their own classrooms and schools, curriculum development and evaluation have already been discussed in Chapter 1. The process of developing the teacher-based assessments was in many ways similar to any curriculum development work and therefore involved teachers at various stages and in different ways. This research was also able to draw upon the experiences of curriculum developers as well as the strong tradition of teacher involvement in assessment and examinations within Britain in the development of the teacher-based assessments. A brief review of teacher participation in the development of assessment procedures follows before considering the role of teachers in this research.

One of the most significant developments in assessment arose from the introduction of the Certificate of Secondary Education in 1963. Three types of teacher involvement were suggested. All of these embodied the principle of teachers playing a major part in the formulation of both the curriculum content and the pattern of assessment and examinations. There was, also, a desire to relate the examination more closely to the work done in schools. Under the auspices of the Schools Council, investigations were carried out to examine alternative modes of assessment; in particular, the continuous assessment of course work by teachers (Eggleston and Kerr, 1969; Hoste and Bloomfield, 1975;



Cohen and Deale, 1977). A number of studies were conducted into the feasibility of defining pupil attainment in such a way as to allow for teacher assessment of a number of specific criteria related to pupil classroom performance. These assessments would then be moderated by the use of a variety of items and tests designed to assess each criterion independently (Eggleston and Newbould, 1969; Eggleston and Lobel, 1969; Kelly and Lister, 1969). Results from these studies suggested that continuous assessment was a viable mode of assessment providing suitable moderation procedures were available and used in conjunction with more traditional forms of summative examinations. As a result, secondary teachers became more involved in course design, specification of objectives, preparation of test items and the assessment of pupil performance. This suggested that similar procedures could be designed to assess behaviourally defined criteria based on each aim of education under investigation within this study.

Teachers participating in the Certificate of Secondary Education became increasingly involved in the processes of both curriculum development and assessment and it is assumed developed some measure of expertise as a result of this involvement. This type of development served as a model for the strategies used within this research for the involvement of teachers in developing and assessing pupil learning in primary schools. The idea of developing short "tests" related to behaviourally defined criteria derived from course objectives was particularly useful in enabling the moderation of teachers' judgements of pupil performance.

In Scotland, research into assessment practices in secondary schools was initiated by the teaching profession. In 1972, a working party was established by the Headmaster's Association of Scotland (now the Headteacher's Association of Scotland) to consider national examinations. They sought the co-operation of Her Majesty's Inspectorate and the Scottish Council for Research in Education to consider

*"the manner and range of assessments in secondary schools that might result in a form of report or certificate applicable to all pupils completing S4."* (Scottish Council for Research in Education, 1977, p.2)

This research into profiling in the secondary school was conducted with regard to two principles. First, meetings with all interested parties such as headteachers, employers' organisations, teachers' organisations, colleges, commercial groups and parents were held so that all would be informed of the work. Second, the decision was made that whatever types of assessment were used these would be based on teachers' knowledge of

their pupils, thus placing control of the curriculum in the hands of teachers through the assessment of pupils. This commitment to teacher involvement was not without difficulties. During initial field trials of the profiling research it became clear that teachers involved in making such assessments required more guidance and support in completing what were quite complex profiles. Teachers were subsequently involved in the preparation of appropriate subject guides and in the informal evaluation of the acceptability, practicality and validity of the profiles (Black and Dockrell, 1981; Black and Broadfoot, 1982). This work has subsequently led to diagnostic assessment research as reported by Black and Dockrell (1984).

The Schools Council continued to provide considerable momentum to the development of the curriculum and assessment practices in England and Wales during the 1970's. Two projects: *Progress in Learning Science* (Harlen *et al.*, 1977a, 1977b) and *Communication skills in early childhood* (Tough, 1976, 1977, 1979) were particularly concerned with curriculum developments related to the appraisal and assessment of pupils in primary schools. The *Progress in Learning Science* project (Harlen *et al.*, 1977a, 1977b) developed as a result of the need to provide appropriate forms of assessment for the evaluation of the Science 5-13 curriculum materials. This project was of particular relevance to this study since it dealt with the development of assessment procedures which were related to some of the extended aims of primary education identified in the Ashton *et al.* survey (1975).

The project team of *Progress in Learning Science* enlisted the help of teachers to devise checklists for other teachers to use, both for picking up information as immediate feedback in helping learning and for recording progress. The involvement of teachers was in two stages. First, several working parties of teachers were established to discuss the goals of inquiry-based work in the age range 5 to 13 years. These groups produced 24 goals which included abilities, concepts and attitudes. In the second stage, discussions with teachers focused on each of these goals in turn and considered how their development showed in the actions and responses of children. Statements of behaviour considered to be representative of progressive levels of development were identified and gradually refined in a series of trials. This resulted in the preparation of two checklists which were intended to structure observation, focusing on significant aspects of children's behaviour and providing a framework for interpreting the observations in terms of progress towards the project's goals.



Harlen (1976b) argued that teacher involvement in such curriculum development projects was effective and desirable. It was also suggested that there was a need to enhance the understanding of other teachers who were not involved in the development process. As a result teachers' groups were established to prepare in-service materials for a larger audience. These materials were later published (Harlen *et al.*, 1977a, 1977b) and were the basis of an extensive in-service programme for teachers taken up by a number of local education authorities.

The Communication skills in early childhood project (Tough, 1976, 1977) looked at developing children's use of language in the infant classroom. This was followed by an extension of the approach to the age range 7 to 13 years (Tough, 1979). This project aimed explicitly to produce in-service materials for teachers which would enable them to appraise children's use of language more effectively. Teachers were also involved in the development process; collecting examples of children's use of language, undertaking case studies of pupils; piloting materials, analysing and discussing their design and effectiveness, putting forward suggestions for their revision and new ideas for inclusion.

It can be argued that curriculum developments such as the Schools Council Progress in Learning Science project (Harlen, 1977a, 1977b), Communication Skills in Early Childhood project (Tough, 1976, 1977), the Scottish Council for Research in Education work on pupil profiling and the introduction of Mode 3 examinations went some way towards developing and using the everyday evaluations of teachers as the basis for monitoring and assessing pupil learning for summative as well as formative purposes. The use of teachers' knowledge of their pupils was therefore seen as an important aspect of countering the criticisms of primary pedagogy and the call for increased accountability prevalent in the late 1970's and early 1980's by providing a focus for professional development. However, where teachers were involved in the development of such assessment procedures the emphasis shifted from the use of more formal tests to the use of checklists, teachers' observations and the assessment of classwork, often in the form of diagnostic, criterion-referenced procedures and teacher assessments with all the attendant problems described in the previous section. The importance of teachers' assessments of pupils should not be underestimated in this context; evidence from studies that have investigated the relationship between teachers' perceptions of pupils and how well these pupils have performed has suggested that the way teachers have viewed their pupils has affected achievement.



Research into teacher expectancy effects has contributed to an understanding of the ways in which teachers' judgements have influenced the performance of pupils. This research was stimulated by the work of Rosenthal and Jacobsen (1968). This study appeared to demonstrate that if teachers were given evidence of a pupil's ability to do well in the form of an inaccurate test score, that pupil would do better than expected in subsequent examinations. This effect was termed a "*self-fulfilling prophecy*". In well over a hundred subsequent studies researchers have argued about the extent to which teacher expectations have self-fulfilling prophecy effects. Attempts to reproduce the original findings have often been unsuccessful and resulted in much controversy on the validity of teacher expectancy effects. A number of reviews of the literature in this area have been published (Brophy and Good, 1974; West and Anderson, 1976; Dusek and Joseph, 1983; Braun, 1976; Persell, 1977; Cooper, 1979; Good, 1980; Brophy, 1983). Brophy concluded from all this research effort that

*"These data imply that teacher expectations do not always or automatically function as self-fulfilling prophecies, but that they can and often do have these effects ... the existence of a teacher expectation for a particular student's performance increases the probability that the student's performance will move in the direction expected, and not in the opposite direction."* (Brophy, 1983, p.633)

The research into expectancy effects, therefore, has shown that what teachers have expected from pupils has been likely to affect the way in which a particular pupil has performed in school. Clearly accurate assessment and evaluation of a pupil's ability and behaviour is needed so that whatever and however such expectations are communicated to the pupil they are at an appropriate level. This has been of particular importance in the light of the suggestion that teachers often under-estimate the ability of their pupils and have not provided curriculum content to match their needs (Bennett, 1978). Studies by Pigeon (1970) and Good *et al.* (1975) have also suggested that teacher expectancies have been communicated through instructional materials.

However, there is still an under-estimation of the amount of time and effort required to implement such improvements in teachers' assessments on a large scale. Elliott (1980a) and Clift *et al.* (1981) have investigated the implications of using such methods to monitor standards. They considered their format, the categories or questions included, intended mode of use and the problems of clarity, ambiguity and time required for their completion. They demonstrated that if such procedures were used as intended there would be a considerable impact on school life, with a heavy commitment on the part of



teachers to acquire new skills in the assessment and monitoring of pupil achievement. In addition, in the previous chapter it has been argued that teachers are also often unaware of the need to change their practice and, if made aware, lack the necessary expertise to carry out such assessments resulting in the need for more professional development.

The improvement of the validity and reliability of teachers' judgements of pupils was, therefore, of crucial concern to the development of this research since the importance of demonstrating pupil achievement over an extended range of aims for primary education, enabling teachers to provide more appropriate curriculum and instructional materials, as well as more equitable treatment of individual pupils was central to effective teaching and learning and thus accountability in its widest sense. As a number of authors have suggested

*"Our behaviour towards other and interaction with them is guided by the evaluations and judgements we make of them."* (Downey, 1977, p. 78)

*"These perceptual differences almost certainly lead to more or less subtle differences in the behaviour towards groups and individuals."* (Morrison and McIntyre, 1972, p.171)

The most appropriate way of involving teachers in the research was considered to be through collaborative research where both the teachers and the researcher had differing areas of expertise. The role of the teacher was to provide an understanding of the problems of implementing the approach in practice. The teachers also had first hand experience of teaching a range of primary school age children and were expected to bring a sense of realism concerning the abilities of the children that were to be assessed so that materials and tasks were at an appropriate level of difficulty. The role of the researcher was to facilitate the working groups of teachers, to provide expert knowledge of assessment procedures and to pilot the materials before introduction to the schools taking part in the Oracle research programme.

During the life of the Oracle research programme the moves towards a more active engagement of teachers in research and the provision of school-focused and school-based in-service was increasingly evident. Thus, in addition to the greater involvement of teachers in developing the teacher-based assessment research within the Oracle research programme, it was suggested that teacher professional development through in-service provision would be desirable and appropriate for further dissemination of the research. In addition, the experience gained from conducting the research was expected to provide a



useful basis for the development of appropriate in-service provision. An in-service development project was thus designed and implemented at a later stage of the research to order to promote the use of such assessments in other schools and classrooms.

This evolved in two different contexts. The first was an attempt to establish a collaborative group of teachers who together with the researcher studied assessment practices in their classrooms as a means of identifying needs and to develop more appropriate assessment methods to meet these needs. The group was self-selected in that the teachers specifically enrolled for a ten week in-service course in assessment in the primary school for which there was no professional award. The second strategy used this pilot course as a basis for developing an in-service course offered as a strand in a full-time in-service B.Ed Honours degree programme based at Worcester College of Higher Education. This strand was developed to facilitate self-evaluation and classroom-based research using the areas of assessment and evaluation as the focus of enquiry. This focus was selected since it was hypothesised that this was a means of improving the quality of teaching by promoting self-evaluation and the quality of pupil learning by ensuring more accurate diagnosis of learning needs and effective evaluation of outcomes.

The ideas embodied within teacher self-evaluation were important to the development and evaluation of this in-service provision. The strategy advocated was one which embodied elements of self-evaluation on the part of the researcher, as well as the teachers, in order to provide a role model for professional development, since the questions of how teachers learn "*professionally*" (whether as the provider or receiver of instruction) and the means whereby "*re-education*" (Ruddock and Kelly, 1970) can be enhanced were central to this part of the study. During the project the researcher monitored the course in progress and on the basis of the data collected modified the nature and direction of the course as a means of self-evaluation for professional development. The course was premised on the need for the researcher as well as the teachers to build into their teaching the ability to change and critically appraise their own and the learners' activities in order to bring theories-in-use in line with espoused theories and to examine critically such theories of action in the light of experience gained and new knowledge acquired during the course. This applied to both researcher (as provider of the course) and the teachers in implementing their learning in their own classroom contexts.

This model was chosen for a number of reasons. First, the research was concerned



with a specific context where the problem of improving the quality of professional learning was central. Second, the collection of data for self-evaluation allowed for the participation and collaboration of the teachers in the process of the research and for modifications to be made as the course progressed. The evaluation of the in-service provision was also designed to investigate whether teachers could implement in practice (theories-in-use) their knowledge and understanding of assessment (espoused theories) in the classroom setting and how far new theories of action developed as a result of their in-service experience. The researcher intended to model this type of self-evaluatory professional development based on practical deliberation within the course.

## 2.4 CONCLUSION

In the preceding discussion various forms of assessment such as *formal* and *informal* methods were reviewed. The selection of assessment procedures within the Oracle research programme was then considered. It was suggested that judging the suitability of any form of assessment should take account of the appropriateness of the procedure for the purpose for which it was to be used. The selection for this research was therefore determined by the need to adequately describe pupil progress and performance in the investigation of the effectiveness of various teaching styles. This selection was based on the application of certain criteria for the evaluation of various methods of assessment; for example, norm referenced and criterion referenced assessments, pupil referenced and diagnostic assessments and teachers' assessments. From this discussion it was clear that any *one* type of assessment procedure did not provide a satisfactory range of information on pupil learning in order to demonstrate the relative effectiveness of different teaching styles. Thus, three different types of assessment were included within the context of the Oracle research programme; the modified Richmond Tests, "*tests*" of study skills and the teacher-based assessments.

By using a range of different techniques drawn from both formal and informal methods of assessment the Oracle research programme provided evidence of pupil achievement on a wide range of cognitive outcomes arising from current primary practice. The use of teachers' assessments complimented other measures so that the whole range of assessments provided a more accurate picture of the nature of pupil performance. In relating these to teaching styles, more valid evidence of their effectiveness for particular pupil types was made available (Galton and Simon, (eds), 1980). In addition, the



particular difficulties associated with the use of teachers' assessments were discussed.

In an attempt to compromise between the conflicting demands of the accountability movement and the professional development of teachers, the research recognised both the importance of using the teachers' knowledge of pupils as a means of assessing an extended range of primary curriculum aims and the benefit of involving teachers in the development of these assessments. However, this commitment to the use of teachers' assessments of their pupils raised certain methodological issues. These have related principally to reliability and validity. As a result of these concerns the research has endeavoured to develop suitable ways of minimising the unreliability of teachers' ratings both within the research and as a basis for developing in-service provision.

Thus, the development of more appropriate assessment procedures has been advocated in the context of criticisms of the methods of assessment used in teacher effectiveness research, curriculum evaluation and accountability. It has been seen as a means of providing teachers with ways of monitoring pupil performance in the classroom in order to provide evidence of their success, or otherwise, in implementing a progressive primary curriculum within the political and economic climate of the late 1970's and 1980's, which was increasingly advocating public accountability. Further it has adopted a methodology which has been designed to promote change in the reliability and validity of teachers' judgements of pupils' classroom performance and facilitate self-evaluation. The purposes of the research can be summarised as follows:

- a) To counter criticisms of previous teacher effectiveness research by developing measures to collect more extensive data on pupil performance. These assessment procedures were to be used in the Oracle research programme to investigate the effectiveness of different teaching styles.
- b) To provide valid and reliable assessment procedures for primary school teachers to monitor aspects of pupil performance in the classroom in line with the changed nature of their aims in the primary curriculum and provide a basis for countering the moves towards externally imposed accountability systems.
- c) To develop and evaluate in-service training procedures designed to improve the quality of teachers' assessments.



**PAGE  
NUMBERING  
AS ORIGINAL**

A discussion of the scope and limitations of the research, its design and the methods used in the study of teachers' assessments of pupil performance is thus the subject of the following chapter.



## **CHAPTER 3: RESEARCH DESIGN, METHODS OF DATA COLLECTION AND ANALYSIS.**

### **3.1 SCOPE, LIMITATIONS AND METHODS OF THE RESEARCH**

In the previous chapter, methodological issues have been addressed regarding the use of teachers' assessments of primary school pupils' achievement in the evaluation of the effectiveness of different teaching styles and as a basis for the development of in-service provision. A number of questions arising from this discussion are now considered in order to clarify the scope, limitations and methods of this research.

#### **3.1.1 How is the primary curriculum to be defined and by whom?**

This first question was of concern since any measures developed were required to match more closely the practice of teachers in primary schools. Alternative strategies for the definition of the primary curriculum were available. One possibility was to start with groups of teachers and to encourage them to describe their own classroom practice in order to define what was taught and then select areas which were inadequately assessed as the basis for the development of alternative methods of assessment. There was one major difficulty associated with this approach. As suggested in the previous chapter, teachers were often found to be unaware of their own practice and therefore to attempt to involve teachers at this early stage of the programme in defining the curriculum and describing their own practice would have meant a major shift of emphasis within the research.

An alternative was to use the survey of primary teachers' opinions (Ashton *et al.*, 1975) as the basis of definition of the primary curriculum. The assumption was made that these aims, because they were formulated by teachers, would provide an accurate reflection of current primary school practice at the inception of this research. In addition, these teachers were likely to be more representative of the profession as a whole than any smaller sample who might have been interested in participating in this research. Given that these seventy two aims of primary education provided an adequate description of the curriculum then some selection from the seventy two aims listed within the survey was necessary since it was outside the scope of this investigation to consider all the areas which included aspects of aesthetic, emotional/personal, intellectual, moral, social,



physical and spiritual/religious development.

This selection was made with regard to four criteria, which were determined by the nature of the overall research programme. First, those aims which could be assessed by established means such as the Richmond Tests of basic skills or the *"tests"* of study skills described in Chapter 2 were excluded. Second, those aims unrelated to intellectual development were not considered since the main focus of the Oracle research programme was to investigate the effectiveness of different teaching styles in terms of pupil achievement. Third, only those aims which were judged by teachers in the original survey to be of *"utmost"* or *"major"* importance were considered. Fourth, the teachers involved in developing the teacher-based assessments made the final selection on the basis of their own individual concerns and interests, in the expectation that these would more accurately reflect the needs of the teaching profession at that time. This allowed for the teachers who would be involved in the development of the teacher-based assessments to make some contribution to determining the agenda for the research in anticipation that this would facilitate the workings of the teacher groups.

### **3.1.2 What type of procedure is appropriate for the assessment of the selected aims of primary education?**

It has been argued that the use of teachers' assessments was the only appropriate method given the nature of some of the aims selected for the purposes of this research. These teachers' assessments would be made with reference to criteria which were to be developed by teachers as unambiguously as possible in order that other teachers would be able to use them effectively. However, it has been noted in the previous chapter that when such criteria are used there are still sources of error which reduce the validity and reliability of teachers' assessments.

It was suggested that a moderating procedure for these assessments was necessary so that sources of error could be investigated and taken into account in the analysis and interpretation of results. This led to the development of checklists of abilities linked to each of the aims selected, supported by structured activities which were used as moderating instruments to facilitate the reliable and valid assessment of pupil performance. It was also proposed to use a variety of means to facilitate the development of these measures so that their validity and reliability were maximised.



### **3.1.3 Do teacher-based assessments provide a valid and reliable alternative form for the assessment of pupil performance?**

The investigation into the reliability and validity of the teacher-based assessments forms a substantial part of this study. The methods used to investigate this area have included:

- \* the correlation of teachers' ratings with pupil performance on the structured activities,
- \* factor analysis and comparison of scores on the teacher-based assessments with performance on similar assessment procedures if available, with the modified Richmond Test results in mathematics, language and reading and study skills exercises,
- \* the degree of inter- and intra-rater agreement on assessments of the same samples of pupil performance and the stability of teachers' assessments over time.

### **3.1.4 What pupil and teacher characteristics influence judgements made by teachers of pupil performance using teacher-based assessments?**

A large number of previous studies have indicated that the assessments of teachers have been subject to inconsistencies linked to aspects of the teachers themselves and the use of a variety of pupil characteristics to reach an assessment of pupil performance. This study was undertaken to investigate a number of factors that have contributed to such assessments by the use of covariance analysis. In particular, an investigation of the relationship between teachers' assessments and their *observed* teaching style was carried out since detailed observational data of classroom events was collected in the Oracle research programme. In this way it was possible to see whether certain styles of teaching presented more difficulties in the achievement of valid and reliable assessments. Previous studies have indicated that certain teacher characteristics which could not be changed, such as personality, contributed to bias in teachers' judgements. However, if assessments were related to teaching style then this would suggest that it was possible to change aspects of a teacher's style resulting in some improvement in the validity and reliability of that teacher's assessments.



The previous discussion of teacher assessments also suggested a number of important variables which have been shown to influence teachers' judgements; of these, pupil sex, age, social class, achievement on tests of basic skills, and classroom behaviour (pupil type) have been studied. Other variables such as race and pupil personality were outside the scope of this study and their impact on the teachers' ratings has not been evaluated.

### **3.1.5 What are the characteristics of pupil performance on the teacher-based assessments in relation to the selected aims?**

Given that some of the data generated from the teacher-based assessments provided valid and reliable information on pupil performance, this could then be used to investigate the relationships between pupil performance in these areas to a number of pupil variables. Again, the availability of extensive observational data on pupil behaviour and the style of teaching experienced by these pupils enabled the investigation of a whole range of processes and skills which have been considered important by primary teachers but which have not been previously documented. Other pupil characteristics which could be related to performance on these extended aims of primary education were also investigated within this study and included: pupil age, sex, social class and achievement on basic skills tests.

### **3.1.6 Can teachers improve their ability to assess pupils and if so in what ways do they need to modify their teaching styles and assessment practices?**

The final area for consideration has been whether such teacher-based assessments have relevance to the day-to-day classroom practice of teachers in general. The opportunity to test out some of the recommendations and suggestions for the involvement of teachers in research and their own professional development discussed in the previous chapters arose on completion of the Oracle phase of the research. As a result of a pilot study based on a collaborative research model, a case study of the effectiveness of in-service provision was carried out where the ability of teachers to acquire more knowledge of assessment methods and incorporate such techniques into their classroom practice was investigated. This study was based on the principles outlined in the previous chapters for the professional development of teachers. This study was necessarily a case study involving small samples of teachers and the results have not, therefore, been generalisable to a wider population. However, the evidence from this case study could suggest future directions for study.



## **3.2 HYPOTHESES**

### **3.2.1 The following working hypotheses are investigated in this research into the validity and reliability of teachers' assessments.**

- \* Valid and reliable teacher-based assessments, consisting of structured activities and teachers' ratings of pupil performance can be developed for the measurement of pupil performance for the aims listed below:**

- a. The child should be able to listen with concentration and understanding.*
- b. The child should develop some inventiveness and creativity in some fields; for example, by painting, music, mechanical things, poetry, movement.*
- c. The child should know how to acquire information other than by reading; for example, by asking questions, by experimenting, from watching television.*
- d. The child should be able to:*
  - i. Compute the four arithmetic rules using his knowledge of number, multiplication tables and different units of measurement.*
  - ii. Use mathematical techniques in his everyday life, i.e. estimating distances, classifying objects, using money.*
- e. The child should know how to convey meaning clearly and accurately through speech for a variety of purposes; for example, description, explanation, narration.*

- \* Teachers' assessments of pupil performance in relation to the five aims listed above are influenced by the following pupil characteristics: sex, age, social class, achievement on tests of basic skills and classroom behaviour (pupil type).**

- \* Teachers' assessments of pupil performance in relation to the five aims listed above are correlated to their teaching style.**

### **3.2.2 The following working hypotheses are investigated in this research into the nature of pupil performance on the teacher-based assessments.**

- \* Pupil performance on the structured activities for each of the five aims listed above is correlated with the pupil characteristics of sex, age, performance on the modified Richmond Tests of basic skills, social class or pupil type.**

- \* Pupil performance on the structured activities for each of the five aims listed is correlated to the teaching style experienced.

**3.2.3** The working hypotheses relating to the evaluation of the effectiveness of in-service provision within the B.Ed. Honours degree programme are detailed below:

- \* Professional learning involves the recognition of theories in use and espoused theories and the nature of any discrepancy between them is necessary for the generation of new theories of action and thus self-evaluation.
- \* The study of assessment procedures facilitates this process of professional learning and leads to teacher self-evaluation and which is then transferred to the classroom context.

The remainder of this chapter details the research design, methods of data collection and analysis used in the study of the validity and reliability of the teacher-based assessments and the analysis of pupil performance in relation to the selected aims. The design, methods of data collection and analysis used in the evaluation of the in-service project are also considered before reporting the results in Chapters 4, 5 and 6.

The next section of this chapter outlines the context of the research reported in this thesis by describing the Oracle research programme. This is of particular importance since the research design, methods of data collection and analysis used to investigate the hypotheses listed above were closely linked to this longitudinal, process-product study. Data derived from this research programme has facilitated this investigation of teachers' assessments by providing evidence related to a number of factors; for example, pupil progress and performance on tests of basic skills, patterns of classroom behaviour and interaction of both pupils and teachers and pupil social class, which otherwise would have been beyond the scope of this study. It should be noted that the following section (pp.103-111) has drawn on the published accounts of the Oracle research in order to provide the discussion of research design, methods of data collection, analysis and descriptions of teaching styles and pupil types included in this chapter. Appropriate acknowledgements have been made to identify the sources of this information.



### 3.3 RESEARCH DESIGN

#### 3.3.1 The Oracle research programme.<sup>1</sup>

The Social Science Research Council programme: Observational Research and Classroom Learning Evaluation (Oracle) was developed in the climate of accountability and debate on the relative merits of different pedagogies in primary schools described in the review of the literature presented in Chapter 1. The main objective of this programme was to study the relative effectiveness of different teaching approaches across the main subject areas of primary school teaching. The programme also included a study of transfer to the next stage of education and the development of new forms of teacher-based assessment which are the subject of this thesis. The programme was designed as a longitudinal, process-product study where data collection was carried out over a period of three years.

The main research method was that of classroom observation since

*"systematic observation of teacher and pupil behaviour would be of help to teachers facing these criticisms, by providing descriptions of current classroom practice against which they could evaluate their own teaching. By collecting information about pupil performance, while at the same time engaging in these observations, it was also intended to present evidence about many of the issues which dominated the debate between supporters of progressive and traditional methods."* (Galton and Simon, (eds.) , 1980, p. 7)

Thus in the Oracle study an attempt was made to remedy some of the deficiencies of earlier investigations of teacher effectiveness research by collecting observational data on both pupils and teachers and by developing new forms of assessment to reflect the extended nature of the primary curriculum. In addition, the research programme was able to replicate the analysis of teaching styles and pupil types carried out in the first year since some of the same teachers and pupils were involved in the study in subsequent years, as well as carry out the investigation of a number of aspects related to transfer from one stage of education to the next.

---

<sup>1</sup>Full details of the research design, methods of data collection and analysis and results of the study have been published in a number of volumes which document all aspects of the programme (Galton, Simon and Croll, 1980; Galton and Simon (eds.), 1980; Simon and Willcocks (eds.), 1981; Galton and Willcocks (eds.), 1983).



### **3.3.2 Selection of schools and classes**

The selection of schools and classes in the Oracle research programme was determined by the series of studies in which pupils from primary and middle schools were observed before and after their transfer to the next stage of education. Three local authority areas were studied, each with a different age of transfer. The age of the pupils in the sample ranged from 8+ to 10+ in the first year of the study. Two transfer schools were selected from each local education authority area. This selection was based on information provided by local advisers as to the nature of the curricular and organisational patterns of each pair of transfer schools so that contrasting schools could be studied which drew pupils from similar catchment areas. Once agreement had been obtained from each transfer school the selection of primary and middle schools was pre-determined.

The sample included all those feeder schools from which pupils were to move to the selected transfer schools. The teachers selected for observation were those who happened to be teaching the appropriate classes during the year. Details of the sample are given in Tables 3.1 and 3.2. It can be seen from Table 3.1 the schools selected were drawn from suburban areas centred on former villages, which had been developed to take city overspill, new private housing estates and urban inner city areas. This sample ensured that the major social class groupings were represented in the pupil population. Table 3.2 indicated that there were approximately the same proportion of male teachers in the Oracle schools as found in junior schools without infant classes. Department of Education and Science figures (1976) indicated 35.4% of teachers in such schools were male as compared to 34.5% in the Oracle schools. A proportion of the schools were open plan (25.9%) and an equivalent percentage were vertically grouped. The majority of pupils in the study were aged 8+ (46.6%).

It was important to ascertain the comparability between this relatively small sample of classes and teachers and the total population from which they were drawn. Comparisons were made between the Oracle sample and figures published by the Department of Education and Science in terms of the age of the teachers and class size. Although the teachers and their classes were not randomly selected it can be seen from Table 3.3 that the characteristics of the teachers and the size of their classes closely matched that of the country as a whole. The major discrepancy arose in the number of small village schools and larger inner city schools being under represented. The average class size of 29.9 also approximated the national average of 29.2.



**Table 3.1 School characteristics<sup>2</sup>**

	Number of schools	Number of classes
<i>Suburban/village</i>		
New: 1965+	1	5
Post-war	—	—
Pre-war	—	—
Mixed	3	14
<b>Total</b>	<b>4</b>	<b>19</b>
<i>Housing Estates</i>		
New: 1965+	8	20
Post-war	2	6
Pre-war	—	—
Mixed	—	—
<b>Total</b>	<b>10</b>	<b>26</b>
<i>Urban inner city</i>		
New: 1965+	1	3
Post-war	—	—
Pre-war	3	7
Mixed	1	3
<b>Total</b>	<b>5</b>	<b>13</b>
<b>Total</b>	<b>19</b>	<b>58</b>

---

<sup>2</sup> Taken from *Inside the primary classroom*, (Galton *et al.*, 1980, p. 25).

**Table 3.2 Characteristics of classes and their teachers by local education authority area.<sup>3</sup>**

		L.E.A.			Total %
		A	B	C	
<i>Sex of teachers</i>	Male	3	12	5	34.5
	Female	6	21	11	65.5
<i>School type</i>	Open plan	1	6	8	25.9
	Box	8	27	8	74.1
<i>Organisation</i>	Single age	6	22	15	74.1
	7+ and 8+	3	7	—	17.2
Vertically grouped	8+ and 9+	—	—	1	1.7
	9+ and 10+	—	4	—	6.9
<i>Age of pupils</i>	8+	8	19	—	46.6
	9+	1	12	—	22.4
	8+ and 9+	—	2	—	3.4
	10+	—	—	16	27.6

---

<sup>3</sup> Taken from *Inside the primary classroom*, (Galton *et al.*, 1980, p.26).



**Table 3.3 Comparison of Oracle and Department of Education and Science<sup>4</sup> data for age of teachers and their class size.<sup>5</sup>**

	No.	Oracle %	D.E.S. (1976) %
<i>Age of teacher</i>			
up to 30 years	23	39.7	38.7
30 - 40 years	17	29.3	22.5
40 - 50 years	13	22.4	23.4
50+ years	5	8.6	15.4
<i>Class size</i>			
up to 20	1	1.7	4.9
21 - 25	9	15.5	11.6
26 - 30	20	34.5	29.9
31 - 35	25	43.1	37.1
36 - 40	3	5.2	15.0
41 and over	—	—	1.6
Average size		29.9	29.2

---

<sup>4</sup> Statistics of education, 1976, volume 4, *Tables*; volume 1, *Schools*.

<sup>5</sup> Taken from *Inside the primary classroom*, (Galton *et al.*, 1980, p.27).

### **3.3.3 Selection of pupil sample**

The selection of the target pupils who were to be observed in the Oracle research programme was made on the basis of their initial achievement as measured by their score on the modified Richmond Tests. The tests were administered to 1,404 children in sixty classrooms at the beginning of the longitudinal process-product study in September, 1976. The results of these tests were used to select a sample of 486 target pupils. During the first year of the study two classes were dropped from the study as the transfer school was changed. Test results were thus available for 1,201 pupils in fifty-eight classrooms. Eight pupils were selected from each class using a stratified random sampling procedure to ensure that a range of initial achievement scores was represented in each unit of analysis, that is, the class. Each group of eight pupils consisted of four boys and four girls, representing the whole range of achievement as a boy and girl were selected from top and bottom quartiles and two boys and two girls from the central quartiles.

## **3.4 METHODS OF DATA COLLECTION USED IN THE ORACLE RESEARCH**

### **3.4.1 Observational instruments.**

Two systematic observation instruments: the Pupil Record and the Teacher Record were used in the study. These had been developed in a series of pilot projects (Boydell, 1974, 1975a, 1975b). The main categories of the schedules are given in Appendices 2 and 3. The Pupil Record which was originally modelled on the Personal Record of School Experience (Medley *et al.*, 1973) was used to examine three areas of pupil behaviour in the classroom: pupil activity, pupil-teacher and pupil-pupil interaction. The Teacher Record focused on teacher interactions by coding questions, statements, silent interactions, no interaction; the nature and composition of the audience and curricular activity. Observations were made every 25 seconds by trained observers who focused on one pupil at a time for a period of approximately five minutes. Eight target pupils were thus observed during a particular observational period. In addition, the observers focused on the teacher for some 20 minutes, and noted the type of interaction which was taking place, the audience and curriculum area. The observational data was collected throughout the school day to ensure a representative sample of pupil and teacher behaviour. The daily observational timetable is summarised in Table 3.4 overleaf.



**Table 3.4     Daily observational timetable<sup>6</sup>**

	A.M.								P.M.										
DAY 1	A	B	C	D	T	E	F	G	H	T	C	D	E	F	G	H	A	B	
DAY 2		T	E	F	G	H	A	B	C	D*	G	H	A	B	T	C	D	E	F
DAY 3	E	F	G	H	A	B	C	D	T	C	D	E	F	G	H	A	B	T	

\* This observational period was carried out before break, the other two morning observational sessions were carried out after break.

Key : T - teacher  
      A - H target pupils

Each pupil column represents just under 5 minutes, and each teacher block about 19 minutes.

Total observation time per session amounts to some 55 minutes.

<sup>6</sup> Taken from Table 1.1, *Inside the primary classroom*, (Galton, *et al.*, 1980, p. 20)

Other measures were used during the observation periods in school. At the end of each observation session the observers completed a summary sheet which detailed the physical layout of the classroom, the seating arrangements, an outline of the curriculum activity, methods and resources used, the order and timing of the observations as well as any particular incidents or conditions which were unusual or noteworthy. A daily summary sheet was also completed to show the class timetable, and an outline of all activities taking place during the day whether observed or not. The observation of pupils and teachers was restricted to classroom activities and excluded such areas as physical education, foreign language classes, music and drama. In this way it was possible to ensure that the systematic observation of pupils and teachers reflected the actual events in the classroom and that the complete range of observable activities was represented in the data collected by the observers.

At the end of each period of observation in a particular classroom the observer wrote a short descriptive account of the time spent with that teacher. This gave an opportunity to the observers to note any aspect of classroom life that might not have been reflected in the other schedules. In addition a grouping instrument was completed at the end of the year which recorded the physical layout, grouping policy and rationale for each of the classes observed. The observers also completed a short questionnaire, the teacher questionnaire which summarised aspects of the teachers' organisation and management policies in a similar way to that collected in Bennett's (1976) study of teaching style. These observations then formed the database from which the descriptions of teaching style and pupil type were derived. Details of the results of this analysis are given in the next chapter. This data was used in this research to explore the relationship between teachers' assessments of pupil performance, teaching style and the classroom behaviour of pupils (pupil type).

### **3.4.2 Assessment procedures: Richmond tests, study skills exercises and teacher-based assessments.**

In order to identify which aspects of teaching and learning were most effective in improving pupil performance, a range of assessment procedures were used as indicated in Chapter 2. The Richmond Tests of basic skills (France and Fraser, 1975) were designed for children aged from 8 - 13 years old, and each test was divided into six levels corresponding to the six year groups (second year juniors to third year secondary pupils in the conventional system of transfer at 11 +). These levels overlapped so that a number



of items were answered by both second and third year pupils. Since the complete test battery would have taken some five hours to complete it was decided to develop a shortened version which was acceptable to the teachers taking part in the project. The shortened version was constructed in the range of difficulty of the first four levels of the tests for Reading, Language skills and Mathematics. The item numbers from the published Richmond tests that were used in these shortened versions of the test are included in Appendix 4. This data was used to investigate the relationship between the teacher-based assessments and pupil performance on these tests of basic skills, as well as providing the basis for the selection of the target pupils (section 3.3.3).

Two other forms of assessment were used in the Oracle research programme; the study skills exercises and the teacher-based assessments. The study skills exercises have been fully documented in *Progress and performance in the primary classroom* (Galton and Simon, (eds.), 1980). These exercises were selected in an attempt to provide pieces of work which might involve common underlying study skills. They involved the interpretation of some pictorially presented data and the construction of a block graph, making a plan of a classroom and a model of a clockface with moveable hands (Willcocks and Jasman, 1980, pp.101 - 123). The selection of these exercises was based on an attempt to evaluate the ways in which the pupils made graphs and other artefacts, as well as the ways in which they made use of them or interpreted them. The teacher-based assessments were also used to provide data on pupil performance in the areas of study skills based on some of the aims of primary education identified in the Ashton *et al.* (1975) survey of teachers' opinions. Detailed descriptions of the development of these measures is presented in section 3.5.

### 3.4.3 Social class of pupils

In addition to these measures of pupil performance, data was collected on the social class of a limited sample of target pupils (Croll, 1981). This data was available for only one of the local education authorities participating in the study. In this area, details of the father's or guardian's occupation were obtained for each of the eight target pupils in the twenty-nine classrooms involved in the second year of the study. This information was obtained by reference to school record cards and was available for 188 of the 240 pupils in the sample. Four social class groupings were initially derived: professional and managerial (11.2% of the pupils), other non-manual (24.5%), skilled manual (35.6%) and other manual (28.7%). For the purposes of data analysis these four groupings were



subsequently amalgamated into two larger groups: non-manual and manual since the numbers of pupils within the four groups did not constitute a large enough sample for a valid statistical analysis to be performed. This information on pupil social class was used to explore the relationship between this variable and teachers' assessments of pupil performance, measured performance on the basic skills and on the structured activities.

### **3.5 DEVELOPMENT OF THE TEACHER-BASED ASSESSMENTS**

#### **3.5.1 Formation and functions of teacher discussion groups.**

Teachers from the Leicestershire region were invited to attend a meeting to discuss the Oracle research programme and obtain information on ways in which teachers could participate in the research. Some sixty teachers attended this meeting, of these twenty four indicated that they were interested in becoming involved in the development of the teacher-based assessments. This sample of teachers represented a range of different types of school background, age and experience, although they predominantly taught junior aged children (7-11 years old). These teachers were asked to list in order of preference three of the fifteen intellectual aims of primary education listed as of "*utmost*" or "*major*" importance in the Ashton *et al.* (1975) survey of teachers' opinions (see Appendix 5). Each of the teacher groups undertook to develop appropriate assessment procedures for the following aims of primary education (Ashton *et al.*, 1975).

- a. The child should be able to listen with concentration and understanding.*
- b. The child should develop some inventiveness and creativity in some fields; for example, by painting, music, mechanical things, poetry, movement.*
- c. The child should know how to acquire information other than by reading; for example, by asking questions, by experimenting, from watching television.*
- d. The child should be able to:*
  - i. compute the four arithmetic rules using his knowledge of number, multiplication tables and different units of measurement,<sup>7</sup>*
  - ii. use mathematical techniques in his everyday life, i.e. estimating distances, classifying objects, using money.<sup>7</sup>*
- e. The child should know how to convey meaning clearly and accurately through speech for a variety of purposes; for example, description, explanation, narration.*

---

<sup>7</sup> These two aims were considered together, as teachers selected them conjointly.



Five teacher groups were formed to work with the researcher to devise and pilot the teacher-based assessments described in the following sections. Each group followed through the stages of the development process summarised in Table 3.5 overleaf.

The purpose of these discussion groups was to devise more appropriate methods for the assessment of those aims of primary education not assessed by more conventional procedures as discussed in the preceding chapter. It has been argued that an important strategy for improving the validity of teachers' assessments was the opportunity for teachers *"to develop behaviourally-anchored scales based on a consensus view of what most teachers considered achievement to consist of"* (Wood and Naphthali, 1975, p. 160). Thus the first stage of the process was to clarify each of the five aims in terms that would provide precise and understandable criteria for the assessment of pupil performance in the five areas selected. In order to facilitate this process each teacher was individually asked to clarify and articulate his/her understanding of the aim under consideration based on his/her own experience. In this way differences in interpretation were expected to emerge and provide the basis for developing a consensus view through later discussions. Similar strategies had been successfully employed in the development of the seventy two aims of education in the Ashton *et al.* (1975) survey of teachers' opinions and in the Progress in Learning Science project (Harlen *et al.*, 1977a, 1977b) to develop criteria for assessment in primary science.

The methods used to facilitate this process were based on the clarification of the aims by the use of contrast and examples, similar to that used in conceptual and philosophical analysis (Chambers, 1983). This process involved the teachers in thinking about what the aim meant, noting the abilities and behaviours expected of a pupil who was considered to have performed well in relation to that aim, thus providing examples of that aim. It was also suggested that the teachers noted examples of what the aim did *not* mean and the abilities which would indicate that the pupil had not performed adequately in relation to that aim, thus providing a contrasting view. In order to help the teachers focus on the realities of classroom practice they were encouraged to observe pupil behaviour in order to identify particular instances of satisfactory and unsatisfactory performance in relation to each particular aim. Finally, the teachers were asked to record any activities that they considered encouraged pupils to exercise their abilities in that area. Letters, written instructions and booklets for completion were posted to each participating teacher (see Appendix 6) which asked them to write down their understandings of each aim in these terms.

**Table 3.5      Stages in the development of the teacher-based assessments**

**Development phase**

Teacher groups set up, notes made on the conceptualisation of the aims.



Discussions leading to the formulation of the pilot checklists of abilities.



Checklist used to rate pupil performance in the classroom.



Checklists revised



**Application in Oracle research programme.**

Checklist of abilities sent to Oracle teachers to use in the process-product study.



Range of structured activities devised to match criteria listed on checklists of abilities.



Pilot of structured activities.



Structured activities revised.



Structured activities sent to and used by the Oracle teachers.



Validity and reliability checks<-----



Feedback from Oracle teachers



Materials scored and data analysed.



The initial response of the teachers to this task was unfavourable, since they seemed unprepared to write about their understandings. However, a visit to each school by the researcher to discuss her expectations and clarify the nature of the task reassured the groups and resulted in all the teachers who had initially volunteered to participate, completing the exercise. Their initial reluctance to complete the task stemmed from a lack of confidence in their ability to contribute to "research" rather than considering the activity inappropriate. The information generated from this open-ended approach to soliciting teachers' understandings of each aim resulted in a wide variety of responses which were content analysed to provide a basis for further discussion and further clarification of the criteria to be used in the assessment of each aim. The discussion documents prepared are presented in full in Appendix 7.

Each aim was discussed with respect to the following categories:

- a. The aim, towards a definition.
- b. The abilities, concepts and attitudes associated with the aim.
- c. Examples of activities where these were observed.
- d. Examples of satisfactory and unsatisfactory pupil performance.

Thus, each discussion paper was a synthesis of the views of the teachers in that particular group and represented the first step in the articulation and clarification of the aim in operational terms.

The second stage in the process involved the teacher groups meeting at regular intervals to discuss, clarify and refine the content of the discussion documents until a "*checklist of abilities*" for each aim was developed. In order to facilitate these discussions, the meetings of the teachers' groups were tape recorded and the transcripts formed the basis of further discussion documents which resulted in the development of the final checklists detailed in Appendix 8.

An initial trial of the practicality of using such checklists in the classroom was conducted during this second phase of development. Each teacher set up an activity where the skills and abilities on the checklist of abilities could reasonably be expected to be observed. These activities reflected the range described in the discussion documents. The teacher and researcher made independent assessments of each pupil in the class and brought their results to the next group meeting. The teachers based their judgements on



their prior knowledge of the pupils and evidence gained from their observations as the pupils engaged in the activity. The researcher made her assessments on the basis of the limited sample of pupil behaviour observed during the activity.

Two common difficulties emerged in conducting these trials. First, the practical problems of observing every pupil during the class activities became apparent and second, the checklist of abilities often did not provide sufficient clarity for the teachers to distinguish the particular skills and abilities included on the checklist so that they were unable to differentiate effectively between the different levels of pupil performance. In order to alleviate this problem the completion of the checklists was based on a rating scale similar to that used with the **Progress in Learning Science** project trial materials (Harlen *et al.*, 1977a, 1977b). Initially, a three point scale was applied to each criterion. These were as follows: A - the child always achieves this level of performance; B - the child sometimes achieves this level of performance and C - the child rarely achieves this level of performance. Intermediate levels for grading were later included between A and B and between B and C thus providing a five point scale. This modification was necessary as many teachers exhibited the tendency to rate all pupils as B. They also appeared particularly reluctant to suggest that pupils always achieved certain levels of performance so this descriptor was later amended to A - the child usually achieves this level of performance.

As a result of the particular problems associated with using the checklists of abilities, the suggestion was made that specific activities would facilitate both the observation of different levels of pupil performance and the differentiation between the different types of pupil behaviour associated with each aim. These activities, called the "*structured activities*" were similar to the type of assessment procedure used in the Certificate of Secondary Education where practical exercises were given to the pupils in the classroom setting and used as a basis for grading individual pupils and for moderating the degree of agreement between teachers' assessments.

Thus, these structured activities were specifically designed to match the checklist of abilities which provided the criteria for the assessment of pupil performance in relation to each of the aims under consideration. They, therefore, served two purposes. First, they provided a specific context for assessing a particular criterion on the checklist of abilities and helped the teachers focus on the specific pupil behaviours. Second, by using a set of common structured activities within the Oracle research programme, it was possible to



moderate the assessments of different teachers in different schools to ensure the comparability of the data generated on pupil performance and to investigate the reliability and validity of teachers' assessments.

However, it became apparent from the diversity of the discussion documents and the constraints of the Oracle research programme that the development of the teacher-based assessments for each aim would have to be phased in stages and modified to fit with the other tests and assessments which were part of the programme. The time-table for developing and implementing the teacher-based assessments has been included as Appendix 9. A brief discussion of the development of the checklists of abilities and structured activities for each aim is now included in order to identify particular problems and concerns which were expressed in relation to each of the aims under consideration and the ways in which these were resolved by the teachers' groups in question. In addition, brief descriptions of the final checklists of abilities, structured activities and scoring systems have been provided. Details of these are fully documented in Appendices 8, 10 and 11.

### **3.5.2 Aim: the child should be able to listen with concentration and understanding**

The teachers' interpretations of this aim focused on two main areas; first, what the pupils were listening to and second, how the pupils showed that they were listening with concentration and understanding. The particular occasions when the teachers considered it was important for the pupils to listen in this way included: story time, when acquiring factual information, as instructions were given, during conversations and when listening to music and sounds.

The teachers identified two strategies for assessing this aim: first, by observing the pupils as they were listening and second, by testing (orally or by written responses) recall and understanding. The teachers' accounts indicated that the following pupil behaviours were used to make an assessment of pupils in relation to this aim. First, pupils were judged to be listening if they sat quietly and attentively and were able to recall what they had heard. Second, pupils showed concentration by listening for longer periods of time, by not showing signs of distractedness or fidgeting and by recalling more detailed information such as the order or sequence of events in a story. Finally, understanding was demonstrated by the pupils using the information, acting upon instructions and by



asking relevant questions.

The pupil behaviours that were given as associated with this aim in particular included: following instructions, recalling and recording verbatim, recalling and recording in their own words, applying, extending and explaining what they understood. As would be expected a large number of activities where such behaviours could be observed were identified by the teachers, such as, physical education lessons, hand work, craft, art lessons, when viewing television programmes, during discussions, debates, stories, school assemblies, lessons where the teacher talked for periods of time, in completing exercises associated with Concept 7-9 (Schools Council, 1973b) and SRA listening skills materials (Science Research Associates, 1972), in music and memory recall games.

A preliminary checklist of abilities was derived from the first discussion document (Appendix 7) and the initial trial carried out. This involved the teachers making a preliminary assessment of each pupil on each criterion based on their existing knowledge of the pupils' performance. This was followed by a period of observation of the pupils by the researcher, who made a judgement based on this limited period of observation, and the teacher, who revised their assessment of the pupils' performance if necessary. This trial resulted in assessments which showed considerable variation as individual pupils were rated under different circumstances and with differing degrees of prior knowledge on the part of the teachers and the researcher.

The teacher group was, at this stage, more concerned with the practical problems of observing the whole class than with any further clarification of the criteria which they claimed to understand and expressed confidence in applying validly when assessing pupils. The group attempted to resolve the practical problems by focusing on small groups of pupils in the first instance, thus spreading the observations over a longer period of time. Although confident of applying the criteria included in the checklist of abilities, the teachers expressed uncertainty as to the "*level*" of pupil performance which could be considered "*satisfactory*", for example, teachers noted that the ability to follow a set of instructions would vary according to the complexity of those instructions.

The checklist of abilities was simplified as a result of the trials to take account of the practical problems experienced by deleting one of the criteria used: "the child should be able to follow the sequence and order of speech" which the teachers were unable to judge on the basis of their observations. In addition, one criterion which included the



assessment of the pupils' ability to apply and extend what had been heard was simplified to consider only the application of what had been heard; for example, following a set of instructions. Thus, the revised checklist included three levels of performance which constituted the criteria for assessment. The teachers were expected to indicate whether the pupils were able to listen and repeat or recall some content, repeat or recall the order or sequence of the content and apply what was heard and understood. If the pupils were able to demonstrate that they achieved all three criteria for the majority of the time then they were judged to have reached a satisfactory level of performance in the area of listening with concentration and understanding. A method of grading the level of performance was also introduced to provide more information on how satisfactory this was, so that the teachers were asked to decide whether the child rarely, sometimes or always achieved a particular criterion.

The teachers who devised these teacher-based assessments, in particular the structured activities, were concerned that pupils with learning difficulties should not be disadvantaged, so that tasks were constructed which reduced the need to read and write to a minimum. This concern precluded the use of existing materials such as the SRA listening skills materials (Science Research Associates, 1972). The development of the structured activities was thus determined by this concern for ensuring that the influence of the pupils' ability in oral expression, reading and writing were minimised. Picture cartoons, taped sounds and instructions were, therefore, used to overcome this problem. The only written response required was the pupil's own name on one of the tasks. The structured activities consisted of three tasks. The first task involved the identification of five musical instruments, recorded on a sound tape singly and in various combinations. Each instrument was identified at the beginning of the tape, and the pupils made their responses by marking the appropriate instruments which were pictorially represented on the answer sheet. The focus of this task was the pupil's ability to listen and recall content.

In the second task, a story was told on tape entirely in sound effects. It dealt with a sequence of events from the time a man was woken by his alarm clock to the time he arrived at work. There were three sections with intervals between them. At the end of each of the first two sections, the pupils chose one picture from each of three pairs which represented the events they had heard, placing their choices in the correct sequence. In the third section four pictures were chosen from six, and again the pupils recorded their choices in a similar way. The focus of this task was the pupil's ability to listen and recall



the sequence of events accurately.

In the third task the pupils were required to follow a series of instructions from a pre-recorded tape. Pupils were asked to write their first name inside a circle, triangle or square on a printed sheet, then to draw another shape in which they wrote their surname. The choice of shape depended on the initial letters of the pupil's first name and surname. The focus of this task was the pupil's ability to follow instructions of the type commonly given in classrooms. The selection of an appropriate scoring system was relatively straightforward since the aim was explicated in terms of particular criteria and the structured activities were devised to match these criteria. All these tasks could be scored right or wrong and considerations such as neatness and spelling were not taken into account. An example of the materials developed for the assessment of this aim and the relevant scoring system can be found in Appendices 10.1 and 11.1.

### **3.5.3 Aim: the child should be able to show some creativity and inventiveness in some fields.**

This proved to be the most difficult aim for teachers. Their articulation of what it meant for pupils to show some creativity and inventiveness lacked clarity. However, three dimensions were identified from the analysis of the notes made by the teachers which were consistently represented in all cases. The first consideration to be born in mind was associated with the context and nature of the activity during which the assessment was to be made, since the teachers were concerned that a pupil might exhibit creativity in doing one activity but would not produce an equivalent performance in another area. The following areas were identified by the teachers as examples where pupils might show evidence of inventiveness and creativity: art and craft, music, movement and drama, creative writing as well as other situations such as engaging in mathematics and science.

Second, the assessment of this aim was considered dependent on the degree of pupil choice allowed. At one extreme, the context was described where the teacher set the topic, provided the materials and suggested ways in which the pupil might respond to these stimuli. In this way the degree of pupil choice was kept to a minimum. At the other extreme, the pupil could have been given complete freedom to choose the topic or theme, the materials and the mode of expression. The teachers considered that a different judgement should be made of a pupil's performance in these two cases. Finally, the



teachers' notes identified a criteria for assessment which corresponded to the degree of originality or uniqueness of the particular pupil's response. This was articulated in terms of whether the pupils had their own ideas, or adapted and extended other people's ideas or had to be told what to do.

The first attempt at organising these dimensions incorporated all these considerations into one checklist. This resulted in a three by three grid, which also recorded the context of that particular assessment. Thus, the topic, the materials used and the type of technique or application were cross referenced against *"pupil chosen"* and *"teacher set/teacher provided"* and in the last instance against *"pupil adapted or extended normal use"* or *"pupil used techniques or materials in a conventional way"* (Appendix 7.2). Three levels of pupil performance were identified. At the first level, the pupil was usually told what to do, copied another's ideas or followed instructions closely. At the second level, the pupils extended or adapted ideas which were given to them but did not think of their own ideas. At the third level, the pupils used their own ideas or produced a different interpretation of a theme in context.

This initial checklist was piloted by the teachers and the researcher for a variety of different activities such as art and craft, creative writing and drama. However, the checklist proved inoperable under classroom conditions. At one level it was a complex document and the level of pupil performance in relation to the criteria proved too difficult to distinguish in practice. An additional factor which contributed to the difficulty in the assessment of this aim was the variation found in what the teachers considered to be a creative activity. For example, some teachers considered situations where children were set a theme, provided with materials, given instructions and help in completing the activity as an appropriate context for the assessment of creativity, whilst others considered that creativity could only be assessed if it was manifest as a spontaneous, independent act on the part of the pupil. This underlying lack of agreement between the teachers clearly had to be resolved as well as the practical problems associated with implementing the teacher-based assessments within the Oracle research programme.

Thus, a much simplified checklist emerged from discussions with the teacher group about the problems associated with the pilot version. Two central criteria were suggested. The first was that of *"originality"* which was easily assessed on whether one or many pupils produced a similar response to a given stimulus. The second was that of *"appropriateness"* which relied on the teacher's judgement of how well the pupils'



responses matched the medium of expression or the parameters of the activity. These criteria were easily applied in a variety of situations such as art, craft, drama, music, dance, science, creative writing and mathematics as well as being independent of the teachers' individual understanding of the term "creativity".

In the case of this aim, the structured activity was a simple picture completion exercise. This technique has been widely used for the assessment of creativity (Foster, 1971). In this task each pupil was asked to draw a picture incorporating a circle and a V-shape which had been printed on sheets of A5 paper in the following manner:



This activity was developed principally to aid in the moderation of teachers' assessments rather than help teachers apply the two criteria of "*originality*" and "*appropriateness*". These criteria also formed the basis of the scoring system. The drawings were sorted so that those with similar contents were grouped together. Marks on originality were given on a four point scale, ranging from one point for the most common response (a person with the circle as the head and the V-shape as legs) to four points for any picture in which the content was unique. The marking scheme for appropriateness was devised by the teacher group through discussion of examples. A picture was judged appropriate if the child made full use of the two printed shapes incorporating them into a single picture and this was awarded three marks. A picture which did not make use of the two shapes was judged inappropriate and given no marks. Two intermediate groups were awarded one or two marks dependent on the judgements of the teachers as to the degree of appropriateness in the use of the two shapes. Further details of the criteria and scoring schemes are included in Appendices 10.2 and 11.2.

#### **3.5.4 Aim: The child should be able to acquire information other than by reading for a variety of purposes.**

The particular difficulty found by teachers in defining what they understood by this aim arose as a result of the breadth of its original formulation. A number of the teachers included skills such as the child should be able "*to express ideas clearly*" which on first inspection would seem to be related to "*conveying meaning clearly and accurately through*



*speech*" which was being considered by another teacher group. This problem arose from the initial formulation of the aims of primary education (Ashton *et al.*, 1975) which were the basis for the development of the teacher-based assessments. These aims were not intended to be mutually exclusive. In fact, some of the aims described what the pupils should be able to achieve in a particular area from only slightly different perspectives, as in the case of knowing *"how to convey meaning clearly and accurately through speech"* and *"how to speak in a clear and fluent manner appropriate to different situations"*. As the researcher was involved in each teacher group it was possible to ensure that there was a minimum amount of overlap between the criteria used for the assessment of "acquiring information other than by reading" and other aims, such as "listening with concentration and understanding" and "communicating clearly and accurately through speech" which were being considered by different teacher groups.

The teachers' notes raised three areas for further consideration: how and where pupils acquired information, and what information the pupils acquired. The comments concerning how the pupils acquired information ranged from listening and observing, which were relatively passive activities on the part of the pupils, to asking questions and experimenting which involved a more active acquisition of information. Sources of information were broadly classified as coming from people; such as teachers, parents, experts, friends, and through the physical environment such as teacher-planned resources: for example, audio-visual materials and experiments; or by observing and asking questions about experiences which had not been planned by the teacher. Two main kinds of information which pupils acquired were identified by the teachers; first, factual information and second, more practical or applied information such as manipulative skills. Many activities were listed where teachers thought that they would be able to observe a pupil's ability to acquire information other than by reading. These included: visits to museums, films, theatre, through discussion, practical experiments, assembly, visitors and other audio-visual materials and resources.

The aim was thus initially described in terms of the following levels:

1. The preliminary acquisition of information: by observing, questioning, examining, listening.
2. The recalling, recording and ordering of information acquired.
3. The extension of information: by experimenting, evaluating, developing hypotheses.

These areas only represented a limited sample of the sixteen abilities originally listed by the teachers (see Appendix 7.3) and did not include the attitudes which teachers hoped to encourage through this aim since these were beyond the scope of this study. As with the other aims taken from the Ashton *et al.* (1975) survey which were considered in this study this checklist of abilities was developed through the experiences of operating the checklist in the classroom, until the following list of skills and abilities was refined.

**1. *Observing and questioning* The child is able to:**

- a) make his/her own independent observations,
- b) ask relevant questions,
- c) use appropriate language and terminology and appear to understand meaning.

**2. *Recording and ordering information* The child is able to:**

- a) report in an appropriate way, e.g. orally, by writing, model-making, pictorial information,
- b) bring together information from a variety of sources, recording this logically and in sequence,
- c) where relevant, extend the activity or add to information by expressing his/her own ideas.

**3. *Extending an activity or information* The child is able to:**

- a) test or try out a suggestion made by another person,
- b) make his/her own suggestions for extending an activity or adding to information,
- c) test or try out his/her own suggestions on his/her own initiative.

Two structured activities were devised to match those criteria which could be assessed in a written form. This was a necessary prerequisite for the moderation purposes for which the structured activities would be used. In the first task the pupils were asked to complete a spot the difference exercise which was related to their observational abilities. This activity was a simple exercise to implement and was also familiar to the pupils. It was scored on the number of differences the pupils were able to identify and mark. For



the second task, the pupils heard a story and were also given a picture which illustrated it. They were then presented with a series of questions, some of which related to information appearing in the story but not in the picture and others relating to information contained in the picture but not in the story. The task was intended to focus on a pupil's ability to acquire information from different sources. Marks were given for each correct response. The pupils were then asked to write down the five questions they would most like to ask if they could meet the person in the story, but they were also encouraged to add further questions if they could think of any. As with the *"listening with concentration and understanding"* materials the teachers were concerned not to penalise pupils with reading or writing difficulties so the pupils were allowed to ask the teacher to record their questions rather than be inhibited by any lack of writing ability. A predetermined scoring system was not used for the analysis of the questions asked by the pupils since it was not possible to predict the type of responses which might arise.

A number of aspects of questioning which could be taken into account emerged as a result of the initial trials of this structured activity. These then provided the basis for a scoring system which related to the various dimensions of the pupils' responses observed which it is suggested provided a representation of the pupils' ability to ask questions. These aspects included the number of questions asked, whether there was any variation in the types of information elicited and the complexity of the information which would result from the question. Thus the questions were scored using three criteria: first, the total number of questions; second, the range of questions, that is, whether they all dealt with one similar theme or covered a variety of topics; and third, the number of *"open"* questions asked. As conceptualised by the teachers, open questions were those framed in such a way as to elicit an explanation or information rather than a one-word response such as yes or no. Examples of the checklist can be found in Appendix 8.3 and of the structured activities in Appendix 10.3. Details of the scoring scheme can be found in Appendix 11.3.

### **3.5.5 Aim: The child should know how to compute in the four arithmetic rules and use mathematical techniques in his/her everyday life.**

Those teachers who were particularly interested in the area of mathematics nominated two aims for consideration: the first dealt with computation, the second with mathematical applications. As the development of the teacher-based assessments progressed the second area was emphasised since many standardised tests were available for the assessment of



computational skills in mathematics. The teachers' conceptions of this aim centred initially on the variety of everyday situations where pupils would be expected to apply mathematical techniques and the wide range of mathematical content, skills and abilities which would be required for pupils to achieve a satisfactory performance in relation to this aim.

The initial discussion document attempted to categorise topic areas so that simple measures such as money, time, length, weight, temperature were separated from topics which involved more than one measurement such as area, volume, speed, rate, force and costings. However, this did not take account of many mathematical concepts and procedures which are applied in everyday life: for example, finding the average, percentages, proportion, drawing to scale and making graphical representations of information.

Since the range of mathematical content which could reasonably be assessed within the scope of this investigation was limited, the decision was made to emphasise the processes which pupils engaged in when solving a small sample of problems associated with everyday life rather than attempt to devise a comprehensive range of activities to cover all possible mathematical content which might be applicable. In this way it was anticipated that any underlying skills would be effectively measured by the structured activities rather than content knowledge. The skills and abilities which a pupil would have to exhibit in order to achieve a satisfactory performance in relation to the two areas under consideration were identified by the teachers as follows.

The child should:

- a) know mathematical symbols, terminology and notation,
- b) indicate that s/he knows the appropriate method for everyday application,
- c) measure accurately, using the appropriate apparatus where necessary,
- d) make reasonable estimates,
- e) distinguish between relative difference (more than and less than),
- f) convert into appropriate units of measurement,
- g) apply the appropriate technique in an unusual situation.

It became clear after the initial trial in the classroom setting that the above list of criteria included both content knowledge as well as process skills. The checklist was, therefore, refined and simplified to focus more specifically on the process of solving a problem. The



final checklist used in this research included the following criteria:

The child is able to:

- a) state what mathematical processes are involved using correct terminology,
- b) state if additional information is needed before a problem can be solved,
- c) carry through the operation and represent his/her solution to others,
- d) arrive at an accurate result.

As has been indicated previously the structured activities devised to aid the assessment of these aims represented a limited sample of the wide range of everyday situations where mathematical techniques could be applied. The structured activities, therefore, were designed to match the process skills described in the above checklist of abilities. Twenty sums were included to provide a baseline of the pupils' abilities at arithmetic computation. Five were included for each of the four rules: addition, subtraction, multiplication and division. These sums were of progressive difficulty, which required the pupils to work with increasing place values and carry out more complex computational operations. The second set of activities required the pupils to insert the correct sign in a series of sums and was designed to tap how well the pupils could state mathematical processes involved in arriving at the given solutions.

The next set of problems were made up of a series of *"impossible"* sums where pupils were asked to identify the additional information that would be required in order to arrive at a solution to these problems. This activity was designed to aid the assessment of the pupil's ability to identify the necessary information for the solution of a problem. The exercises chosen to match the third criterion on the checklist: *"the child should be able to carry through this operation and represent this solution to others"*, required the pupils to select from four alternative arithmetic sums the one equation which would provide the solution to a written problem of the type often used in standardised achievement tests of mathematics. Pupils were also required to write the appropriate equation for a second series of similar written problems but not the solution to the problem. In this way the processes of solving such written problems were to be assessed.

The final criterion on the checklist was assessed by presenting six problems relating to everyday mathematics which the pupils were asked to solve. These problems included constructing and using a bus timetable, drawing a sketch map from written instructions, answering a series of questions related to a school time-table, a classification exercise,



drawing a block graph and finally reading from a series of meter dials, to deduce the amount of electricity used. Examples of all these activities are included in Appendix 10.4. The penalisation of pupils who may have had difficulty with reading the questions or who required some form of assistance was countered by allowing the teachers to help pupils in a number of ways. Keys were devised for each of the questions on this set of structured activities. These described the type of help that had been offered and enabled this to be taken into account in scoring individual pupil responses. Details of the keys used by the teachers and the scoring system for these activities are given in Appendix 11.4.

### **3.5.6 The child should be able to convey meaning clearly and accurately through speech**

In clarifying this aim three areas were identified by the teachers as important to the assessment of the pupil's ability to convey meaning clearly and accurately through speech. First, there were those abilities associated with the mechanics of speaking, such as voice control and articulation. Second, the accuracy with which pupils could repeat or convey information from another source; such as when they were asked to relay a message or relate information from a book or other medium. Third, those occasions when the pupils were asked to convey meaning by expressing themselves in their own words; for example, by describing, narrating, explaining or instructing. The initial checklist focused on the particular types of performance that would be associated with conveying meaning through expression in the pupils' own words. These criteria are listed below:

The child is able to:

- a) use appropriate vocabulary,
- b) translate other sources into his/her own words,
- c) distinguish between relevant and irrelevant information, for a particular purpose or context,
- d) order information logically and in sequence,
- e) convey meaning so that the purpose or context can be understood.

Many classroom situations were identified by the teachers which lent themselves to assessing the pupil's ability to convey meaning clearly and accurately through speech; for example, describing objects, teaching a skill, telling a well known story, summarising a film, explaining the rules of a game. The overwhelming practical problem that faced the teachers here was the necessity of ensuring that individual children should have the



opportunity of demonstrating their ability to communicate through speech. Thus, it was suggested that it might prove necessary to set up a specific situation for the assessment of this aim and record the pupils' responses on audio tape for later assessment since it was considered unlikely that the teacher would be able to complete the checklist for all the pupils without some form of recorded material.

There followed a series of meetings of the teachers' group. Examples of pupils talking on a variety of themes were collected by the teachers and these were discussed and assessed using the initial checklist. The checklist of abilities was continuously refined during these meetings as particular problems arose in applying the criteria. This resulted in the following checklist being used as the basis of assessing a pupil's ability to communicate clearly and accurately through speech.

The child is able to:

- a) vocalise audibly and articulately,
- b) use language and vocabulary appropriate to the situation or context,
- c) talk fluently without undue pausing, hesitation or repetition,
- d) maintain logical connections when describing objects, relating events and during discussions,
- e) respond with reasonable speed and accuracy to questions and other contributions,
- f) modify and develop his/her own responses,
- g) link his/her own experiences to the present situation.

The use of structured activities of the type described for the preceding four aims was inappropriate in the case of this aim since no written product could provide a valid check on the assessments made by teachers of pupil performance in this particular area. Thus an alternative strategy was adopted to moderate the assessments made by teachers in relation to this aim. The absence of a structured activity also meant that there was little opportunity for the teachers in the Oracle research programme to receive the help provided by the structured activities in focusing on specific aspects of pupil performance as they completed particular exercises. A tape was therefore prepared which illustrated each of the criteria on the checklist and included a set of examples for each teacher to assess. This provided a basis for moderating teachers' assessments in order to ensure their comparability within the research as well as facilitating the understanding of each of the criteria on the checklist of abilities by the teachers participating in the Oracle research



programme.

The "conveying meaning" tape consisted of two separate sections. The first part provided illustrative examples of children talking and the assessments which were made by the development group of teachers. Each of the extracts was assessed on one or more of the criteria using a five point scale. A performance well above average was rated A, an above average performance was rated B, an average performance was rated C, a below average performance was rated D and a performance well below average was rated E. In the case of each extract the assessment made by the teachers was given together with their reasons. For example, one extract gave the pupil Anthony a rating of D on the first three criteria for the following reasons: some of the words were difficult to hear, there was a limited range of vocabulary and frequent use of words such as "and then" with pauses which inhibited the fluency and flow of the speech. A second extract of Kate was assessed in the following way. B was given for vocalising clearly and articulately, since the speed and running together of words was thought to preclude awarding an A. The remaining criteria were rated as A, as the language and vocabulary were appropriate and there was no undue pausing and repetition. In all seven extracts were included which covered each of the criteria and provided a range of examples to assist in discrimination between different levels of performance.

In the second part of the tape five different children were heard recounting the story of the railway children which they had seen the previous day. Teachers were asked to rate these examples on each of the seven criteria on the checklist of abilities using the rating scale A - E which had been illustrated on side 1 of the tape. In this way a basis for moderating the teachers' assessments was available since comparisons could be made between their assessments of these five examples and if discrepancies were found it was possible to adjust their assessments of their own pupils in order to take account of any over or underestimation of performance. An example of this tape has been included as Appendix 12.

### **3.5.7 Summary of teacher-based assessments: checklists of abilities and structured activities.**

Some of the variation in the procedures used to develop the teacher-based assessments can be ascribed to the differences in the original formulations of the aims. Some aims were stated in almost behavioural terms, for example, "the child should be able to listen



with concentration and understanding". Whilst others such as "the child should be developing inventiveness and creativity in some fields" were stated in more expressive terms. Thus the opportunity to develop a consistent approach across each of the aims under consideration was limited by their initial variation in format and expression. However, some measure of consistency was ensured between the different teacher groups by the researcher being party to all the discussions, development of each checklist and associated structured activities and the initial trials of the materials. Thus each aim was described in terms of observable pupil characteristics and where practicable a specific activity was designed to assess particular criteria. As has been noted previously this was not possible for all criteria, however, an alternative strategy was developed to assist in the moderation of the teachers' assessments in the case of "conveying meaning clearly and accurately through speech". Where structured activities were devised to assess particular criteria the descriptions in themselves do not necessarily convey the appropriateness of the activity for the assessment of the particular criterion. The way in which they were scored (see Appendix 11) provides a more useful basis for judging the appropriateness of each activity as interpreted through the scoring scheme for the criterion under investigation. A summary of the criteria used for each aim and the associated structured activities is presented in Table 3.6 overleaf.

### **3.5.8 Introduction of materials to schools in the Oracle programme.**

The original intention was to introduce all five teacher-based assessments during the first year of the study and repeat the exercises in the second year in order to investigate the progress of pupils in these areas. However, a research programme of this type made considerable demands on the teachers involved so that the materials had to be phased in their introduction to schools. The introduction of two teacher-based assessments in the autumn term of 1976 and a third in the spring term of 1977 proved to be as much as the teachers in the study were prepared to complete. This was largely due to the demands of the testing programme using the modified Richmond Tests of basic skills and the study skills exercises which were carried out in the summer term of 1977. This imposed a limitation on the research so that it was only possible to record pupil performance in these areas once and thus no evidence relating to pupil progress was available. The materials were distributed to the teachers in the schools by the observers who were conducting the systematic observation in the schools (see Table 3.7). These included details for completion of the checklists of abilities, administration of the structured activities and a questionnaire requesting feedback on the appropriateness of the materials (Appendix 13).



**Table 3.6 Summary of teacher-based assessments**

<i>Criteria for assessment</i>	<i>The child is able to:</i>	<i>Structured activity</i>
<i>Listening with concentration and understanding.</i>		
1.	listen and repeat or recall some content,	Musical instruments Sound story Follow instructions
2.	repeat or recall the order or sequence of the content,	
3.	apply what was heard and understood.	
<i>Creativity.</i>		
1.	show originality in response to a stimulus	Picture completion exercise.
2.	show appropriateness in response to a stimulus	
<i>Acquiring information other than by reading.</i>		
1a.	make his/her own independent observations,	Spot the difference Writing five (or more) questions.
1b.	ask relevant questions	
1c.	use appropriate language and terminology and appear to understand meaning.	
2a.	report in an appropriate way, e.g. orally, by writing, model-making, pictorial information.	Questions on picture and story activity.
2b.	bring together information from a variety of sources, recording this logically and in sequence.	
2c.	where relevant, extend the activity or add to information by expressing his/her own ideas.	
3a.	test or try out a suggestion made by another person	No structured activity
3b.	make own suggestions for extending an activity or adding to information.	
3c.	test or try out own suggestions on own initiative.	
<i>Everyday mathematics.</i>		
1.	state what mathematical processes are involved using correct terminology,	Sum completion
2.	state if additional information is needed before a problem can be solved,	Impossible sums
3.	carry through the operation and represent solution to others,	Six problems related to everyday maths.
4.	arrive at an accurate result.	
<i>Conveying meaning through speech.</i>		
1.	vocalise audibly and articulately,	Moderation of teachers' assessments by use of tape of five examples of pupils' talking.
2.	use appropriate language and vocabulary	
3.	talk fluently without undue pausing, hesitation or repetition,	
4.	maintain logical connections	
5.	respond with reasonable speed and accuracy to questions,	
6.	modify and develop his/her own responses,	
7.	link his/her own experiences to the present situation.	



**Table 3.7 Schedule for the introduction of the teacher-based assessments into the Oracle research programme schools.**

**Autumn term, 1976.**

- a. The child should be able to listen with concentration and understanding.**
- b. The child should develop some inventiveness and creativity in some fields; for example, by painting, music, mechanical things, poetry, movement.**

**Spring term, 1977.**

- c. The child should know how to acquire information other than by reading; for example, by asking questions, by experimenting, from watching television.**

**Autumn term, 1977.**

- d. The child should be able to:**
  - i. Compute the four arithmetic rules using his knowledge of number, multiplication tables and different units of measurement.**
  - ii. Use mathematical techniques in his everyday life, i.e. estimating distances, classifying objects, using money.**

**Spring term, 1978.**

- e. The child should know how to convey meaning clearly and accurately through speech for a variety of purposes; for example, description, explanation, narration.**

Despite initial difficulties associated with the unfamiliarity of these forms of assessment all the teachers completed the checklists of abilities and the majority provided feedback by completing the open-ended questionnaire designed to review the validity of the checklists of abilities and structured activities. The analysis of data proceeded on completion of each teacher-based assessment. Feedback on pupil performance and the responses made by the teachers on the questionnaire were also provided at regular intervals. In this way it was anticipated that the teachers involved in using the teacher-based assessments would continue to cooperate with the research and it was hoped be able to use the evidence provided by the research for formative purposes in developing their skills in the assessment of pupils and thus their own classroom practice.

In the following section the methods used to ascertain the validity and reliability of the teacher-based assessments are presented. This discussion relates to the validity of the checklists of abilities developed by the teacher groups (3.6.1), the structured activities designed to measure the criteria on these checklists (3.6.2), validity of the teacher assessments using the checklists of abilities (3.6.3) and the practicality and reliability of the teacher-based assessments (3.6.4).

### **3.6 VALIDITY AND RELIABILITY OF TEACHER-BASED ASSESSMENTS**

Validity is a relative concept in that the purpose of the assessment must be taken into account when determining the type of validity which is to be demonstrated. Thus in the case of the teacher-based assessments, their purpose was to provide a broader range of product measures for use in the Oracle research programme to compliment the other assessments used in the description of pupil performance and progress. It was important to establish the validity of the teacher-based assessments for both the checklists of abilities and structured activities to ensure that they accurately represented those aspects of pupil performance on the extended aims of primary education which were considered important by teachers. Thus the content validity of the checklists of abilities as adequate descriptors of the aims under consideration needed to be established. In addition, it was necessary to consider the construct and content validity of the structured activities in order to ensure that they reflected pupil performance in areas other than those traditionally assessed by tests of the basic skills and provided a representative sample of the behaviours exhibited in the performance domains set out in the aims under consideration.

Strategies were therefore devised which made use of the feedback obtained from



teachers who were part of the development process (development group) and those who were involved in the application of the teacher-based assessments in the Oracle research programme (application group). In this way evidence of the validity of the checklists of abilities and the structured activities was available based on the notion of logical review (Black and Dockrell, 1984) in addition to the more traditional empirical review procedures which were incorporated wherever practicable.

The validity and reliability of the teachers' assessments were also considered since, as has been previously indicated, these assessments have been found to be subject to error. It was, therefore, important to demonstrate that it was possible for teachers using the checklists of abilities to apply the criteria appropriately and for different teachers to show a good measure of agreement when assessing the same pupils or over time. Correlations of teachers' assessments with the structured activity scores obtained for each criterion and scores on the tests of basic skills were used to establish whether teachers discriminated between the abilities of pupils as assessed by these different methods and thus the validity of their ratings. The methods used to ascertain the validity of the teacher-based assessments including the structured activities, checklists of abilities and the teachers' assessments are summarised in Table 3.8 overleaf. Assessments of the same pupils by different teachers and by the same teachers over time were used to ascertain reliability.

### 3.6.1 Validity of the checklists of abilities

The validity of an assessment procedure can be described in a number of ways, such as *face*, *predictive*, *construct* and *content* validity. Various types of validity are now considered in more detail with regard to their appropriateness for establishing how well the teacher-based assessments measured those areas of the curriculum which were selected for study. In the case of the checklists of abilities, face validity was to some extent ensured by the process of development. The areas used were taken from a large-scale survey of teachers' opinions on the aims of primary education in England and Wales (Ashton *et al.*, 1975), so that they were representative of the understandings of teachers of the nature of the primary school curriculum and what pupils were expected to be able to achieve at that time. The selection of the intellectual aims listed in Table 2.1 has already been justified in terms of the nature of the Oracle research programme in that it was designed to measure pupil achievement in terms of progress and performance under different teaching styles rather than those aims related to spiritual/religious, aesthetic, personal/social or emotional development.

**Table 3.8** Summary of methods used to ascertain the validity of the teacher-based assessments.

<i>Teacher-based assessment component</i>	<i>Methods used</i>	<i>Type of validity</i>
Checklists of abilities associated with each of the five aims under consideration.	<i>Logical review procedures:</i> open ended questionnaire given to development and application groups of teachers.	Content
Structured activities designed to assess criteria on the checklists of abilities.	<i>Logical review procedures:</i> open ended questionnaire given to application group of teachers.	Content
	<i>Empirical review procedures:</i> Factor analyses of the structured activity and Richmond test scores, intercorrelations between structured activity and study skills components.	Construct
	Correlation of relevant structured activity scores with tests of listening skills and creativity.	Construct
Teachers' assessments using the checklists of abilities.	Correlation of teachers' assessments on relevant criteria with associated structured activity scores. Analysis of covariance by pupil age, sex, achievement in basic skills, social class, type and teaching style.	Construct



The final selection of the aims considered in this research was also made by a sample of teachers and thus represented the most important curriculum areas as perceived by these teachers. In this way the actual aims under consideration were those which were of relevance to primary school teachers at that time and thus had face validity in that they represented the most important aspects of the primary curriculum for assessment.

Content validity is normally established by showing that the behaviours demonstrated in the assessment "*constitute a representative sample of behaviours to be exhibited in a desired performance domain*" (American Psychological Association, 1974, p. 98) so that an assessment has normally been considered to be content valid if the measure has been shown to be sufficiently representative to allow generalisations to be made about pupil achievement on the basis of that particular measure. Thus evidence of content validity was required for the checklists of abilities in relation to each aim in order to establish whether or not they provided a representative sample of pupil behaviour shown in relation to the performance domain described by each aim. This was provided by the use of logical review procedures. Both the development groups and the application groups of teachers were asked to complete an open-ended questionnaire (Appendix 13), giving their opinions of the appropriateness and range of criteria for each aim. The results of this questionnaire are detailed in the following chapter (section 4.1) and constituted an attempt to use the type of logical review procedures advocated by Black and Dockrell (1984, p.67-70) to establish the content validity of the checklists of abilities.

### **3.6.2 Validity of the structured activities**

Much of the material and methods used in the structured activities were not susceptible to traditional forms of validity testing since they constituted a form of criterion referenced diagnostic assessment. Alternative strategies such as those advocated by Black and Dockrell (1984, p.67) were, therefore, used to review their validity. These strategies included the use of "*logical review*" and "*empirical review*" procedures. Logical review procedures involved the demonstration of the "*fit*" or "*quality*" of items used in the structured activities when compared to the checklists of abilities. This was achieved by asking both the development and application groups of teachers to provide feedback on the appropriateness of the items and any questions that they considered problematic. The views of the development group of teachers were also reflected in the discussions leading to the development of the teacher-based assessments particularly with respect to the structured activities which were designed to match the checklists of abilities. The feedback obtained from the application group of teachers is discussed in the following



chapter (section 4.2.1). It is argued that the results of these analyses when taken with those of the empirical review procedures now discussed presented a more complete picture of the validity of the structured activities than would have been possible if only traditional statistical methods were used.

Empirical review procedures involving the analysis of the structured activity scores made use of more traditional statistical techniques where applicable. In the case of the structured activities both construct and content validity were considered using methods of statistical analysis. Two methods were used to demonstrate construct validity. First, factor analyses were carried out to provide evidence that the constructs measured by these exercises were different from those measured by the Richmond Tests of basic skills and the study skills exercises described earlier in this chapter and that each of the structured activities assessed aspects of pupil performance related to the aim under consideration.

Second, the relationship between the scores obtained on the structured activities and those on measures broadly covering a similar area was investigated (4.2.2). This strategy was appropriate in relation to two of the aims only; *"listening with concentration and understanding"* and *"inventiveness and creativity"*. Materials from the Schools Council Oracy project (Wilkinson *et al.*, 1976) were available for the area of listening comprehension and provided some measure of overlap with the area of *"listening with concentration and understanding"* although they were not designed to assess the criteria listed on the checklists of abilities specifically. These tests were administered to pupils who also completed the structured activities for *"listening with concentration and understanding"*. In addition, data was collected on creativity using other techniques advocated by Foster (1971, p.113) and compared with pupil performance on the structured activity associated with the aim related to *"inventiveness and creativity"*. The criteria for assessment used in these alternative techniques for the assessment of creativity differed from those developed by the teachers' group, in that they defined creativity in terms of novelty, fluency and flexibility of response rather than originality and appropriateness. However, if both measures were tapping creativity then some measure of correlation and thus construct validity could be assumed if the teachers' understanding of this concept matched that described in the literature.

The demonstration of content validity using traditional procedures such as ensuring the test contained homogenous items was clearly inappropriate for the analysis of the structured activities, since these were designed so that each component of the structured activity matched particular criteria on the checklist of abilities and thus measured different



aspects of pupil performance. The demonstration of the content validity of the structured activities in relation the checklist of abilities was also limited. For practical reasons, the structured activities were designed so that the evidence of pupil performance was in a written form. This was to ensure that the data could be used to moderate the assessments made by teachers since the large size of the sample precluded other checks such inter-rater agreement. However, this meant that certain abilities on the checklist were not tapped by a structured activity. As the content validity of the checklist depended on the inclusion of all the abilities thought to represent the aim, and the structured activities did not tap all these criteria they, therefore, were designed with limited content validity.

However, teachers in the classroom were able to observe and question pupils in order to make assessments of the abilities which were not measured by the structured activities thus ensuring that all aspects of the checklists of abilities were assessed either by the structured activities and/or by the ratings made by the teachers in relation to each criterion. In this way the performance domain was assessed either by some aspect of the structured activities or a teacher assessment. It has been noted previously that the teachers' assessments have been found to be inconsistent over time and between individuals, as well as being influenced by a number of characteristics which were unrelated to pupil performance on the areas of the primary curriculum being assessed. It was, therefore, even more important that the teachers' assessments were valid.

### **3.6.3 Validity of teachers' assessments**

As has been noted previously the validity of teachers' assessments has generally been taken for granted. However, in the development of the teacher-based assessments a key feature was the design of the structured activities to match the criteria listed on the checklist of abilities. Although it was not possible to provide structured activities to match all criteria on the checklists of abilities the data did enable a number of factors which it was hypothesised might bias a teachers' judgement of pupil performance to be investigated. This was achieved by using analysis of covariance to identify the contribution made by a number of factors such as pupil age, sex, achievement in the basic skills, social class and type, as well as teaching style in determining the kind of bias reflected in teachers' assessments. In addition, comparisons between teachers' assessments of pupils on the checklists of abilities were made with pupil scores on the Richmond tests of basic skills and structured activities in order to ascertain the degree of agreement between them and thus the construct validity of the teachers' assessments. The methods used to ascertain the validity of the teacher-based assessments have been



summarised in Table 3.8. The strategies used to establish the practicality and reliability of the teacher-based assessments are now considered.

#### **3.6.4 Practicality and reliability of the teacher-based assessments**

The first priority was to establish the practicality of making assessments of this kind. It would have been desirable to pilot all the checklists of abilities to ascertain their practicality, as well as their reliability and validity prior to distribution to schools. Given the time constraints operating within the research programme this was precluded. As was noted previously the schedule for the development of the teacher-based assessments and their introduction to the Oracle schools was demanding. This meant that the checklists of abilities and the structured activities were developed and piloted within six months of setting up the development groups. Given this schedule and the fact that the teachers were working full-time in the classroom it was not possible to complete full reliability studies in relation to each of the aims under investigation. However, it has been assumed that if a particular aspect of reliability was demonstrated for one aim then the same principles apply in relation to each of the other aims. This strategy reduced the demands made on teachers in the expectation that the reliability of their assessments would be improved.

The reliability of teachers' assessments was investigated in three ways in this study. First, the "*listening with concentration and understanding*" checklist was used in order to ascertain the degree of intercorrelation between different teachers' assessments of the same pupils. This group of teachers had not previously used the checklist and this data provided evidence of the practicality of the checklist and inter-rater reliability. This checklist was one of the first to be developed and was used for this reason. Second, a similar procedure was used by the development group of teachers who assessed taped examples of pupils talking for the "*conveying meaning clearly and accurately through speech*" checklist in order to establish that this format would be useful for teachers and provide reliable data for the moderation of teachers' assessments in the Oracle study. Third, teachers made independent assessments of their pupils using the "*acquiring information*" checklist on two occasions separated by a four week interval. The intention here was to ensure that the ratings given by teachers using the checklists were comparable over time. This activity was selected for this exercise since the checklist was one of the most complex and provided less opportunity for the teachers to remember their assessment of a particular pupil on each of the criteria. In addition, a test-retest reliability check was made of the "*acquiring information*" structured activity. The reliability and validity checks that were carried out prior to the use of the teacher-based assessments in



the Oracle schools are now described in detail in the following section.

### **3.6.5 Reliability and validity checks**

As has been indicated a number of reliability and validity checks were carried out prior to the administration of the teacher-based assessment materials in the Oracle research schools. Each set of checklists and structured activities were used for a different reliability or validity check since it was impractical given the time span for development and administration to have carried out a full and detailed study of each set of teacher-based assessments with regard to all the aspects of reliability and validity discussed previously. The results of these studies are presented in Chapter 4. Details of these checks are now described in relation to each of the five aims considered in this study.

*a. The child should be able to listen with concentration and understanding.*

The checklist of abilities associated with this aim was given to three teachers not associated with the Oracle research programme who team taught their classes in a open plan area. Each teacher completed the checklist rating the pupils from all three classes on each of the criteria. The purpose of this check was to ensure that different teachers when assessing the same pupils reached a good measure of agreement so that the inference could be drawn that other teachers would be able to use the checklist reliably. The correlation coefficient of agreement between the teachers as a group and with reference to each other was calculated as an indication of the comparability of their assessments. In addition construct validity was checked using the standardised test materials for the assessment of listening comprehension developed in the Schools Council Oracy Project (Wilkinson *et al.*, 1976) for comparison. These tests were administered to pupils who also completed the structured activities for "*listening with concentration and understanding*" in order to evaluate the degree of comparability of these two types of assessment using Pearson r product moment correlation coefficients.

*b. The child should develop some inventiveness and creativity in some fields; for example, by painting, music, mechanical things, poetry, movement.*

The structured activities associated with this aim were given to pupils in three classes which were not involved in the Oracle research programme. In addition, they were given tests of creativity based on the picture completion exercise devised by Torrance (1967).



The correlation between pupil scores on the structured activity for this aim and the other creativity tests was then calculated using Pearson  $r$  product moment correlation coefficients to determine the degree of overlap between these two types of assessment of creativity.

- c. The child should know how to acquire information other than by reading; for example, by asking questions, by experimenting, from watching television.*

Three teachers were given the checklists of abilities associated with this aim to complete and asked to repeat their assessment some four weeks later. The assessments on each of the criteria on the checklists for each pupil were then correlated in order to ascertain the stability of the assessments over time. In addition, the scores on the structured activities associated with this aim were compared with reading scores obtained on the modified version of the Richmond Tests of basic skills used in the Oracle research programme. This evidence was used to ascertain whether these activities were assessing skills other than those of reading.

- d. The child should be able to :*

*compute the four arithmetic rules using his knowledge of number, multiplication tables and different units of measurement,  
use mathematical techniques in his everyday life, i.e. estimating distances, classifying objects, using money.*

The structured activities related to this aim were given to a class of pupils who also completed the modified Richmond tests of basic skills in the mathematics. An analysis of variance was carried out on the data in order to determine the extent to which the structured activities assessed different areas of mathematical competence from those assessed by more conventional standardised tests of mathematical achievement.

- e. The child should know how to convey meaning clearly and accurately through speech for a variety of purposes; for example, description, explanation, narration.*

A group of four teachers were asked to rate the performance of three pupils according to the criteria listed on the checklist of abilities. The level of agreement was calculated using the Pearson  $r$  product moment correlation coefficient to ascertain the degree of comparability between these teachers' assessments and hence the probability of other teachers being able to use this procedure effectively.



### **3.7 PUPIL PERFORMANCE ON THE STRUCTURED ACTIVITIES**

The working hypotheses regarding pupil performance which were investigated within this study have been detailed in section 3.2.2. These are given below.

- \* Pupil performance on the structured activities in relation to each of the five aims investigated is correlated to the following pupil characteristics: sex, age, social class, performance on the modified Richmond tests of basic skills and pupil type.
- \* Pupil performance on the structured activities is correlated to the teaching style experienced.

These hypotheses were posed as it was expected that the pattern of pupil performance in relation to the extended aims of primary education would reflect that found with basic skills achievement. It was suggested that for this age range (7 - 9 years old) there might be different levels of performance between boys and girls similar to those differences observed in mathematical and language achievement. For example, boys would be expected to perform better than girls in relation to everyday mathematics and acquiring information other than by reading whilst girls might be expected to do better in relation to listening skills and conveying meaning clearly and accurately through speech. Similar differences in performance might also be expected with regard to age and social class. In this way older pupils would be expected to perform better than younger pupils and pupils from middle class families to achieve better results than those of working class backgrounds.

It was also hypothesised that although the structured activities were designed to assess skills not measured by the Richmond tests of basic skills some overlap in ability to perform successfully on both types of assessment could occur. For example, it was more likely for a pupil who was a successful learner to do well on both tests than not, although it was also possible to conceive of a less successful pupil in terms of achievement on the basic skills tests showing a better level of performance on the structured activities than expected. In addition, given that different pupil types were identified in the Oracle research then any observed differences in performance on the structured activities might be understood with reference to aspects of pupil classroom behaviour which would seem to favour successful learning in relation to the extended aims of primary education which were under investigation.



Data available from the Oracle research programme for the variables of pupil sex, age, social class, performance on the Richmond tests of basic skills, type and teaching style experienced for each pupil who completed the structured activities (see sections 3.3 and 3.4). This data enabled comparisons to be made between the performance of pupils on the structured activities for each of these variables. Analysis of variance controlling for performance on the Richmond tests of basic skills was used to account for differences in initial levels of achievement. It was proposed to account for any measure of correlation between pupil performance on the tests of basic skills and the structured activities by the use of covariance analysis which would enable differences in specific skills related to pupil performance on structured activities to be distinguished from any general skills or pupil attributes which contributed to pupil performance as measured by the Richmond tests of basic skills.

It was also possible for any residual differences in pupil performance on the structured activities to be a result of the teaching style they had experienced. In order to investigate the relationship between teaching style and pupil performance on the structured activities it was necessary to consider the possibility that some students might do well on the structured activities because they were in general successful students as indicated by their performance on the Richmond tests of basic skills rather than as a result of the different teaching style experienced. It was, therefore, important in the analysis of the influence of particular teaching styles to allow for any effects of proficiency in the basic skills upon the structured activity scores. For this purpose the total score from the pre-test administration of the Richmond test was obtained by adding together each individual pupil's score on the tests of mathematics, language skills and reading. By determining the regression line between this total pre-test score and the scores on the structured activities it was possible to allow for the effects of successful performance in the basic skills by adjusting the structured activity scores. The difference which remained, if any, between the observed and predicted scores was then attributed to the different teaching style experienced by the pupils in the classroom. The results of these investigations into pupil performance on the structured activities and the effects of different teaching styles are presented in chapter 5.

### **3.8 METHODS OF ANALYSIS USED IN THE STUDY OF TEACHER-BASED ASSESSMENTS**

A number of statistical procedures and methods of analysis have been referred to in the preceding descriptions of the methods used to investigate the validity and reliability of



the teacher-based assessments and the study of pupil performance on these assessments. These are now considered in more detail and their applications in the research described.

### **3.8.1 Correlational analysis**

This statistical procedure provided a measure of the degree of agreement between two variables. It has been used to ascertain the level of agreement between the ratings made by different teachers of pupil performance on the checklists of abilities which formed part of the teacher-based assessments. In addition it has been used to identify the degree of relation between the scores of pupils on the structured activities, Richmond tests and teachers' assessments in order to investigate the content and construct validity of the teacher-based assessments.

### **3.8.2 Factor analysis**

This statistical procedure has been designed to show in quantitative terms the pattern of correlations between variables. In the case of the teacher-based assessments it has been used to determine the degree of overlap between items on the structured activities, study skills exercises and the Richmond Tests in order to ascertain whether the teacher-based assessments measured significantly different elements of pupil performance than did the Richmond Tests and the tests of study skills. It was noted in the previous chapter that these tests of study skills related to aspects of pupil performance concerned with independent study. The tasks which comprised these tests were selected on the basis that a pupil might be expected to complete such exercises as part of their normal classroom activities and that they might involve common underlying study skills. It was, therefore, important to determine the extent to which the structured activities assessed similar underlying study skills when compared with these tests and to what extent they assessed different aspects of the aims under consideration.

If, as a result of the factor analysis, new variables were produced, then each factor would represent a cluster of inter-related items. The clusters of items generated by this analysis have no meaning in themselves since they simply represent mathematical relationships. However, by inspection of the clusters of items it is possible to identify related areas and thus label each factor in such a way as to describe its attributes as accurately as possible. In this way it was possible to determine the content and construct validity of the structured activities which formed part of the teacher-based assessments since the aspects of pupil performance which were tested by the items contributing to



particular factors could be used to describe pupil attributes, in this case related to their abilities in selected areas of the curriculum other than the basic skills.

### 3.8.3 Analysis of covariance

This statistical procedure has been used in cases where the sample under investigation has not consisted of independently chosen random groups, where analysis of variance would have been the appropriate method of analysis. In the case of this study it was not possible to control variance through the selection of a random sample of teachers or pupils. Similarly, it was not feasible to set up an experimental situation where the performance of a control group was compared with that of an experimental group. Analysis of covariance has thus enabled the within group variance to be controlled statistically in a non-experimental situation. The procedures followed in the Oracle research programme were those given in *Statistical Package for the Social Sciences* (Nie *et al.*, 1975, pp. 405-8).

By using this technique it was possible to investigate the effects of pupil age, sex, social class and type on the assessments made by teachers even though the initial performance of pupils with respect to basic skills and the structured activities were different for each group of pupils. The assumption was made that there was a general and specific set of skills associated with each of the teacher-based assessments. Thus the scores obtained from the structured activities could be adjusted to look at the specific skills associated with each aim by eliminating the general component through the covariance analysis of the scores on the Richmond tests of basic skills and the study skills exercises.

In addition, the working hypothesis that certain teaching styles could result in more effective assessment of different pupil types could be tested by the use of a two factor analysis of covariance since one of the problems which arises when considering the effects of more than one variable is that of a statistical interaction between those variables. In the case of teaching style and pupil type it was possible to conceive of a particular teaching style achieving better results with one pupil type than another. This was indeed the case and the relationship between teaching style and pupil type (Galton *et al.*, 1980, p.150) was controlled for in order to investigate the particular features of teaching style associated with more valid and reliable assessments of pupil performance. A two factor analysis of covariance was used to ascertain the joint effects of teaching styles and pupil types on teachers' assessments with the covariates of basic skills and structured activities



being controlled for in the analysis. By carrying out this analysis it was possible to treat any differences found between the assessments of different pupil types or different teaching styles with confidence.

#### 3.8.4 Content analysis

This method of analysis has been used as a means of analysing documents in historical research, in linguistic analysis and in the interpretation of focused interviews and is a multipurpose research method developed specifically for investigating a broad spectrum of problems in which the content of communication serves as a basis of inference. It has been defined by Holsti as

*"any technique for making inferences by systematically and objectively identifying specified characteristics of messages."* (Holsti, 1968, p.601)

Originally the skills needed to analyse documents or transcripts required the researcher to collect, classify, order, synthesise, evaluate and interpret the content of the verbal material. Developments in content analysis have resulted in the superimposition of a quantitative mode of analysis over these essentially qualitative descriptions of verbal material. The basic premise of this method of analysis has been the creation of suitable categories and units of analysis which reflected the nature of the document and the purpose of the research. Categories were normally constructed from an inspection of the document to reflect the main areas of content, further units of analysis may then be included, such as the use of particular words, themes, sentences or paragraphs.

Within this study the categorisation of verbal material has been used in a number of different contexts. First, it was used in the development of the teacher-based assessments to analyse the transcripts collected for each of the teachers' meetings and formed the basis of the discussion documents which have been summarised in Appendix 7. Second, in the structured activities associated with the aim *"acquiring information other than by reading"* pupils were required to write the questions they would like to ask the person in the story. These responses were categorised using the techniques of content analysis. Third, the responses of teachers to the open ended questionnaire administered to review the appropriateness of the teacher-based assessments were also analysed using this technique. Finally, the technique was used extensively to study the data generated from the evaluation of the in-service course now described.



### **3.9 DEVELOPMENT AND EVALUATION OF AN IN-SERVICE EDUCATION PROGRAMME.**

In the review of the literature given in Chapter 1 a number of considerations were discussed which have relevance to the development of such a curriculum for the in-service education of teachers. It has been suggested in previous discussions that teachers should contribute more to research and this was reflected in the involvement of teachers in developing the teacher-based assessment research within the Oracle research programme. This in-service programme was developed as a result of the observations made in the development of the teacher-based assessments and the findings of the research presented in the following chapters, in addition to the increasing emphasis on classroom and school-based in-service education which was being advocated as a means of improving the quality of teaching and ultimately pupil learning.

It has already been argued that any development in the ability of the teacher to accurately diagnose the needs of pupils, to provide appropriate instructional materials and to assess pupil performance validly and reliably might lead to improvements in the quality of teaching and learning in the classroom and of educational standards in general (Bennett *et al.*, 1984). In previous discussions the involvement of teachers in research and the desirability of in-service provision to help teachers in their task of diagnosing individual learning needs and assessing pupil performance was also advocated as a response to previous studies of the invalidity and unreliability of teachers' assessments. This part of the study has, therefore, focused on the provision and evaluation of an in-service curriculum which would provide for this need taking account of the recommendations for increasing teacher involvement in their own professional development.

The issues which are addressed in the evaluation of in-service provision within the B.Ed. Honours degree programme offered at Worcester College of Higher Education are detailed below:

- \* Professional learning involves the recognition of theories in use and espoused theories and the nature of any discrepancy between them is necessary for the generation of new theories of action.
- \* The study of assessment procedures facilitates this process of professional learning and leads to teacher self-evaluation, which is then transferred to the classroom context.



The first hypothesis was derived from the findings of previous studies, in particular, those of Day (1981) who applied the theories of professional learning proposed by Argyris and Schon (1976) to teachers' practice. The basic premise of this theory of professional learning was that theories of action determine practice and these theories can be attributed to explain or predict a person's behaviour. Within any theory of action, two components were distinguished: espoused theories and theories-in-use. Espoused theories were used by an individual to justify or describe behaviour, theories-in-use reflected what people actually did or how they operationalised their espoused theories. Argyris and Schon maintained that profession practice was made up of a number of inter-related theories of action. Thus to change practice the provider of in-service education would have to change the teachers' theories of action. This could be achieved by providing situations where individuals gained insight into the conditions under which their theories-in-use inhibit and facilitate their own growth and the growth of others. Information should also be provided from which individuals could design programmes for self-improvement, gain help from others and evaluate their own progress as well as help individuals learn how to discover their own theories-in-use and generate new ones; that is, learn to generate directly observable data, infer theories-in-use, alter them and test new theories of action.

This theory of professional learning has some commonality with the ideas associated with the self-monitoring and self-evaluating teacher advocated by Elliott (1978a). In both cases teacher professional development was related to changing practice, where teachers were expected to confront their own theories-in-use (practice) by the analysis of events, to reflect on the significance of these events in relation to their intentions (espoused theories) in order to change their practice so that they achieved their intentions and thus developed new theories of action. However, research has also indicated that teachers were often unaware of discrepancies between their theories-in-use and their espoused theories (Keddie (ed), 1973; Day, 1981; Woods, 1979). Thus within any in-service provision there must be the opportunity for these discrepancies to be made explicit and for new theories of action to be acquired so that theories-in-use become more compatible with espoused theories. Previous studies have also indicated that the way in which the such discrepancies were identified has made a significant difference in the effectiveness of the change which occurred.

Day (1981) argued that self evaluation was promoted by encouraging teachers to judge their own levels of effectiveness and set their own goals for what they wished to attain. It has also been suggested that the motivation for change and professional learning arises



out of the identification of a problem by the client, since it is argued that long-term changes are more likely to occur when learning is self-motivated and knowledge is self-appropriated and self-actualised (Rogers, 1967).

The facilitation of professional learning can therefore be based on a number of principles. First, effective learning only occurs in response to the confrontation of problems by the learner (Day, 1981); second, decisions about teaching should stem from reflection on the effects of previous actions (Sutcliffe and Whitfield, 1976; Calderhead, 1984); third, effective confrontation of problems requires the maximizing of valid information (Day, 1981). These principles provided the basis for developing the in-service course provision described in the following section.

The catalyst for confronting the teachers with their own practice was different from that advocated by Day (1981) and Elliott (1979b) in that the data collection was focused on describing pupil learning rather than the collection of data on the teachers' own performance. This strategy was used since it has been argued previously that one of the more important factors in the improvement of teachers' practice lay in being able to make valid and reliable assessments of pupils' abilities and thus provide more appropriate learning experiences. Previous experiences of working with teachers had also suggested that they examined their own practice as a result of identifying the qualities in pupil learning. The effectiveness of this approach was evaluated within this research. It was thus hypothesised that this focus on assessment, in theory and practice, was a means of improving both the quality of teaching by promoting self-evaluation and the quality of pupil learning by ensuring more accurate diagnosis of learning needs and effective evaluation of outcomes of particular teaching styles.

### **3.9.1 Strategies for the development of in-service provision**

The strategies for the development of appropriate in-service education evolved as a result of a number of different types of involvement by the researcher in providing in-service courses for practising teachers as well as evidence provided by the research literature. The first experience of working with groups of teachers occurred in the development phase of the research into teacher-based assessments which has been described in section 3.5. This experience was important in formulating the way in which the materials were presented to the application groups of teachers in the schools participating in the Oracle research programme. In addition, the researcher hypothesised that in terms of promoting change in the practice of teachers the most important part of the



study was the involvement of teachers in the processes of developing the teacher-based assessments. This was further reinforced by conversations with teachers who had been involved in this part of the study.

As a result of these experiences I attempted to establish a collaborative group of teachers to study assessment practices in their classrooms as a means of identifying their needs and to develop more appropriate assessment methods to meet these needs (Jasman and Ashby, 1980). The materials and notes for teachers generated within this course are summarised in Appendix 14. This collaborative group of teachers was self-selected in that they specifically enrolled for a ten week in-service course in assessment in the primary school during 1979 for which there was no professional award. This in-service course was not formally evaluated at the time but the experiences of the teachers further reinforced the idea that more teachers should have the opportunity for learning the skills of observation and acquire knowledge of assessment in order to facilitate the improvement of their classroom practice. Thus the need was perceived to incorporate such perspectives into both initial teacher education and in-service education courses.

During the period from September 1980 until July 1982, I was involved in lecturing on a full-time B.Ed (Honours) in-service degree programme based at Worcester College of Higher Education which also contributed to the development of ideas implemented in the in-service programme described here. This was begun in September 1984 when the researcher was again lecturing on the full-time B.Ed.(Honours) in-service degree programme<sup>1</sup>. This course was developed to facilitate self-evaluation and classroom-based research.

These experiences were the basis of the design of the particular course structure, content and processes which constituted one of four strands which comprised Part 1 of the degree programme (Appendix 15.3) and occurred during the first term of the course. Teachers enrolled on the course were subsequently required to implement some form of professional study (Appendix 16) within the context of their school or institution when they returned to their previous positions or took up a new appointment. A second theory-practice strand was developed in Part 2 of the degree which addressed the questions of how best to implement change within such contexts through the use of teacher-based research, curriculum development and evaluation. This second strand was completed in the final term of the taught course.

---

<sup>1</sup> Referred to as the course hereafter.



There were a number of significant features in the programme which suggested that it offered teachers a relevant and challenging course of study and also provided the necessary conditions for professional development as outlined in previous discussions. One of these features was the opportunity for placement in a local school one day a week for the duration of the taught course. During this placement the teachers were expected to identify needs within the school and to investigate these with reference to the theoretical and methodological issues which formed the basis of the college-based course. This facilitated the application of theoretical constructs in the context of a "real" situation with reference to a particular pupil, classroom or school and the practical concerns of the teachers as they arose.

Another significant feature was the school-based study which each teacher completed in their own school. This also facilitated their professional development and that of the school or institution as a whole on completion of the college-based part of the programme. It was, therefore, possible for the teacher to apply the ideas, concepts, principles and techniques which were encountered during the course and explored through the school placement, to their own context by means of this school or classroom-based project.

As has been noted previously the development of this strand of the course drew on my experiences of working with teacher groups in the area of assessment, particularly in relation to the classroom performance of primary school pupils. My intentions and purposes at the time of running the course were shaped by these experiences. The most important feature of my espoused theory at the start of the course was that teachers should be actively engaged in the college course and be involved in the observation and assessment of pupils in the classroom. This was linked to my belief that this would provide the appropriate context for the teachers to appreciate and become aware of the discrepancies between what they intended the pupils to learn, the opportunities that were provided for learning and ultimately what the pupils did actually learn.

This central assumption was based on previous experience which suggested that teachers were generally more prepared to look closely at pupil learning than to be videotaped in the classroom or collect data on their own performance. Focusing on pupil behaviour and learning appeared to be a more acceptable way of investigating the events in the classroom than immediately concentrating on teacher performance and style. However, I had also noted that once the teachers began to observe and assess pupils more closely they were inevitably confronted by the importance of the classroom context



and their own teaching style in providing appropriate learning opportunities for pupils in order to achieve their own curricular intentions. This then provided the impetus for the teachers to examine more closely the nature of their theories-in-use and investigate any discrepancies between these and their espoused theories.

This basic assumption was important to the development of the in-service curriculum in two ways. First, the course was designed in such a way as to facilitate the observation and assessment of pupils in the classroom as well as provide knowledge and skills within the college context. Second, this process of observation and assessment of pupils was equally applicable to my own situation where I planned to monitor and assess the teachers as learners during the course in order to promote my own self-evaluation. Thus the content and processes of the course were based on these considerations.

My strategy in planning the course was to create as many opportunities as possible for the active involvement of teachers in the sessions in an attempt to facilitate professional learning. These sessions were set up as workshops (see Appendix 15.4) with opportunities to discuss issues, analyse and solve particular problems through the use of newly acquired skills and knowledge. The teachers were expected to participate in the workshop sessions and then apply their learning in the placement school. This was to culminate in the preparation of case studies focusing on the learning needs of a particular pupil or group of pupils. The content was developed over a number of years working with teachers enrolled in this course and in previous in-service activities. Particular content was used in order to promote certain processes since some selection of material was necessary with only 24 hours contact time allocated to this strand.

The initial stages of the strand were designed to develop observation skills through the use of the Pupil Record (Galton, 1978b) and introduce the students to a range of assessment procedures using as examples the materials produced in the Schools Council projects *Communication skills in early childhood* (Tough, 1979) and *Progress in learning science* (Harlen *et al.*, 1977a). The introductory sessions focused in the first instance on pupil observation and methods of recording the information which had been collected. The teachers were encouraged to use a range of methods and modify them on the school placement. In this way they were given the opportunity to determine at first hand the strengths and weaknesses of particular types of observation, informal appraisal and assessment, in addition to collecting information on individual pupils for the case studies.



As the course progressed more formal methods of assessment were introduced such as the Richmond tests of basic skills (France and Fraser, 1975) where the issues of reliability and validity were considered. Many of the teachers on previous courses had used these tests in their schools and discussions normally drew upon their experiences. This activity was also designed to help the teachers come to terms with the range of different purposes which educational assessment has served and to recognise the importance of judging the appropriateness of particular methods of assessment for these differing purposes and for various audiences.

This introduction to methods of assessment was then followed by workshop activities. In these sessions the teachers were expected through discussion and reflection to conceptualise assessment for themselves and then to devise some method of assessment which would meet their needs in terms of assessing some aspect of pupil performance in the school placement. The teacher-based assessments served (Appendixes 8,10,11) as a model for this type of material which the teachers might produce. This workshop also served to highlight any misconceptions which the teachers may have had and encouraged them to share their understandings and reach some consensus about terms and issues related to assessment.

There then followed a session where the links between what and how teachers teach and the pupils learn were explored with reference to the findings of the Oracle research programme (Galton and Simon (eds), 1980), Bennett's (1976) study of teaching styles and pupil progress and the Schools Council's project that evaluated different styles of science teaching (Eggleston, Galton and Jones, 1975). This activity served two purposes. First, it was designed to introduce the teachers to research which was classroom-based. Second, comparisons could be made by the teachers between the various research methods, particularly the relative merits of the different forms of assessment used in these studies for measuring pupil achievement and progress.

Finally, the teachers presented a case study of a particular pupil who they had been working with on school placement. This activity was intended to provide the opportunity for the teachers to synthesise their knowledge and practical skills in the study of a particular case. The teachers were expected to describe their findings, evaluate the methods they had used and make recommendations for future learning activities. In this way the teachers were prepared for the subsequent school-based work within the course and the completion of a professional project on their return to a teaching position. The intended course content is summarised in Table 3.9.



**Table 3.9 Intended course structure**

<i>Date</i>	<i>Content</i>
27.9.84	Introduction to observational techniques, concept of appraisal, methods of recording information: for example, Tough (1979), <b>Talk for teaching and learning.</b>
4.10.84	Pupil Record category system with practice observations of videotaped examples of pupil classroom behaviour.
11.10.84	Informal assessment methods: Progress in learning science checklists.
18.10.84	Structured questions and activities in the assessment of pupil performance in the classroom: for example, assessment of project work, teacher-based assessments
25.10.84	Formal assessment methods: for example, Richmond tests of basic skills, Quality of Listening materials, 16 plus examinations.
1.11.84	Issues in monitoring and assessing pupil behaviour, performance and achievement.
8.11.84	Conceptualisation of "assessment".
15.11.84	Teacher record category system and analysis of teaching style.
22.11.84	Teaching style, pupil type and influences on pupil progress and performance.
29.11.84	Teachers to present case study of one pupil observed and assessed during
4.12.84	school placement.
<hr/>	
7.5.85	Nature of the professional project.
14.5.85	Research methodologies and strategies.
21.5.85	Research methods.
28.5.85	Curriculum evaluation: for example, evaluation of in-service course.
4.6.85	Methods of data collection: questionnaires, interviews.
11.6.85	Observation and assessment of learning: review.
18.6.85	Reading week - preparation of professional project proposal.
25.6.85	Statistical analysis, microcomputer usage.
1.7.85	Analysis of questionnaire responses: for example, teacher prepared questionnaires for the evaluation of the in-service course.
8.7.85	Presentation of professional project proposals.



Thus the above description has indicated my intentions in selecting the particular course content and teaching strategies in the in-service course. As noted previously, these activities were based on my experiences with two previous cohorts of teachers attending the course in 1980-1981 and 1981-1982. Teacher feedback on previous courses completed as part of the internal college programme of evaluation indicated that they were successful in achieving the programme objectives. However, I had not conducted any self-evaluation and there was no evidence to suggest that these strategies and curricular content were an effective way of promoting teacher self-evaluation. Thus there was a need to evaluate this course provision with regard to my own professional development as well as the effectiveness of this approach for facilitating teacher development, in particular, their professional learning related to assessment.

### **3.9.2 Research design and methods used for the evaluation of in-service provision.**

This research was, of necessity, a case study of a particular situation since the work was carried out within my own professional context. The evaluation of in-service provision was conceived as a project where I endeavoured to monitor the processes of the course and on the basis of the data collected (or the research) modify the course (or the action) as a means of enhancing professional development through self-evaluation. The course was premised on the need for teachers to build into their teaching the ability to change and critically appraise their own and the pupils' activities in order to bring theories-in-use in line with espoused theories and to examine critically such theories of action in the light of experience gained and new knowledge acquired through the course. Thus both the teachers and myself were involved in similar processes. By conducting this self-evaluation I intended to model the process as well as collect evidence of its effectiveness for change for myself and the teachers.

In the Ford Teaching Project (Elliott and Adelman, 1975, 1976) it was noted that teachers needed particular help in reflecting on their classroom practice and were encouraged to check their understandings of events and actions with how these were perceived and received by the pupils and by an outside observer; as well as recording events by photographs, slides, audiotape and videotape. The results obtained from this process of triangulation illustrated how often the perceptions of participants differed and emphasized the need to take account of differing perspectives of events. Thus the principle research method used was that of triangulation which has been described as a technique for the validation of data by collecting information in a number of different



ways or from a number of sources in order to cross check for the degree of congruence between these. This method has been favoured by Elliott (1977) as a means of validating any judgements that might be reached or conclusions drawn during teacher self-evaluation. Triangulation was, therefore, used to provide a number of different perspectives on the course processes and professional learning of teachers within the course and on the school placements. The timetable for the collection of this data is summarised in Table 3.10.

Elliott (1978b) has also argued that implicit in the notion of the self-monitoring teacher was the rendering of an explanatory account. The results of the course and self evaluation have been produced in the form of such an account drawing upon evidence provided by audiotape recording of the college sessions, tutorials with teachers and a personal diary. The evaluation of the processes and outcomes of teachers' professional learning drew on these sources of data as well as teacher self-reports, oral and written presentations of pupil case studies, school and course files and comparisons between line-labelled concept maps produced in the first and third terms of the course.

### 3.9.3 Methods of data collection

A number of different methods of data collection were available which would provide evidence relevant to the issues previously identified in accordance with the principals of triangulation. These methods are described in more detail and their relationship to the various working hypotheses discussed.

#### *Teacher self reports*

The completion of teacher self reports enabled the collection of information relating to the starting points of the teachers in order to provide contextual data for the modification of course content to meet individual needs. During the second session of the course the group of eighteen teachers were asked to provide details of their school context. These personal accounts provided information of school type and teaching responsibilities. The teachers were also asked to give a general description of assessment practices within their own schools and information on the methods of assessment which they had themselves used in the past. This information was used to provide a context for the development and modification of the course materials presented in order to make these more relevant to the individual needs of the students.

**Table 3.10** Schedule for the collection and analysis of data

<i>Date</i>	<i>Data collected</i>	<i>Data analysis</i>
27.9.84	Teacher self reports	Content analysis
4.10.84	Interviews with teachers	Analysis of interviews
27.9.84	Taught course sessions commenced	Audiotape recording and transcription on a week by week basis.
6.10.84	School placement commenced	
18.10.84	First line labelled concept maps completed	
29.11.84	Case studies of individual pupils presented	
6.12.84	End of part 1 taught course	√
<hr/>		
20.1.85	Submission of written accounts : school and course files	Assessment of course work and school files
<hr/>		
7.5.85	Part 2 course commenced: preparation for the professional project	
18.6.85	Submission of course and school files for Part 2 studies: containing personal account of learning in assessment and evaluation as an overview of the component of the files.	Content analysis of course files, school placement material and overview
1.7.85	Second line labelled concept maps presented.	Analysis of line labelled concept maps
8.7.85	Part 2 taught course completed.	

---



The teachers were also asked to provide a personal account of their understanding of assessment procedures, identifying those in which they considered they had some strength and those where they would welcome some help. This was used to provide data on the teachers' own priorities and further information for course modification. In addition, the teachers were asked to reflect on areas relating to the monitoring and assessment of pupil performance which they considered particularly important in the school or classroom.

A relatively unstructured approach was used to collect this data since it was considered important at the outset to recognise and take into account the individual concerns of the teachers. At the end of the course the teachers were asked to provide a similar account in the context of their assessed work by completing an overview of their learning related to assessment procedures. An alternative approach would have been to provide a questionnaire for the teachers to complete, however, this might have inhibited the depth of response made by the teachers and would have resulted in a less complete picture of their concerns. The self reports and overviews were content analysed and data resulting from this is presented in Chapter 6.

### *Tutorials*

In order check the analysis and authenticity of the self reports the issues raised by the teachers were discussed in more depth during individual tutorials. These provided an opportunity for the teachers to express any concerns they may have felt regarding the evaluation and how they perceived my intentions since I was also the lecturer and assessor for this strand of the course. They also facilitated the elaboration of points raised in the teachers' self reports and an exploration of significant previous experience that the teachers considered had influenced their thinking on assessment. I also asked the teachers to consider ways in which I could help them to monitor the way they assessed pupils. Cooperation in the application of particular assessment techniques including the teacher-based assessments was explored and whether the teachers were interested in developing their own methods of assessment which would meet their needs during the school placement. These tutorials were taped and subsequently transcribed for content analysis.

### *Tape transcripts*

An audiotape recording of each of the workshop sessions was made in order to



provide a record of actual events. This was to facilitate the analysis of the content actually provided within the course and compare this with the intended course structure. In this way I was able to identify changes in strategy and by reflecting on the patterns of interaction within the sessions identify key issues and concerns which resulted in changes within the structure and content of the taught material. This information was important in the context of analysing the congruence between my intentions or espoused theories and my theories-in-use. The transcripts also served as a way of monitoring the course processes. The teachers' contributions within the sessions provided valuable information on the processes of their learning, developing understandings and areas of concern during the course.

### *Assessed course materials*

The assessed components of the course provided a valuable source of evidence with respect to the outcomes of teachers' professional learning. These components comprised a school file and a course file which were designed to reflect the development of the teachers' thinking and application of course material, in other words the links between theory and practice. Details of the assessment criteria are given in Appendix 15.4. Of particular importance was the requirement during the first term of the course to identify the needs of a group of pupils in the placement school. The teachers were asked to carry out a needs assessment and make suggestions as to an appropriate curricular programme to meet these needs. The findings of each teacher were presented orally at the end of the first term. These were then submitted in written form in the school file whilst theoretical issues arising from this activity were considered in the course files. This material was content analysed so that evidence relating to the teachers' professional learning and development could be gathered.

### *Diary*

I kept a diary of reflections and ideas which occurred during the implementation of the course during the first term of the programme. This method of recording enabled me to trace the development of my thinking and reflections on events to support the subsequent analysis of data. It served a variety of purposes which included: providing a record of changes in the presentation of material, questions that arose during the course, issues of importance for further investigation and a means for facilitating reflection on events as they occurred so that these could be incorporated into the development of the course and its ongoing evaluation.



### *Content structure diagrams.*

The concept of cognitive learning style has been extensively explored in the literature but those studies which related to students in higher education were of particular relevance. Marton and Saljo (1976) identified two distinct approaches in their study of students' reading of texts: surface level processing and deep level processing. Pask (1976) found similar strategies with students who were asked to draw up classification systems. The "holists" tended to adopt a broad perspective and look for a variety of relationships whilst "serialists" were typified by their attention to detail and the pattern of learning by increments.

The development of the in-service provision described here was based on the assumption that if teachers were to become self-evaluative they should move from surface level processing and serialist thinking towards deep level processing and holistic thinking. In this way it was hypothesised that they would be able to ask questions which would result in the identification and solution of professional problems encountered which were characterised by a wide variety of contributing factors and a complexity of interrelationships between them. The professional learning of teachers would, therefore, be facilitated if the opportunity for more complex ways of thinking about classroom contexts and processes were promoted through the close observation and evaluation of classroom events particularly focused on the assessment of pupil learning. This suggestion was considered in the in-service evaluation.

The use of concept maps and content structure diagrams as a method of analysis arose from the field of cognitive science which involved the study of problems concerned with the representation of meaning, the structural and information processing aspects of language, knowledge and problem solving (Rumelhart and Ortony, 1977). It was suggested that this method of data collection was of potential use in curriculum research and development where it was necessary to represent the structure of a discipline or a body of knowledge (Posner, 1978). This method was, therefore, used to analyse the course content and as a means of assessing teachers' learning. This was of particular importance if it were shown that the teachers who were able to move from a surface type of processing to deep level processing also showed a greater ability to link theory and practice and engage in self-evaluation.

Within the context of the development and evaluation of the in-service provision for teachers this technique was applied in analysing the content and concepts of the course



itself. This type of analysis required that individuals ordered and represented in diagrammatic form their understandings of the relationships between certain curriculum content, concepts and ideas by the use of line labelled diagrams. The diagrammatic representations of cognitive structure derived from such data have been used for a number of purposes. These have included the analysis of curriculum materials, the assessment and evaluation of observed learning outcomes and the investigation of stages of cognitive development and learning style.

Thus a content structure diagram was a representation of the "teacher's" intentions in terms of the content and relationships which were to be taught. A labelled line concept map was a configuration of concepts and relations constructed by the learner. By using this method of data collection it was possible to investigate the way in which teachers were understanding the course content and to identify the degree of consistency between these and the course content as represented by the content structure diagram prepared by myself. In addition, by asking the teachers to complete line-labelled concept maps at the beginning and on completion of the course it was possible to investigate the development of complexity in their thinking which it has been suggested related to the holistic, deep level processing necessary for the problem-solving approach necessary for self-evaluation.

### 3.9.4 Methods of data analysis

#### *Analysis of line labelled concept maps.*

The qualities of the configurations derived from line labelled concept maps can be described in terms of *intensity* or the amount of structure determined by the total number of relations shown. A number of different types of relations can be represented in the concept maps. These relationships have been described in the following way (Ghaye, 1984).

- a) *Composite relationships* were those where two or more relationships were used to make one meaningful expression. They represented clusters of ideas and relationships. The whole expression incorporated conjunctives, prepositions, or pronouns, were descriptive and dependent on each other as functional parts. A concept or memory cue also served to link one concept to another.
- b) *Dependent relationships* were those which represented a procedural or logical



relationship between concepts.

- c) *Descriptive relationships* were those which represented a structural, functional or locational relationship between concepts.
- d) *Erroneous relationships* were those links which generated a false, misleading or confused expression.

In general descriptive type relationships were found to be more simple than dependent type relationships and the most complex way that the learner, in this case the teachers on the course, could display their thoughts and understandings was represented by composite type relations. Examples of some of the types of relationship and forms of expression used for the analysis of line-labelled concept maps are given in Table 3.11.

Each teacher line-labelled concept map was subsequently analysed using this classificatory system. Each relationship was scrutinised and classified; the number of relationships were then counted, and the total used to describe the intensity of the structure. The proportions of each type of relationship were then calculated and enabled comparisons to be made between the learning of each teacher at different times within the course. The results from this analysis are presented in Chapter 6.

### 3.9.5 Summary

The in-service degree programme was evaluated in order to investigate the processes and outcomes of teachers' professional learning. More specifically the study was designed to investigate whether teachers involved in this type of in-service provision could implement their knowledge and understanding of assessment (espoused theories) in practice (theories-in-use) in the classroom and how far new theories of action developed as a result of their in-service experience. The underlying assumption was that teachers' professional learning would be facilitated by studying assessment in theory and practice. The evaluation was a case study using triangulation methods. Various types of data were collected as evidence of the processes and outcomes of the course: for example, audiotape transcripts of the course and tutorials; line-labelled concept maps, school and course files; teacher self-reports and a diary. The results of this study are presented in chapter 6.

**Table 3.11 Examples of the types of relationship and forms of expression used in the analysis of line labelled concept structure maps.**

<i>Types of relationship</i>	<i>Forms of expression</i>
<b>DESCRIPTIVE:</b>	
Structural types express a taxonomic or hierarchical relationship between concepts such as subset, inclusion, coordinate, "kind" or "parts" relationship.	is a (kind/sort of) has/have includes concerns
Functional types of relationship express a function, purpose or use.	needed for, used by /for, for, gives, uses.
<b>DEPENDENT:</b>	
Procedural types express an order, sequence of steps, progression, pre-condition, process or pre-requisite relationship between concepts.	first, then, then becomes, makes.
Logical types express a logical or conditional relationship between concepts.	is sometimes, can be, could be.
<b>COMPOSITE AND ERRONEOUS:</b>	
See previous definitions	



### **3.10 SUMMARY OF RESEARCH DESIGN AND METHODS USED IN THE STUDY OF TEACHER-BASED ASSESSMENTS AND THE EVALUATION OF IN-SERVICE PROVISION**

In this chapter a number of questions relating to the development of teacher-based assessments have been raised and details of the research design, methods of data collection and analysis presented. The first question addressed was whether teacher-based assessments could provide a valid and reliable assessment of pupil performance in relation to the aims of primary education selected. It was proposed to investigate this by a number of means. The methods described included: first, the correlation of teachers' ratings with pupil performance on the structured activities; second, factor analysis and comparison of scores on the teacher-based assessments with performance on similar assessment procedures if available, and with the modified Richmond Test results in mathematics, language and reading and "tests" of study skills. Finally, the degree of inter- and intra-rater agreement on assessments of the same samples of pupil performance was calculated.

The question of how far certain pupil and teacher characteristics influenced the judgements made by teachers of pupil performance on the teacher-based assessments was also raised. It was proposed that factors that may contribute to such assessments be investigated by the use of covariance analysis. The variables which were considered included pupil sex, age, social class, achievement on tests of basic skills and classroom behaviour (pupil type). In addition, the question of the effectiveness of the assessments made by teachers with different styles was addressed. This was investigated using two factor analysis of covariance as described previously. It was also proposed to investigate the characteristics of pupil performance on each of the selected aims. In particular, the relationships between pupil performance in these areas to a number of pupil variables such as pupil age, sex, social class, achievement on basic skills tests, classroom behaviour and style of teaching experienced were suggested for study. Finally the question of whether such teacher-based assessments have relevance to the day-to-day classroom practice of teachers in general or not was addressed. This led to the development and evaluation of in-service provision where the professional learning of teachers has been investigated. The results of the study of the validity and reliability of the teacher-based assessments are presented in chapter 4, the results of the investigation of pupil performance on the selected aims in chapter 5 and the evaluation of in-service provision detailed in chapter 6.



## **CHAPTER 4: PRESENTATION AND DISCUSSION OF RESULTS: THE VALIDITY AND RELIABILITY OF TEACHER-BASED ASSESSMENTS**

The results of the study into the reliability and validity of the teacher-based assessments comprising teacher ratings of pupil performance on the checklists of abilities and associated structured activities related to the selected aims of primary education are now presented. These results have been considered in the following sections: validity of the checklists of abilities (section 4.1), validity of the structured activities (section 4.2), reliability (section 4.3) and validity of teachers' assessments (section 4.5).

### **4.1 VALIDITY OF CHECKLISTS OF ABILITIES**

#### **4.1.1 Logical review procedures**

As has been described in the previous chapter the process of developing the checklists of abilities was designed to facilitate their validity. As part of the logical review procedure the application group of teachers who took part in the Oracle study also completed an open ended questionnaire related to the validity of the checklists of abilities for each of the five aims. Details of the teachers' responses to the questionnaire are included in Appendix 17. The following analysis of the questionnaire responses was made with reference to the questions relating to the clarity and usefulness of the criteria on the checklist of abilities (Appendix 13) and formed the basis of the logical review of the checklist of abilities associated with each aim. Of the sample of 58 teachers in the first year of the study 50% responded to this questionnaire.

*Teacher feedback on the checklist of abilities associated with the aim "the child should be able to listen with concentration and understanding".*

Approximately 50% of the teachers who responded to the questionnaire considered the checklist was clear enough to assess class activities. For the remaining teachers the main difficulty arose from the procedure for indicating the level of performance of the pupils by indicating this as A-usually achieves this level of performance, B-sometimes achieves this level of performance, or C - rarely achieves this level of performance. This procedure was subsequently modified for the remaining teacher-based assessments to account for this criticism by introducing intermediate levels between A and B, B and C thus



producing a five-point scale. Specific criticisms which were made of the checklist included one comment that the abilities listed on the checklist were *"too broad to be used"*. However, the development group had considered it inappropriate to provide more detailed criteria since these had already been found difficult to operate in practice. Another teacher commented that *"no precise instructions were given about the nature of the assessment during the activities"* suggesting that this teacher had experienced some difficulty in using the checklist so that more detailed elaborations of each criterion might well have been appropriate. Another teacher commented that there *"was insufficient time to assess every child during the structured activity"*. This reflected the difficulties that the development group had faced in the trial of the checklist but given the practical considerations of using these activities was not unexpected.

*Teacher feedback on the checklist of abilities associated with the aim "the child should show some creativity and inventiveness in some fields".*

This aspect of the teacher-based assessments posed the most problems for the teachers. Concern was expressed that the criteria of originality and appropriateness would be interpreted differently by different teachers even though a number of teachers had reported that the criteria for the assessment of creativity were relevant and appropriate. One teacher clearly identified the problem in the following terms *"no researcher working in this field had yet given an adequate definition of creativity"*. This quite clearly reflected the uncertainty that the teachers felt in judging pupil performance in this area given that no consensus regarding the definition of creativity has been reached. Reservations were expressed concerning the accuracy of such assessments based on the structured activity since the teachers viewed a "one-off" assessment of this type with suspicion. The evidence from the teachers' comments suggested that this was a concern for the apparent summative nature of their use in the context of the Oracle research programme. The method of indicating a pupil's level of performance was also criticised as was noted in the comments relating to the previous aim.

The assessment of creativity and inventiveness was also difficult for a number of practical reasons. Those teachers who took several different classes during the course of the day and where teaching took place in an open plan area did not feel that they knew the pupils sufficiently well to make a valid assessment. However, since the original intention of specifying these criteria was to aid such formative assessment in the classroom, these comments indicated that the teachers were perhaps less familiar with their pupils than had been anticipated and thus found it more difficult to use these criteria than other checklists where



the abilities appeared at first sight more familiar.

It might also be suggested that in the case of creativity and inventiveness the teachers were able to voice their uncertainties regarding the assessment of pupils since there was as little consensus outside the classroom as to the meaning of creativity as between teachers themselves. There also appeared to be some confusion amongst the teachers regarding the role of the structured activity in aiding their assessment of pupils' abilities in this area. It appeared that several teachers thought that it was necessary to base their assessment on the performance of the pupil on that *one* structured activity alone rather than using it as a means of checking the performance of the pupil on this example against their existing knowledge of the pupil.

*Teacher feedback on the checklist of abilities associated with the aim "the child should be able to acquire information other than by reading".*

The general comments of the teachers indicated satisfaction with the criteria for assessing this aim. Reservations were expressed in relation to the clarity of criteria associated with the pupil's ability to extend an activity. Five teachers commented that the checklist was too complicated and involved some repetition of criteria, although several other teachers commented that the checklist helped them to focus their thoughts in making assessments in this area.

Teachers also noted that other pupil characteristics should have been included in the criteria for assessment: for example, expressive ability through oral speech and drama and acquiring information by logical deduction. Other pupil characteristics related to attitudinal variables such as persistence were also mentioned. Some of the additional criteria did not match the development group's perception of the aim as in the case of speech and drama. This did not, however, preclude the inclusion of additional criteria if appropriate as in the case of "*acquiring information by logical deduction*" but given the opinion of some of the application group of teachers that the checklist was already too long and complicated this was not attempted. Any attitudinal pupil characteristics such as persistence were originally excluded by the development group of teachers since the focus of these assessments was on cognitive dimensions of pupil performance, thus these had not been included in the checklist.

The importance of intensive observation of the pupils was also stressed by the application group of teachers, which suggested that they were more used to this type of



informal assessment and had internalised some of the expectations implicit in the teacher-based assessments from the previous term. Overall the comments from the report forms suggested a general satisfaction with this checklist as relevant and appropriate for the area despite some limitations as expressed previously.

*Teacher feedback on the checklist of abilities associated with the aim "the child should know how to compute in the four arithmetic rules and use mathematical techniques in his/her everyday life".*

The checklist was generally considered to be appropriate although it was suggested that it was rather too general to cover so many different aspects of mathematical competence particularly where a pupil's ability might vary with different activities. The criteria were found to be relevant and were assessable during the structured activities. The teachers appeared to be more at ease with assessment in the area of mathematics than with the assessment of the previous three aims. This was perhaps a function of the emphasis placed on the basic skills in the primary classroom and an apparent familiarity with assessment in this area. However, one teacher did not refer to the checklist at all, whilst another made the following detailed comment on the relationship between the checklist of abilities, structured activities and classwork assessment.

*"We were instructed to complete the checklist before giving any tests and to assess the children on classwork. As a lot of the material in the test bore little relationship to the classwork we are doing at the moment, it was rather unfair to relate the children's assessment with the performance in the structured activities. If asked to assess how well each child would complete the proposed tests I would have graded them all slightly lower."*

This comment was particularly constructive since it illustrated an essential confusion that some of the teachers exhibited in relation to the checklists of abilities. This teacher viewed the assessment of mathematics in terms of content not processes. Thus, she was concerned that the content of the structured activity was not related to the mathematical content which the pupils were doing in class. As a result the "test" was unfair because the content was unfamiliar. This teacher did not recognise that the criteria on the checklist of abilities were process-orientated and therefore applicable to a wide range of content. This particular problem of focusing on the content of the structured activity rather than the process-orientated criteria on the checklists of abilities would suggest a possible source of error in teachers' assessments.



*Teacher feedback on the checklist of abilities associated with the aim "the child should be able to convey meaning clearly and accurately through speech".*

All teachers reported that the criteria on the checklist of abilities were clear enough with the exception of one teacher who found the criteria related to the pupil's ability to link his/her own experiences to the current situation difficult to apply. As a result this teacher did not complete an assessment of this criterion. All teachers considered the criteria appropriate to the aim. Suggestions made for other abilities to be included tended to relate to variables such confidence which had been excluded from the checklist because of the emphasis on observable pupil behaviours and cognitive abilities. In addition, one teacher found even seven criteria difficult to keep in mind.

#### **4.1.2 Conclusions**

Despite the obvious limitations associated with introducing a relatively unfamiliar form of assessment to teachers with little experience of assessing in this way, there were few instances where teachers reported the criteria used as inappropriate to the aims under consideration. Those teachers who found particular problems in applying the checklists tended to leave the assessment blank. Feedback from the teachers did, however, suggest ways in which the checklists could be further improved. These were taken into account when possible. For example, the introduction of the five point scale resulted from the difficulties experienced by the teachers completing the first two checklists of abilities. In later assessments this did not arise as an issue for the teachers.

Other criticisms tended to relate to some misunderstanding regarding the assessment procedures and their use in the research. For example, the assessment of creativity and inventiveness was viewed as a one-off test rather than as an aid to assessment. Similarly the problem which arose when teachers focused on the content of the structured activity in order to assess the process-orientated criteria on the checklists of abilities related to a misunderstanding of the basis of the teachers' assessments. This was particularly noted in the assessment of everyday mathematics, with the conclusion that the checklists did not assess the same qualities as the structured activities. This suggested that some of the teachers still placed greater emphasis on the written product in the form of the structured activities than their own knowledge of the pupils and the process skills as assessed by the checklists of abilities. It would have been desirable to spend more time preparing the application groups in order to overcome these difficulties in implementation but this was beyond the resources of the research team. However, this finding suggested that there



was a general need for dealing with the difficulties of implementing such assessments by provision of appropriate support and in-service for teachers in this area. Thus the validity of the checklists of abilities was supported in a limited way by this logical review since the teachers considered the criteria representative and appropriate for the assessment of the aims, but problematic to implement.

## **4.2 VALIDITY OF STRUCTURED ACTIVITIES**

### **4.2.1 Logical review procedures**

Logical review procedures were used in the first instance to establish the validity of the structured activities. The process of development for the teacher-based assessments, as has been noted previously, was intended to maximise the appropriateness of the structured activities for the assessment of the four aims where structured activities were devised. The development group of teachers were involved in piloting the initial versions of the materials and suggestions made during discussions were incorporated into the materials that were sent to the participating schools and the application groups of teachers. This group of teachers were also asked to respond to a number of questions related to the use of the structured activities as a means of assessing pupil performance for the Oracle research programme and in facilitating their assessments of pupils on the checklists of abilities.

The validity of the structured activities was thus considered in relation to the two purposes for which they were intended: formative purposes, associated with facilitating teachers' assessments of pupils in the classroom and summative purposes associated with measuring pupil achievement and ensuring the comparability of teachers' assessments.

The structured activities were reviewed in relation to: first, their "*appropriateness*" for the age and ability range of the pupils who were assessed and second, with regard to the "*fit*" of the structured activities to the criteria on the checklist of abilities and their "*usefulness*" in facilitating the assessment of pupils on the criteria listed on the checklist. The responses made by teachers to the open ended questionnaire have been documented in Appendix 17. A discussion of these responses related to each of the aims under consideration is now presented.



*Teacher feedback on the structured activities associated with the aim "the child should be able to listen with concentration and understanding".*

The majority of teachers found the structured activities appropriate to both the age and ability levels in their classes. Inevitably with children aged from eight years old to eleven plus some parts of the structured activities were thought to be too easy or too difficult for some of the pupils. The musical instrument and following instructions sequences were generally considered appropriate, if a little easy for older pupils. However, the sound story sequence presented some difficulties for the pupils. This was not in relation to the actual activity itself but rather a function of the quality of reproduction of the sound story on the cassettes sent to the teachers. This was not a particularly general criticism and some teachers adopted strategies such as repeating or replaying the tape to ensure that the pupils were able to hear it clearly whilst completing this part of the structured activity. These structured activities which were designed to facilitate the assessments using the checklist of abilities were not as effective as had been expected. This was probably due to the fact that the teachers were not involved in the marking of the structured activities and therefore did not have available to them the information on pupil performance which would have facilitated their assessments relative to the checklist of abilities. There was also the practical difficulty of observing all of the pupils as they completed the exercises. This was perhaps to be expected since none of the teachers had received particular training in observational techniques or assessment and were therefore unfamiliar with the procedures involved. The teachers, however, considered the activities appropriate for the assessment of this aim.

*Teacher feedback on the structured activities associated with the aim "the child should be able to show some creativity and inventiveness in some fields".*

The materials for this structured activity were considered appropriate for the age and ability range of the pupils. The teachers considered the criteria appropriate, as was noted previously but concern was expressed about the use of only one exercise to make an assessment of the pupil's creativity and inventiveness. The teachers considered this might lead to a misrepresentation of the pupils' achievements when compared to their normal class performance. This was particularly true of this area since teachers were concerned that pupils, although not artistic, may well have shown evidence of creativity and inventiveness in other aspects of the primary curriculum such as drama and music. The content validity of this structured activity must therefore be considered limited since the sample of pupil behaviour that was assessed did not, in the opinion of the teachers,



adequately described the domain under investigation.

*Teacher feedback on the structured activities associated with aim "the child should be able to acquire information other than by reading".*

The majority of teachers considered the materials were appropriate to the age and ability range of the class. Reservations that were expressed centred on the part of the structured activity which required the pupils to write down questions they would like to ask the person in the story. This concern was related to the difficulties experienced by pupils with learning difficulties in completing this part of the structured activity despite the suggestion that the teachers could help with the mechanical aspects of completing the task by writing the pupils' responses themselves. However, the main concern of the teachers was the limited coverage of the criteria listed on the checklist of abilities by the structured activities. Those which were covered were thought appropriate but the teachers found it difficult to assess all these criteria in the time allocated. The suggestion was made that intensive observation was necessary if the pupils' abilities were to be accurately reflected, however, the teacher concerned recognised that if this had been requested by the researcher the amount of time required to carry out the teacher-based assessments would have been impractical.

*Teacher feedback on the structured activities associated with the aim "the child should know how to apply mathematical techniques in everyday situations and compute accurately in the four rules".*

The responses to this set of structured activities were the most extensive. The opinions of the teachers as to the appropriateness of the materials varied greatly depending on the age and ability of the pupils they taught. This was perhaps to be expected particularly in relation to the first set of sums which ranged from simple addition of two numbers such as 3 and 7 to sums which required the pupils to carry out long division operations. The remaining sections of the structured activity used unfamiliar types of activities to assess process and application skills. The "impossible sums" appeared to present the most difficulty for the pupils but as a result provided a clear distinction between those pupils who understood the necessity of appropriate information in the solution of such problems. The remaining parts of the exercise were, in general, considered by the teachers appropriate for the more able, reasonably well attempted by those in the middle range of ability but too difficult for the less able.



The relatively low use of the keys to identify the amount of help given to the pupils could be accounted for if the teachers did not consider this an appropriate strategy when assessing pupil performance. However, the responses to the questionnaire suggested that many teachers found the problems in section C of the structured activity took longer to do than anticipated. Thus they were unable to give the degree of individual attention that this strategy would have required. In addition, if this constraint was coupled with a view that the assessment of mathematics should be in the form of a standardised mathematics achievement test, then the teachers would not have used the key as intended as an aid to diagnostic assessment.

Most teachers found the criteria appropriate to the areas being assessed but did not comment on how well the structured activities matched the criteria on the checklist so no inference can be drawn as to the appropriateness of the structured activities in relation to the checklist of abilities. However, some comments made by the teachers would suggest that they did not see much relationship between the two components; for example, "*we don't know what is being assessed*", "*I think one uses one's overall assessment of the individual*". These comments would suggest that these teachers had not attempted to operate the teacher-based assessments as intended which would be reflected as unreliability in their assessments.

#### **4.2.2 Empirical review procedures: structured activities**

Empirical review procedures made use of the techniques of statistical analysis commonly employed to establish the construct and content validity of norm and criterion referenced standardised tests such as factor analysis and correlational procedures. These analyses were carried out on the achievement test scores of basic skills, tests of study skills and the three sets of structured activity scores obtained in the first year and subsequently on the everyday mathematics structured activities and the scores obtained for the modified Richmond tests in the second year of the study.

A preliminary analysis of the intercorrelations between the structured activity components from the first year of the study was carried out in order to investigate the degree of overlap between particular parts of these exercises. In this way the null hypothesis that the structured activity scores were not drawn from different populations of items testing different traits was investigated. Thus, if the intercorrelations between the different components of the structured activities were low then the null hypothesis could be rejected. The data obtained from this analysis is presented in Table 4.1 overleaf.



**Table 4.1** Inter-correlation of structured activity scores on the aims: the child should be able to: "listen with concentration and understanding", "show some inventiveness and creativity" and "acquire information other than by reading".

<i>Key:</i>		<i>Scores for each structured activity component.</i>							
<i>Creativity and inventiveness</i>		Originality Appropriateness						Orig App	
<i>Listening with concentration and understanding</i>		Identification of musical instruments Sequencing of sound story Following instructions						Mus Seq Foll	
<i>Acquiring information other than by reading</i>		Observation (spot the difference) Questioning- number and range Comprehension-looking and listening Open questions						Obs Ques Comp Open	

	<i>Creativity and inventiveness</i>		<i>Listening with concentration and understanding</i>			<i>Acquiring information other than by reading</i>			
<i>Structured activity scores</i>	<i>Orig</i>	<i>App</i>	<i>Mus</i>	<i>Seq</i>	<i>Foll</i>	<i>Obs</i>	<i>Ques</i>	<i>Comp</i>	<i>Open</i>
<i>Orig</i>	1.00								
<i>App</i>	0.29*	1.00							
<i>Mus</i>	-0.09	-0.002	1.00						
<i>Seq</i>	-0.16*	0.03	0.16*	1.00					
<i>Foll</i>	-0.04	0.08	0.33*	0.24*	1.00				
<i>Obs</i>	-0.03	0.08	0.17*	0.17*	0.16*	1.00			
<i>Ques</i>	-0.16*	-0.02	0.10	0.17*	0.13	0.21*	1.00		
<i>Comp</i>	-0.11*	0.05	0.17*	0.20*	0.17*	0.02	0.30*	1.00	
<i>Open</i>	-0.13	-0.02	0.13	0.21*	0.15*	0.05	0.51*	0.71*	1.00

N= 405

\* Significant at or beyond the 1% level

It should be noted that the "listening with concentration and understanding" and "creativity and inventiveness" activities were completed in the Autumn term of 1976 whilst the "acquiring information other than by reading" activities were completed in the Spring term of 1977. Although the majority of pupils (86% of the sample) completed all sets of structured activities, the three sets of data are not directly comparable since different pupils were absent during each of these assessments.

The data in Table 4.1 shows that positive intercorrelations occurred between components of the same structured activity. The originality and appropriateness scores were found to have a Pearson  $r$  correlation coefficient of 0.29. Similarly, significant correlations were obtained with the components of the "listening with concentration and understanding" activities (Pearson  $r$  correlation coefficients = 0.16, 0.33, 0.24) and for the "acquiring information other than by reading" components (Pearson  $r$  correlation coefficients = 0.21, 0.30, 0.51, 0.71).

However, there were also correlations found between components of different structured activities. The activities associated with the assessment of creativity and inventiveness showed significant negative correlations with all other components of the structured activities. It was inferred from this evidence that these activities were assessing a separate dimension of pupil performance. Sequencing on the listening activity was found to be positively correlated with all components of the acquiring information activities, whilst performance on the listening activities related to the musical instruments and following instructions were found to correlate positively with observation and comprehension skills associated with looking and listening which were components of the acquiring information structured activities. This suggested that the structured activities were assessing different aspects of pupil performance but that they tapped a number of underlying skills in different ways.

The next stage of the investigation was to test whether these underlying skills as measured by the structured activities were significantly different from those measured by the modified Richmond tests of basic skills. Thus a similar correlational analysis was carried out. The underlying assumption of this analysis was that if the structured activities could be demonstrated to assess significantly different aspects of pupil performance when compared with scores on the modified Richmond tests of basic skills then the null hypothesis that the items on the structured activities were assessing the same traits as measured by the Richmond tests of basic skills could be rejected. The results of this analysis are presented in Table 4.2.



**Table 4.2 Pearson r correlation coefficients of structured activity scores with average scores on the Richmond tests of basic skills: total, mathematics, language and reading.**

<i>Basic skills scores:</i>	<i>Total</i>	<i>Mathematics</i>	<i>Language</i>	<i>Reading</i>
<hr/>				
<i>Structured activity scores:</i>				
<i>Creativity and inventiveness</i>				
Originality	-0.09	-0.04	-0.11	-0.09
Appropriateness	0.09	0.08	0.06	0.08
N=	373	383	385	387
<i>Listening with concentration and understanding</i>				
Musical instruments	0.21*	0.17*	0.20*	0.23*
Sound story	0.32*	0.32*	0.27*	0.27*
Following instructions	0.17*	0.17*	0.15*	0.20*
N=	378	388	388	392
<i>Acquiring information other than by reading</i>				
Observation	0.27*	0.26*	0.25*	0.20*
Questions	0.41*	0.38*	0.37*	0.37*
Comprehension	0.22*	0.21*	0.19*	0.23*
Open questions	0.19*	0.24*	0.15*	0.17*
N=	371	383	384	386

\* Significant at or beyond the 1% level.

In Table 4.2 the structured activity scores for the first year of the study have been correlated with the total basic skills score derived from averaging pupil scores on the pre-test carried out in September 1976 and the post-test carried out in July 1977. In addition, the correlation between each structured activity score and component tests of the basic skills in mathematics, language and reading are also presented since it was suggested that certain structured activities might correlate more highly with language and reading tests than with the test of mathematics. The analysis of the data relating to this hypothesis is now discussed.

The data from this analysis showed that the structured activity scores for originality and appropriateness were not significantly correlated with performance on the basic skills and therefore it was inferred that they assessed a different aspect of pupil performance related to this aim. The remaining structured activities showed varying degrees of positive correlation with the basic skills which were all significant at or beyond the 1% level. Some measure of correlation between these activities was to be expected since they made use of the basic skills of reading and language. This was particularly noticeable with the high correlations obtained between the structured activity which required the pupils to write down the questions they would most like to ask the person in the story. This component of the structured activities clearly was related to the basic skills achievement of the pupils. By comparing the correlations found in Tables 4.1 and 4.2 it can be seen that, in general, the correlations of the structured activities with the basic skills were higher than those found with the unstructured activities. This feature of the data may well be accounted for by the lack of variation in the structured activity scores presenting a similar problem as found in attempts to calculate the coefficients for sub-tests of achievement tests.

Thus it was important to ascertain whether there was more commonality between the achievement of pupils on the tests of basic skills and the structured activities or between the various components of the structured activities. If the evidence supported the former proposition the inference could be drawn that the basic skills and structured activities were tapping the same underlying skills and abilities, whereas if the second proposition was supported, the validity of the structured activities as measures of different aspects of performance in relation to these extended aims of primary education could be inferred with some confidence.

In order to investigate these propositions the structured activity scores and the average scores for the pre-tests and post-tests of basic skills were factor analysed using the



Statistical Package for Social Sciences, Version 6.50 (Nie, *et al.*, 1975). Two analyses were executed. The first comprised a factor matrix using principal factor with iterations on the modified Richmond tests of basic skills scores and the structured activity scores collected in the first year of the study. The breakdown in the structured activity scores used here was different from that used in the previous analyses of the intercorrelations between structured activity scores and the correlation between structured activity scores and the basic skills scores. The sound story sequence score from the "listening with concentration and understanding" structured activity was broken down into its component parts since each score was based on the performance of a slightly different operation by the pupils. In the comprehension component of the structured activity associated with "acquiring information other than by reading" the elements related to comprehension by listening and comprehension by looking were considered separately since there was some evidence to suggest from the intercorrelations between the structured activities that the listening comprehension score would be more closely related to "listening with concentration and understanding" than "acquiring information other than by reading".

This factor analysis yielded five factors (Table 4.3). The principal factor was made up of the achievement test scores. The remaining factors related to different elements of the structured activities; formulation of questions, sequencing, comprehension by looking and listening, and creativity. This evidence supported the construct validity of the structured activities, in that the parts of the exercises designed to measure different abilities emerged as separate factors. It also supported the validity of the checklist itself as these activities were measuring different qualities from those measured by the achievement test scores.

The second factor analysis (Varimax rotated factor matrix with Kaiser normalisation, Statistical Package for Social Sciences, Version 6.50, Nie, *et al.*, 1975) was used for scores on the Richmond tests and structured activities collected in the first year of the study (Table 4.4). This factor analysis yielded similar results in that five factors emerged, the first of which comprised the basic skills scores. The remaining four factors again reflected differing components on the structured activities: number and range of questions on the "acquiring information other than by reading" structured activity; comprehension by listening and looking, which included aspects of both the "listening with concentration and understanding" and "acquiring information other than by reading" structured activities; sequencing and finally, originality and appropriateness.



**Table 4.3** Factor matrix using principal factor with iterations on modified Richmond tests of basic skills and structured activities scores collected in the first year of the study.

	<i>Factor 1</i>	<i>Factor 2</i>	<i>Factor 3</i>	<i>Factor 4</i>	<i>Factor 5</i>
<i>Basic skills:</i>					
Average total score	.95713*	-.21388	-.18401	-.07067	-.00640
Mathematics average score	.87894*	-.15554	-.11382	-.22127	.07411
Language average score	.85747*	-.21155	-.18678	-.00991	-.05706
Reading average score	.84563*	-.21126	-.19307	.04443	-.03924
<i>Structured activities</i>					
<i>Creativity</i>					
Originality	-.15423	-.11104	-.07254	-.01663	.44932*
Appropriateness	.07912	-.06615	.01992	.08348	.46293*
<i>Listening</i>					
Musical instruments	.31303	.02203	.10412	.26780	.00856
Sound story 1	.26261	-.02020	.39478*	-.19286	.01285
Sound story 2	.30493	.02155	.33003*	-.23528	-.01339
Sound story 3	.42456	-.03564	.32239*	-.31959	.01439
Following instructions	.29613	.04713	.23622	.14914	.12739
<i>Acquiring information</i>					
Spot the difference	.32158	-.15663	.18830	.17071	.04749
Comprehension - oral	.53843	-.00868	.15836	.31901*	-.08187
Comprehension - pictorial	.41244	.00523	.21640	.35695*	-.03962
No. of questions	.55633	.65665*	-.12571	.00263	.05677
No. of open questions	.13576	.25250	.02546	-.06547	-.03361
Range of questions	.46669	.71972*	-.05463	-.02677	.06020

\* Principal components loading on factors

	Eigen Value	% of variation	
Factor 1	4.80605	61.7	Basic skills test scores
Factor 2	1.21844	15.6	Number and range of questions on the acquiring information other than by reading structured activity
Factor 3	0.68679	8.8	Sequencing
Factor 4	0.61525	7.9	Comprehension by listening and looking;
Factor 5	0.46188	5.9	Originality and appropriateness



**Table 4.4 Varimax rotated factor matrix after rotation with Kaiser normalisation for scores on modified Richmond tests and structured activities collected in the first year of the study**

	<i>Factor 1</i>	<i>Factor 2</i>	<i>Factor 3</i>	<i>Factor 4</i>	<i>Factor 5</i>
<b><i>Basic skills:</i></b>					
Average total score	.91612*	.15613	.30326	.21238	.00059
Mathematics average score	.82594*	.19409	.17761	.32847	.06960
Language average score	.83051*	.11524	.30490	.14162	-.04623
Reading average score	.81356*	.10946	.33977	.10246	-.02654
<b><i>Structured activities</i></b>					
<b><i>Creativity</i></b>					
Originality	-.05443	-.10248	-.10880	-.06420	.46273*
Appropriateness	.04332	-.00145	.11233	.01674	.46640*
<b><i>Listening</i></b>					
Musical instruments	.13003	.09094	.39445*	.01793	.00342
Sound story 1	.05056	.03487	.13392	.49065*	-.01194
Sound story 2	.10850	.09746	.09076	.47611*	-.04315
Sound story 3	.23982	.09571	.07685	.56130*	-.01269
Following instructions	.06121	.11157	.34876*	.18461	.11082
<b><i>Acquiring information</i></b>					
Spot the difference	.17744	-.07070	.36348*	.15303	.05278
Comprehension - oral	.28091	.11908	.55976*	.09607	-.08880
Comprehension - pictorial	.14207	.08229	.55718*	.07960	-.04695
No. of questions	.24603	.81624*	.17709	.03806	-.01163
No. of open questions	.01283	.27804	.01189	.08176	-.06289
Range of questions	.12472	.83744*	.14174	.07686	-.01615

\* Principal components loading on factors

Factor 1 Basic skills test scores

Factor 2 Number and range of questions on the "acquiring information other than by reading" structured activity

Factor 3 Comprehension by listening and looking

Factor 4 Sequencing

Factor 5 Originality and appropriateness



It was concluded from these various analyses of the scores obtained on the structured activities and modified Richmond tests of basic skills that the component elements of the structured activities measured different underlying skills than those assessed through the use of standardised tests of the basic skills. Thus the null hypothesis that valid assessments of the extended aims of the primary school curriculum can not be developed was rejected. However, the nature of these underlying skills suggested by the factor analyses did not conform to the conceptualisation of the aims as generated by the teacher groups. Pupil performance on the different sets of structured activities showed that although high intercorrelations within each set of structured activities were found there were also correlations between aspects of each set of structured activities.

It was, therefore, appropriate at this stage to investigate the relationship, if any, with the tests of study skills developed within the Oracle research programme to ascertain whether there were more general traits related to independent study skills underlying the extended aims of primary education being investigated. The inter-correlations between the scores of the pupils on the various tests of study skills and the structured activities are summarised in Table 4.5. The tests of study skills have been previously described in chapter 2 (section 2.2.2). These activities involved the pupils in drawing a block graph from a picture of a number of different sea animals, a plan of the classroom using a number of mapping conventions and the construction of a clockface with moveable hands. Preliminary analysis of the study skills exercises (Willcocks and Jasman, 1980, p.104 -105) yielded a number of scores. For example, the block graph was scored on the pupil's understanding of the concept, layout, accuracy and presentation; the plan of the classroom was scored on the basis of the use of conventions and positioning of items within the plan and the clockface was given one score based on general impression since this exercise was judged to be represented by a single skill.

The structured activity scores used in this analysis were based on the four factors which were identified in the previous factor analysis. Thus scores based on the following components of the structured activities were included: sequencing, creativity (where the scores for originality and appropriateness were considered separately because of their negative correlation), comprehension and formulating questions. The correlations between structured activity scores and study skills exercises have been represented as positive or negative correlations for ease of reading. This data has been adapted from Figure 6.3 in *Progress and performance in the primary classroom*, (Willcocks and Jasman, 1980, p.120).



**Table 4.5**      **Inter-correlations between tests of study skills and structured activity components.<sup>1</sup>**

	<i>Block</i> <i>C L A P</i>	<i>Plan</i> <i>C P</i>	<i>Clock</i> <i>-face</i>	<i>Listening</i> <i>S</i>	<i>Creativity</i> <i>O A</i>	<i>Acquiring</i> <i>C Q</i>
<i>Block graph</i>						
Concept (C)						
Layout (L)						
Accuracy (A)	+					
Presentation (P)		+	+			
<i>Plan</i>						
Conventions (C)		+	+			
Positioning (P)			+	+		+
<i>Clock face</i>						+
<i>Listening</i>						
Sequencing(S)		+				
<i>Creativity</i>						
Originality (O)						
Appropriateness (A)					-	
<i>Acquiring information</i>						
Comprehension(C)						
Questions(Q)						+
+ = Positive correlation at or beyond the 1% level - = Negative correlation at or beyond the 1% level						

<sup>1</sup> Adapted from Figure 6.3, "The assessment of study skills" in Galton and Simon, [eds.], *Progress and performance in the primary classroom*, 1980, p.120.

From the data presented in Table 4.5 it can be seen that about half of the significant inter-correlations were between pairs of skills derived from single sets of tasks. However, not all the study skills within each task correlated highly with each other particularly in relation to the "tests" of study skills. The skill most highly correlated with the most skills from other tasks was the one which involved the correct positioning of items on the mapping task. The pupils who did well on this task generally performed well at model-making, sequencing, comprehension through looking and listening and asking relevant questions. Apart from these skills there was no general tendency for pupils to be successful or unsuccessful across all these tasks. This would suggest that together with the results of the factor analysis data reported previously, the structured activities reflected real component skills which were assessing substantially different aspects of pupil performance than those assessed by the modified Richmond tests of basic skills and the tests of study skills.

The data collected in the second year of the study related to the "everyday mathematics" structured activities. Similar empirical review procedures were employed to investigate the validity of these structured activities. It should also be noted that it was not possible to devise structured activities to match the criteria on the checklist of abilities associated with "conveying meaning clearly and accurately through speech", so there was no opportunity to apply similar empirical review procedures to this aim. The following section presents the results of the data relating to "everyday mathematics".

The first analysis conducted correlated the components of the structured activities with each other in order to ascertain if the different tasks were tapping different process skills as identified on the checklist of abilities. The data from this analysis is presented in Table 4.6. As can be seen from this table there were significant positive correlations between all the scores obtained for the various tasks comprising this structured activity. This result was similar to that found with the structured activities completed in the first year of the study where high inter-correlations were found between each set of tasks associated with a particular structured activity.

In the case of "everyday mathematics" it was not possible to correlate the scores on this structured activity with pupil performance on other structured activities or on the tests of study skills since this pupil sample was not assessed using these measures. Thus, it was not possible to ascertain the degree of overlap between the everyday mathematics structured activities and these other assessments of pupil performance.



**Table 4.6** Inter-correlations between component structured activity scores associated with the aim: the child should be able to apply mathematical techniques to everyday situations and compute in the four arithmetic rules.

<i>Structured activity component</i>	<i>Code</i>
Four rules sums	Rules
Insertion of correct sign	Insert
Impossible sums	Imp
Matching problem to sum	Match
Correct equation for problem	Corr
Bus timetable	BTT
Mapping task	Map
School timetable	STT
Classification	Class
Graphing exercise	Graph
Meter reading	Meter

	<i>Rules</i>	<i>Insert</i>	<i>Imp</i>	<i>Match</i>	<i>Corr</i>	<i>BTT</i>	<i>Map</i>	<i>STT</i>	<i>Class</i>	<i>Graph</i>	<i>Meter</i>
<i>Rules</i>	1.00										
<i>Insert</i>	0.50*	1.00									
<i>Imp</i>	0.47*	0.35*	1.00								
<i>Match</i>	0.44*	0.30*	0.52*	1.00							
<i>Corr</i>	0.52*	0.34*	0.60*	0.49*	1.00						
<i>BTT</i>	0.44*	0.28*	0.40*	0.34*	0.46*	1.00					
<i>Map</i>	0.32*	0.15*	0.20*	0.31*	0.25*	0.47*	1.00				
<i>STT</i>	0.40*	0.27*	0.36*	0.30*	0.39*	0.52*	0.38*	1.00			
<i>Class</i>	0.33*	0.21*	0.30*	0.14*	0.26*	0.39*	0.25*	0.29*	1.00		
<i>Graph</i>	0.14*	0.13*	0.20*	0.16*	0.22*	0.30*	0.36*	0.36*	0.31*	1.00	
<i>Meter</i>	0.21*	0.10*	0.23*	0.16*	0.19*	0.26*	0.25*	0.22*	0.33*	0.36*	1.00

\* Significant at or beyond the 1% level

However, a strong correlation between pupil performance on the structured activities associated with "everyday mathematics" and pupil achievement on the modified Richmond tests of basic skills, particularly that of the mathematics sub-test, was found. This was not a surprising result given that the way in which parts of the structured activity were devised by the teachers bore a strong relationship to types of items found in the Richmond tests of basic skills in that they dealt with arithmetic computation. The Pearson  $r$  correlation coefficients were, therefore, calculated for the component tasks of the "everyday mathematics" structured activities against the average total score on the modified Richmond tests and for each sub-test total. These correlation coefficients are presented in Table 4.7.

The inter-correlations were all positive and of varying degrees of magnitude. Not all were significant which suggested that in part the structured activities tapped a different skill or competence than some aspects of the language and reading sub-tests. However, significant correlations were found between all components of the structured activities and the average total score and mathematics sub-test score on the modified Richmond tests of basic skills. In order to investigate whether there was more than one aspect of mathematical ability being tested by both the standardised tests and the structured activities a factor analysis was carried out in a similar manner to that described for the data collected in the first year of the study.

The structured activity scores were factor analysed with the achievement test data and two factors emerged (Tables 4.8 and 4.9). The first consisted of items related to arithmetic competence, and the second consisted of those exercises dealing with the manipulation, organisation and presentation of data. Arithmetic competence was highly correlated with the modified Richmond test of mathematics whilst the manipulation and presentation of data factor was unique to the structured activities. This indicated that part of the structured activities tapped a different skill from that assessed by the Richmond tests and supported the construct validity of these exercises.

Comparisons were also made between pupil results on the structured activities and other tests which tapped related areas of the primary curriculum. This analysis was carried out in order to ascertain whether the structured activities duplicated existing tests or whether the structured activities assessed different aspects of pupil performance which had not been previously tapped. This data was used to further investigate the validity of the structured activities. Results of the correlational analyses carried out are now presented.



**Table 4.7** Inter-correlations between component structured activity scores for "the child should be able to apply mathematical techniques to everyday situations and compute in the four arithmetic rules" and average scores on the modified Richmond tests of basic skills.

<i>Structured Activity Component</i>	<i>Average scores obtained on the tests of basic skills.</i>			
	Total	Mathematics	Language	Reading
<i>Rules</i>	0.58*	0.46*	0.43*	0.38*
<i>Insert</i>	0.34*	0.34*	0.26*	0.22*
<i>Imp</i>	0.50*	0.52*	0.44*	0.30*
<i>Match</i>	0.43*	0.47*	0.35*	0.26*
<i>Corr</i>	0.63*	0.61*	0.48*	0.47*
<i>BTT</i>	0.46*	0.49*	0.39*	0.28*
<i>Map</i>	0.23*	0.21*	0.18*	0.20*
<i>STT</i>	0.38*	0.41*	0.32*	0.26*
<i>Class</i>	0.28*	0.20*	0.27*	0.21*
<i>Graph</i>	0.17*	0.19*	0.09	0.12
<i>Meter</i>	0.21*	0.18*	0.19*	0.13

\* Significant at or beyond the 1% level.

**Table 4.8** Factor matrix using principal factor with iterations on modified Richmond tests of basic skills and structured activities scores collected in the second year of the study.

<i>Structured activity components</i>	<i>Factor 1</i>	<i>Factor 2</i>
<i>Rules</i>	0.71202*	-0.16429
<i>Insert</i>	0.48976	-0.09296
<i>Imp</i>	0.64163*	-0.08144
<i>Match</i>	0.61115*	-0.08729
<i>Corr</i>	0.74691*	-0.19971
<i>BTT</i>	0.68926*	0.22346
<i>Map</i>	0.46681	0.34282*
<i>STT</i>	0.61968*	0.20827
<i>Class</i>	0.42379	0.31655*
<i>Graph</i>	0.35063	0.49810*
<i>Meter</i>	0.34070	0.30482*
<i>Basic skills components</i>		
<i>Mathematics</i>	0.77401*	-0.34838
<i>Language</i>	0.61235*	-0.17414
<i>Reading</i>	0.47906	-0.12986
	<i>Eigen value</i>	<i>% of variance</i>
<i>Factor 1</i>	4.78908	84.00%
<i>Factor 2</i>	0.91039	16.00%

\* Principal components loading on factors.

Factor 1 Arithmetic competence

Factor 2 Manipulation, organisation and presentation of data



**Table 4.9** Varimax rotated factor matrix after rotation with Kaiser normalisation for scores on modified Richmond tests and structured activities collected in the second year of the study.

<i>Structured activity components</i>	<i>Factor 1</i>	<i>Factor 2</i>
<i>Rules</i>	0.69177*	0.23543
<i>Insert</i>	0.46527	0.17898
<i>Imp</i>	0.58829*	0.26877
<i>Match</i>	0.56546*	0.24775
<i>Corr</i>	0.74007*	0.22370
<i>BTT</i>	0.46816	0.55303*
<i>Map</i>	0.21620	0.53730*
<i>STT</i>	0.41702	0.50346*
<i>Class</i>	0.19347	0.49231*
<i>Graph</i>	0.03565	0.60809*
<i>Meter</i>	0.12902	0.43857*
<i>Basic skills components</i>		
<i>Mathematics</i>	0.84143*	0.11161
<i>Language</i>	0.61224*	0.17455
<i>Reading</i>	0.47561*	0.14198

\* Principal components loading on factors.

Factor 1 Arithmetic competence  
 Factor 2 Manipulation, organisation and presentation of data

The Schools Council tests of listening skills (Wilkinson *et al.*, 1976) were administered to 74 pupils who also completed the structured activities for the aim "listening with concentration and understanding". A Pearson  $r$  correlation coefficient of 0.39 was obtained which was significant at or beyond the 1% level. Although the skills assessed by the School Council tests did not match the checklist criteria exactly nor were they specifically designed to tap the area as conceived by the teachers, this correlation was considered to provide further evidence of the construct validity of the structured activities in that they appeared to measuring similar listening skills to those assessed by the Schools Council materials.

Inter-correlations of the structured activities associated with the aim "the child should show some inventiveness and creativity" and other measures of creativity such as the Torrance picture completion exercise were also carried out. This analysis did not yield significant positive correlations. This was attributed to the differences in the criteria used to score these two types of exercise. The structured activities were scored on two criteria: originality and appropriateness, whilst the Torrance picture completion task was scored with respect to criteria such as fluency, which related to the number of responses the pupil completed and/or gave a title; flexibility which was calculated with reference to the different types of responses given and originality which was the sum of the weighted scores based on whether the response was symmetrical or asymmetrical. These criteria did not match those of the structured activities and therefore the lack of positive correlation was not considered to imply any lack of validity of the structured activity since this was designed to assess a different conceptualisation of creativity.

#### 4.2.3 Conclusions

The evidence presented here regarding the validity of the structured activities and the checklist of abilities was encouraging in that a number of study skills were identified and assessed by these activities which were related to the extended aims of education as described by the survey of teachers' opinions (Ashton *et al.*, 1975) investigated during this study.

The findings from both the logical and empirical reviews of the structured activities and the checklist of abilities would suggest that there were particular difficulties associated with the validation of such methods in practice. For example, the constraints imposed by the time frame of the study and the need to minimise the amount of "testing" so that the teachers were not overburdened within the Oracle research programme placed a



limitation on the amount of data which was collected. In addition, the evidence of validity of the structured activities was also limited by the lack of suitable alternative assessments to compare with the performance of pupils on the structured activities.

These findings do, however, support the idea that study skills can be discriminated, but most important, that they are linked to particular content rather than reflecting specific underlying study or process skills across the structured activities and tests of study skills. The modest but reasonable correlations found in the results from the comparison of structured activity scores with those of the Richmond tests of basic skills lend further support to this view. It is suggested that these correlations represent an underlying general academic ability rather than specific study or process skills such as observation, questioning and communication skills.

The results of this study of the validity of the structured activities are of particular interest in the context of debates on the basis for curriculum planning and progressive practice in primary schools where an emphasis on process skills has been advocated. If no such process skills can be identified independently from the content areas also being assessed then there would appear to be no justification to support a process-orientated curriculum which did not take account of the content areas used in the mediation of these skills. These findings also illuminate the debate on the transferability of particular skills located within a content domain in that the lack of any underlying skills being identified in this study would lend support to the view that the problem of transfer is still a real issue for any educator.

#### **4.3 RELIABILITY OF TEACHER-BASED ASSESSMENTS**

As noted in chapter 3 (section 3.6) there were particular problems associated with demonstrating the reliability of the teacher-based assessments, in particular the time frame of the research and the need to minimise the demands which were made on the teachers contributed to the limited amount of data available. It was also necessary to balance the competing demands for demonstrating the validity of the checklists of abilities, structured activities and teachers' assessments with the time which would be involved in completing the reliability study. On balance more weight was given to establishing validity since it was assumed that if one type of reliability check was completed for each area then it might reasonably be concluded that this would also apply to the other areas under study. A summary of the reliability checks carried out is thus presented in Table 4.10.

**Table 4.10 Summary of reliability checks carried out on the teacher-based assessments**

<i>Teacher-based assessment</i>	<i>Reliability check</i>
<i>Listening with concentration and understanding</i>	Intra-class agreement: 3 teachers each assessing the same pupils at the same time on all five criteria. No. of pupils = 97
<i>Creativity and inventiveness</i>	Inter-rater reliability on application of criteria in the assessment of the structured activities. No. of teachers = 4 No. of pupils = 50
<i>Acquiring information other than by reading</i>	Stability of assessments: 3 teachers assessing class on two occasions separated by four weeks. No. of pupils = 67 Test-retest reliability of the structured activities. No. of pupils = 60
<i>Everyday mathematics</i> <i>Conveying meaning</i>	None Comparability of teacher ratings in applying criteria of assessment. No. of teachers = 4 5 examples assessed on 7 criteria.



The reliability of the structured activities associated with the aim "acquiring information other than by reading" was established by calculating the test-retest correlation coefficient for this structured activity. The resulting correlation coefficient for this activity was 0.96. In this case, the inference could be drawn that the structured activity was a reliable measure in that pupil performance did not change significantly over the period of four weeks between structured activity administrations. It was not possible to repeat this type of reliability analysis for each of the structured activities within the time frame of this research and this must limit the confidence with which the results can be regarded. However, the structured activities were shown to be valid measures of the aims being assessed which, in the context of this research, was a more important consideration since reliable measures are not necessarily valid measures but valid measures may still provide valuable data within the limits imposed by the unreliability or errors of measurement which are represented in any assessment.

The reliability of the teachers' assessments was investigated in two ways in this study. First, different teachers assessed the same pupils or structured activity using the "*listening with concentration and understanding*", "*conveying meaning*" and "*creativity and inventiveness*" checklists to ascertain whether they could arrive at similar assessments of the same pupils. Interclass correlation coefficients for the "*listening*" checklist were 0.81, 0.87 and 0.83 respectively. For the four criteria on the "*conveying meaning*" checklists the interclass correlation coefficients were 0.93, 0.96, 0.95 and 0.89. For creativity the intraclass correlation coefficients for each criterion were 0.9 for originality and 0.77 for appropriateness. Second, teachers made independent assessments of their pupils using the "*acquiring information*" checklist on two occasions separated by a four week interval. The correlation coefficients obtained on the ten criteria assessed by the teachers ranged from 0.58 to 0.96 with most of the values lying between 0.70 and 0.88. These correlation coefficients were relatively high and indicated that teachers were able to use the checklists reliably over time.

These results indicated that the teachers involved in these small scale studies when given the checklist of abilities with little prior preparation were able to reach relatively high levels of agreement with each other and provide consistent assessments over time. It appeared that some of the difficulties normally associated with the use of such ratings had been overcome by the involvement of teachers' in the development process and also in the way in which the materials had been presented to the teachers concerned. The indication was that the checklists could be introduced to the application groups of teachers with the expectation that these teachers would also apply the checklists reliably.



Thus, it was concluded in section 4.2 that the structured activities were reasonably valid measures of the abilities they were designed to assess and in the small scale studies described in section 4.3 which investigated the reliability of teachers' assessments that teachers were able to reach agreement with each other and be consistent in their assessments.

However, the working hypothesis that pupil characteristics other than achievement contributed to the teachers' assessments of pupil performance on the five aims under investigation thus reducing their validity was also investigated. This study into the factors which might influence teachers' assessments of pupils considered teaching style and pupil age, sex, social class and pupil type (section 4.5). In order to aid the interpretation of this evidence details of the teaching styles and pupil types derived from the observational data collected during the Oracle research programme have been included in the next section.

#### **4.4 TEACHING STYLES AND PUPIL TYPES<sup>3</sup>**

The observational data collected from the first year of the study was used to derive descriptions of pupil and teacher behaviour which resulted in the identification of four pupil types; solitary workers, quiet collaborators, intermittent workers and attention seekers, and six teaching styles; individual monitors, class enquirers, group instructors, and a group of three styles categorised as style changers; infrequent changers, rotating changers and habitual changers.

This differentiation of teaching style, and also of pupil type, was carried out through the use of cluster analysis. An understanding of the characteristics of each teaching style and pupil type is necessary since the discussion of teachers' assessments and their relationship to these variables draws heavily on these descriptions.

---

<sup>3</sup> The following sections on teaching style and pupil type are taken from *Inside the Primary Classroom*, Galton and Simon (eds), 1980, p. 36 - 38 and p.143 - 147 respectively. The substantive content has not been altered although some change in style and presentation has been adopted to match more closely that of this thesis.



#### **4.4.1 Teaching styles**

##### ***Style 1: Individual monitors (22.4% of the sample).***

This group had the highest level of individualised one-to-one interaction with pupils and the lowest levels of interaction with pupils as members of a group or of the class as a whole. They engaged in very little group or class teaching, but concentrated on individuals.

Their teaching was largely didactic and was concerned with telling pupils what to do, while their questioning was mainly factual rather than probing or open-ended. With the high level of individualised work, marking with the pupil present was a significant feature of this group in that they spent a higher proportion of time on this than any other group.

##### ***Style 2: Class enquirers (15.5% of the sample)***

This group devoted the highest proportion of time to class teaching of any group (31%). The teachers combined this approach with individualised interaction although they spent a substantial proportion of time in this way (42.5%), this was less than any other style. Very little time was spent interacting with pupils in groups.

They had the highest proportion of questions when compared with other groups particularly those related to task work. Both open and closed questions occurred in the highest percentage for this group of teachers, as well the highest number of statements of ideas and problems. The emphasis on problem solving and ideas, together with teacher control through class teaching gave them their title.

##### ***Style 3: Group instructors (12.1% of the sample)***

This group maximised interaction with pupils as members of a group. Although this constituted less than 20% of all interactions it was still on average three times greater than the rest of the sample. Their teaching appeared largely didactic but with a high level of oral feedback to the pupils. The main emphasis was on the informational aspects of task work.

#### *Style 4: Style changers (50% of the sample)*

This large group of teachers had the second highest level of individual, group and class interaction; certain features of their teaching style were also associated with one or other of the remaining groups. Examination of the data and material from the descriptive accounts (S3) suggested that this group tended to adopt the characteristics of a specific style when they changed to its pattern of organization. The way in which this occurred differed and three main types of style changer were identified.

##### *Style 4a: Infrequent changers (10.3% of the sample)*

This group appeared to make a deliberate change in style during the observational year, although not always in the same direction or for the same reasons. This sub-group reached the highest overall level of questioning of all types, including higher order cognitive questions. They also had a relatively high proportion of statements of ideas as well as feedback levels. This group reached the highest proportion of interaction of all styles (86%).

##### *Style 4b: Rotating changers (15.5% of the sample)*

This group was identified by the organizational form of its teaching strategy. All except one used pupil grouping but in this case each group of pupils although working on the same curricular area did not work in the same curricular area as other groups in the class. Thus when the activity was due to be changed the pupils or the materials were rotated. This resulted in the whole class being involved in changing places and redistributing materials at given times during the day. There was, as a result, a large number of task supervision questions as well as control statements. It also appeared that this type of style led to control and management problems for the teachers.

##### *Style 4c: Habitual changers (24.2% of the sample)*

This group was comparatively large and was characterised by making regular changes between class teaching and individualised instruction. These seemed to be unplanned as the teachers responded thus when the pupils' behaviour caused them some difficulty. This group used questioning and statements of ideas relatively infrequently. Time spent on interaction with the pupils was the lowest of the whole sample (71.4%).



#### **4.4.2 Pupil types**

##### ***Type 1 pupils: attention seekers (19.5 % of the sample)***

These pupils either cooperated on task or routine work during 66.6% of the time they were observed. The main characteristic of this group was the way in which they interacted with adults. They sought attention or were the focus of the teacher's attention for more of the time that they were observed than any other pupil type. Most of these interactions were concerned with task or routine work. These pupils sustained average amounts of interaction with other pupils but tended to respond rather than initiate. This pupil type was also more likely to be found waiting for the teacher, out of the base area, moving around the classroom or leaving the room than any of the other groups.

##### ***Type 2 pupils: intermittent workers (35.7% of the sample)***

This group of pupils exhibited the lowest levels of interaction with the teacher. However, they had the highest levels of contact with other pupils. They engaged in task or routine work for 64.4% of the time, but significantly were distracted for some 20% of the observations made. They avoided contact with the teacher but were observed watching the teacher during interactions with other pupils.

##### ***Type 3 pupils: solitary workers (32.5% of the sample)***

This group of pupils received very little individual attention from the teacher, their contact appeared to be largely as part of the class or listening and watching the teacher interact with other pupils in the class. They showed little inclination to interact with other pupils. This type cooperated on task for the highest proportion of the observations (77.1%) remaining in their base for much of the time.

##### ***Type 4 pupils: quiet collaborators (12.3% of the sample)***

This pupil type was characterised by their interactions with the teachers, in that they tended to be part of a group or the class. These interactions were also usually about the task. This group did not have such high levels of verbal interaction but showed the highest proportion of observations sharing materials and interacting non-verbally. They had the highest levels of work interaction (72.6%), with relatively more routine activity than other types.



## 4.5 VALIDITY OF TEACHERS' ASSESSMENTS

The validity of the teachers' assessments has been investigated in four ways. First, the degree of discrimination between assessments on the checklists of abilities was studied by correlating the teacher ratings. Second, a factor analysis was conducted of all teacher assessments and structured activity scores on the data from the first year of the study in order to ascertain whether the assessments and structured activity scores for each aim loaded on the same factor or were represented in different factors. Third, it was important to see whether the teachers' assessments were more highly correlated with the structured activity scores or the scores on the tests of basic skills since this would give some indication of the aspects of pupil achievement which the teachers were taking into account when they made their assessments. Intercorrelations were, therefore, calculated for the teacher assessments with the structured activity scores and the basic skills achievement test scores. Once this was determined it was then possible to portion out the extent to which there was an overlap in the teachers' assessments of pupils in each activity and in the scores on the Richmond tests through analysis of variance.

The data presented in Table 4.11 shows the intercorrelations found between the teachers assessments on the aims "listening with concentration and understanding", "inventiveness and creativity" and "acquiring information other than by reading" which were collected in the first year of the study. It can be seen from this table that, although these assessments were made at different times, they are all significantly correlated with each other. The correlations found with criteria linked to the same aim were generally higher than those found between different aims.

These results would suggest that the teachers were making a global assessment of some underlying skill or ability rather than discriminating between the criteria and the aims. Such high intercorrelations were not found between the structured activity scores where separate content dependent study skills were identified. This would suggest that if the structured activities were valid measures of pupil performance then the teachers' assessments were invalid. This was further supported by the evidence obtained from a factor analysis of all teacher assessments and the structured activity scores from the first year of the study which are presented in Table 4.12. The teachers assessments were found to load heavily on the first factor, whilst the structured activity scores loaded on the remaining factors. In the second year of the study similarly high intercorrelations were found between the teachers' assessments on "everyday mathematics" and "conveying meaning clearly and accurately through speech" as can be seen in Table 4.13.



**Table 4.11     Intercorrelations between teachers' assessments on all criteria on the checklists of abilities for the first year of the study.\***

<i>Key:</i>	
<i>Listening with concentration and understanding*</i>	
Recalling	REC
Sequencing	SEQ
Following instructions	FOLL
<i>Inventiveness and creativity*</i>	
Originality	OR
Appropriateness	APP
<i>Acquiring information other than by reading*</i>	
Observation	OBS
Comprehension	COM
Questioning - Number	QN
- Range	QR

	OR	APP	REC	SEQ	FOLL	OBS	COM	QN	QR
OR	1.00								
APP	0.47	1.00							
REC	0.48	0.56	1.00						
SEQ	0.48	0.52	0.76	1.00					
FOLL	0.48	0.49	0.70	0.76	1.00				
OBS	0.43	0.40	0.60	0.51	0.52	1.00			
COM	0.41	0.33	0.55	0.51	0.44	0.79	1.00		
QN	0.48	0.42	0.56	0.61	0.56	0.63	0.65	1.00	
QR	0.43	0.32	0.53	0.59	0.50	0.61	0.66	0.77	1.00

\* All significant at or beyond the 1% level.

**Table 4.12** Factor analysis using principal factor with iterations for the teachers' assessments and structured activity scores from the first year of the study.

<i>Key:</i>	
TA ORIG	Teacher assessment - originality
SC ORIG	Structured activity score - originality
TA APP	Teacher assessment - appropriateness
SC APP	Structured activity score - appropriateness
TA REC	Teacher assessment - recall
SC MUS	Structured activity score - musical instruments for recall
TA SEQ	Teacher assessment - sequencing
SC SS	Structured activity score - sound story for sequencing
TA FOLL	Teacher assessment - following a set of instructions
SC INST	Structured activity score - following a set of instructions
TA OBS	Teacher assessment - observation
SC DIFF	Structured activity score - spot the difference for observation
TA QUES	Teacher assessment - asking questions
SC QUES	Structured activity score - asking questions
TA COM	Teacher assessment - comprehension
SC COM	Structured activity score - comprehension

	Factor 1	Factor 2	Factor 3	Factor 4
TA ORIG	.59028*	-.09649	.04972	-.03402
SC ORIG	-.15250	-.03470	.21644	.06406
TA APP	.54695*	-.04720	.10494	-.18779
SC APP	.10502	.12446	.17540	.12472
TA REC	.78807*	-.12648	.03776	-.29793
SC MUS	.31449	.19226	.30695*	.16940
TA SEQ	.81736*	-.04186	.06972	-.33023
SC SS	.27922	.30588	.13124	-.03376
TA FOLL	.75607*	.01521	.07379	-.26343
SC INST	.27552	.34095	.40238*	.12180
TA OBS	.73825*	-.26872	-.10208	.18238
SC DIFF	.30620	-.02577	.21818	.10798
TA QUES	.75138*	-.30299	-.21319	.24173
SC QUES	.43118	.69867*	-.28020	.01256
TA COM	.80787*	-.15828	-.03368	.17255
SC COM	.42215	.2289	.12602	.20006*

Factor		Eigen Value	% Variation
1	Teachers' assessments	5.73536	67.8
2	Questioning	1.47194	17.4
3	Following instructions	0.66568	7.9
4	Comprehension	0.58965	7.0



**Table 4.13**      **Intercorrelations between teachers' assessments on all criteria on the checklists of abilities for the second year of the study\***

<b>Key:</b> <i>Everyday mathematics*</i>					<b>The child should be able to demonstrate:</b> 1. computation 2. finding additional information 3. problem solving 4. calculating the correct answer						
<i>Conveying meaning clearly and accurately through speech*</i>					5. vocalise audibly and articulately, 6. use language and vocabulary appropriate to the situation or context, 7. talk fluently without undue pausing, hesitation or repetition, 8. maintain logical connections when describing objects, relating events and during discussions, 9. respond with reasonable speed and accuracy to questions and other contributions, 10. modify his/her own responses, 11. link his/her own experiences to the present situation.						
	1	2	3	4	5	6	7	8	9	10	11
1	1.00										
2	0.77	1.00									
3	0.79	0.78	1.00								
4	0.76	0.75	0.85	1.00							
5	0.44	0.35	0.41	0.42	1.00						
6	0.55	0.44	0.52	0.54	0.72	1.00					
7	0.42	0.27	0.40	0.40	0.71	0.73	1.00				
8	0.56	0.41	0.49	0.53	0.69	0.80	0.73	1.00			
9	0.50	0.39	0.52	0.50	0.74	0.76	0.73	0.77	1.00		
10	0.53	0.38	0.50	0.50	0.67	0.80	0.68	0.82	0.81	1.00	
11	0.48	0.35	0.46	0.42	0.63	0.74	0.66	0.76	0.77	0.83	1.00

\* All Pearson r correlation coefficients were found to be statistically significant at or beyond the 1% level.

Correlation coefficients were then calculated for teachers' assessments on each of the factors generated in the analysis of pupil performance on the structured activities in order to determine whether the teachers' assessments were related to performance on the structured activities or the basic skills. The basic skills scores used in this analysis was the average of pre- and post-test scores for the year of the study in question. The results are given in Table 4.14 and were statistically significant in all cases except the two elements of the creativity exercise. In this case the teachers' assessments were clearly invalid since they were negatively correlated with performance on the structured activities and positively correlated with the achievement test score for the basic skills.

Teachers' assessments were also correlated more highly with the pupils' scores on the basic skills than with their scores on the specific structured activity associated with that assessment, except in the case of arithmetic competence. From this it can be inferred that the teachers were more likely to assess some general academic ability which was reflected in the basic skills achievement test scores rather than according to the criteria from the checklists of abilities.

Furthermore, the teachers' assessment appeared to be a global assessment which did not discriminate between different criteria. The results from the first and second years of the study for the teachers' assessments were, therefore, aggregated in all subsequent analyses where teachers' assessments were compared with performance on the structured activities and basic skills scores as well as pupil characteristics such as age, sex, social class and type which are described in the following section.

#### **4.6 FACTORS INFLUENCING THE VALIDITY OF TEACHERS' ASSESSMENTS**

If it is accepted that the structured activities were valid measures of the aims of primary education under consideration and, as found in the previous section, teachers' assessments were intercorrelated and shown to be more highly correlated with basic skills scores than those obtained on the structured activities then teachers' assessment can be considered invalid. The following data relates to those factors which may have contributed to bias teachers' judgements of their pupils. In the review of the literature a number of pupil and teacher characteristics were suggested which might have accounted for some of this bias which it was hypothesised influenced the way teachers made their assessments of pupils. These are now considered in more detail.



**Table 4.14** Correlation coefficients for teacher assessments with total achievement test score and structured activity score.

<i>Teacher assessments</i>	<i>Achievement test score</i>	<i>Structured activity score</i>
Originality	0.30*	- 0.07
Appropriateness	0.30*	0.14
Sequencing	0.40*	0.26*
Following instructions	0.40*	0.21*
Formulating questions	0.32*	0.17*
Comprehension	0.46*	0.17*
Arithmetical competence <sup>2</sup>	0.44*	0.50*
Organisation and presentation of data <sup>2</sup>	0.41*	0.36*

\* Statistically significant at or beyond the 1% level.

<sup>2</sup> Data relating to everyday mathematics was collected in the second year of the study and represented a different group of teachers from those involved in the first year of the study, therefore, these correlation coefficients were not directly comparable.

#### **4.6.1 Results and discussion of data collected in the first year of the study (September 1976 - July 1977)**

The following presentation and discussion of results has drawn on the evidence collected during the first year of the study. The teacher-based assessments for the aims related to *"listening with concentration and understanding"*, *"inventiveness and creativity"* and *"acquiring information other than by reading"* were investigated to ascertain the influence of pupil age, sex, social class and pupil type as well as teaching style on teachers' assessments. It was necessary to take pupil achievement into account before the significance of these other factors could be ascertained since the pupils' performance on the basic skills and structured activities correlated significantly with teachers' assessments (Table 4.14). Covariance analysis (section 3.8.3) was used to control for those factors which could not be held constant under these non-experimental conditions. In this way, comparisons were possible between the assessments made by teachers when pupils were grouped by their age, sex, social class and type, even though initial performance on basic skills or structured activities was likely to be different for each group of pupils. In this covariance analysis the assumption was made that there was a specific set of skills associated with each of the teacher-based assessments. Thus the scores obtained from the structured activities could be adjusted to look at the specific skills associated with each aim by eliminating the general component assessed by the Richmond tests through the covariance analysis of the scores on the Richmond tests of basic skills and structured activities. Although each ability on the checklist had been assessed separately by the teachers the high intercorrelations between these ratings (Table 4.11 and 4.13) suggested that teachers did not discriminate between them and thus the teachers' ratings were aggregated for this analysis.

An analysis of covariance was carried out on the assessments made for the three aims assessed in the first year of the study. Mean teacher assessment scores for pupils of different age, sex, social class and type were adjusted to take account of the different levels of achievement on basic skills and the structured activity scores for the four factors previously generated (Tables 4.3 and 4.4). The resulting differences were expressed as residual assessments (Table 4.15). As can be seen from the first column of this table, the raw teacher assessment scores were found to differ significantly for pupil age, social class and type, but not for pupil sex. These differences did not necessarily reflect real differences between the performance of these groups of pupils. However, when pupil achievement on the tests of basic skills and performance on the structured activities were taken into account there was still a significant difference in the residual assessments.



**Table 4.15** Analysis of covariance of teacher assessments by sex, age, social class and type of pupil with basic skills and structured activity scores as covariates for the data collected during the first year of the study.

<i>Pupil characteristics</i>		<i>Raw teacher assessment</i>	<i>Residual assessments</i>		
			<i>Basic skills as covariate</i>	<i>Structured activities as covariate</i>	<i>n</i>
<i>Sex</i>	Boys	39.07	0.39	0.71	155
	Girls	38.35	-0.37	-0.68	165
<i>Age*</i>	9+	40.33	7.29	2.33	140
	10+	39.78	-1.22	0.73	87
	11+	35.40	-9.66	-4.20	93
<i>Social class*</i>					
Non-manual		44.45	4.33	5.30	42
Manual		35.00	-2.33	-2.90	79
<i>Pupil type*</i>					
Attention seekers		37.58	1.36	-0.39	63
Intermittent workers		36.03	-2.92	-2.49	123
Solitary workers		40.75	0.65	1.53	89
Quiet collaborators		45.00	4.76	4.31	45

\* Significant at or beyond the 1% level

It can be inferred from this data that teachers were over- and under-estimating pupils relative to their actual performance. From this analysis it can be seen that younger pupils tended to be over-rated whilst older pupils were under-rated relative to their actual performance. One interpretation of this result can be made in relation to the way these assessments were made. Each teacher assessed a particular group of pupils who were generally of the same age. The achievement tests and structured activity scores were, however, scored for the whole age range. The abilities of the younger pupils may have been over-estimated by the teachers, if they compared individual pupil performances with that of their peers rather than the whole age range where the scores of the younger pupils would have fallen in the lower range overall thus inflating the teachers' assessments. Similarly, older pupils were under-estimated since the teachers rated performance without reference to the lower scores of younger pupils on the structured activities and achievement tests of basic skills. Alternatively, it may be argued that the teachers of older pupils under-rated them because they had some expectation of what these pupils should be able to achieve and concluded that they had not met this expectation, whilst the teachers of younger pupils over-rated their performance because their expectations of the pupils' abilities were inflated in relation to their actual performance.

Teachers over-estimated the performances of pupils whose parents were in non-manual occupations, and under-estimated pupils whose parents were in manual occupations. Previous work has suggested that middle-class pupils tended to do better than expected from their IQ scores, while working-class pupils tended to do worse than expected. The teachers also reflected this trend in their assessments. This result can be interpreted in the light of the suggestion that middle-class pupils more closely match the stereotype of a *"good pupil"* and that teachers often equated their social competence with success on tests and, therefore, over-estimated their cognitive abilities. At the same time they tended to under-estimate those pupils from working-class backgrounds for similar reasons.

The influence of pupil behaviour on teachers' assessments was investigated by considering the differences between the various pupil types described in section 4.4.2. Details of the overall assessment given by teachers for each pupil type are given in Table 4.15. As in the previous analysis the teacher assessments on each individual structured activity were highly intercorrelated and have been aggregated. These results showed that the quiet collaborators received the highest ratings overall followed by solitary workers, attention-seekers and intermittent workers in that order. Where basic skills were the covariate the aspect of pupil performance which was represented by the residual



assessment was that due to general academic ability as measured by the Richmond tests. Thus, when actual achievement in the basic skills was taken into account, attention seekers came second to the quiet collaborators; and the abilities of these two groups were over-estimated as were those of the solitary workers. Intermittent workers, however, were under-estimated with respect to their performance. Here teachers appeared to make their judgements of pupil performance on the basis of the amount of interaction with pupils. Thus, pupils who have a high degree of interaction with the teacher either individually or in groups were over-estimated.

Quiet collaborators who had the highest levels of interaction with the teacher either individually or in a group were over-estimated to the greatest degree. Attention seekers were second in terms of contact time and similarly their abilities were equivalently over-estimated. Solitary workers did not have as much individual or group contact with the teacher and were only slightly over-estimated, whilst the intermittent workers who avoided contact with the teacher were under-estimated in relation to their performance on the basic skills. Thus, it appeared that contact with the teacher was equated in the teacher's mind with pupil success and resulted with the over-estimation of pupil ability relative to their measured achievement in the basic skills.

This result is of particular significance given the importance of teachers being able to accurately assess pupils in order to match learning activities with individual pupil needs. Clearly if the degree of contact is increased then the teachers are more likely to over-estimate the abilities of these pupils in relation to their measured performance. The question of whether such over-estimation contributes through teacher expectations and the provision of more demanding materials to improved pupil learning or not remains unanswered. However, this evidence suggested that if the teacher has more contact with a pupil it did not necessarily mean that their assessment became more accurate. This opens up the question of how teacher professional learning may be facilitated so that their assessments become more valid and reliable. This evidence would suggest that recommendations for increased contact and thus opportunity for assessment do not necessarily improve the validity of teachers' assessments. Perhaps a more appropriate focus is the development of skills and knowledge regarding assessment so that teachers are able to maximise the available contact time with pupils to collect valid and reliable evidence of pupil performance.

When performance on the structured activities was the covariate, residual assessments of the various pupil types differed from those obtained when basic skills were used. In



this case by controlling for the effect of the structured activities the element which was represented in the teachers over and under-estimation of pupil performance related to the general study factor proposed in section 4.3. In this analysis, attention-seekers and intermittent workers were under-estimated. Given that the structured activities were devised to relate to aspects of the extended aims of the primary curriculum, particularly sustaining independent study and enquiry, the teachers appeared to have assumed that these pupil types were ill equipped to pursue their work independently in such areas. Both the quiet collaborators and the solitary workers were over-estimated and it would appear that their behaviour, which was characterised by quiet concentration on the task in hand, was perceived by the teachers as closely related to success in areas other than the basic skills and related to aspects of independent study and enquiry.

Although teachers' assessments were related to various pupil characteristics, it was also hypothesised that certain features of particular teaching styles might also affect the bias of the teachers' assessments. However, it was demonstrated in the Oracle research programme that teaching style and pupil type were related, in that particular teaching styles were associated with particular pupil types, such as the group instructors and quiet collaborators (Galton *et al.*, 1980, p.150). An appropriate analysis to control for this relationship between teaching style and pupil type is a two-factor analysis of covariance which was used to ascertain the joint effects of teaching styles and pupil type on teachers' assessments with the covariates of basic skills and structured activities being controlled for in the analysis. As before the two covariates were pupil performance on the basic skills and structured activities and the joint effects of pupil types and teaching styles on teachers' assessments were considered. No significant interactions were found between the variables so that the analysis of the assessments of different pupil types and teaching styles can be treated with confidence.

The results are summarised in Table 4.16. This table showed that with the basic skills scores as covariate, the significance of the pupil type effect was reduced and the degree of over- and under-estimation was altered when teaching style was taken into account. With the structured activity scores as covariate the differences between assessments of the pupil types were no longer significant. Teaching style thus reduced the significance of the previous results in Table 4.15. When basic skills were used as covariate, attention-seekers and quiet collaborators were over-estimated but solitary workers and intermittent workers were under-estimated in relation to their performance.



**Table 4.16 Two-way analysis of covariance of teacher assessments by pupil type and teaching style related to the data of the first year of the study.**

<i>Pupil type (adjusted for teaching style)</i>	<i>Residual assessment (basic skills scores as covariate)</i>	<i>Residual assessment (structured activity scores as covariate)</i>
Attention seekers	3.03	0.34
Intermittent workers	-1.45	-1.82
Solitary workers	-0.79	0.61
Quiet collaborators	1.41	3.27
<i>Significance</i>	*	N.S.

<i>Teaching style (adjusted for pupil type)</i>	<i>Residual assessment (basic skills scores as covariate)</i>	<i>Residual assessment (structured activity scores as covariate)</i>
Individual monitors	-7.66	-1.71
Class enquirers	7.60	4.28
Group instructors	7.64	2.64
Infrequent changers	3.74	2.18
Rotating changers	-0.51	-2.18
Habitual changers	-4.20	-2.68
<i>Significance</i>	**	**

\* Significant at the 5% level  
\*\* Significant at and beyond the 1% level  
N.S. Not significant

This difference in teachers' assessments of different pupil types has previously been interpreted with reference to the amount of teacher interaction with individuals and groups. Removing the influence of teaching style effectively accounted for the teacher-pupil interaction in groups so that the range and order of teacher assessments of the pupil types in Table 4.15 appeared to match the quantity of teacher interactions with individual pupils. Attention-seekers and quiet collaborators interacted with the teachers as individuals on more occasions than the solitary workers and intermittent workers and were over-estimated in relation to their performance on the basic skills tests. In some way, these teachers were equating successful pupil performance with the amount of interaction they had with pupils as individuals.

In the second part of the analysis teachers were grouped by style and pupil type controlled for, instead of grouping by pupil type and controlling for teaching style. From this analysis, two groups of teaching styles emerged, those teachers who over-estimated pupil performance: class enquirers, group instructors and infrequent changers, and those who under-estimated pupil performance: rotating changers, habitual changers and individual monitors. The teachers in the first group have been shown elsewhere (Galton and Simon (eds), 1980) to place greater emphasis on basic skills and were more successful in terms of pupil progress in these areas. It was also suggested that this led to a devaluation of the skills associated with the extended aims of primary curriculum and independent enquiry which resulted in an expectation that pupils would acquire them without specific teaching.

The second group of teachers placed more emphasis on individualisation and independent study and enquiry, but under-estimated pupil performance in these areas. It was also demonstrated that this latter group had difficulty in effectively organising their classes and were not the most successful styles in terms of their pupils' performances on structured activities or on basic skills. Nevertheless, they used more independent work and were more aware of the problems involved, so that their assessments of their pupils appeared generally to be rather harsh.

#### **4.6.2 Results and discussion of data collected in the second year of the study (September 1977 - July 1978)**

In the second year of the study, the aims assessed by teachers were concerned with *"everyday mathematics"* and *"conveying meaning clearly and accurately through speech"*. The composition of the sample of teachers in this case was different from that in the first



year and the results, therefore, cannot be directly compared. In addition, the materials associated with "*conveying meaning clearly and accurately through speech* " did not include structured activities as was the case with all the other aims described so that a similar type of analysis was not possible. However, in the case of "*everyday mathematics*" pupils completed a set of structured activities related to this aim as well as the Richmond test of mathematics. The significant intercorrelations (Table 4.13) found between teacher assessments on all the criteria for both "*everyday mathematics*" and "*conveying meaning*" suggested that the teachers were making a global assessment of some aspect of pupil performance and these assessments were again aggregated. The correlations between the teachers' assessments and the pupils' scores on the modified Richmond tests and structured activities (0.44 and 0.40 respectively, Table 4.14) suggested that the teachers' assessments might be more accurate in this case and less subject to the influence of pupil characteristics.

An analysis of covariance was thus carried out to investigate the differences between teachers' assessments of pupils of different age, sex, social class and type in relation to "*everyday mathematics*". Details of this analysis are given in Table 4.17. Although significant differences were observed between the raw teacher assessments, when basic skills scores and structured activity scores were taken into account these differences disappeared in all cases except pupil age. It can again be suggested that this over- and under-estimation of younger and older pupils respectively was a function of the same circumstances that applied to these assessments in the first year of the study. Teachers made their judgements in relation to the performance of the pupils' peers rather than in relation to the whole age group. No other significant differences were obtained from this analysis so it may be inferred that this group of teachers besides being more accurate in their assessments were not taking such pupil variables into account in the same way. This difference could be accounted for if the teachers were able to assess pupil performance in "*everyday mathematics*" more accurately because achievement was more readily quantified in terms of the right answers for this curriculum area and the teachers thus made more valid assessments because of the similarity between the structured activities and achievement tests of basic skills.

An analysis of covariance was again used to investigate whether there were significant differences between assessments when teachers were grouped by teaching style and for different pupil types. Basic skills scores and the structured activity scores were used as covariates. The variations in teaching styles and pupil types were taken into account, and the results are summarised in Table 4.18.

**Table 4.17** Analysis of covariance of teacher assessments by sex, age, social class and type of pupil with basic skills and structured activity scores as covariate related to the second year of the study.

<i>Pupil characteristics</i>		<i>Raw teacher assessment</i>	<i>Residual assessments</i>		
			<i>Basic skills as covariate</i>	<i>Structured activities as covariate</i>	<i>N</i>
<i>Sex</i>	Boys	13.47	0.39	0.71	126
	Girls	13.12	-0.37	-0.68	129
<i>Significance</i>		<i>N.S.</i>	<i>N.S.</i>	<i>N.S.</i>	
<i>Age</i>	10+	12.65	0.97	0.53	102
	11+	13.25	-0.37	-0.60	69
	12+	14.06	-0.76	-0.15	85
	<i>Significance</i>	*	**	<i>N.S.</i>	
<i>Social class</i>					
Non-manual		11.88	1.05	-0.03	49
Manual		14.10	-0.65	0.01	76
<i>Significance</i>		*	<i>N.S.</i>	<i>N.S.</i>	
<i>Pupil type</i>					
Attention seekers		14.25	0.49	0.38	62
Intermittent workers		12.74	-0.38	-0.41	40
Solitary workers		13.68	0.54	0.40	69
Quiet collaborators		12.85	-0.82	-0.42	91
<i>Significance</i>		<i>N.S.</i>	<i>N.S.</i>	<i>N.S.</i>	

\* Significant at the 5% level  
 \*\* Significant at and beyond the 1% level  
 N.S. Not significant



**Table 4.18 Two-way analysis of covariance of teacher assessments by pupil type and teaching style for data related to the second year of the study.**

<i>Pupil type (adjusted for teaching style)</i>	<i>Residual assessment (mathematics basic skills score as covariate)</i>	<i>Residual assessment (structured activity score as covariate)</i>
Attention seekers	0.27	0.22
Intermittent workers	-0.52	-0.51
Solitary workers	0.48	0.37
Quiet collaborators	-0.52	-0.23
Significance	N.S.	N.S.

<i>Teaching style (adjusted for pupil type)</i>	<i>Residual assessment (basic skills scores as covariate)</i>	<i>Residual assessment (structured activity scores as covariate)</i>
Individual monitors	1.58	1.90
Class enquirers	1.59	-0.13
Group instructors	-0.96	-0.89
Infrequent changers	-0.28	0.51
Other changers	0.21	0.36
Significance	**	**

\* Significant at the 5% level  
\*\* Significant at and beyond the 1% level  
N.S. Not significant

In the case of this sample of teachers the habitual changers and rotating changers were not differentiated since the size of the clusters were relatively small compared to the previous year and data on classroom organisation was not available to confirm the distinctions between these two groups. The results obtained for different pupil types adjusted for teaching style showed no significant differences suggesting that these teachers made more accurate assessments and were not influenced by different pupil behaviours. Two groups of teaching styles emerged in the analysis of co-variance when adjusted for pupil type. The teachers who over-estimated pupil performance in relation to the structured activities included the individual monitors, infrequent and other changers, and those who under-estimated pupil performance included class enquirers and group instructors. This supported the view that those teachers who placed greater emphasis on particular skills, whether in mathematics or in independent work, tended to under-estimate their pupils' performance.

No structured activity data was obtained for the aim "*conveying meaning clearly and accurately through speech*" so no comparisons could be made between the teachers' assessments and structured activity scores. However, the teachers' assessments were found to be highly intercorrelated (Table 4.13) and the degree of correlation between these assessments and performance on the basic skills in the second year of the study was also high (Table 4.19). It was, therefore, inferred from this data that teachers made a global assessment of pupils in relation to achievement in the basic skills as noted in relation to the evidence from the first year of the study. Therefore, any interpretation of these results in terms of pupil performance in "conveying meaning" would not provide valid evidence relevant to this aim and further analysis of this data was not attempted.

#### 4.7 SUMMARY OF MAIN FINDINGS

The first hypothesis investigated in this research into the validity and reliability of teacher-based assessments was concerned with the development of such assessments, consisting of checklists of abilities, structured activities and teachers' ratings of pupil performance. The working hypothesis was stated that valid and reliable teacher-based assessments could be developed for the measurement of pupil performance for the extended aims of primary education related to the areas of listening with concentration and understanding, inventiveness and creativity, acquiring information other than by reading, application of the four arithmetic rules and mathematical techniques to everyday life and conveying meaning clearly and accurately through speech.



**Table 4.19** Pearson r correlation coefficients of teachers' assessments on the checklists of abilities associated with the aim "conveying meaning clearly and accurately through speech" and post-test scores on the Richmond tests of basic skills.\*\*

<i>Post-test basic skills scores</i>	<i>Total*</i>	<i>Mathematics*</i>	<i>Language*</i>	<i>Reading*</i>
<i>Checklist of abilities criterion</i>				
1. Vocalisation	0.47	0.41	0.37	0.41
2. Appropriate language	0.57	0.50	0.47	0.48
3. Fluency	0.43	0.37	0.32	0.39
4. Sequencing	0.58	0.52	0.48	0.48
5. Responding	0.49	0.46	0.39	0.37
6. Modifying responses	0.49	0.43	0.44	0.40
7. Linking and extending	0.45	0.38	0.41	0.34

- Key:** The child should be able to:
1. vocalise audibly and articulately,
  2. use language and vocabulary appropriate to the situation or context,
  3. talk fluently without undue pausing, hesitation or repetition,
  4. maintain logical connections when describing objects, relating events and during discussions,
  5. respond with reasonable speed and accuracy to questions and other contributions,
  6. modify his/her own responses,
  7. link his/her own experiences to the present situation.

\* Post-test scores were used in this analysis since they were collected closest in time to the teachers' assessments.

\*\* All Pearson r correlation coefficients were significant at or beyond the 1% level.

A number of methods were employed to investigate the validity and reliability of the teacher-based assessments. These have been described in detail in chapter 3 and the results reported in this chapter in sections 4.1, 4.2, 4.3, and 4.5. These strategies incorporated both logical and empirical review procedures applied to the various components of the teacher-based assessments; that is, the checklists of abilities, structured activities and teachers' ratings of pupil performance based on their assessment of each criterion listed on the checklists of abilities.

In the light of the data collected by both the logical and empirical review procedures the inference was drawn that the structured activities, which formed part of the teacher-based assessments, provided a valid measure of some aspects of the extended aims of primary education. Given the limitations of the study, in particular, the time constraint in developing and validating the structured activities these results were encouraging and indicated that some element of pupil performance was being assessed which was different from that measured by existing instruments. The results of the correlational studies and factor analyses of the scores obtained by pupils on the structured activities, modified Richmond tests of basic skills and the study skills exercises indicated that the structured activities measured a component of pupil performance which differed significantly from the data generated on the achievement tests of basic skills and other study skills exercises. The element which was assessed here appeared to be content dependent but not representative of underlying study skills across the curriculum, such as listening, questioning or communication skills.

There were, however, some significant correlations between pupil performance on each of these methods of assessment which would suggest that some underlying skills and abilities were tapped by all three types of measure: Richmond Tests of basic skills, study skills and the structured activities. This element can be described as a general ability component more closely related to achievement in the basic skills than a generic study skills element.

The reliability of the teacher-based assessments was established in three ways. First the degree of correlation between the assessments made by different teachers using the checklists of abilities to assess the same groups of pupils was calculated. Second, the degree of correlation between assessments made on different occasions was established and third the stability of the structured activities over time was calculated. The results from this study indicated that teachers could reach a good level of agreement with each other and over time. In addition, the test-retest reliability for structured activity scores for



*"acquiring information other than by reading"* showed high levels of correlation indicating that they were reliable means of assessing the variables that they were designed to measure. Thus, the structured activities and teachers' ratings on the checklists of abilities were considered to be reliable methods of assessment.

However, the reliability of teachers' assessments did not ensure that such assessments were valid. The results from the investigation of the validity of teachers' assessments of pupil performance using the checklists of abilities indicated that these were more closely related to the performance of pupils on the modified Richmond tests of basic skills than to the performance of pupils on the associated structured activities except in the case of arithmetic competence. The teachers' assessments also showed high degrees of intercorrelation between the ratings given on different criteria and on different structured activities. These assessments were clearly invalid in that they did not discriminate between pupil performance in relation to the particular skills identified within the checklists of abilities or the structured activities. However, the teachers appeared to be able to provide a fairly accurate global assessment of pupil performance in terms of general academic ability since the ratings correlated highly with scores on these assessments particularly the Richmond tests of basic skills.

The subsequent investigation of the validity of teachers' assessments was carried out to test the working hypothesis that teachers' assessments of pupil in relation to the five aims of primary education under consideration were influenced by pupil sex, age, social class, achievement on tests of basic skills and classroom behaviour. The conclusion was drawn that even when achievement on the basic skills and performance on the structured activities were taken into account that the teachers were over- and under-estimating pupils relative to their actual performance. This was true for pupil age, social class and type but not for pupil sex for the data collected in the first year of the study. Similar trends were also observed in the data collected in the second year of the study but the results did not reach statistical significance beyond the five per cent level.

The over-estimation of younger pupils and under-estimation of older pupils was explained in terms of the teachers appearing to use the age group which they taught as a basis for judging the performance of their class. Thus the teachers of younger pupils did not take into account the older pupils in the whole sample so that their ratings appeared to be an over-estimation of the pupils abilities when compared with the higher scores obtained by older pupils. Similarly, teachers of the older pupils under-estimated the performance of their pupils since they did not take into account the younger pupils in the



sample.

Teachers were also found to over-estimate the performance of pupils whose parents were in non-manual occupations and under-estimated pupils whose parents were in manual occupations. This trend reflected that found in previous studies and was interpreted in the context of pupil stereotyping and the notion that teachers appeared to equate cognitive abilities with social competence.

The behaviour of pupils in the classroom might also be expected to have some impact on the way in which teachers perceive them. This was found to be the case with pupils who were on task more often than others in the class such as the quiet collaborators and solitary workers who were over-estimated whilst the attention seekers and intermittent workers were under-estimated relative to their measured performance in relation to study skills components. In addition those pupils who had more contact individually with the teacher tended to be over-estimated in relation to basic skills achievement. The data also showed that when teaching style was controlled for in the two way analysis of covariance the pupil type effects noted previously were diminished although similar patterns existed. This can best be interpreted in the light of the Oracle findings which suggested that teaching style tended to determine pupil type rather than vice versa (Galton and Simon (eds), 1980). For example, when the effect of the class enquirers who tended to encourage solitary workers and achieved greater pupil progress scores on the basic skills was controlled for, the impact of those classes on the residual assessments was reduced. These results are further discussed in Chapter 7.

The next stage of the research was to investigate the differences if any in the performance of pupils on the various components of the structured activities. These findings are described in the following chapter in relation to pupil characteristics such as age, sex, social class and type as well as the impact of experiencing different teaching styles.



## **CHAPTER 5: PUPIL PERFORMANCE ON THE STRUCTURED ACTIVITIES**

In this chapter pupil performance on the structured activities is described in order to ascertain whether this differed significantly with regard to the pupils' age, sex, social class, and teaching style experienced. Pupil achievement on the modified Richmond tests of basic skills has been taken into account in subsequent analyses since it has been shown that there was some measure of correlation between pupil performance on the structured activities and on the modified Richmond tests of basic skills. In each case the data relating to pupil performance on the structured activities is based on the scores derived from the components of the structured activities which made up each of the factors described previously (Tables 4.3, 4.4, 4.8 and 4.9) since the structured activities assessed content dependent study skills. Comparisons have then been made between the scores obtained by pupils by their sex, age, social class and teaching style experienced.

### **5.1 Pupil performance on the structured activities associated with the aim "the child should be able to listen with concentration and understanding".**

The structured activities designed to assess this aim consisted of three parts; musical instrument identification, sound story sequencing and following a set of instructions. Two scores were derived from this exercise on the basis of a factor analysis which generated two factors; one for following a set of instructions and the other for placing the events of a sound story in the correct sequence. Differences in pupil performance by sex, age and social class are presented in Table 5.1. The only significant difference noted in these results was when the scores of pupils of different ages on the sequencing activity were compared. This activity required the pupils to listen and remember information related to a sound story, choose from a selection of pictures provided and then place these in the correct sequence to match the sound story. This task clearly required considerable concentration and the ability to remember and recall the necessary information. The major difference was noted between pupils aged nine plus and those aged ten plus. This result suggested that the development of this particular skill was associated with the development of pupils' general abilities to cope with more complex tasks such as this. However, the results of this preliminary analysis must be treated with some caution since it has already been demonstrated that sequencing was highly correlated with achievement on the basic skills.

**Table 5.1** Mean structured activity scores for "listening with concentration and understanding" by pupil sex, age, social class and pupil type.

<i>Pupil Characteristics</i>	<i>Structured activity component</i>		
	Instructions	Sequencing	N
<i>Sex</i>			
Male	10.50	6.48	231
Female	10.65	6.70	236
All pupils	10.56	6.59	467
<i>Significance</i>	<i>N.S.</i>	<i>N.S.</i>	
<i>Age</i>			
Nine+	10.52	6.09	218
Ten+	10.67	7.00	116
Eleven+	10.52	7.07	113
All pupils	10.56	6.59	467
<i>Significance</i>	<i>N.S.</i>	** [between 9+ and 10+]	
<i>Social class</i>			
Non-manual	11.05	6.30	118
Manual	10.55	7.00	66
All pupils	10.73	6.60	184
<i>Significance</i>	<i>N.S.</i>	<i>N.S.</i>	

*Overall variance*

\* Significant at the 5% level

\*\* Significant at or beyond the 1% level



A subsequent analysis controlling for basic skills achievement was carried out. In this case the pre-test score was used as the controlling variable since this was the closest in time to the administration of the structured activities. The results obtained from this analysis indicated that when achievement on the Richmond tests was taken into account a significant part of the variation in pupil performance on the structured activities was accounted for. This was perhaps not surprising given that the requirements of the structured activities included an ability to concentrate and recall information. This ability was also likely to be a pre-requisite for success on the basic skills as well as in relation to the extended aims of primary education relating to independent enquiry. However, the difference by age for sequencing did remain and was statistically significant (at or beyond the 1% level). However, these results did not take account of teaching style influences. These were investigated using similar statistical procedures as before. The relationship between teaching style and the mean scores on following instructions and sequencing were studied controlling for basic skills achievement and age since these were shown to be significant factors in the level of pupil performance in these areas.

In this analysis no significant differences in pupil performance were found when teaching style differences were considered in relation to following instructions (Table 5.2). This may be accounted for by the fact that all pupils experience following instructions independent of teaching style since it was found that all styles were characterised by substantial amounts of time spent on managerial and task information. For sequencing, the major influence was again the pupils' scores on the modified Richmond tests. Differences in the raw score means between the teaching styles were significant at the 5% level, and when allowance was made for both the age and pre-test scores of the pupils the effect due to teaching style remained. Group instructors and class enquirers had the highest adjusted mean score and the rotating changers the lowest. This result indicated that teaching style experienced had some influence on pupil performance on the sequencing dimension of the structured activity and that the experience of pupils differed in relation to particular teaching styles. This result can be interpreted in relation to the patterns of interaction exhibited by these teaching styles. Both styles were characterised by teacher-pupil interactions which more commonly involved more than one pupil. Class enquirers had the highest amount of interaction with pupils as a class and group instructors the highest for interaction in a group. Such interactions often involved the use of task instructions and longer periods of sustained interaction between pupils and teachers. It would appear that those pupils who experienced such teaching styles in the classroom had learnt to listen carefully and remember what had been said and be able to act upon that information.

**Table 5.2    Relationship between teaching style and mean scores on "listening with concentration and understanding" structured activities.**

<i>Structured activity factor</i>	Instructions			Sequencing			
<i>Controlling variable</i>	None	Pre-test	Pre-test & age	None	Pre-test	Pre-test & age	N
<i>Teaching styles</i>							
Individual monitors	10.41	10.25	10.25	6.90	6.54	6.53	94
Class enquirers	10.84	10.98	11.06	6.88	7.07	7.25	62
Group instructors	10.91	11.17	11.07	7.04	7.44	7.31	50
Infrequent changers	10.85	10.89	11.12	6.30	6.36	6.41	40
Rotating changers	9.69	9.81	9.83	5.66	5.85	5.77	64
Habitual changers	10.75	10.70	10.74	6.83	6.74	6.77	93
<i>Significance level</i>							
Age							
Pretest		**	**		**	**	
Style				*	*	*	

\*    Significant at the 5% level  
\*\*   Significant at or beyond the 1% level



## **5.2 Pupil performance on the structured activities associated with the aim "the child should show some creativity and inventiveness in some fields".**

The structured activity associated with this aim assessed two aspects of creativity: originality and appropriateness. The mean scores for these dimensions are listed in Table 5.3. It can be seen from this table that there were no significant differences between pupil performance on either criterion.

The differences in average raw scores between teaching styles were not found to be significant for either of these exercises (Table 5.4). For appropriateness, neither the Richmond pre-test scores nor age accounted for a significant proportion of the variation in this score on the structured activity. However, for originality, when the means were adjusted to take account of the inequalities in both pre-test and age, there were significant differences between styles. In particular the pupils who were taught by individual monitors showed the highest levels of performance in relation to originality.

This result is of particular interest since research into progressive primary school practice has suggested that pupils who were taught by progressive methods, with an emphasis on the child's needs and interests through individualised instruction, did better in such tests of creativity (Haddon and Lytton, 1968). The individual monitors engaged in the highest levels of such individualised interaction. It would appear that a relationship existed between such tactics and the performance of pupils on these activities. The nature of this interaction with individual pupils did not suggest, however, that these teachers placed any great emphasis on the development of a pupil's creativity. In contrast to this ideal the pupils as individuals had relatively little contact with the teacher and it may be that this relative isolation of the pupils resulted in their success on this activity.

These pupils characteristically received little oral feedback on their work and there seemed to be a reduced pressure to conform to the teachers' expectations on the presentation of classwork or how they approached a particular exercise. This relative freedom appears to have resulted in the pupils taught by individual monitors performing well on this exercise since they worked on their own and completed the exercise without reference to the teacher or to other pupils. The pupils taught by class enquirers did less well since the converse appeared to have occurred; that is, the pupils checked with each other and needed guidance from the teacher on how to complete the task which resulted in less creativity and more conformity in the pupils' responses.

**Table 5.3** Mean structured activity scores for "creativity and inventiveness" by pupil sex, age and social class.

<i>Pupil Characteristics</i>	<i>Structured activity component</i>		
	Originality	Appropriateness	N
<i>Sex</i>			
Male	1.86	1.81	210
Female	1.75	1.56	226
All pupils	1.81	2.00	437
<i>Significance</i>	<i>N.S.</i>	<i>N.S.</i>	
<i>Age</i>			
Nine+	1.90	1.62	218
Ten+	1.64	1.72	116
Eleven+	1.79	1.73	133
All pupils	1.81	1.68	467
<i>Significance</i>	<i>N.S.</i>	<i>N.S.</i>	
<i>Social class</i>			
Non-manual	1.84	1.67	107
Manual	1.85	1.88	58
All pupils	1.84	1.74	165
<i>Significance</i>	<i>N.S.</i>	<i>N.S.</i>	

\* Significant at the 5% level

\*\* Significant at or beyond the 1% level

N.S. Not significant



**Table 5.4    Relationship between teaching style and mean scores on "creativity and inventiveness" structured activities.**

<i>Structured activity factor</i>	Originality			Appropriateness			
<i>Controlling variable</i>	None	Pre-test	Pre-test & age	None	Pre-test	Pre-test & age	N
<i>Teaching styles</i>							
Individual monitors	1.92	1.92	1.95	1.71	1.69	1.64	90
Class enquirers	1.77	1.56	1.61	1.73	1.74	1.73	57
Group instructors	2.13	1.65	1.55	1.63	1.65	1.72	41
Infrequent changers	1.59	1.41	1.37	1.86	1.88	2.00	39
Rotating changers	1.71	1.57	1.50	1.55	1.56	1.62	61
Habitual changers	1.69	1.57	1.50	1.69	1.69	1.58	89
<i>Significance level</i>							
Age							
Pretest		*	**				
Style			*				

\*    Significant at the 5% level  
\*\*   Significant at or beyond the 1% level

### **5.3 Pupil performance on the structured activities associated with the aim "the child should be able to acquire information other than by reading".**

This structured activity was scored in relation to two aspects: comprehension by looking and listening and questioning. The data on pupil performance in relation to age, sex and social class are presented in Table 5.5. Significant differences were noted for pupil scores obtained by boys and girls and at different ages for questioning. However, this was one of the study skills which was significantly correlated with performance on the basic skills. Thus, when pre-test scores were taken into account the difference between boys and girls was no longer significant. However, statistically significant differences for pupil performance at different ages remained suggesting that this was related to some developmental process.

In the analysis of teaching style differences (Table 5.6), the scores of pupils on the Richmond pre-tests accounted for a significant proportion of the variation in the structured activity scores for both comprehension by looking and listening and questioning. Differences between the style means were not significant for the comprehension activity but were significant for the task in which pupils were required to pose suitable questions.

However, the raw score means of the group instructors were only the fourth highest average score but after taking account of the differences in the Richmond test scores the value rose to 5.13. There was a further increase to 5.52 when age was also taken into account. This made them second only to the class enquirers, the most successful style. These style differences can be interpreted in relation to the quality of the teacher-pupil interactions.

Both class enquirers and group instructors were characterised by the relatively high degree of questioning in their interactions, when compared with other teaching styles. It would appear that where pupils experience this type of interaction they learn to apply it in their own work. This resulted in significantly different levels of performance for those pupils taught by these styles.



**Table 5.5** Mean structured activity scores for "acquiring information other than by reading" by pupil sex, age and social class.

<i>Pupil Characteristics</i>	<i>Structured activity component</i>		<i>N</i>
	<i>Comprehension</i>	<i>Questions</i>	
<i>Sex</i>			
Male	8.41	4.49	231
Female	9.00	5.14	236
All pupils	8.69	4.81	467
<i>Significance</i>	<i>N.S.</i>	*	
<i>Age</i>			
Nine+	8.71	4.36	218
Ten+	9.00	4.83	116
Eleven+	8.38	5.52	133
All pupils	8.69	4.81	467
<i>Significance</i>	<i>N.S.</i>	** [between 10+ and 11+]	
<i>Social class</i>			
Non-manual	8.75	4.26	118
Manual	8.82	4.38	66
All pupils	8.78	4.30	184
<i>Significance</i>	<i>N.S.</i>	<i>N.S.</i>	

*Overall variance*

\* Significant at the 5% level

\*\* Significant at or beyond the 1% level

**Table 5.6**      **Relationship between teaching style and mean scores on "acquiring information other than by reading" structured activities.**

<i>Structured activity factor</i>	Comprehension			Questions			
<i>Controlling variable</i>	None	Pre-test	Pre-test	None	Pre-test	Pre-test	N
	& age			& age			
<i>Teaching styles</i>							
Individual monitors	8.40	8.01	8.10	5.01	4.54	4.26	94
Class enquirers	8.81	8.98	9.53	5.57	5.78	6.25	62
Group instructors	8.80	9.14	9.19	4.70	5.13	5.52	50
Infrequent changers	8.09	8.13	8.91	3.61	3.66	4.06	40
Rotating changers	8.73	8.88	8.90	4.27	4.47	4.73	64
Habitual changers	9.07	9.00	9.30	5.20	5.10	5.31	93
<i>Significance level</i>							
Age							
Pretest		**	**		**		
Style			**	**	**		

\* Significant at the 5% level  
\*\* Significant at or beyond the 1% level



#### **5.4 Pupil performance on the structured activities associated with the aim "the child should know how to compute in the four arithmetic rules and use mathematical techniques in his/her everyday life".**

Pupil performance on the structured activities associated with this aim was highly correlated to performance on the basic skills, thus, any significant differences between pupil performance on the two components of the structured activities were found to be accounted for by basic skills achievement (Table 5.7). The effects of teaching style were again considered by the use of analysis of covariance to investigate the performance of pupils in the use of the four rules compared with the performance that would have been expected on the basis of their score on the modified Richmond tests of mathematics at the end of the first year of the study. For this sample of pupils, age was not a significant variable once achievement on the basic skills was accounted for so it was not necessary to control for this in the analysis conducted on this data. The results are presented separately for the two structured activity components: competence in the use of the four rules and manipulation, organisation and presentation of data (Table 5.8).

This data confirmed the findings in the first year of the study in that the pupils who were taught by class enquirers did better in exercises requiring competence in the use of the "four rules". In the case of the structured activity component requiring the manipulation and organisation of data, the pupils taught by class enquirers and infrequent changers did very much better than might have been expected. However, these results were only statistically significant in the case of the four rules and not with regard to the practical mathematics exercises which required the manipulation, organisation and presentation of data.

#### **5.5 SUMMARY OF MAIN FINDINGS**

The working hypothesis that pupil performance was correlated to pupil characteristics such as age, sex and social class was investigated. Some significant differences in pupil performance were found even when initial achievement on the basic skills was taken into account. In addition, the second working hypothesis which was tested in this part of the study stated that teaching style was correlated with pupil performance on the structured activities. This was also found to be the case. These results are now summarised in relation to each of the structured activities.

**Table 5.7** Mean structured activity scores for "everyday mathematics and computation in the four rules" by pupil sex, age and social class .

<i>Pupil</i>	<i>Structured activity component</i>		
<i>Characteristics</i>	Four rules	Organisation of data	N
<hr/>			
<i>Sex</i>			
Male	21.19	34.05	149
Female	21.36	31.95	152
All pupils	21.27	32.98	301
<i>Significance</i>	N.S.		
<hr/>			
<i>Age</i>			
Nine+	17.11	30.20	115
Ten+	22.66	35.46	82
Eleven+	24.96	30.74	104
All pupils	21.27	32.98	301
<i>Significance</i>			
<hr/>			
<i>Social class</i>			
Non-manual	22.34	37.35	59
Manual	18.41	30.33	86
All pupils	20.01	33.14	135
<i>Significance</i>			
<hr/>			

*Overall variance*

- \* Significant at the 5% level
- \*\* Significant at or beyond the 1% level



**Table 5.8    Teaching styles and performance in the structured activities associated with "everyday mathematics"**

<i>Residual score using previous year's modified Richmond test score as covariate</i>			
<i>Structured activity factor</i>	Use of the four rules	Manipulation and organisation of data	
<i>Teaching style</i>			
Individual monitors	-0.90	-2.90	30
Class enquirers	1.54	6.00	19
Group instructors	-0.80	-0.80	78
Infrequent changers	0.10	6.10	22
Other changers	0.50	-0.90	120
<i>Significance level</i>	*		

\* Statistically significant at the 5% level

Pupil performance on the structured activities associated with the aim "the child should be able to listen with concentration and understanding" varied with age, with a significant shift taking place between nine and ten year old children for the sequencing component. Teaching style differences were observed for this part of the structured activity also. Group instructors and class enquirers were the most successful styles whilst the rotating changers were the least successful in terms of pupil performance in this area.

No statistically significant differences were found between the scores obtained by different groups of pupils for the aim associated with inventiveness and creativity. However, the analysis of the impact of teaching style showed that one style that of the individual monitors was by far the most successful in relation to this aim.

Statistically significant differences in the performance of pupils on the structured activities associated with the aim "the child should be able to acquire information other than by reading" were found with respect to age and sex on the "questioning" part of the exercise. Teaching style differences were also noted with pupils in the classes of group instructors and class enquirers performing significantly better in the questioning activity.

In the second year of the study, any difference in initial performance on the structured activities associated with the aim "the child should be able to apply mathematical techniques to everyday situations" was accounted for by performance on the basic skills. Similarly teaching style differences although showing a similar trend to that found in the first year of the study were also accounted for by performance in the basic skills.

Thus pupil performance on the sequencing and questioning components of the structured activities was found to vary significantly with age. Whilst pupil performance on the questioning component also varied with sex of the pupil. In the second year of the study significant differences observed were accounted for by achievement in the basic skills. Teaching style effects were found in relation to the following aspects of the structured activities investigated: sequencing, questioning and originality. However, for following instructions; comprehension by looking and listening; appropriateness; organisation, manipulation and presentation of data no statistically significant teaching style effects were found. The most successful styles overall for pupil performance on the study skills were the class enquirers and group instructors, followed by the individual monitors, then the habitual and infrequent changers, and finally the least successful style that of the rotating changers.



Thus it was found that pupil performance on certain dimensions of the structured activities was influenced in some way by the teaching style experienced. In addition, the results from the previous chapter also indicated that particular teaching styles either over- or under-estimated pupil performance in relation to their measured performance on the basic skills and the structured activities. These findings led to a further consideration of the question of how best to improve the validity of teachers' assessments and to promote the professional development of teachers. These problems were investigated through an evaluation of in-service provision which is described in the following chapter.

## CHAPTER 6: REPORT OF THE IN-SERVICE EVALUATION STUDY

### 6.1 INTRODUCTION

This chapter focuses on the presentation and discussion of the results obtained during the development and evaluation of the in-service degree strand related to assessment and evaluation described in Chapter 3. This evaluation of in-service provision explores the nature of teachers' professional learning through the examination of the processes and outcomes of the course. It was suggested previously that one way that teachers, including myself, can be encouraged to change their practice is through self-evaluation. In order to promote this, however, it is necessary for any discrepancies between theories-in-use and espoused theories to be recognised and documented. The intervention which was proposed to stimulate the recognition of any such discrepancies was the observation and assessment of pupil learning. Once these discrepancies were made explicit then new espoused theories and theories-in-use could be learnt resulting in more appropriate theories of action or professional learning.

In order to evaluate the course provision a case study was carried out where I monitored the processes of the course and assessed the teachers' professional learning. This data also provided the basis for my own self-evaluation and thus further evidence of the effectiveness of this strategy for facilitating teachers' professional development, since I viewed myself as a teacher engaging in the same processes as the teachers in the group. The principle research method used was that of triangulation. Thus the investigation of teachers' professional learning drew on a variety of sources of data including teacher self-reports; their contributions within the tutorials, lectures and workshops; line-labelled concept maps; oral and written presentations of pupil case studies, school and course files. This evidence was used to explore the development of teachers' professional learning in relation to the theory and practice of assessment. My self-evaluation report also drew upon this evidence as well as a personal diary.

The accounts which follow present the evidence related, first, to the processes whereby the teachers acquired expertise in assessment through the course and applied this knowledge to the classroom context (6.2). Second, the outcomes of the teachers' professional learning are presented with reference to line-labelled concept maps, pupil case studies, school and course files in order to appraise the quality of their professional learning within the course (6.3). Finally, the impact of assessing teachers' learning on my



own professional development is considered with reference to how far I achieved my own aims and intentions within the course and developed new theories of action (6.4).

## **6.2 PROCESSES IN TEACHER PROFESSIONAL LEARNING**

The focus of this discussion is to explore the suggestion that the study of assessment within the B.Ed. (Honours) In-service degree programme would promote self-evaluation and provide a means of facilitating professional development for myself and the teachers on the course. In particular, the findings of this study are considered with reference to four questions. First, what were the concerns and issues identified by these teachers at the start of the course since it has been suggested previously that for teachers to engage in professional development the focus of study should be relevant to the clients' needs (Day, 1981)? Second, since it has been suggested that there are different forms of teacher self-evaluation (Elliott, 1983), was there evidence that the teachers went through defined stages in their professional learning which reflected these different modes of self-evaluation? Third, what evidence was there to suggest that the assessment of pupil learning promoted self-evaluation and finally, were other aspects of the course equally effective in promoting teacher self-evaluation, as has been suggested by others (Elliott and Adelman, 1976; Day, 1981)?

### **6.2.1 Characteristics of the teachers at the start of the course.**

Given that I had no prior knowledge of the client group one of the first activities which I introduced was the teacher self-report in order to ascertain the starting points of these teachers and identify what they perceived as their needs. Thus during the second session of the course, the following information was requested from the teachers:

- a) a personal account of their previous school context which included general information of school type, teaching and other responsibilities,
- b) a general description of assessment practices within their schools and those which the teachers particularly used,
- c) a personal statement from each teacher identifying areas of strength and areas where they would welcome some help, particularly in relation to profiles and records of achievement,



- d) identification of any aspects of monitoring and assessing pupil performance in the school or classroom which were considered by the teachers to be important.

The self report responses have been summarised in the form of a pen portrait for each teacher and these can be found in Appendix 18. An analysis of these reports yielded the following information on the composition of this cohort of teachers. There were twelve females and seven males in the group. Of the nineteen teachers enrolled in the course the majority (twelve) were previously employed in the secondary sector in schools which were high schools (11 - 14 and 11 - 16) and upper schools (13 - 18 and 14 - 18). Of the remaining seven teachers two were from first schools (5 - 9), two from middle schools (9-13), one from a primary school (5 - 11), one was employed in a remand centre and one came from the private sector (5 - 16). Areas of subject expertise represented in the group included: physical education; remedial and basic skills teaching; home economics; science; mathematics; rural studies; history; geography; social studies; health education; humanities; art; craft, design and technology.

The number of years spent in teaching ranged from three to over twenty. Two teachers had less than five years teaching experience, whilst three had between five and ten years. The remaining fourteen teachers had spent between 10 and 25 years in teaching. This group had a wide range of experience where the majority of the teachers had spent considerable time in schools as well as holding posts of responsibility and so represented very different starting points.

A summary of the concerns of this group of teachers is now presented to illustrate their perceptions and experience of assessment in their schools at the start of the course. Two important issues emerged with respect to school assessment practices in secondary schools. First, most of the secondary school teachers were concerned about the use of norm-referencing as a basis for determining the proportion of pupils who could be awarded a particular grade for achievement. Second, they saw little connection between the assessments they made of pupils in the classroom, laboratory or games field and the way in which these teacher assessments were formalised in profiles, reports and achievement grades. Of particular concern in the context of teacher expertise in assessment was the response of one secondary school teacher who had spent considerable time in the development of a leaving certificate which was to be validated by one of the examining boards as a record of pupil achievement. He particularly noted that teachers' assessments of a pupil's class performance did not appear to feed into these profiles, further emphasising the need to explore how teachers' assessments could be



used effectively in describing pupil achievement. This was particularly so when teachers in areas such as physical education and craft worked in schools where the assessment policy involved regular internal examinations, class rankings and less flexible arrangements for setting pupils. This suggested that there would be a need to focus on the relationship between teachers' assessments and more formal methods of assessment which were used by these teachers in their schools.

The teachers were also asked to identify areas of strength and weakness in relation to their knowledge and practice of assessment in school. Eight of the secondary teachers who had a particular responsibility for working with the less able, either as subject teachers or teaching basic skills within the remedial department, expressed a lack of confidence in assessing pupils particularly in relation to subject knowledge. There was also a general dissatisfaction with the methods of assessment used in such departments for the identification of pupils with special needs. Much emphasis was placed on the use of particular standardised tests and little on more informal methods of assessment: for example, Richmond tests were used for selection, banding and setting by six schools with varying degrees of acceptance by the teachers involved. The use of Schonell (1963), Daniels and Diack (1970) was noted by two secondary teachers working in the remedial department. However, other subject-based teachers did not make reference to the use of such tests.

The three teachers in first and primary schools were more familiar with informal methods of assessment based on classroom observations, examples of the pupils' work and anecdotal comments. Their concerns regarding assessment centred on developing more effective ways of record keeping and the need to develop expertise in assessment generally but with particular reference to language, number and social development. The two middle school teachers were more aware of the importance of both formal and informal methods of assessment. These concerns were more closely linked to relaying the evidence of pupil learning in a form which was acceptable both within the school and to outsiders. One teacher in particular noted the confusion he experienced with regard to reporting pupil achievement given the different expectations of the local education authority, parents and the headteacher.

Thus the material derived from the teacher self reports provided valuable background information as well as identifying some areas of concern for the group. In particular, the differences in the composition of the group from previous years suggested that more emphasis should be placed on informal methods of assessment for the secondary



teachers, whilst the group as a whole might benefit from a consideration of the problems associated with more formal methods of assessment. As a result of collecting this information it was also possible to group the teachers in such a way as to maximise the opportunities for discussions which crossed the traditional divide between primary and secondary teachers, although the absence of a significant number of middle school teachers was disappointing since in previous years they had served as a bridge between primary and secondary concerns.

### 6.2.2 The process of teachers' professional learning: the course events.

This section details events which illustrate processes related to teacher professional learning which were documented during the autumn term of 1984. The data in this section has been selected from the course description and edited extracts from the tape transcripts of each session which are detailed in Appendix 19. A number of significant themes arose during these lectures and workshops which provided some clues to the development of the teachers' professional learning from the various starting points described in the previous section.

During the first three weeks of the course the teachers concentrated on getting to know each other, coming to terms with the demands of the course and what was required of them. One of the first tasks which the teachers engaged in was the completion of their self reports. It was here that signs of uncertainty regarding the course and lack of confidence in their own knowledge and experience began to emerge as the following dialogue illustrates. This teacher (L) was unclear about the type of assessment which "counted" and undervalued her own judgements of pupil performance.

AJ *"...I don't want to constrain the response that you give me, because I want it to be the way you feel rather than the way you think I want you to write - the way you interpret is very important."*

L *"...with some assessments practices - obvious forms - exams, gradings on reports, C.S.E.'s etc, do you want us to go into confidential records as group tutors?"*

AJ *"Do you want to?"*

L *"As a P.E. teacher you may find that that's the only thing I know anything about - nothing to do with formal assessment in that sense."*

AJ *"O.K. - you've answered your own question."*

The teachers also sought reassurance that they were doing it "right" by expressing a



concern for "answers", "is this what is expected of me" and "are we giving you what you want". This supported my previous experiences which suggested that the teachers lacked confidence in expressing their own ideas, perceptions and anxieties. The teachers also highlighted that they themselves were learners and that the course was providing an uncomfortable experience of what it felt like to be a pupil in their own classes as illustrated in the following extract.

*AM "...doing it in sentences,"*

*AJ "... whatever you might want to give me."*

*AM "... now we know how the children feel, isn't it? Are we doing what she wants us to do?"*

*AJ "Which is not what I'm asking you to do."*

*AM "I know but that's what children must feel."*

[Audiotape recorder was switched off for short period of time to change cassette.]

*AJ "I don't want to tell you what to write or how much to write. I want you to respond in the way that you feel is appropriate to address the particular points that I've raised. I've not phrased them as questions either just made them fairly brief areas - I've left them fairly open-ended intentionally."*

*H "Well I've finished does that mean I haven't done it right - nobody else has."*

Once the self-reports had been completed the first part of the course was introduced. This dealt with a number of informal methods of assessment. The first three sessions were devoted to improving the teachers' observation skills by introducing the Pupil record (Galton, 1978b) and the concept of appraisal (Tough, 1976) so that they would be better equipped to operate such informal methods of assessment. In learning how to use the Pupil Record, the teachers were particularly concerned with the problems of introducing an observer into the classroom and the effect that this might have on the teacher and pupils who were being observed. This led to a discussion of some of the strategies used by observers to minimise the amount of interference in classroom events. Their concerns about interfering with classroom events suggested that the teachers were viewing observation as something which was done to them (and their pupils) rather than a technique which they could use to increase their knowledge and understanding of their own and the pupils' learning.

The teachers also raised a number of points with respect to this type of category system. The first concern was that nothing was being recorded about the quality of the



interactions. The second concern was with the checking of agreement between observers within the context of the Oracle research programme (Galton *et al.*, 1980). This suggested that the teachers were not understanding the course content as I had intended. At this point they did not see the relevance of such methods of recording pupil behaviour for their own practice.

The teachers were also concerned about the difference between pupil intentions and observable behaviour. This issue arose during discussion of the disruptive category on the Pupil Record. As one teacher put it *"if it's in their mind to be disruptive then they are disruptive"*. At this point I emphasised that the category system was based on observable behaviour and not on the observer inferring pupil intentions. The elaboration of this point was in terms of such category systems providing a more balanced view of pupil behaviour by focusing on the routine events in the classroom rather than noting significant events alone. It was important for me to emphasise this since one of my intentions was to help teachers observe more effectively by recognising the bias in their own observations, since I believed that this would result in them being able to make more valid assessments of the pupils' classroom performance. The following extract illustrates my position.

*"I understand what you are saying to me - classroom life is more complicated than this - this is not going to give me information about particular pupils that I think I need but what it might do is focus your attention on patterns of behaviour. For example, I've known people who are surprised at the amount of disruption exhibited by a pupil they thought of as being cooperative - they'd worked out a very good way of concealing their non-cooperation so there's that dimension. On the other hand there are other pupils that we think of as being very distracted or disruptive or aggressive, we suddenly find their behaviour isn't mirrored by the observations. There might be one instance in ten minutes, the rest of the time they may be cooperating on their task. All it's designed to do is balance the immediacy of certain events which kind of are lifted into high relief because they're very powerful events. It's saying 'yes these events do occur but what's happening the rest of the time?'."*

However, as this discussion continued, it became clear that the teachers were still seeing this type of observation schedule as a research tool used by other people not as a technique which might be of value to them in the classroom. This was particularly true for those teachers from secondary schools who did not see the value of this work to their own classroom practice. This was expressed quite clearly by one of teachers (C) who said *"I don't see how this, interesting though it may be, is going to help me in the future."* The focus of the discussion then shifted back to a consideration of the value of such research in general.



The following justification of the course content illustrates how I conceived the value of this activity.

*"How I expect it might help you ... two things ... I don't agree that a secondary school teacher does not have time to make these records because I think, not in the way it's used here, but by devising a checklist of behaviours that you want to observe. Pupils are not sitting listening to you all the time. You monitor what pupils do when you're doing practical work, when they are writing, answering questions, doing various types of activity. Therefore, the first reason for giving you this kind of way of looking, and that is all it is, is to highlight the problems of observing what pupils are actually doing, actually categorising the types of behaviour, to be confident about the judgements you are making about pupils. We all observe all the time..."*

The idea of categorising observations as a means of making the task of recording pupil behaviour and performance more manageable was then discussed in response to the teachers concerns for creating the time to do this. There were a number of significant contributions at this stage regarding the feasibility of using such methods which continued the debate as to their value for practising teachers. These are now detailed in full together with my responses since this illustrates the negative views of the teachers prior to trying out the materials on school placement and also the connection which was made in at least one case between pupil assessment and teacher self-evaluation.

*O "How can you observe when..."*

*AM "... remedial kids jumping around with the possibility of .. just cannot do it."*

*? "Only time I could do it was when they cooperated."*

*J "It depends on how you see your role as a teacher if you see yourself there as an observer or getting stuck in and if they don't understand helping the individual children especially low ability kids. I feel personally that this type of tick list ... example of social sciences ... individuals are all individual. I don't think you can classify them in little boxes, much better being in with kids for a year getting to know them ... things that happen at playtime, things that they tell you, the work they've done, things about home. I think being involved with them you get much more out than sitting down ... the kids there ... and you ticking little boxes."*

*AJ "O.K. I'm not suggesting that you should operate as an observer in the classroom first of all. I agree with a lot of what you are saying from a personal standpoint but I don't see that that should stop me from providing you with alternative ways of looking and that is what I'm trying to do. What you are saying is 'I get to know my kids, how do you get to know them?' ... [inaudible response] ... so what you are saying to me is that your observations are going to be more reliable than the information I'm going to collect in this way."*

*R "I don't agree with that from experience .."*

*G "... different teachers have different things get on their nerves so there is no*



*objective way of saying what the [pupil] was actually doing with all the emotion cut away."*

*AJ "Do you think it is important that one should try to work towards some ... closer approximation to what that child is really about?"*

*C "What this is to me [is] more help monitoring my own performance than the kids' performance because what they are doing depends on what I'm doing if they are distracted, that is not their fault that is my fault."*

This teacher (C) was clearly alert to the point I was trying to make. He was aware that through the assessment of pupils he would become more aware of his own practice (theories-in-use). However, this was not pursued by other teachers at this stage of the course since they had already expressed a generally negative response to the course content thus far, particularly in terms of school placement activities. The teachers were apparently reluctant to apply more systematic methods of observing, collecting and analysing data related to classroom events, particularly those related to the assessment of pupils. In addition, the teachers rejected the notion that they could engage in "research" which would be of value to them in the classroom and perceived this to be something which was done to them and to pupils by others.

Further evidence from the discussions suggested that the teachers linked research activities and the assessment of pupils to an increasing emphasis on external accountability. This can best be understood in relation to the impact of the particular pressures exerted on a number of teachers in the group. Seven of the group had been seconded by the same local authority which was undergoing reorganisation from a selective system to a comprehensive system. This was taking place at a time of falling rolls and reduced expenditure in education. As a result the local education authority had redeployed a large number of teachers at the end of the previous year and the possibility of redundancies within the profession was being seriously considered. The teachers who were seconded obviously felt considerable pressure to succeed and upgrade their qualifications so that they would not become redundant. The following transcript clearly illustrates the feelings of the teachers regarding the political and economic climate of the time where funds were restricted and there was little valuation of the worth of teachers by the public or politicians. In retrospect these teachers may well have been reflecting the build up of frustration and anger concerning the lack of resources, financial reward and loss of professional status which culminated in the protracted teachers' strike during 1985 in the United Kingdom.

*AJ "They want to get an answer, because ultimately an effective teacher depends on what values you've got."*



AM *"But they've already done that. In .... in the reorganisation they decided on who was effective and who wasn't. Some got promoted but some went down an awful long way ... [inaudible] ... that's exactly it they didn't know what was effective."*

[General laughter]

B *"[You] see their role as being a twenty four hour social worker, twenty four hours a day, seven days a week."*

AJ *"No that would be unrealistic, I know what you're saying I mean I'm parodying but the nature of the professional role of a teacher is changing."*

B *"Undoubtedly, but what I'm saying is .. whoever reads what we say and what we think about or makes effective use of the kinds of observations we have .. [inaudible] .. teachers come out with ideas .. bolshei radical should be booted out."*

This extract clearly illustrated the frustration experienced by at least one teacher who had tried to promote change in his own school. Another teacher went on to describe her particular experiences with the local education authority which was in the process of reorganising from grammar and secondary modern schools to the comprehensive system. The following transcript illustrates how much apprehension had been produced for this teacher by insensitive handling during the change over. Examples of problems experienced by other members of the group followed so that their feelings of frustration were openly expressed and this appeared to be a necessary catharsis for the teachers concerned. However, it was also noted that these teachers did not reject the need for change but were angry at the constraints which existed if they tried to extend this beyond their own classrooms.

AM *"I wasn't even interviewed. I was obviously so dynamic he remembered my interview, gave me a job and I'd never even been in the school before and that shows you how ineffective they are ... shows how ineffective I am, if I'm so ineffective they didn't even notice if I had an interview or not. What good is it me saying what teachers are effective. I know if I'm effective or not, I think this is useful for seeing how effective I am but beyond me and my classroom..."*

B *"Are you measuring the pupils? Are you measuring the fact they all get eight 'O' levels that's great that makes somebody who gets the pupils through ... I teach within the school system and that to me is the important crucial point we're talking about modifying the system.. working with this within our school [reference to the professional project]. We're going back to this in a system that doesn't really work and we can all see, we all know when we think about it but who the hell.... I try within my own little small way through professional associations then I get branded as a trouble maker and a stirrer. 'Oh! he's off on his union again!' I really resent that because what I'm doing is to benefit my colleagues and ultimately the pupils. People are defending the status quo. What we're in great danger of doing this morning is defending a system .. they're part of the system anyway it's in their interest to defend their*



*... because it's the system that put them there in the first place so if you change the system .. [inaudible] .. so what, no good comparing my school to your school, because there's a totally different set of teachers and pupils, syllabi and exams ... so totally different that it's almost impossible to compare two places."*

? *"... reading the results"*

B *"that's the interpretation people, so called experts, who interpret educational research statistics, with the local rag that I happen to read occasionally, they just can't get anything right. I know because when I was a trade union member I used to be misquoted. Totally opposite to what I actually said but they'd interpreted it wrongly ... so until we get down off our ... and look at ways of interpreting things together ... [laughter] ... O.K. teacher's pet ... we've got a system here of observing pupils or teachers that we all can research to some degree."*

We then discussed what constituted a successful teacher in terms of pupil learning. The group recognised that they gave different priorities to different aims dependent on a number of factors. It was at this point that I introduced the notion of mismatch between aims and intentions and classroom practice in the hope of shifting the discussion towards the idea of self-evaluation for professional development. However, the teachers resumed the previous discussion focusing on external accountability, advocating the desirability of assessing pupil achievement across the curriculum not just with regard to success in public examinations. Finally the discussion drew to a close on a rather pessimistic note as illustrated by the following transcript.

AM *"There's no future in teaching. The majority of teachers know that they are not upwardly mobile, they know they are going to spend the next twenty years doing the same job. The majority have got to learn to live with that. They are going to be in that job for a long time."*

B *"People do what they see as being their job, keep the head down, don't rock the boat. I'll be alright."*

AM *"We're completely governed by the system. People feel they're paying off their mortgage, I don't agree with that. I'm really there because I really basically like teaching but the trouble is the system squashes you so much you end up being cynical."*

Given this sense of frustration with the "job" of teaching and the reduced mobility within the profession, the formal qualification which would be acquired by successfully completing the course was one way which the teachers could see of sustaining and possibly improving their position within the system. In the case of the most vocal teachers (AM, B and C) they were prepared to embark on self-evaluation in the hope of improving their classroom practice even though they expressed anger and concern about the way external accountability had influenced them in the past. It was at this point that I



began to think that the teachers were coming to terms with my intentions. As I concluded at the end of the first three weeks:

*"I think one of the things we've managed to do this morning is to air a lot of the hidden agenda that was around and coming up in previous discussions. That's great because it gives me a lot more insight into where each of you individually is coming from and your concerns. For that reason I suspect I might well modify the way in which I'm going to approach the next few weeks so can you leave that with me and not hold me to my outline just yet."*

I, therefore, had to shift my emphasis in the course to overcoming their resistance in general rather than focus on the study of pupil assessment as an intervention for change. Thus the next major activity (week 5 of the course) related to the study of the nature of assessment so that more general issues relating to the validity, reliability and practicality of various methods of assessment could be considered. The teachers were asked to address a series of questions. These were elaborated as follows:

AJ *"What do you understand by assessment? What are its characteristics? What makes an assessment an assessment? What are the purposes you would be required to assess for in school? There are quite a number you are assessing for. How might you assess for these different purposes? This workshop will draw on your knowledge of school assessment procedures but it should not be framed by your knowledge of particular assessment procedures."*

This session continued with a discussion of the concept of assessment and the construction of line-labelled concept maps to illustrate their current understandings of assessment. An initial appraisal of the line-labelled concept maps suggested that the teachers found some difficulty in conceptualising assessment, although particular organising concepts were emerging. For example, the idea of formal and informal methods of assessment, evaluation based on a range of assessment methods as an end point and different purposes of assessment were identified in some of the line-labelled concept maps.

The session continued with a consideration of different methods of assessment and their appropriateness for different purposes. In particular continuous assessment and the practical difficulties associated with marking and operating Mode 3 syllabi and examinations were considered. In the ensuing discussions the teachers identified the changing nature of their professional role in terms of the lack of time to reflect, develop and evaluate the curriculum as they would wish, which suggested a more rationale and positive approach to recognising the need for change than had been in evidence during the first three weeks of the course.



- C "... they haven't got time to change. If you've got to change all your system, you need so much more time to assess what you're doing and put it into practice ... most teachers haven't got that time."*
- ? "It is possible to do. I've done it, my school is changing at a phenomenal rate and the examining system is changing at an enormous rate ... continuous assessment has been a major part of all the exams we've done in school for the past eight /ten years. It's becoming increasingly important in other subjects, isn't it? Other subjects are being asked to be taught in a more practical way."*
- D "Over the last ten years we get less and less time allowed in schools."*
- AJ " and more and more is expected of you, of the teaching profession, which is why you're here, one of the reasons you're here: to have time out, to reappraise, to look again, to have time to think about what you're doing and why you're doing it. There is already evidence that you are beginning to think about what you're doing and why you're doing it."*
- H "That's definitely why I came."*
- D "You're not given it otherwise. You can't redirect your career or anything. You're never given time to sort it out, to think about it ..."*

Thus halfway through the course the teachers again expressed concern regarding their professional status. This coincided with discussions of changes in the 16+ examination system which the teachers argued were being implemented without sufficient time and resources being made available, despite the demands which the teachers experienced in schools they still acknowledged the need to change both the school system and their own practice. They may not have been able to see how to achieve this but the concern for the needs of individual pupils was obvious. However, whilst recognising this need they were also very aware of the demands that this extended role made on them as individuals and the often impossible task this presented in the schools where lack of time, money and resources restricted many attempts to improve the quality of the educational provision that these teachers believed should occur. It appeared that the constraints of the school system operated against change but once in the college environment the teachers became more open to adopting new methods and implementing these as part of their practice as evidenced through the school placement.

At this stage of the course I was collecting and analysing three different types of data, which included the concept maps produced by the teachers, information from the tutorials as well as the evidence of the group processes and content covered within the college-based part of the course. These developing understandings of the teachers provided valuable feedback as to the particular misconceptions and possible directions for course discussions and workshops, in particular I proposed to place more emphasis on



the application of the college-based workshops to the school placement since the evidence from the tutorials (section 6.2.3) suggested that they were experiencing some difficulty in coming to grips with the school placement requirement to apply their skills and knowledge in the practical context of the classroom. It was at this stage that I changed the emphasis in the college-based course away from large group discussions and lectures towards small group work and individual tutorials. The course content was focused explicitly on the links between the course material and school placement activities which are more fully documented in the following section.

The next activity was designed to focus the teachers' attention on the wider applications and purposes of assessment, particularly at the school level. In addition, I hoped that it would provide some useful insights which would help the teachers select and organise the information they were collecting on pupils for the case studies since this need had been expressed during the tutorials. These intentions are illustrated in the following extract of the tape transcript.

*AJ "How do we describe pupils overall in order to get some picture of their various abilities. We've all talked about what criteria we use to make a judgement of the pupil. So in order to open up the exercise away from the classroom level, to think about the overall achievement and experience of pupils, I'm going to suggest that you work in groups of your own choosing, to draw up profile categories appropriate to either pupils leaving school or one of the other transfer points prior to leaving school. So if you are a secondary school teacher you might want to work with a primary orientated teacher to determine what kinds of information would be appropriate to transfer. One or two of you might be interested at looking at transfer into the system. I want you to be aware of the implications of the four issues I've already identified: the purpose, validity, reliability and practicality of that particular profile. The idea is to try and support the school files where you have been asked to produce a profile of a particular pupil, to alert you to the kinds of things you might be looking for. Can you keep in mind what instructions or arrangements might be necessary to help teachers complete such a profile, given that you know what some of the problems are ... a pupil profile for a particular purpose which you will define.*

Thus the pupil profiles which were being developed were intended to help towards the work they were doing on school placement and in the preparation of the case study presentations. It became clear that the teachers were experiencing considerable anxiety as the submission date for the course and school files approached as well as the presentation of their pupil case studies, particularly as these were the first pieces of graded and formally assessed work which the teachers had to complete successfully in order to progress to the next stage of the course. The session continued with a discussion of the problems and possible strategies for dealing with these. The following extract illustrates



the way in which the teachers were making sense of the information derived from the school placement through the college-based discussions.

*AJ "You've got a lot of information and you don't know what to do with it. How to organise it? How to make sense of it? Perhaps I could address you to some questions: why have you collected that information?"*

*D "... because you told us to."*

*AJ "Have you any other purpose for collecting it? "*

*D "... to find out if they can do it."*

*AJ " Can you express that in appropriate language which will help you to fit it into what we've been talking about this morning."*

*D "What the pupils actually gained from all the teaching-learning processes."*

*AJ "To assess what they've gained from their previous experience. Why are you doing it though?"*

*D "So you know what to go on to?"*

*AJ "Right! Now that's why you're collecting the information - to make a diagnosis of their present position, what they've learnt previously, where they're at now so that you can make decisions about where they go to next."*

*"It enables you to explore the range of techniques that are available by focusing on one pupil in particular but also to gain a general assessment of the remaining pupils you are working with in the group. Clearly, each of you have focused on different areas of the curriculum, because you are working with different ages who are at different stages of development, where there are different priorities of what you think they ought to know or you feel you need to know about in order to plan their learning. It would be inappropriate for me to give a list of headings for organising the information that you have collected because that should arise from the work that you've been doing ... I think it's up to you to devise the categories, the organising concepts."*

The session proceeded with further clarification of the nature of the school and course files. The teachers also talked about the feelings they associated with being assessed by different tutors and their uncertainty about our expectations. They did not, however, press me as they had done in an earlier session to provide detailed guidelines for completing the assessed work suggesting that they were more prepared to work out their own individual solutions to these difficulties. This was not however applicable to the whole group since one teacher raised the issue of not having any data, stating that she did not have test results. This developed into a consideration of the value of different types of data in the files and the organisation of this material.

*J "Certain tests I've done I haven't found useful at all and I've got them in my*



*folder. But what does that really tell me, it tells me absolutely nothing ... it doesn't tell me anything. I thought I was going to learn something about ... if they can do this, this and this then they are at this level of thinking as far as I'm concerned it was a waste of time."*

*AJ "But you know now that perhaps this is an assessment procedure which wasn't appropriate for the information that you wanted to know about the pupils in order to develop a diagnostic teaching approach. See that's what this term's about. It's about being critical of the way in which you collect and interpret and reach conclusions about individual pupils' performance but it's within the context of the classroom, the professional situation."*

The implication of this discussion was that some of the teachers were still coming to terms with the basic requirements of the course, operating at a technical level rather than reflecting on events and engaging in self-evaluation. It also clearly highlights that for this teacher the links between theoretical components and their application in the classroom were not operationalised in practice. This raises the question of whether this was true for all the teachers or whether as suggested earlier there were defined stages in professional learning which reflected the different modes of self-evaluation proposed by Elliott (1983). This question is addressed in the following section where the evidence on the ability of teachers to link theory and practice during the school placement is presented. In addition, it was also suggested that the assessment of pupil learning was an appropriate stimulus for encouraging teachers to examine their own practice and thus engage in self-evaluation. The effectiveness of this focus is also examined in the following section with reference to the tutorials conducted from October 25th and November 13th.

### **6.2.3 Linking theory and practice: teachers' experiences of the school placement**

Throughout the course individual tutorials were conducted so that the teachers could discuss issues related to both the college component and the school placement. I was thus able to get immediate feedback on the problems the teachers were experiencing. The tutorials also provided an insight into the individual teacher's needs. The most important area to emerge in these tutorials related to the application in the school placement of the knowledge and skills they acquired in the course. This enabled me to explore the interaction between teachers' practice as evidenced in the school-based work and their theoretical understandings in relation to the college-based work. The questions and discussions which arose during the tutorials varied from teacher to teacher as might be expected in the context of providing for individual needs. In most cases our discussions was prefaced by a statement which took the following form:



*"As part of the course evaluation I'm interested in the ways in which your understanding of assessment changes and the impact of this on the way in which you assess pupils in the school placement."*

The following extracts from the tutorials are presented in chronological order so that the processes of teacher professional learning with particular reference to linking the theory of assessment and the teachers' practice in school as reported by them can be examined. The first tutorials were conducted in week 5 of the course when the teachers had made at least two visits to their placement schools.

During the tutorials with teachers on October 25th, teacher O indicated that it was very difficult to assess the science group that she was teaching. In particular, she was experiencing problems in trying to *"assess them on the experiments they do, not on understanding, but on the processes"*. She decided to make explicit the criteria she was beginning to formulate for assessing the pupils' practical work and use this as a basis for assessment. Teacher V elaborated the concerns expressed in her self report about the difficulties she encountered in assessing individuals in relation to some notion of a group standard. I suggested ways I might be able to help her develop assessment skills with reference to individual pupils by using some of the materials developed in the teacher-based assessments described previously. In particular, she chose to use the checklist designed to assess the pupil's ability to convey meaning clearly and accurately through speech. Teacher B expressed a need for more appropriate methods of assessment and teaching resources for working with primary pupils since his previous experience was in a secondary school. He undertook to use some of the structured activity materials associated with the aim that the child should be able to listen with concentration and understanding. These early tutorials were very practically orientated, the teachers wanted to know of specific materials which they could use on school placement to collect evidence of pupil learning since they were still unfamiliar with the procedures for implementing particular tests and more informal methods of assessment.

The following tutorials were conducted on 30th October 1984. The issues which emerged here were rather different from those which arose during the previous session. I asked H if she had made any initial assessments of the pupils she was working with in the school placement. She reported that she had observed and talked with the pupils and had drawn up a language checklist. At this point I suggested she might complete the checklist on the basis of her initial impressions of the pupils and review her assessments over time to see if there was any change in her perceptions of the pupils. During the ensuing discussion H indicated that she was undergoing some changes in the way she



was thinking about her role as a teacher.

*"We've come out of school and are forced to think about things we might not have been thinking about had we been in school ... we are made to think about things ... some things you do automatically in the classroom, don't stop to think, forced to take a new look."*

Teacher H was clearly feeling some pressure to reappraise her classroom activities as a result of her initial experiences on the course. She was one of the first teachers to express such views and this suggested that in practice the idea of assessing pupil learning as a stimulus for self-evaluation had some merit despite the initially negative response in course discussions from the majority of the teachers.

Teacher D was undergoing a similar process but in this case she identified the unfamiliar context as the stimulus for reappraisal. Her previous experience was in physical education, she was now working in a primary school on placement and felt like a student teacher in the classroom. She identified her needs largely in terms of finding appropriate materials and resources for teaching in this new situation. However, her comments on the work she was doing suggested that she was beginning to incorporate some of the ideas encountered in the course into the school placement. She intended to use the Pupil Record to monitor the interactions of a small group of pupils in the Slow Learner Unit as her initial unstructured observations seemed to suggest that one particular pupil had problems in making relationships with others in the class.

Teacher L was experiencing similar problems of adjusting to teaching a different age range in the curricular areas of mathematics, English and humanities. Previously she was involved in teaching physical education. She expressed an interest in exploring oral communication skills since she had noted that many of the pupils she worked with were *"incredible in speech, appalling on paper"*. She undertook to look more closely at this aspect of pupil performance using the checklist developed for the assessment of the pupil's ability to convey meaning clearly and accurately through speech.

Teacher G who had previously taught home economics noted that there was a very different way of approaching things in the school placement, with new schemes and ideas. She also commented that the course and the various approaches to assessment were having an impact on the way she approached the school placement and on her judgements of pupils. As she said during the interview:

*"I've certainly found that with the boy I'm using for my child study I watch him*



*much more closely than I would have ever done in class or watched any child."*

*"I used the Aston Index and taped, I don't think I'm going to draw a great deal from the results but the fact that I taped the conversation ... it was another revelation to me. I don't think I shall use the test. I shall use the tape .. most revealing."*

*"I've already shifted it [my assessment] from what I was told the day I arrived."*

This teacher was thus beginning to report changes in her understandings of the pupils at this early stage in the course. Teacher J was a secondary school teacher who had specifically requested to work in a primary school. Her self report identified considerable concern over the labelling of children and the role of assessment in schools. She commented during the interview that

*"I've got some charts of what they've been assessed on in school, then I'm trying to make my own personal judgement by just watching them and then nearer the end I'll give some tests as I go along, I've made out my own profile plan. I was totally against the one I was given and I'm going to try and use that one .. make notes every day and fill it up at the end."*

A number of comments which followed suggested that the confidence she thus expressed in her own understanding of assessment was not as secure as I first thought. She identified a need for more material, asked if it would be worth recording a conversation as part of the school file and if it would be possible to see a school file from the previous year. I interpreted these comments as indicating some uncertainty about the appropriateness of the work she was doing on placement and the need for reassurance that she was in line with the course requirements and my expectations.

Teacher AB was experiencing considerable difficulties in the course as evidenced below. He was a self-financing student which placed an additional burden on him. He expressed considerable confusion about the course, in particular the preparation of the school file. He was also finding some difficulty in organising and using his time profitably on the placement.

*"I've only got them for two periods a day. The first hour was basically an acquaintance session. I want to find out as much about them, their background which I recorded. Just talked to them about their interests. In the afternoon, we got talking about something which took up the whole afternoon. I did a bit of observation with the other years which wasn't very constructive at all."*

*"A lot of it was due to the school's interpretation as the Head wasn't there and nobody knew why I was there or what I was supposed to be doing even though I did try to explain. So I mingled around, looked at some of their books, talked to one or two of the children and last week I took in some Ginn books and I did some*



*reading with them which worked O.K. but I'm not really sure how that ought to be interpreted."*

All his previous teaching experience was in a small private co-educational school which would suggest that he was unfamiliar with the demands of a large state primary school. He was the only teacher to express such concerns about school placement. His responses when asked about the plans he had for assessing pupils were indicative of a fairly rigid approach to pupils.

*"On my first visit I did a very brief summary and I recorded on tape what they were telling me about themselves and I worked out a pattern in my own mind as to what kind of individual they were. I've elaborated on that and I don't think I've deviated from what my original opinion of them was."*

This teacher also expressed considerable anxiety at the thought of the tutors visiting him on the school placement. He was unusual in this regard and all the indications were at this stage that he had not yet come to terms with the aims and objectives of the course. I anticipated that he would require considerable assistance in building his confidence as a classroom teacher in the first instance before he would be able to assess pupils effectively and evaluate his own performance. This particular teacher illustrated some of the difficulties which have been in evidence when planning for self-evaluation in this way. His relative inexperience and limited competence in the classroom was such that he was likely only to move from this starting point to a basic level of competence rather than becoming an "extended professional" as conceived by Stenhouse (1975).

A number of different types of teacher response to the school placement were in evidence from these first tutorials. All the teachers were confronted in some way with the need to reappraise their activities in the school. Those teachers who had little or no previous experience of working with a particular age group on school placement were concerned to develop appropriate strategies for teaching and improve their classroom competence, whilst those who had previous experience with that age group focused more immediately on the assessment of individual pupil needs. In both cases the teachers asked for appropriate materials and resources to meet their requirements either to teach or assess the pupils they were working with.

Particular anxiety was expressed by the younger less experienced teachers, such as V, J and AB who wanted more information and reassurance about what was expected of them. Other teachers reported changes in the way they viewed the pupils which would suggest that the course and/or the school placement provided a stimulus for reflection on



events in the classroom particularly in those cases where they did not have to cope with adjusting to a very new situation, for example, teachers H and G.

The following tutorials were carried out on the 6th November. By this stage the teachers had spent a least three days in their placement schools and were developing their strategies for meeting the school file requirements. Teacher U had begun to keep a diary for each pupil. This recorded his observations of pupil behaviour, their use of spoken language, significant conversations, what the pupils drew and pictogram work. He had also completed a miscue analysis and used the teachers' materials from the Communication skills in early childhood project (Tough, 1979). On suggesting that it might be helpful to try and make some sense out of the information that he had available, he commented that

*"if I worked it out it would, if it's something imposed that would be different."*

He also made some interesting comments on the way his views of the pupils were changing as a result of some of the informal assessments he had made using the Tough (1979) materials, which suggested that, in this case at least, the initial reluctance to use such methods was being overcome.

*"The work with the picture was quite interesting, quite surprising. Three girls who usually chatted so much, they didn't do so well really, didn't seem to be able to use their imaginations or project into it and the boy who usually sits very quiet was able to cope with all aspects of it really."*

I then interviewed teacher K who described the school placement and children he was working with. He had also taped some of his work with pupils and was beginning to make sense of the information he was acquiring about individual needs and developing strategies for teaching the group.

*"I've written up reasons for doing the miscue and what I've discovered last week and have put down where I think we go next and what I think the three readers should be doing."*

He was also using a variety of methods in the classroom setting and using a diary approach to record his observations. This teacher appeared to have come to terms with the methods that were advocated within the course since he was implementing them in practice.

Teacher M previously worked in a secondary school teaching craft, design and



technology and was allocated to a primary school for placement. He commented that

*"Yes, it's different ... I'm in a completely foreign environment. I'm finding that difficult to cope with. I'm trying to keep one step ahead there ... this is so new to me. It's a different game, it's nice, the children are no hassle at all. To me they've got no problems. They're reading well and it's difficult to find any problems to study."*

This provided further evidence to support the idea that in an unfamiliar situation teachers responded initially by focusing on their own teaching rather than appraising the pupils' learning. However, he went on to describe some of the assessment procedures he had been using with the class, although he was finding difficulty in making sense of the information which he was collecting as indicated below.

*"Miscue analysis, one boy perhaps has a few problems. They think he's dyslexic ... not tested in any way ... I'm all at sea with what to do with it now."*

This teacher identified a number of factors which he saw as contributing to the difficulties he was facing during the subsequent discussion. These included the unfamiliarity of the situation he was working in, the ability level of the pupils, his lack of knowledge of appropriate curricular materials and requirements of the school placement. These factors were creating for this teacher the confrontation with his own assumptions about teaching that I had expected would be found when the teachers began assessing pupils in depth. In addition, he was finding the organisation of the assessment data he had already collected rather difficult.

P was in a similar situation to the above teacher in that he had experience of working in the prison system with young men between the ages of 16 and 21. However, he did not express the same anxieties and concerns about working with pupils in a primary school. He was operating at a different level as the following extract indicates.

*"I am observing one group and then the other, observing D [a pupil] who seems to do very little. She keeps a very low profile, strategies are interesting. This is the first time I've had time to do this. I'm going to utilise some of your strategies and then come back and ask you what to do with it. Although I'm getting a feel for the five kids."*

This teacher continued to outline his intended strategies for working with the pupils and conveyed a sense that he welcomed the opportunities of working closely with pupils of this age and did not appear to be overawed by the newness and unfamiliarity of the situation in which he had been placed.



The following tutorials were conducted on November 8th by which time the teachers had spent at least four days in their placement school. An increasing familiarity was reflected in their responses when describing their work in school. Teacher F was a middle school teacher whose self report about assessment practices indicated that her level of understanding and awareness of the issues regarding assessment was more complex than the majority of the teachers in the group. During the interview she was confident in her evaluations of different assessment procedures seeing them as tools for different purposes. She was beginning to reflect on the appropriateness of particular methods of assessment for different purposes as indicated in the following extract which referred to the use of the Richmond tests of basic skills.

*"I think we used it more to confirm what we thought any way, didn't use it to set children initially, we'd set them already before the tests. If we thought he'd been doing fairly well after a month, we'd look at his Richmond tests and if it sort of correlated with what we thought we gave him a trial and put him up ... very flexible ... I think that's the best way to use the Richmond tests."*

In her description of the placement school and the pupils she was working with F commented on the grounds which the teacher seemed to have selected the group of pupils with which she was working.

*"The others are not really that bad, verbally they're quite good, they can read quite well, I've done the miscue analysis and they're not that bad ... she's chosen two of them certainly on the grounds that they're lazy. They don't manage to do the quantity of work that the others do. One of the other girls is not very bright but she scored on the Goodenough draw-a-man, age eight and a half. She's not really that bad. It shows she's probably capable but she gets herself tied in knots."*

This teacher went on to give detailed descriptions of all the pupils she had been working with during the school placement indicating at the close of the interview how her methods had developed.

*"To start off with it was general observation, I suppose it's been more specific as time has gone on, preparation to deal with those problems and evaluation of the procedures."*

This comment suggested that she was applying the course content appropriately. However, there was no evidence to suggest in this case that she was engaging in self-evaluation or had confronted any need to change her practice as a result of the assessments she had made of pupils.



Teacher W was another who had been placed in an unfamiliar situation. She was working with primary children whereas her previous experience had been in a junior high school. She expressed her anxieties in the following way and also indicated that the perception of her role as a teacher was grounded in interacting with the pupils all the time. She did not seem to accept that it was important to create time and space for reflecting on what was actually happening in the classroom.

*"There's so many children, ages and stages all needing attention. I'm falling down on testing on everything and I don't really know what's required over this profile business. I can't really sit back and take a look at her because I'm involved. I don't really get a chance to observe. I'm an extra pair of hands I just can't sit there and do nothing."*

As the interview progressed it appeared that this teacher was beginning to come to terms with the situation although there was still room for development of her skills in assessment and their application in the classroom.

*"Well to begin with I had no idea of criteria at all. Now I'm aware that it's necessary but not very adept at doing it .. I'm not saying what I want the pupils to learn, I'm just saying how I see them."*

Teacher AL found that she had a particular pupil who was presenting her with "insurmountable problems". She had spent considerable time trying to come to grips with the nature of the specific learning difficulties which he was experiencing. She had used a range of assessment techniques including observation, tests in number and reading as well as appraising motor development. At this stage she had reached no firm conclusion about remediation strategies. Her focus was on the organisation of the information at her disposal in order to facilitate the development of a suitable programme for this pupil. As with teacher F this teacher had focused on the assessment of pupils effectively but this had not as yet prompted any self-evaluation.

Although some of the concerns expressed previously for finding appropriate teaching materials and methods of assessment were mentioned by those teachers who were interviewed during the period November 6th - 8th, there was also evidence that some were starting to consider how they were going to organise and make sense of the information they were collecting.

Two teachers were interviewed on November 13th during the workshop time available. N indicated during her interview that she had undergone a number of changes during the school placement. This teacher gave some of the strongest indications of the



way in which the course was affecting her way of approaching the teaching-learning process. She had spent some five days working with the pupils and was beginning to make sense of the complexity of the individual children, their needs and the way in which this was influencing her teaching.

*"I think that's changed already, the sort of things I was looking at, I seem to be seeing more now. I had set ideas about what I was looking for and through the course it's made me look for other things as well ... mainly in behaviour, the way kids cough sometimes when they're nervous, just made me more observant about the kind of things they do when put in certain situations. I detected with one child that she was coughing every time she didn't know a word. I probably wouldn't have noticed that before. It would never have occurred to me. I would have given her the word and carried on but I seem to be looking more."*

*"What I've done is organise myself some checklists because I found it was all too airy fairy. I wasn't really getting out of it ... I wasn't making any sense of it so I had to have something to follow. Obviously adding to that, not taking that as read but adding to it. Each child is different, each one isn't absolutely geared to each child but I thought it would be advantageous to follow something as a guide and then put in extra things."*

*"What I've assessed before has never been sufficient but now I'm thinking more along the lines of trying to understand them and what they need. I need more material to help me do it. It needs to be deeper and I can't get the answers unless I do it like this, maybe the answers I want are changing. Before the answers I got were probably enough but now they don't seem to be enough, so I've got to try other things and extend it to get the answers I want otherwise I'm not going to be able to do anything for them."*

*"At first you don't see the relevance, but when you start to use it then it becomes more relevant, if it's working it's even better."*

This teacher had dealt with the problem of what information to collect and resolved how to organise this information to provide a basis for developing her teaching strategies. In addition she confirmed for me the importance of trying out something before dismissing it as irrelevant.

C was on school placement at a special school which was outside his previous experience of teaching in a comprehensive upper school. He was constantly being surprised by the pupils' lack of ability in coping with very simple tasks. The assessment materials which he had used were also inappropriate for these children. He indicated that he *"had already worked out where they were at"* based on the very detailed assessments made by the school. He was unprepared to acknowledge that this view might change.

*"I've already made my mind up and it's difficult once you've made up your mind to sway it."*



This teacher had already expressed views in his self report which suggested that he had a fairly inflexible view on what could be achieved with the less able. He had not made much progress towards self-evaluation, despite recognising that assessment of pupils was *"more helpful in evaluating my own performance"*. This teacher was also very involved in the development of a record of achievement for a consortium of secondary schools and examining boards. However, this involvement did not appear to have much impact on his own teaching style. I perceived this teacher as exemplifying the mismatch between espoused theories and theories-in-use, since he stated that self-evaluation was important but was unable or unwilling to look at his own practice at this stage.

The teachers interviewed on November 13th were thus very different. One (N) had accepted the need for change whilst the other (C) was still resisting any change in practice despite a recognition of the self-evaluation process. Both were secondary teachers in new situations, but their responses were quite different. This emphasised for me the importance of talking to individuals in order to identify their needs and subsequently attempting to provide for their differing responses and stages of thinking.

The last teacher (AM) did not wish to be interviewed because of the pressure of time, but submitted a written account reflecting the changes she had made thus far during the course. In her account she described the class with which she had been working with since the start of the school term. The majority of the placement had occurred before she began the in-service course. However, she reported that once the college course begun that her use of assessment had *"naturally changed"*. As she noted

*"Firstly, my perceptions of general classroom activity were heightened considerably. I began to assess the children's interaction with each other not just with me, their involvement with the activities I had set. In the final session, I was interested in any feedback that they may give without any pressure or request from me, just knowing that they were moving on to a different group. This change in assessment or rather knowledge of assessment methods enabled me to build up a more complete profile of each of the pupils I was studying."*

This teacher like others on the course were clearly implementing the assessment strategies which they had learnt during the term but there was little evidence to suggest that this had made a significant impact on their ability to evaluate their own performance in the light of these assessments. This conclusion and other aspects of the processes in teacher professional learning which have been documented are now discussed in more detail to evaluate the effectiveness of the course in promoting self-evaluation.



#### **6.2.4 Processes of teachers' professional learning: discussion of results and some conclusions.**

A number of questions were raised at the start of this discussion which relate to the effectiveness of the course and school placement in promoting professional learning and self-evaluation. First, what were the concerns and issues identified by these teachers since it has been suggested previously that for teachers to engage in professional development the focus of study should be relevant to the clients' needs (Day, 1981)? Second, since it has been suggested that there are different forms of teacher self-evaluation (Elliott, 1983), was there evidence that the teachers went through defined stages in their professional learning which reflected these different modes of self-evaluation? Third, what evidence was there to suggest that the assessment of pupil learning promoted self-evaluation or were other aspects of the experience equally effective as has been suggested by others?

In consideration of the first question the identification of client needs through the teacher self report was only partially successful. The secondary teachers identified two issues of particular concern: first, the use of norm-referencing for selection, banding and streaming and second, ways of ensuring that a teacher's classroom assessments were reflected in assessments of pupil achievement, such as profiles or external examinations. Whereas the primary school teachers identified their priorities in terms of formal methods of assessment, record keeping and the communication of their assessments to outsiders. The course content was such that these issues were addressed through the lectures and workshops, as well as the school placements and were, therefore, relevant to the needs perceived by the teachers at the start of the course.

However, there were clearly other issues which emerged during the first three weeks of the course which had not appeared in the teacher self reports despite the opportunity that was offered to comment on more personal concerns. These were related to more affective dimensions of engaging in the in-service course, such as coming to terms with the demands of the course, getting to know each other and trying to work out what was required of them both on campus and during school placement. These concerns were manifest as a generally negative response to the course content and the school placement activities. The teachers were reluctant to apply more systematic methods of observing, collecting and analysing data related to classroom events, particularly those related to the assessment of pupils. In addition, the teachers initially rejected the notion that they could engage in "research" which would be of value to them in the classroom and perceived this



to be something which was done to them and to pupils by others. The teachers linked both the research activities described and the assessment of pupils to an increasing emphasis on external accountability rather than as a stimulus for self-evaluation. As noted by Elliott

*"when 'self-evaluation' gets advocated as a matter of policy in a context of public mistrust teachers may suspect an implication that they are incompetent and in need of correction - albeit self-correction - and that the real intent underlying self-evaluation policies is one of subtly increasing external control over their activities to the detriment of their professional autonomy." (Elliott, 1983, p.226)*

This view was clearly held by this group of teachers. These teachers were coping with economic cuts in education, reduced resources, redeployment and the threat of redundancy in school which made them very conscious of external control and accountability. When these factors are linked with the formal qualification which would be acquired by successfully completing the course and was, therefore, of material value to the teachers in sustaining their position within the system, it is hardly surprising that they should view the notion of self-evaluation for professional development with some suspicion.

The teachers were also experiencing considerable pressure to succeed and thus to "play the game" of professional development. However, the requirements of the course were threatening to the teachers' professional "front" (Woods, 1979) since they were required to change attitudes and behaviour in order for them to be successful. As a result the initial stages of the course tended to focus on making explicit these concerns. However, as Argyris noted, this recognition of anxiety and frustration was instrumental in preparing the teachers for a closer examination of their own practice as a means of professional development.

*"if people are not led to be at all dissatisfied with their current performance and experience no anxiety whatsoever, the learning process may not proceed at all." (Argyris, 1953, p.105)*

The strategies employed allowed for the expression of the teachers' concerns and anxieties so that the initial rejection of the course was followed by a more accepting response as the teachers put into practice the suggestions made during the lectures and they found that there was no hidden agenda regarding their competence and professionalism. It was observed that during the following three weeks (sessions 4, 5, 6) the teachers began to accommodate the expectations embodied within the course indicating that the course content and processes were in some measure meeting their



needs, even though these were perhaps engendered by the demands of the course. They tried out some of the materials and suggestions within their placement schools and found that they provided some useful evidence related to the pupils' performance, behaviour and achievements. It would appear that the level of anxiety induced was such that the teachers were prompted to act and at least attempt to engage in the course processes.

This did not mean, however, that the teachers accepted uncritically the underlying assumptions of the course since about half way through there was another expression of concern regarding the professional status of teachers. This coincided with a discussion of the changes in the 16+ examination system which the teachers considered were being introduced with insufficient time and resources to produce effective results. This led to a more general debate on the demands placed on teachers if they fulfilled an extended professional role in school and also reflected their anxieties on a more personal level as the first major assessment point approached.

By this time I had made an initial assessment of each teacher's understanding of the concept of assessment through the line-labelled concept maps as well as conducting tutorials with individuals. I was, therefore, able to gain considerable feedback related to each teacher's needs. I found through the construction of line-labelled concept maps that there were significant misconceptions regarding the use of various assessment methods in schools so that considerable time was spent on questions concerned with their purpose, validity, reliability and practicality. In this case, the identification of the teachers' needs resulted from my own assessments and interpretation of events which suggested that emphasis should be placed on certain aspects of the course content for the group as a whole. In addition, the need to focus on the application of the knowledge and skills within the college-based workshops to the classroom situation became increasingly apparent.

As the course progressed I concentrated more on meeting the teachers' individual needs through group work and tutorials particularly in relation to the school placement. It was apparent that each teacher had different priorities depending on their previous experience and current situation. The individualised nature of their difficulties was such that it was impossible for me to provide materials within the course which would have been equally appropriate and relevant to all the teachers. However, these discussions provided some measure of feedback on the individual needs which were reflected in the course provision. For example, a selection of the teacher-based assessment materials were introduced to a number of the teachers in addition to other course materials. I also



gave consideration to the use of different methods of assessment for different purposes, various ways of recording and analysing the data which was collected during school placement so that discussions focused on the integration of theory and practice and thus met the needs of individual teachers in relation to their particular contexts.

These tutorials highlighted the importance of organising time for the identification of individual needs and for providing for particular concerns within the context of professional learning. I was aware that the opportunity for such discussions was created by the provision of weekly tutorial time as well as organising small group activities within the workshop sessions. The more usual lecture type presentation would not have provided the time to explore individual needs in this way. Also, it would not have been possible to discuss the particular problems of implementation of the knowledge and skills acquired within the college-based components to the school placements which it is argued aided the integration of theory (espoused theories) to practice (theories-in-use).

In conclusion it can be seen that there were three distinct phases in the course component described here. The first was characterised by an initial rejection of the course and its methods by the teachers, which was coupled with feelings of insecurity and being threatened by the new and unfamiliar demands which resulted from the need to observe closely, assess and evaluate in classrooms and schools. The next phase was more open, the teachers began to experiment and try out new materials and methods on school placement and gained in confidence as they acquired new knowledge and skills. However, during the third phase feelings of insecurity were again expressed as the first formal assessment point for the teachers approached. This resulted in a concentration on individual teachers' needs and an increasing emphasis on relating theoretical concerns to the practical issues which arose during the school placement so that the teachers were better equipped to meet the assessment requirements of the course.

The three questions posed earlier regarding the processes of teachers' professional learning and what served to prompt self-evaluation within the course are now considered in more detail. The suggestion that the assessment of pupils might sensitise these teachers to the need for self-evaluation was, to some extent, pre-empted by the existing pressure on them to account for their right to employment. At the start of the course this particular group of teachers was all too aware of the type of self-evaluation which was part of the movement to increase the accountability of the teaching profession to external bodies such as the local education authority and central government. Thus the majority of the teachers did not view the assessment of pupil learning as an activity which might help them to



engage in self-evaluation for their own professional development. They were locked into the model of *unreflective self-evaluation* proposed by Elliott (1983) which is based on the acceptance of pedagogic tradition and external validation. The majority of the group saw the assessment of pupil learning as one way that external agencies were attempting to measure the effectiveness of different teachers and did not recognise it as a means of achieving autonomy through self-evaluation and thus professional development.

However, evidence from the tutorials suggested that there was a shift in the nature of the problems described by the teachers indicative of particular stages in the development of their professional learning. At first the major concerns related to mastering the skills necessary to use various assessment methods in the classroom context. This concern was also reflected in the college-based workshops as previously described. There was a concern for right answers and measuring success against my expectations and the course requirements. This matches the notion of self-evaluation based on an explicit knowledge of technical rules proposed by Elliott (1983) where the teachers monitor their performance in the light of explicit rules and criteria. This conception of what was required within the course was largely manifest during the tutorials where the younger and more inexperienced teachers particularly, sought reassurance that they were doing it correctly and asked for materials and techniques which would help them to fulfill the course requirements. There was no attempt on the part of these teachers to develop their own criteria or judge their success in relation to their previous experiences.

For the more experienced teachers the desire for precise rules to follow was not in evidence to the same extent. Here the teachers were more interested in developing their knowledge of assessment methods and how to use them in the classroom context. This stage can be described as one where the desire for practical knowledge "*which enables the teacher to assess and evaluate what actions are required in particular circumstances*" (Elliott, 1983, p.228) was most important. This practical knowledge was evaluated by the teachers with reference to their previous experience and those of others, in other words *a tradition* of tacit practical knowledge which had developed over time. There was still a need, as Elliott (1983) argued, for these teachers to have their actions validated by some external authority which arbitrated on the value of their practice in relation to this tradition. It was only when they had mastered this practical knowledge and no longer sought external validation that they might be considered autonomous professionals.

Thus the degree of autonomy and teacher professional development which arises from this type of unreflective self-evaluation depends largely on the degree to which the



teacher has mastered the pedagogic tradition. Thus, those teachers who were confronted with a new and unfamiliar pedagogic tradition; for example, when they moved to a different age group or subject area, found that they did not have this tacit practical knowledge to draw upon. As a consequence, they were more concerned with building up a body of tradition than becoming self-evaluative and autonomous professionals.

Once the teachers had used a number of different methods and collected data on the pupils and thus built up some tacit practical knowledge they then expressed difficulties in organising the data which they had collected during the school placement. They individually were confronted with the need to make sense of their observations and to find solutions to the problems they had identified. This stage in the process of their professional development can be characterised as being consistent with practical deliberation as defined by Elliott (1983) where teachers are faced with problems which can no longer be resolved with reference to their tacit practical knowledge. They, therefore, have to recognise the nature of the problem and deliberately choose to change their behaviours in order to accommodate the new situation.

Thus some of the teachers, when confronted with a different age group or subject emphasis, went through a process of reappraisal of their teaching skills which was expressed as a need to acquire new knowledge and skills for teaching in such new situations. Others when confronted by new situation were confident enough of their practical knowledge to engage in the process of practical deliberation focusing on pupil learning and thus moved more quickly to the next stage of professional development. This was particularly in evidence as the teachers tackled the problem of presenting the case study material and became aware of the shifts in their perceptions and understandings of individual pupil needs. It was only at this stage of their professional development that it could be argued that they were moving beyond competency to begin a process of self-development through self-evaluation.

At least three teachers in the group saw the possibilities that this emphasis offered even though the recognition of the need for change was accompanied by feelings of anger and frustration at the constraints which operated within the system which worked against this in practice (Calderhead, 1984). As Elliott noted

*"Teachers' capacities to realise their educational ends depend upon the extent to which the institutional structures provide them with the freedom to select appropriate means."* (Elliott, 1983, p.244)



The importance within the course of providing teachers with a different context and role during school placement did help to overcome this problem. However, it was still an issue for them on returning to their schools and attempting to implement the professional case study which dealt with an issue of relevance to that context. This dimension of the course would be an interesting one to evaluate in order to assess the degree to which teachers can engage in what Elliott (1983) defined as consciousness-raising self-evaluation within their own school contexts. However, such an investigation was beyond the scope of this study but is worthy of further research.

The evidence thus presented would suggest that the course provided for professional learning as intended through the use of the particular teaching strategies and curricular content detailed previously. However, for triangulation purposes this evidence should be supported by other data related to the course and teachers' professional learning. For example; it would have been desirable to observe these teachers during their school placements but this was not practicable given the demands of my other teaching commitments. However, further evidence is provided in the following section which describes the nature and quality of the teachers' professional learning as evidenced by their assessed work, in particular, the line-labelled concept maps, pupil case studies, course and school files.

### **6.3 TEACHERS' PROFESSIONAL LEARNING: OUTCOMES OF THE COURSE**

The evidence related to teachers' professional learning reported here draws on a number of sources of data in line with the principles of triangulation. These include the line-labelled concept maps completed in October 1984 and in June 1985, the individual pupil case studies, school and course files which were presented for assessment in December 1984. The teachers also wrote a commentary on changes in their understanding of assessment after completing a second line-labelled concept map in June 1985. This also served as a self-report on their own professional development. The particular issues which are addressed in the presentation of these results are as follows:

- \* an exploration of teachers' professional learning as evidenced through the assessed components of the course in order to investigate the congruence between course processes as previously described and outcomes,
- \* the quality of teachers' understandings regarding the concept of assessment identified through the use of line-labelled concept diagrams, since it is



suggested that this knowledge is the basis on which practical knowledge can be developed.

### **6.3.1 Case studies, school and course files: evaluation of assessed work.**

The final assessment of Part 1 of the course included the presentation of school files and course files based on the work completed in the first term. This assessed work also included the pupil case studies. As indicated in previous discussions teachers experienced a conflict between the theoretical insights they acquired during the course (espoused theories) and the practical implications of these in relation to their own practice (theories-in-use). The results described in this evaluation of the teachers' learning provide evidence of some aspects of their performance in order to ascertain whether the concepts, ideas and skills developed during the course and school placement were in evidence in the assessed work. The grades awarded to each of the teachers are summarised in Table 6.1. These reflect the criteria used in the assessment of the course and school files outlined in Appendix 15.4. The school file grades were distributed in the following way: A = 1, B+ = 2, B = 4, B- = 2, C+ = 3, C = 4, C- = 1, D+ = 1; whilst the course file grades were distributed as follows: B+ = 3, B = 2, B- = 3, C+ = 3, C = 5, D = 2 indicating a fairly wide spread of ability and slightly better performance on school files when compared with course files.

The differential in the performance of each teacher is also documented to provide a view of whether they performed equally well (=) in relation to the school placement and course work, or performed better (+) or worse (-) on the school placement when compared with the course work. In this way some comparison of the teachers' assessed work is possible in relation to the "practical" and "theoretical" components of the course. This analysis shows that five teachers (AL, B, C, G and K) performed equally well on school and course files; ten teachers (AB, AM, D, H, J, M, N, O, U and W) performed better on the school file whilst three teachers did better on the course file when compared with the school file (L, P and V).

Each of the teachers also completed a school and course file overview. Extracts taken from these teacher self reports and a selection of my comments on assessed work are summarised in Appendix 20. The following discussion focuses on the ways in which these teachers stated they were coming to terms with the integration of theories-in-use and espoused theories through the course and the relation between these and the school and course file grades.

**Table 6.1   Summary of grades for assessed work during Term 1**

<i>Teacher</i>	<i>School file:</i>	<i>Course file:</i>	<i>Differential</i>
AB	D+	D	-
AL	B	B	=
AM	B	C+	--
B	C	C	=
C	C	C	=
D	B+	B-	--
F	. Did not submit since course not completed due to pregnancy.		
G	B+	B+	=
H	C+	C	-
J	C+	C	-
K	B	B	=
L	B	B+	+
M	B-	C+	---
N	C	D	---
O	C+	C	-
P	C-	B-	+++
U	A	B+	--
V	C	B-	++
W	B-	C+	-



For the following analysis the teachers have been grouped into two categories: those who achieved grade B or above (AL, AM, D, G, K, B, and U) on their school files and those who achieved B- or below (AB, B, C, H, J, M, N, O, P, V and W). The school file grade has been used since this was where the teachers were required to link theory and practice and thus provided the most appropriate evidence of success in terms of the course. The comments made by the "successful" group of teachers indicated that they were all aware of the need to examine their own practice and the course had met their needs since it had provided an opportunity for them to ask questions, try out and evaluate new approaches in a relatively non-threatening way. This was substantiated through their comments on the course, school placement and overall coherence of their assessed work. As one teacher (AM) commented

*"The course has given me time to begin to consider the various ways I could start to improve my teaching. I have to value the part that knowledge can play in the professional work of every teacher. As a result of this knowledge I would now set myself pedagogic aims, which would have been impossible prior to coming on this course."*

Another "successful" teacher (K) noted that *"if a child is not apparently learning, then we should examine our teaching methods critically."* There was ample evidence to suggest that he had done exactly this within his own context. Other teachers expressed similar sentiments with reference to the course and school placement suggesting that for these teachers the experience had promoted their professional learning and moved them to at least question some of their accepted practices (tacit professional knowledge).

The more "unsuccessful" teachers merit further consideration. It would appear from the course and school files that these teachers experienced particular problems in adjusting to the course requirements. Particular reasons emerged from the overviews which illuminate some of their difficulties. Teacher AB, who performed inadequately on the course, and teacher V identified their lack of experience as the main contributory factor in their failure to meet the requirements. Teachers M and W found that adjusting to new situations took up most of their resources so that they were unable to develop effectively what they identified as a need for their own professional development. Interestingly, these two teachers were only ones who achieved a B- on school file work.

Teachers B, C, H, J, N, O and P produced very descriptive accounts which did not reflect any aspect of personal or professional learning in the context of the school. They produced the kind of academic essay or descriptive case study which they appeared to think the course tutors expected on the basis of their previous experience in initial



training. They did not reflect on their own performance in any way and did not make recommendations for changes in practice. In these instances, the starting point of the teachers as well as the nature of the school placement contributed to their lack of success. In those cases where significant adjustments needed to be made to a new situation, the professional learning of the teachers was not as much in evidence in the course work. In those cases where the teachers were still developing a basic competency in the classroom this was even more pronounced.

In this exploration of teachers' professional learning it was found that the outcomes of the course were varied with a group of "successful" teachers emerging on the basis of the assessed work who reported significant changes in their perceptions, knowledge and skills in relation to the professional context. These findings further support the conclusions drawn in the previous section regarding the facilitation of teachers' professional learning since the successful teachers were those who had reached a point where their technical expertise was sufficient to sustain them through the initial period of uncertainty as they addressed the need to change; in other words they had achieved a level of tacit practical knowledge which they could use in the solution of their "problems". However, where teachers, for whatever reason, felt that their professional competence was under some form of threat the assessed course work did not provide evidence of professional development. Thus it can be concluded that for this group of teachers the course was of limited success in promoting professional learning on the basis of their performance at the end of the first term. The degree of success was however closely related to their starting point and degree of difference between their previous experience and their current situation.

### **6.3.2 Teachers' understanding of the concept of assessment: an analysis of line-labelled concept diagrams.**

The purpose of the following analysis is to explore whether teachers who have shown a theoretical understanding of the concept of assessment reflect this in their professional context. Line labelled concept maps were used to assess teachers' understandings of the concept. These were constructed by the teachers in the Autumn and Summer terms. The teachers worked in groups and generated a number of key words they associated with assessment. The following list gives all the key words which were used: profiling, assessment, measurement, judgement, quantification, formative, summative, appraisal, observation, monitoring, examination, continuous, final, testing, formal, informal, norm-referenced, criterion-referenced, ratings, diagnosis, checklists, interpretation,



records, screening, reports, coursework, evaluation, internal, external, process, product, quantity and quality.

The instructions given to the teachers required that they should group and organise these key words in such a way as to clarify their meaning in relation to assessment and link them with labelled lines to show the nature of the relationship, if any, between them. The difficulties encountered by the teachers in constructing their first line-labelled concept diagrams were well illustrated in the course discussion as detailed below.

*AM "Understanding some of the words, if you didn't understand them you could leave some of them off, O.K.?"*

*D "I found it very complicated, I did a simple one first and I did a more complicated one and in the end I went back to the very simple one for me. I got lost with all the words and how you could in fact link them up. [I] abandoned a lot of them."*

*L "There's a similarity with a lot of words, I don't know if you work all the time with the terminology you know, its specific meaning for, just off the top of your head seem very similar."*

*J "The main trouble is that most of the words should be used and yet having room to actually put them in and relate them ..."*

This transcript indicated that the teachers, at this stage of the course had not clearly worked out their understanding or conceptualisation of the term assessment. This was expected since many of the teachers had expressed unfamiliarity with the field in their self reports. However, there were some indications here of the ways in which the teachers tried to make sense of the large number of terms which had been generated. One response involved simplifying the task by reducing the number of words included in the concept map. Another response was to group similar ideas to clarify understandings of terms. The results of the analysis of the first line-labelled concept diagram constructed in late October 1984 are presented in Table 6.2 overleaf. This table includes the number of key words used by the teachers, the structure of the line-labelled concept map, the number and labelling of the links and any organising concepts which the teachers used in presenting the line-labelled concept map. Several different approaches were found in the line-labelled concept maps. The most common approach was the construction of a linear or linear branching arrangement where the teachers differentiated between formal and informal methods of assessment but used less than two-thirds of the key words (D, G, K, L, M, F). These most closely approximated the course content and indicated that these teachers were beginning to make some sense of the quite complex concepts they were being asked to address.

**Table 6.2 Analysis of line-labelled concept maps (Autumn Term, 1984)**

Teacher	Number of key links words used		Structure	Line labels					Organising concepts, key words, comments
				DS	DF	DP	DL	C	
AB	Did not complete line-labelled concept map.								
AL	31	0	List no links	-	-	-	-	-	Evaluation, assessment methods, purposes.
AM	27	28	Composite	-	-	-	-	-	Five groups of key words
B	32	42	Network	-	-	-	-	-	Assessment, testing, records.
C	Did not complete line-labelled concept map.								
D	12	13	Linear branching	-	-	-	-	-	Assessment: formal and informal methods.
F	21	18	Linear branching	-	-	-	-	-	Assessment: formal and informal methods
G	15	12	Linear	3	2	3	-	-	
H	24	21	Circular	-	-	-	-	-	Evaluation: formal and informal methods.
J	27	24	Network	-	-	-	-	-	
K	17	10	Linear branching	-	-	-	-	-	Some linking of key words.
L	31	34	Linear branching	-	-	-	-	-	Assessment to evaluation.
M	31	21	Linear branching	-	-	-	-	-	Assessment: formal and informal methods.
N	Did not complete line-labelled concept map.								
O	32	35	Network	1	-	-	-	-	Assessment
P	32	0	Unlinked balloons	-	-	-	-	-	Novel presentation: cartoon format.
U	31	40	Network	-	-	-	-	-	Formal/ Informal
V	18	8	Circular	1	3	1	-	-	Assessment of individual needs.
W	31	19	Network	1	-	5	-	-	Grouping of key words.



The second common approach was the construction of an elaborate network which used most of the key words but did not identify the relationship between them (B, J, O, U, W). These teachers appeared to be lost in the detail of the task without any conceptual framework with which to make sense of the course content. They were not aware of the underlying concepts and had not developed a clear understanding of assessment. It was likely that they would experience more difficulty in applying the knowledge they were acquiring in the course since they had not developed a set of general principles to guide their practice.

A range of individual responses characterised the remaining teachers. Two (H, V) produced circular arrangements. Teacher P grouped the key words in balloons but the relationship between these groupings was not elaborated. AL listed key words under evaluation, assessment methods and purpose but did not identify the relationships between the key words with any links. AM produced a composite map grouping key words under organising concepts and linking these to assessment as the key concept which was then linked through a series of steps to evaluation. Three teachers did not complete the line-labelled concept map (AB, C, N).

Only three of the teachers made an attempt to solve the problem of simplifying the map by grouping the key words of a similar nature under organising concepts. One of these teachers (AL) did not link these groups of key words, although the groupings reflected the key concepts appropriately, whilst a second (P) produced a disparate collection of key words which were grouped erroneously. However, the third teacher (AM) produced a composite line-labelled concept map which incorporated both network and linear branching arrangements. This was perhaps the closest to my conceptualisation of assessment.

An interesting comparison can be made between those teachers who performed well in the course and school file assessments and those who completed the line-labelled concept maps which most closely resembled my conceptualisation of assessment and the taught course component. The teachers who achieved a grade B- and above on their course work (with two exceptions - U with grade A and W with grade B-) all constructed linear or linear branching arrangements or appropriately grouped, composite diagrams. The remaining teachers tended to produce complex network arrangements or circular diagrams. This would suggest some congruence between the level of understanding as shown in the line-labelled concept maps and success as demonstrated by course work. This relationship did not, however, appear to be as strong in relation to the school files.



This would seem to suggest within the limitations of this very small sample of teachers that conceptual understanding alone does not necessarily promote application within the school context.

A second line-labelled concept map was completed during the third term. The analysis of these concept maps was done by the teachers themselves (Table 6.3). They also made comparisons between their understandings of assessment at the outset of the course and towards the end of the taught course, completing a second self report. In this way the teachers themselves analysed and evaluated the differences between each of the maps in order to provide feedback on the extent of their professional learning during the course. This activity formed part of the taught course which focused on introducing the teachers to a range of methods of data collection and analysis.

Similar structures emerged in the second line-labelled concept maps as had been found previously. Six teachers (AL, AM, D, H, N, O) completed linear branching concept maps but in contrast to their previous attempts more complexity was introduced by the use of more key words, organising concepts and the use of line-labelling. The linear branching arrangement was the most common format and these teachers tended to differentiate between informal and formal methods of assessment leading to some form of evaluation or judgement. This clearly reflected the course content discussed in section 6.2. Only two teachers used a linear representation (B, G). These teachers had overcome the problems of representation by reducing the number of key words used in their maps in their previous analysis of line-labelled concept maps. Both of these teachers now used more of the key words than previously and attempted some form of grouping to simplify the number of links in the concept map.

Four teachers used the composite structure format (AB, K, L, W) which was arguably the most complex way of showing understanding of the concept of assessment since it involved grouping similar terms under organising concepts and showing relationships between these organising principles which allowed for both feedback loops and linear sequences. None of the teachers who produced this type of line-labelled concept map had done so previously. It is suggested that the use of this form of representation showed a development of understanding for these teachers and some congruence between the taught course and their conceptualisation of assessment. However, a different pattern emerged with three of the concept maps which was a co-concentric representation involving one linked circle inside another with links between each of these circles.



**Table 6.3 Analysis of line-labelled concept maps (Summer Term, 1985)**

<i>Teacher</i>	<i>Number of :</i>		<i>Structure</i>	<i>Line labels</i>					<i>Organising concepts , key words, comments.</i>
	<i>key words</i>	<i>links used</i>		<i>DS</i>	<i>DF</i>	<i>DP</i>	<i>DL</i>	<i>C</i>	
AB	27	4	Composite	-	-	-	-	-	How, why, interpretation.
AL	21	9	Linear branching	2	5	4	2	-	Assessment to evaluation informal and formal.
AM	25	25	Linear branching	-	-	2	-	-	Assessment methods: informal and formal.
B	29	9	Linear	2	4	2	1	-	10 groups of key words, no organising concepts.
C	Did not complete line-labelled concept map.								
D	31	29	Linear branching	6	5	7	2	2	Process v. product to evaluation.
G	20	11	Linear	-	3	2	1	5	Key words linked in prose statement.
H	23	21	Linear branching	10	5	1	5	-	Assessment methods: formal and informal
J	22		Co-concentric	-	-	-	-	-	Unclear and confused map.
K	22	16	Composite	-	13	2	1	-	Assessment central concept
L	28	17	Composite	-	1	6	-	5	Assessment and evaluation
M	30	20	Linear branching	-	7	4	-	2	Assessment to school and teacher accountability.
N	Did not complete line-labelled concept map.								
O	22	20	Linear branching	1	3	6	3	1	Assessment, evaluation, interpretation.
P	30	31	Co-concentric	-	-	-	-	-	Evaluation, profiles.
U	26	45	Co-concentric	8	7	3	-	-	Assessment methods: formal and informal.
V	19	12	Circular	5	2	5	-	-	
W	31	32	Composite	5	3	4	3	4	5 groups of key words: assessment, formal and informal methods, ways of recording and reporting.

Codes:

DS - Descriptive structural  
DF - Descriptive functional

DP - Dependent functional  
DL - Dependent logical

C - Composite

The teachers who used this format previously represented the key words in network form with many cross over links which led to lack of clarity (J, U) or in the case of P as a series of unrelated "balloons". It is suggested that these teachers were still trying to make sense of the relationships between the key words. The complexity of the structures indicated a holist perspective (Marton and Saljo, 1976) even if there was a lack of clarity in their understandings.

Only one teacher (V) retained a circular representation although on this occasion she used more links and labelled these more fully. Two teachers did not complete the line-labelled concept map. Teacher F had withdrawn from the course, teacher C did not submit either of his concept maps. Interestingly this teacher was one identified in previous discussions as being the most resistant to the need to change his practice in the light of school placement experiences. The relationship found between the course files completed at the end of the first term and type of line-labelled concept maps constructed in the third terms did not reflect that which was found previously. Data on assessed work for the third term was unavailable so it was not possible to make a similar comparison between course and school file grades and the understandings of assessment in evidence in the third term.

The teachers' comments and evaluation of the line-labelled concept maps provided another source of information on the way in which they saw their understandings develop over time. AL found that as she worked through her second line-labelled concept map the ways of grouping key words which were similar or synonymous became clearer and different forms of representation emerged. AM identified that there were several unattached collections of unrelated ideas in her first line-labelled concept map and this was judged by the teacher to indicate a *"lack of understanding of the basic concept of assessment"*. The second line-labelled concept map showed more links and in the opinion of the teacher showed *"something of a development in understanding the concept"*.

Teacher B described his first line-labelled concept map as *"a rather confused view of assessment"*. His second line-labelled concept map showed a clearer view with assessment leading to some kind of *"evaluation via various methods and techniques leading to certain conclusions for a variety of reasons"*. Teacher D commented that the first assessment diagram meant very little to her, *"as shown by the limited links and related comments used"*. She reported that for the second line-labelled concept map *"the light had dawned on me and I saw assessment in two main areas, one process-based and one product-based, with related evaluation taking on a process or product view."*



Teacher L described her first line-labelled concept map as "*a helpful exercise although the result may not be all that coherent*", whereas her second attempt represented a more coherent combined picture. Teacher M identified that although his second line-labelled concept map was more complex, it also showed clearer distinctions between different forms of assessment. Teacher O reported that her second line-labelled concept map showed more structure and reflected more thought and understanding than before.

Teacher P described his first line-labelled concept map as an attempt at a novel presentation in that he drew balloons labelled with some of the key words. He described this as resulting in an "*artificial grouping of concepts and a crude composite assembly*". In the second line-labelled concept map he described his attempt as "*far more flexibly inter-related*". The analysis of this map was difficult since no line-labelling was in evidence but the conception of inter-relationships was described with reference to more effectively grouped concepts.

Teacher U identified that his first attempt distinguished between formal and informal methods of assessment whilst the second focussed more on observation. This echoed the development of this teacher's skills in this area and his greater use and familiarity with appraisal and monitoring techniques. Teacher V did not consider that she had changed in her understandings of assessment, expressing a preference for her first line-labelled concept map since it emphasised aspects she valued such as the assessment of individual pupil needs rather than embracing other purposes of assessment. This teacher was one who had also experienced particular difficulties in adjusting to the demands of the school placement.

Teacher W's first line-labelled concept map was based on "*careful deduction of meanings from books and lectures*" since she reported that she "*had no real conception of meanings and links at a practical level.*" In compiling her second line-labelled concept map she made no reference to books and reported that she spent very little time on this exercise. However, the second attempt was clearly more complex and reflected the development of her own conceptualisation. Thus the teachers were able to view the development of their own understandings at the end of the course with some measure of objectivity and critical self-appraisal and, with the exception of teacher V, viewed their second attempt as representing a development in understanding as might be expected.

These results lend some support to the view that the two types of thinking: holist and serialist, identified in previous studies (Marton and Saljo, 1976) were reflected in the two



predominant ways in which this group of teachers constructed the line-labelled concept maps. Those teachers who used the network approach attempted to include all key words and produced many links between them, thus reflecting a holist view of assessment. This often resulted in a confused picture which did not appear to clarify their understanding of the field. The teachers who produced a linear or circular arrangement appeared to reflect serialist thinking with one key word leading on to the next in some logical order.

However, in order to produce such linear or circular arrangements the teachers often disgarded key words since it was difficult to fit all of these into an ordered sequence. The most complex form of line-labelled concept map was the composite structure which incorporated elements of both holistic and serialist thinking in the solution of this problem. There was clearly a development in the teachers conceptualisations over time but whether this facilitated the integration of theoretical knowledge of assessment with the practical work carried out in school is open to debate.

It would appear from this small sample of teachers that those who were successful in the school placement were not necessarily those with the highest level of theoretical conceptualisation of the course work. For example, the more successful teachers on the school files included: teachers AL, AM, D, G, K, L and U. On the basis of the first term's data teachers AL and AM produced some grouping of concepts thus representing an attempt at producing a composite structure; K, L and D produced linear branching line-labelled concept maps; U produced a network and G produced used a linear representation. There would, therefore, appear to be little relationship between the two forms of outcome. Similarly those teachers who were more successful on the course work in the first term AL, G, K, L and U produced an equally diverse range of line-labelled concept maps. There did not, therefore, appear to be much congruence between these forms of outcome. The possible reasons for this were not evident from the data which had been collected but this did present a particular problem for future course work since it appeared that theoretical knowledge of this type was not of immediate relevance to the work of the teachers in schools.

This evidence together with the data generated by my own self-evaluation which is described in the following section would suggest that there was a need to reappraise this part of the course as a means of facilitating the professional development of teachers if this was to be seen in terms of practical knowledge which was of benefit to their work in schools. In the following section the account of my self-evaluation is presented as evidence of the outcomes of this study for my own professional development.



## 6.4 SELF-EVALUATION: AN ACCOUNT OF PROFESSIONAL LEARNING

The following account documents chronologically significant events in my own professional learning as evidenced through the course transcripts and diary.

### *Session 1: 27th September 1984*

I considered it was important to enlist the support and cooperation of the teachers participating in the course before embarking on monitoring events during the workshop sessions. Since the evaluation proposal was not prepared for this session and this was my first opportunity to meet the teachers, I decided not to record it on audiotape. It was not possible, therefore, to check for the match between my intentions and actions in this case. The session was planned to introduce the concept of appraisal as described by Tough (1977). This content was reviewed on 4th October in the following way.

*"If you remember last week we talked about the concept of appraisal and we looked at the way in which you could write about a pupil at various times, we looked at the ways you could make that a little bit more systematic by dividing it into talk with other pupils, the nature of that talk, general attitude and behaviour at various times and the final example I gave you which was the checklist approach, where particular categories were identified and ticks made against particular times. So those were three methods of recording information about a pupil, in this case devised to make an appraisal of their use of language."*

*"I did not give you the purposes of language that you would be looking for with this particular method. I did that intentionally because what I was interested in getting across to you was the difficulties there are in actually recording information about the pupils based on observation and the different summaries or conclusions that people reach using different kinds of methods because they have different perceptions, different ways of responding to individual children, that they come to the situation with as a result of previous experience ... own ideas, knowledge, expectations and so on."*

This extract indicated that the planned activity was introduced and formed the substantive part of the session. It also illustrated one of the basic premises which I incorporated into my espoused theory which was to emphasise the need for accurate observation and appropriate methods of recording so that teachers could make more valid and reliable assessments of pupils. This central consideration is reiterated throughout the sessions which follow.

### *Session 2: 4th October 1984*

In the original course outline it was planned that the teachers should develop their



observation skills through the use of the Pupil Record as an example of a systematic observation schedule (see Galton, 1978b). However, it was necessary to start the session by explaining the evaluation project since I wanted to inform the teachers of my intentions from the outset. Also, in order to match their needs more closely it was necessary to collect some preliminary information on their previous experience, understandings of assessment and assessment practices within the schools where they had previously worked. As detailed previously information was requested in the following areas:

- a) an account of the teachers' previous school experience,
- b) a general description of assessment practices within the school and those which the teachers particularly used,
- c) a personal statement identifying areas of strength and weakness,
- d) identification of important aspects of assessing pupils' school performance.

Previous experience in working with both the development groups of teachers and the application groups of teachers during the Oracle research programme had suggested that using such written communications and requests alone was often an unsatisfactory way of gaining teacher cooperation. In addition, as noted previously in chapter 3, teachers were often lacking in confidence and unwilling to ask questions or seek clarification if asked to do something called "research". Therefore, I anticipated that the teachers might need time to discuss the project and the opportunity to clarify their uncertainties. By using course time for the completion of the self reports I was able to discuss any points of concern during the session. The audiotape transcript of this session suggested that setting aside this time was necessary. This may have been due, in part, to lack of clarity in the written statement presented to the teachers but a number of issues did arise as the teachers completed their self reports. I asked if there were any questions about

*"what my intentions are, any problems you might have with this particular proposal?"*

One teacher [B] responded with

*"This is presumably a two way thing you would help us as we help you."*

I took the opportunity at this point to spend some time explaining and clarifying my intentions. The following extract is quoted verbatim from the tape transcript.

*"I would hope so because what I am doing is evaluating what I provide for you in terms of the materials and the work to see if there is any impact on the work and the way your ideas, your perceptions of pupil achievement, pupil performance*



*(develop) and also if it makes any difference in the way you operate within the classroom. That's what I'm interested in looking for - those are my areas of concern. Clearly it has an impact on you because I have to be prepared for the course, I have taught it before - but I'm going to be monitoring my own inputs closely. I haven't put that on the end but obviously my commitment is to actually make a record of all the course provision, the handouts I prepare, the nature of the contact we have during the taught element, finally perhaps videotaping it and using the transcripts from the tapes in order to get a feel for what the course was like. My purpose is to try to match your experience of the course with my intentions, my aims, what I did - and whether my intentions, my aims were actually fulfilled by the way I delivered the programme for you and whether you perceived it in the way I intended."*

[Inaudible teacher contribution]

*"I am doing, in a sense, what we are expecting you to do in terms of monitoring your own intentions, the way in which you deliver and whether in fact it was received in the way it was intended."*

This clarification was intended to alert the teachers to my espoused theories and theories-in-use related to self-evaluation, offering what I proposed to do as a model of this approach. I then identified the use of course time for the completion of the self reports within the session as a problem and suggested that the teachers might have to finish their self reports later. Since I had made no allowance in planning the in-service course for time to introduce the evaluation project or complete the self reports, this caused me some initial conflict in relation to my dual role as lecturer and as researcher. In the former role I had a responsibility to provide the teachers with the appropriate skills and knowledge as described in the course outline. In my role as researcher it was necessary to set aside some time to explain the nature of my research and elicit the support and cooperation of the teachers. I also needed time for the teachers to provide the additional information which would enable me to investigate their perceptions of their needs and the course in progress. A diary entry also supported this view since I stated that *"I felt concern for the use of course time in conducting this research"*.

It was an unexpected realisation that I was concerned about the amount of time it took for the teachers to provide this additional information on their understandings of assessment, previous experiences and areas of concern. My espoused theory would have suggested that this initial data gathering exercise was valuable since I believed in the importance of observation, diagnosis and assessment of learners in order to match their needs more closely, whereas in practice I was more concerned with meeting the course requirements as outlined in the programme summary. This provided some insight into the conflicts which might also be experienced by teachers in using informal methods of assessment and engaging in classroom-based or school-based evaluation.



### **Session 3: 11th October 1984**

The original intention for this session was to introduce the teachers to techniques of informal assessment. The checklists developed within the Schools Council project, **Progress in Learning Science** (Harlen *et al.*, 1977a) were selected as an example of such methods of assessment. However, this was not covered within the session as planned. Two factors contributed to the changes in this session. First, the teachers revised their self reports since I considered it important to provide them with an opportunity to read, correct and add material to the summaries which I had made. The time constraint identified during the previous week was thus further exacerbated by this activity. Second, the loss of teaching time resulting from the introduction of the evaluation and completion of self reports meant that the majority of the Pupil Record categories were still to be covered. My diary again reflected this concern since I noted on 16th October the following comments.

*"My separation of the roles of lecturer and researcher may well parallel the teachers' problems associated with the use of observation and assessment procedures based on the informal model. For example, how valid is this use of time, how much certainty can I place on the results, can't I go with my own judgements, why do I need to check them out? Does this help my professional development?"*

Thus the teachers spent the first ten minutes of this session correcting their self reports. I then continued the session with a review of the work on the Pupil Record which I had introduced briefly the previous week. The time constraint was again identified in the transcript of the workshop since I stated that the previous week was *"fairly rushed"*. The main points to arise during this review was the discussion of time sampling as a way of selecting observations to record, given that it was impossible to record all aspects of the classroom and pupil behaviour. After reviewing the pupil activity categories the teachers then coded written descriptions of pupil behaviour using all the Pupil Record categories. As this activity was in progress, I took the opportunity to reflect back to the teachers my interpretation of their concerns regarding the uncertainty of what was expected of them as learners within the course.

*"Somebody commented at the beginning of last session - now we know how the pupils feel - one of the things I think is quite important is that we should be able to learn to reflect on our own learning and use that to inform the way in which you plan and structure your teaching - you are learners in this situation, you're coming into this situation probably not having been involved at that end - if you like - for a period of time and you may find that the adjustment is quite traumatic and painful - having to get used to a new situation - that's one of the things I think you ought to*



*think about ."*

I selected this extract since it illustrates my espoused theories well, in particular, the importance of taking into account the impact of such a course on the feelings of the teachers in line with the recommendations noted previously (Day, 1981). However, the shift from we to you through the passage was an indication that I still had to come to terms as provider of the course with the trauma of self-evaluation which I was identifying for the teachers. My theories-in-use were not as congruent with my espoused theories as I would have liked to believe. This continued in the following way emphasising the importance of processes as well as content.

*"Keep in mind you can learn a lot from your own experience of being on the course, as a course member, as a learner, as much as the course work we do."*

As noted previously the teachers had responded negatively to the course in the first instance. This was unexpected as I noted in my diary at the time.

*"I seemed to be fighting a hidden agenda initially which finally came out. Issues I perceived were first accountability and the publication of exam results. Do 'they' know what they are looking for. Second, we know our pupils too well, therefore we don't need any system for organising our thinking. Third, the teacher's professional role is changing, too many demands on time, how do you cope? Fourth, resources are not available to do the job."*

These concerns were equally applicable to my own professional development where severe time constraints were operating as the college was undergoing considerable change as a result of government policy and the full-time in-service B.Ed. (Hons.) programme was being evaluated for revalidation by the Council for National Academic Awards. As a result of the discussion during this session and my reflections on events, the course was reorganised to take account of the concerns expressed by the teachers and reduce my own concerns regarding the time constraint and conflict between my roles as lecturer and researcher. My intentions in modifying the course were to raise a number of questions centred on the issue of effective teaching such as "how do you know you are effective" and "what measures are there of effectiveness"? I also intended to direct the following discussions to look at the purposes of assessment and provide the teachers with the opportunity to devise their own assessment methods.

#### *Session 4: 18th October 1984*

This session began with an introduction which outlined for the teachers the decisions I



had made during the intervening week and the changes to the course for the rest of the term. This introduction is given in full since it describes the nature of the changes I proposed, my feelings about the previous sessions and the reasons I offered to the teachers for these changes.

*AJ "As a result of the discussion we had last week I've reviewed the programme that I had intended we should follow. I think what you're addressing are certain important issues relating to, firstly, your role in assessment as a teacher, what part you should play in that and secondly, what I became aware of as I was going over what we were talking about last session, was a comparable problem to the one that was facing me in talking about and working with you in doing my research. So I was reflecting on the fact that I was experiencing much the same kind of conflict as I saw you experiencing in terms of assessing pupils and your role as a teacher.*

*Now my justification for researching in this area is that I believe it makes a difference to the way that I teach or lecture. In other words the way I programme what I'm doing, my expectations and so on are affected by reflecting on what happens in these sessions. Could I just suggest to you that perhaps you might be aware of that kind of conflict in me and be aware of that conflict within yourselves ... it may be that you don't actually agree that formalising and making explicit, in the kind of way I was suggesting you do in the assessment of pupils would actually make a difference. You may actually believe it will actually confirm what you already know. I suppose I'm asking you to suspend your judgement of that until you've actually done it, to give it a try.*

*In the sense that I'm finding I'm having to review and reappraise what I'm doing in the light of things which come up during the sessions I'm suggesting that you might like to do that too and suspend your disbelief and have a try in the light of that reflection on my part. I do understand what it feels like, the conflict, should I really be doing this or should I discontinue. I believe I ought to be doing it. I've reworked the plan of the term, if you'd like to have a look at these and see if there are any comments you would like to make."*

This statement was a turning point for me and supported the suggestion that the identification of the learners' needs through the close observation and reflection on events could provide a more appropriate way of confronting any discrepancies between my own espoused theories and theories-in-use and that this might also apply for the teachers' professional learning. The major change in programming related to the organisation of groups and workshop sessions to consider a number of issues related to methods of assessment linked to a consideration of purposes, including accountability. However, before we could proceed with the new programme it was necessary to complete the work already started two weeks previously on the Pupil Record. This took up the majority of the session.



### *Session 5 and 6: 27th October and 1st November 1984*

During session 5 the teachers constructed the line-labelled concept maps and began the associated workshop session. In session 6 the line-labelled concept maps were reviewed. I presented my conceptualisation of assessment to the group as an example of how I saw the content of the course and explained how I had constructed this. I then asked for questions or comments on the content structure diagram. None were forthcoming, but the teachers did write down my conceptualisation of the course for reference.

As the teachers were completing this I explained the relationship between observation, interpretation and evaluation with reference to different assessment methods. This was directly related to the issues of the validity and reliability of assessment procedures which I intended to discuss during this session in order to help the teachers in the application of different assessment methods to the school placements. The following definition was given "*the validity of a test can only be judged in relation to its purpose.*" My elaboration of the concept of validity continued with a consideration of construct validity. The comments that ensued indicated that the teachers were surprised by my statement that different tests which were designed to measure the same aspect of pupil performance or achievement could yield different results. This led to a discussion of error in educational measurement and assessment.

This was not my intention at the outset of the session but clearly the need was being expressed by the teachers for clarification of the terms used in the discussion of formal assessment procedures, particularly in interpreting standardised test scores. The teachers pursued this discussion with reference to why we assess and test pupils which were considerations directly related to the questions that had been raised in the previous workshop session. In addition, some of the teachers had raised the issue of norm-referenced tests in the tutorials. This was clearly an issue for the teachers and I interpreted this as a response to the school file requirement to use, review and evaluate different methods of assessment as well as a concern about assessment in general.

Different purposes of assessments were identified in the ensuing discussion. Distinctions were made between norm-referenced and criterion-referenced assessments and the following extract illustrates the concerns that these teachers expressed about the self-fulfilling prophecy and the impact of their expectations on the pupils. My interpretation of this was that the teachers were beginning to articulate the relationship between their own teaching and the pupils' learning and questioning the match between



their expectations and the abilities of the pupils they worked with as a result of their recent experiences on school placement.

- AJ *"In order to make a judgement about quality you have to compare it with something ... in order to make sense of it, in order to know where to go. We all have standards implicitly. We all have expectations of what children should be able to do at particular ages, we may not agree about that, but we do have that underlying notion about what children should be able to do at particular stages."*
- D *"When we draw up schemes of work we're thinking of what we hope them to be capable of, aren't we? Our areas may be different but our expectations come through, what we expect the children to be able to do in that year."*
- AJ *"Whatever you plan to teach has embodied in it expectations of what the pupils should be able to cope with ..."*
- AM *"...which is compared with what you've taught before ..."*
- D *"We say this year's better than that year, that's a marvellous year they've reached that standard now ..."*
- J *"I don't think that's always a good thing. If you set standards children only work to those standards and depending on your teaching experiences your standards might be too low so you might be pulling the children down by doing this sort of thing."*

In the following extract various purposes of assessment, which were previously identified, were linked with the practical classroom situation where the teachers were confronting the problems of collecting, recording and using assessment data in their case studies. This aspect of my espoused theory is illustrated by the following statements.

- AJ *"... What are appropriate levels of expectation? What are appropriate goals? My feeling is you can't actually work out what an appropriate goal is until you've got some picture of what the child is doing at the moment."*
- AJ *"The issue you are raising is an important one, which is, why are we assessing in the first place? So this notion of why we're actually doing it is very important in understanding and making judgements about the quality of the assessment procedure you're using."*

The session continued with a consideration of different methods of assessment and their appropriateness for different purposes. In particular, the developments with graded tests and their associated problems were considered. This led to a discussion of the value system underlying selection within the education system. This discussion was very wide ranging and drew out a number of different perspectives regarding the relative merits of the selective system and comprehensive schooling. This was an issue which was close to



this group of teachers as was suggested in previous discussions regarding accountability.

My interpretation of this concern was in the context of the effects of secondary reorganisation on one third of the teachers in the group. As noted previously some of the teachers were on the course because they were given secondment to upgrade their qualifications as a result of the redeployment of staff by particular local education authorities. Although this part of the workshop was not planned, I believe the teachers found it valuable to discuss the issue of selection. This also enabled me to introduce the notion of predictive validity to the group with reference to a problem arising from the teachers' concerns. I then reviewed the two different forms of validity: construct and predictive which had arisen in these discussions. The concept of content validity was then introduced and we considered the validity of the 'O' level examinations. The session was summed up in the following way:

*AJ "There's some important issues that have come up so far. I showed you my conceptualisation, my structure diagram which tried to relate the various bits and pieces together, that's the way I see it. We talked a bit about the problems of making judgements of quality, that's really where most of our discussion has centred today. It's about making judgements of quality, what criteria you use to judge quality. I think a lot of the discussion we've been having centres on the fact that some of us value different things or rather than valuing different things, we have different priorities. We have a different set of priorities in our aims, in what we think we should be doing. So that when somebody says 'this is a bad teacher' it may actually be a statement which says that 'I don't like the values of that person'. If someone says 'that's an ineffective teacher' it may be they don't like what they're doing but they may be quite good at doing it. I think that's something that's quite important when we think about evaluating anything is making clear the criteria we're using to make the judgement because then people can argue with you about it. Up to that point it's not a debate, not a communication ... if the criteria are explicit you can argue about it. Let's have a look at the issues that arose out of the second workshop: choosing the mode of assessment and devising assessment procedures."*

The teachers had not had an opportunity to discuss this part of the assessment workshop materials previously so I explored this area with them and gained feedback from some of teachers who had used the Pupil Record on school placement the previous day. They were surprised by how easy it was to observe and record using this type of systematic observation in practice.

#### *Sessions 7 and 8: 6 and 13th November 1984*

Session 7 was devoted to the workshop activities on profiling and conducting tutorials with some of the teachers. The teacher discussions during the workshop were not



recorded on audiotape. However, the substantive content of their work was reflected in the review carried out in session 8. In the diary entry made after this session I noted that there appeared to be some difference in the way the secondary and primary teachers viewed assessment. The secondary teachers were more confident in talking about formal methods of assessment such as public examinations and testing for grouping, whilst the primary teachers were more familiar with informal methods of assessment such as checklists and the use of anecdotal comments on record cards. Interestingly those teachers who worked in middle schools seemed to have more knowledge of assessment than either group. This observation has also been reflected in the teacher self reports and tutorials.

### *Session 9: 20th November 1984*

The teachers became concerned as the first formal assessment point approached about the expectations for school and course files. In order to clarify this further I outlined what I considered characterised a "good school file".

*AJ "What categorises a 'good school file' is your ability to identify a central thread for yourselves. It's one where you have begun to make an analysis of the information, begun to synthesise it in a way that makes sense to you and that you have the ability to evaluate it and make a realistic judgement and appraisal of what you've done. It's not how much data you've collected, which would be worthless unless you used it in a worthwhile way to inform what you do ... it's where you try to make sense of your own experience, your own way of working, your own decisions in terms of making assessments and judgements about what pupils can do and helps you to make decisions about what to teach next."*

In reflecting on this statement my own understanding of the self-evaluation process and the difficulties of representing this to others as in this chapter were highlighted. The parallels between the experiences of the teachers on the course and my own experience in conducting this evaluation were again reinforced by the process of self-evaluation.

In the remaining thirty minutes of the session I introduced the Oracle research programme to the teachers as an example of a study that investigated the complex interrelationships between teaching style, pupil type and pupil performance. The discussion of this research was continued in the following session and the teachers presented their case studies in the last two sessions of the course. I did not, as I had planned, review the pupil profile data that the teachers had prepared during this session because of the time constraint and thus felt that I had had insufficient time to put into effect all aspects of the course as I had intended.



## 6.5 SELF-EVALUATION: DISCUSSION AND CONCLUSION

In the previous discussion of the processes and outcomes of teachers' professional learning and my own self-evaluation I selected data from a large body of evidence to reflect the main issues which arose. In the following discussion, I have summarised these observations in order to highlight the main aspects of my professional learning in conducting the self-evaluation. This is presented as reflections on the anticipated and unanticipated outcomes of the course for myself and for the teachers.

### *Anticipated outcomes*

I anticipated a number of the factors which emerged during the course. In particular, the initial uncertainty of the teachers was expected and I had ensured that there was ample opportunity for the teachers to question and clarify the aims and expectations which were embodied in both the structure of the course and in the evaluation. A number of difficulties arose for teachers in coming to terms with particular areas of the course content. For example, the initial reluctance of the teachers to try out new techniques and to reject any method which was unfamiliar to them was a usual reaction to the Pupil Record. However, this was followed, as in the past, by surprise at how easy it was to implement such procedures in the classroom during school placement and at the quality of the information which was generated.

Teachers previously had found it difficult to articulate what criteria they were using to assess and evaluate pupil performance. This was no different in this group. The difference between observation and assessment as opposed to interpretation and evaluation also took some time for the teachers to apply in practice. The length of time which was needed for the teacher to begin to make use of the college-based work was also anticipated. Those teachers who had begun to address the questions of improving classroom practice before attending the course were just beginning to make these links in the final sessions. However, a number of the teachers who felt insecure in their subject knowledge or competency in the classroom still had to come to terms with learning from their experiences both in college and school placement. As the course progressed, however, there were many more unanticipated outcomes which reflected the climate of schools at that time. In particular the teachers responded negatively to any implied threat to their professional competence. These unanticipated outcomes are now considered in more detail.



### *Unanticipated outcomes*

There were a number of unanticipated difficulties which can best be understood in relation to the particular impact of the accountability pressures exerted at the time by both local and central government on a number of teachers in the group. Approximately a third of the group had been seconded by the same local authority which was undergoing reorganisation from a well established selective system to a comprehensive system. This was taking place at a time of falling rolls and reduced expenditure in education. As a result the local education authority had redeployed a large number of teachers at the end of the previous year and the possibility of redundancies within the profession was also a serious consideration. The teachers who were seconded obviously felt considerable pressure to succeed and upgrade their qualifications so that they would not become redundant.

This anxiety and pressure to succeed had not been clear to me at the start of the course. My expectation that the assessment of pupils might sensitise these teachers to the need for self-evaluation was pre-empted by the existing pressure on them to account for their right to employment. It was on October 11th that I had to reappraise my basic assumption about the need to introduce the teachers to the idea of self-evaluation through the assessment of pupils. It was in this session that it became apparent that the previous experiences of teachers in the group had already sensitised them to this concern. The accountability issue was at the forefront of their minds and this had clearly made an impact on them which was reiterated in a number of subsequent sessions.

This particular group of teachers was all too aware of the type of self evaluation which was part of the movement to increase the accountability of the teaching profession to the external bodies such as the local education authority and central government. They understood that assessing pupil performance was one way that external agencies were attempting to measure the effectiveness of different schools and different teachers. As was noted in the description of the course these teachers did not agree with the way in which their performance as teachers was being judged by people who often did not even see the teacher concerned in the school or classroom. As Elliott (1983) argued this type of externally imposed accountability did not engender a conducive climate for self-evaluation that might lead to increasing professional autonomy.

Another unexpected outcome was the willingness of these teachers to at least acknowledge the need to change both the school system and their own practice. They may



not have been able to see how to achieve this in practice but the concern for the needs of individual pupils was obvious from the discussions and tutorials. However, whilst recognising this need they also described the demands that this extended role made on them as individuals and the often impossible task this presented in the schools where lack of time, money and resources restricted many attempts to improve the quality of the educational provision that these teachers believed should occur.

Perhaps the most unexpected outcome for me in conducting this evaluation was the relevance of what I was trying to teach the teachers to my own learning and professional development. The impact of the changes in the perceptions of the teachers which resulted from the pressures of the accountability movement influenced me greatly in the way I developed the case study. My expectations of the teachers became more realistic as I dealt with the additional demands of transcribing and analysing weekly workshops, tutorials with the teachers and preparation of materials to meet the changing needs of the teacher group in addition to the demands of full-time employment within an institution which was itself under considerable pressure to change. The apparent hostility to ideas and methods which was expressed at the beginning of the course when these had previously been accepted without question also surprised me. However, the impact of confronting the real needs of the teachers at both an affective and cognitive level provided the necessary stimulus for the reappraisal of my own strategies which prompted self-evaluation.

One of the most important aspects of my own professional development was the realisation that I had the strategies to solve my problems but the way in which I was able to articulate these came from the need to explain and clarify for the teachers my intentions, strategies and reasons for doing certain things in a particular way. Often I articulated these strategies in terms which were directly related to the teachers' needs not as solutions to my own problems. It was only when I was placed in the position of identifying the teachers' needs, reviewing events within the course and the feedback given by the teachers that I became aware of the significance of the recommendations inherent in self-evaluation. I consider now that I had only just begun to challenge my own and the teachers' beliefs and assumptions, espoused theories and theories-in-use at the end of the first term of the course. In terms of my intentions within the course as a whole I achieved some success with regard to facilitating my own professional development. There was, however, one other source of data available to help in my evaluation of the course effectiveness which was based on the teacher evaluation of the B.Ed. in-service programme which was routinely conducted at the end of each year. This evidence is now presented.



## 6.6 TEACHERS' EVALUATION OF THE COURSE

As part of the normal course evaluation conducted by the college each year the teachers completed a questionnaire on my input to the course, rating my performance on seventeen criteria using a 5 point scale. The mean ratings in rank order for the fourteen teachers who completed this evaluation are presented overleaf in Table 6.4. They provide some interesting insights into those aspects of my teaching that the teachers rated highly and those which they did not. The highest ratings were given for my knowledge and interest in the subject matter I was teaching (mean rating = 4.64). This was followed by my ability to make students feel free to question, disagree and express opinion (mean rating = 4.57). The lowest rating was given for my ability to present material in a well ordered fashion (mean rating = 3.00).

In terms of my intentions and espoused theories for the course I considered that this evaluation reflected my priorities well. I wanted the teachers to express their own opinions, to challenge and question the material and ideas I was presenting. The methods that I used were intended to maximise the opportunities for individual concerns to be dealt with as they arose rather than adhere to the pre-planned sessions. If the teachers perceived the course in this way then I had achieved my aims within the course.

## 6.7 SUMMARY AND CONCLUSIONS

The proposition was made at the start of this discussion on the evaluation of the in-service provision that the development of a teacher's ability to assess pupil learning provided an appropriate stimulus for facilitating teachers' professional learning. The evidence that emerged from this case study does not in general support this contention. It is suggested that this was largely due to the change in political and economic climate between the late 1970's and mid-1980's when this study was conducted. The pressures for accountability which were external to the school and unrelated to teacher professional development became part of the system, reduced resources and the threat of redundancy were the everyday experience of these teachers. They were placed in a position where they had to change or lose their place within the system. The intervention was not, therefore, appropriate for their needs which were initially expressed in terms of providing materials and ideas which would help them to at least perform competently in the classroom and thus they expected on the course. Later their expectations changed as they realised that there was something to be gained from challenging their assumptions because success on the course involved change and professional development.



**Table 6.4** Rank order of evaluation criteria with mean rating for theory into practice strand on assessment.

<i>Rank</i>	<i>Criterion</i>	<i>Mean rating</i>
	<i>The course tutor:</i>	
1	appeared to know the subject matter well	4.64
1	appeared interested in the subject matter	4.64
3	made students feel free to question, disagree and express opinion	4.57
4	was available outside class sessions for advice	4.36
5	showed interest in teaching the course	4.14
6	expected an appropriate amount of out of class work	4.00
7	maintained a good atmosphere in the group	3.93
7	followed up student contributions and questions effectively	3.93
9	gave helpful advice on studying the course	3.86
10	showed preparation and planning in teaching	3.71
10	stimulated thinking in the group	3.71
12	guided my assessment work helpfully	3.5
12	made good use of teaching aids and materials where appropriate	3.5
14	helped course members to be aware of their level of achievement	3.41
14	used sufficient examples to clarify issues and material	3.41
16	put material across in an interesting way	3.14
17	presented material in a well ordered fashion	3.00

Rating scale:	Strongly agree	5
	Agree	4
	No particular feeling	3
	Disagree	2
	Strongly disagree	1

The study of assessment in this context did, however, provide an appropriate forum for the teachers to air these issues and discuss implications which went some way towards alerting them to the need for self-evaluation. However, the ability of any teacher to embark on the process of practical deliberation was determined not by the course content and processes but more by the entry point of that teacher in terms of experience and the degree of congruency between this and the situation in which they were placed for school placement. Thus those experienced teachers who were in more familiar situations made most progress towards self-evaluation. The stimulus for this did appear to arise from the course and its rationale appeared effective for this small group. For the less experienced teachers or those who felt their professional competency was threatened by unfamiliar demands the course did not provide an appropriate stimulus for them to move beyond the stage of focusing on basic competency and developing their teaching skills. For myself the process of conducting the evaluation, in particular, my own contribution to the teachers' learning, challenged my assumptions and provided many unexpected insights into the nature of professional learning.



## CHAPTER 7: CONCLUSIONS AND IMPLICATIONS OF THE STUDY.

In this chapter the results of the research into teacher-based assessments are summarised and considered with reference to the hypotheses previously stated. The implications of these findings are discussed with reference to practising teachers. The data collected on the development and evaluation of an in-service programme for teachers is also discussed with reference to the current debate about in-service provision and the need for more school-based professional development based on self-evaluation.

The first working hypothesis investigated in this research into the validity and reliability of teacher-based assessments was concerned with the development of valid and reliable teacher-based assessments, consisting of checklists of abilities, structured activities and teachers' ratings of pupil performance. The hypothesis was stated that teacher-based assessments could be developed for the measurement of pupil performance for the extended aims of primary education identified by Ashton *et al.* (1975) related to the areas of listening with concentration and understanding, inventiveness and creativity, acquiring information other than by reading, application of the four arithmetic rules and mathematical techniques to everyday life and conveying meaning clearly and accurately through speech.

A number of methods were employed to test this hypothesis. These have been described in detail in chapter 4. These strategies incorporated both logical and empirical review procedures applied to the various components of the teacher-based assessments; that is, the checklists of abilities, structured activities and teachers' ratings of pupil performance based on their assessment of each criterion listed on the checklists of abilities. The evidence relating to the validity and reliability of each part of the teacher-based assessments is now discussed in turn before drawing a general conclusion as to their overall validity and reliability.

The validity of the checklists of abilities was considered by the use of both logical and empirical review procedures since it has been argued that the use of traditional statistical procedures alone was inappropriate for the validation of the teacher-based assessments. The methods used to develop these checklists of abilities were also designed to facilitate their validity. Recommendations such as those proposed by Wood and Naphthali (1975) were also taken into account.



The logical review data was drawn from the comments made by the application group of teachers who completed a questionnaire on the appropriateness and use of the various components of the teacher-based assessments. The evidence indicated that in general the checklist of abilities provided clear and appropriate criteria for the particular aim under consideration. However, there were a number of difficulties which teachers identified which reduced the validity of the checklists when used by the teachers for rating pupil performance on the criteria given for each area. These are considered in more detail in the discussion of the validity and reliability of teachers' ratings of pupil performance but the following are of particular interest since they provide some insight into the particular problems which were recognised by the teachers themselves in using such forms of teacher-based assessment.

First, the teachers identified that there was a need to provide more detailed criteria for the assessment of some of the aims under consideration. This was particularly true for those areas such as "acquiring information other than by reading" which covered a wide range of skills, abilities and attitudes in the minds of teachers. To construct a single checklist of abilities relevant to all possible aspects of the aim would have resulted in a document so long that it would have been impossible to complete. Even if such data was collected the information would have been in such an inaccessible form that it would not have provided a useful basis for decision-making as was found in the pilot studies conducted by the development groups of teachers. This problem is further confounded by the variety and complexity of all the other aims described within the Ashton *et al.* (1975) survey of teachers' opinions. Thus, the task of the teacher appeared to be an impossible one if they were expected to provide appropriate learning experiences, monitor and assess individual pupil performance in relation to all these aims.

A second difficulty was also identified by the teachers in addition to that of describing the complexity of the primary curriculum and the variety of skills, concepts, content and attitudes pupils should learn in school. This related to the particular difficulties of observing, monitoring and assessing individual pupils in a class of some thirty or more children. The practical difficulties of seeing each child were great and the difficulty exacerbated by the uncertainty teachers expressed concerning their skill in assessing pupil performance on the basis of their own observations rather than relying on the results of the type of test which produced a number or grade to describe how well the pupils performed.



The comments of the teachers relating to checklists of abilities designed to assess "everyday mathematics" indicated that most confidence was expressed in this area and least confidence in the assessment of "inventiveness and creativity". This suggested that familiarity with a subject area and the various components which were to be learnt by the pupils was an important factor for teachers if they were to carry out effectively an informal assessment of pupil performance.

These particular problems have important consequences for teachers operating within a progressive primary classroom with its emphasis on individualisation and the use of enquiry and discovery learning. If the teacher is unable to accurately observe and record important aspects of pupil performance relevant to the extended aims of primary education as exemplified by the Ashton *et al.* (1975) survey then the opportunities for providing appropriate learning experiences for pupils of differing levels of ability will be much reduced. Similarly if the teachers lack confidence in articulating and clarifying what it is to be learnt by the pupils then they are unlikely to be able to make valid assessments of pupil learning in relation to such extended aims of primary education. This clearly has implications for teacher effectiveness research and the professional learning of teachers. These issues will be considered in more detail later in this chapter when the findings relating to the validity of teachers' assessments and the results of the evaluation of in-service are considered.

The validity of the structured activities was considered using both logical and empirical review procedures in relation to the following intentions of the research:

- a) to facilitate the teachers' assessments of pupils in the classroom,
- b) to provide a measure of pupil performance for use in the evaluation of the effectiveness of different teaching styles in relation to the extended aims of primary education identified in the Ashton *et al.* survey of teachers' opinions (1975),
- c) to ensure the comparability of different teachers' ratings of pupils within the Oracle research programme.

The data was collected from the teachers' responses to the questionnaire on the *appropriateness* or the *suitability* of the materials for the age and ability range of the



pupils assessed and on the *fit*, or how well they matched the checklists of abilities and their *usefulness* in helping teachers assess the abilities of the pupils for each of the structured activities. The evidence suggested that most of the structured activities were considered appropriate. Nevertheless, the teachers commented that some sections of the structured activities were too easy for the more able and too difficult for the less able pupils in the class. Most teachers also commented that although the structured activities did assess some of the criteria on the checklists of abilities they did not tap all the criteria. Most teachers did not find them useful in helping them to assess the classroom performance of the pupils.

This would appear to indicate that the teachers viewed these types of structured activity as a "one off" exercise and as the application groups of teachers did not have the opportunity of developing the skills of observation and their understanding of assessment in the same way as the development groups of teachers they found it difficult to use in their own classrooms as an aid to diagnosis. It has been noted previously that ideally more time and preparation of the application groups of teachers for conducting this type of assessment would have been desirable. However, the timing and load on the teachers participating in the Oracle research programme was such that this was not possible. The results indicated that the structured activities, although appropriate and therefore valid measures of those criteria assessed by them, did not effectively help teachers to make such assessments in the classroom. This necessitated the development of a model of in-service provision later in the research in order to provide an alternative means for teachers to acquire such skills of observation and assessment. It was not a sufficient condition for the validity of teachers' assessments that the materials were developed by teachers for use by other teachers. This seemed to indicate that some aspects of the process of developing the teacher-based assessments were important if the teachers were to apply and use such procedures as part of the teaching and learning process in their own classrooms.

The empirical review procedures used to ascertain the validity of the structured activities have been described in detail in Chapter 4. The results of the correlational studies and factor analyses of the scores obtained by pupils on the structured activities, modified Richmond tests of basis skills and the study skills exercises indicated that the structured activities provided a measure of pupil performance which differed significantly from the data generated on other assessment procedures. There were, however, some significant correlations between pupil performance on each of these



methods of assessment which would suggest that some underlying skills and abilities were tapped by all three types of measure. In the light of the data collected by both the logical and empirical review procedures the inference was drawn that the structured activities, which formed part of the teacher-based assessments, provided a reasonably valid measure of some aspects of the extended aims of primary education.

A number of factors emerged from the analysis of the structured activity data which included: an assessment of the pupil's ability to ask questions, comprehend through listening and observing, sequencing of information, manipulation and organisation of mathematical data, originality and appropriateness. In particular, the method used to assess creativity appeared to tap a significantly different aspect of pupil performance when compared with other assessment procedures which were content specific but related to some general academic ability. These skills were subsequently used to investigate the nature of pupil performance to ascertain if variations occurred with respect to the age, sex, social class and type of pupil as well as studying the impact of different teaching styles in relation to pupil performance.

The reliability of the teacher-based assessments was established in three ways. First the degree of correlation between the assessments made by different teachers using the checklists of abilities to assess the same groups of pupils was calculated. Second, the degree of correlation between assessments made on different occasions was established and third the stability of the structured activities over time was calculated. The results from this study indicated that teachers could, with little preparation, reach a good level of agreement with each other and over time. In addition, the structured activity scores showed high levels of correlation indicating that they were reliable measures of the variables that they were designed to measure. Thus, the structured activities and teachers' ratings on the checklists of abilities were considered to be reliable methods of assessment.

However, the reliability of teachers' ratings did not ensure that such assessments were valid. Evidence from previous studies has suggested that teachers take into account a variety of pupil characteristics when they have been asked to rate pupil performance on academic, intellectual and cognitive dimensions. The results from this part of the study indicated that teachers' assessments of pupil performance using the checklists of abilities were more closely related to the performance of pupils on the modified Richmond tests of basic skills than to the performance of pupils on the relevant structured activities.



structured activities. The teachers' assessments also showed high degrees of intercorrelation between ratings given on different criteria and on different structured activities. These assessments were clearly invalid in that they did not reflect pupil performance in relation to the extended aims of the primary curriculum. However, the teachers appeared to be able to provide a fairly accurate global assessment of pupil performance on the basic skills since the ratings correlated highly with performance on the modified Richmond tests of basic skills.

As has been noted in previous discussions this has important implications for teachers in primary schools. If they choose to organise pupil learning experiences in the progressive tradition then the information on which they take decisions relating to pupil learning in the areas of the curriculum represented by the aims investigated within this study must be suspect. This is likely to result in the kind of mismatch between the capabilities of the pupils and the learning experiences provided which have been noted by Bennett (1984). In order to understand more fully the sources of invalidity in teachers' assessments, this research explored a number of variables which previous research suggested might have influenced teachers' judgements of pupil performance on the criteria listed on the checklists of abilities. Data relating to a number of pupil and teacher characteristics were available, these included teaching style, pupil age, sex, social class and type. This investigation was carried out in order to further explore the nature of teachers' assessments and address some of the questions which have arisen in the literature relating to invalidity in teachers' assessments of pupils.

The results of this investigation of the validity of teachers' assessments was designed to test the working hypothesis that teachers' assessments of pupils in relation to the five aims of primary education under consideration were influenced by pupil sex, age, social class, achievement on tests of basic skills and classroom behaviour. The evidence relating to this hypothesis has been considered in detail in chapter 4. The conclusion was drawn that even when achievement on the basic skills and performance on the structured activities were taken into account that teachers were found to over- and under-estimate pupils relative to their actual performance. This was true for pupil age, social class and type but not for pupil sex.

The over-estimation of younger pupils and under-estimation of older pupils was explained in terms of the teachers appearing to use the age group which they taught as a basis for judging the performance of their class. Thus the teachers of younger pupils did



not take into account the older pupils in the whole sample so that their ratings appeared to be an over-estimation of the pupils abilities when compared with the higher scores obtained by older pupils. Similarly, teachers of the older pupils under-estimated the performance of their pupils since they did not take into account the younger pupils in the sample.

This result would suggest that on the whole the teachers were making their judgements in relation to the particular age range which they taught. Their expectations and assessment of their pupils' performances seemed to be adjusted to take account of the developmental differences between pupils of varying ages and they considered performance in relation to relatively narrow age range. This was perhaps to be expected since in general teachers are not required to assess pupils in relation to a broad age spectrum. This result does, however, suggest that where pupils in a class are grouped vertically or on the basis of family relationships as has been the case in some schools that the expectations of teachers and their assessments of pupil performance may not take account of such differences in pupil development which are a function of age. As a consequence it is likely that mismatches occur between the learning experiences offered to pupils and what they are capable of attempting if the teacher relies on a narrow conception of the abilities of the pupils in his/her class based on a notion of what is appropriate for that age group.

With respect to social class, teachers over-estimated pupils from non-manual occupational backgrounds and under-estimated pupils from manual occupational backgrounds relative to their performance on the basic skills and structured activities. This result reflected that of previous studies where pupils from middle class backgrounds have tended to be considered more able than their actual performance would suggest. The evidence was also in line with similar results where pupils from working class backgrounds have been under-estimated relative to their measured performance.

This result has been interpreted in the light of previous research which has suggested that the process of typification and stereotyping of pupils is based on an ideal type of pupil. Teachers, it has been suggested, perceive pupils from middle class backgrounds as more closely matching this ideal type, and hence tend to over-estimate their cognitive abilities, whilst pupils from working class backgrounds do not match this stereotype and thus are under-estimated relative to their actual performance. This over- and under-



estimation has been demonstrated in relation to performance on intelligence tests, reading ability, arithmetic competence and in this case some aspects of the extended aims of primary education investigated within this study.

Previous research has not investigated the relationship between actual pupil behaviour and the bias of teachers' assessments in any systematic way. The data available on pupil types which was derived from extensive observations of over 400 pupils in the age range 8-12 has enabled this study to look more closely at those patterns of classroom behaviour which seem to be favourably and unfavourably assessed by teachers. This was not investigated by asking teachers to rate pupil classroom behaviour in terms of "good" or "bad" and comparing these judgements with performance on tests as was carried out by Nash (1976). In this study it was possible to investigate the relationship between observed classroom behaviour and the degree of over- and under-estimation of pupil performance as compared with the pupils' actual achievement and performance which were measured by the modified Richmond tests of basic skills and by the structured activities. This analysis was conducted separately for the first and second years of the research.

The results of this part of the study have provided some interesting hypotheses regarding the aspects of pupil classroom behaviour which appear to be taken into account by teachers when making their assessments of pupils. These patterns were different when achievement on the basic skills was controlled for and when performance on the structured activities was taken into account. The following discussion draws on the data generated in the first year of the study when teacher-based assessments were conducted for "listening with concentration and understanding", "inventiveness and creativity" and "acquiring information other than by reading".

When achievement in the basic skills was controlled for the teachers over-estimated the performance of the quiet collaborators, attention seekers and solitary workers and under-estimated that of the intermittent workers. This would suggest that some aspect of the interaction patterns of these pupils has influenced the judgements the teachers make of the abilities of different pupil types. If teachers make their judgements on the basis of the frequency of interactions with pupils; that is, they tend to rate achievement on the basic skills most highly for those pupils with whom they interact most then this would be reflected in the assessments they make of their pupils. This was indeed the case, since the quiet collaborators had the highest level of interaction with the teacher, attention-



seekers the next highest, followed by the solitary workers. These pupils were all over-estimated relative to their performance on the basic skills. Intermittent workers differed from the other pupil types in that they tended to avoid contact with the teacher and have been shown to be under-estimated in relation to their performance on the basic skills.

These results have implications for the classroom teacher since the patterns of interaction exhibited by the different pupil types appear to influence the assessments made by the teachers. If this finding is placed in the context of research on teacher expectancy effects then this may have serious consequences for those pupils who interact less frequently with the teacher. If these pupils are assessed less favourably by the teacher then it is possible that this expectation of the pupils' performance may be communicated to the pupil and result in under-achievement. The teacher in this case can ameliorate the situation by altering the patterns of interaction and thus reduce the opportunities for intermittent workers to avoid contact with him or her. The result would be more equitable treatments of individual pupils and the provision of more appropriate learning experiences and less under-achievement with respect to the basic skills.

The over- and under-estimation of pupil abilities differed when performance on the structured activities was taken into account. In this case the teachers over-estimated the performance of the quiet collaborators and solitary workers and under-estimated the performance of the intermittent workers and attention-seekers. The structured activities were designed to assess aspects of the extended aims of primary education in particular they appeared to be measuring factors such as the pupil's ability to ask questions, comprehend through listening and observing, appropriateness and originality. These skills it can be argued be related to the pupils' ability to sustain independent study and pursue self-directed work. The pupils who performed well on the structured activities would not be expected to seek the teacher attention nor to exhibit the patterns of interaction characteristic of the intermittent workers where they spent less time on task than any of the other pupil types.

Thus teachers appear to have assumed that sustained effort and cooperation on task work as shown by the quiet collaborators and solitary workers resulted in the ability to carry out work independently and achieve success. However, this did not appear to have been the case since pupil performance on the structured activities did not differ significantly with pupil type so that the teachers' expectations regarding pupil



performance were based on an invalid assumption linking classroom behaviour, characterised by quietness and cooperation on the task, with independence and the ability to pursue self-directed enquiry.

Evidence from previous studies has also suggested that the way in which teachers perceived their role, organised the classroom, planned and implemented the curriculum and interacted with pupils, in other words the style of teaching was related in some way with kinds of judgements they made about pupils (Alexander, 1984). The analysis of the impact of teaching style on the bias of teachers' judgements was also investigated within this research. However, because the analysis of teaching styles and pupil types conducted within the Oracle research programme had demonstrated that they were inter-related it was necessary to account for this interaction by analysis of co-variance techniques. For example, teachers characterised as individual monitors were found to have significantly more intermittent workers in their classes than other styles and group instructors more quiet collaborators. Class enquirers were characterised by the presence of a larger proportion of solitary workers than other styles (Galton and Simon (eds), 1980, pp.150). Thus the analysis of the influence of teaching style on teachers' judgements had to take account of this relationship.

The results of this analysis modified the interpretation of the pupil type effects previously discussed in terms of the quantity of pupil-teacher interactions. It appeared that a qualitative difference in the interaction of different teaching styles was also related to the bias observed in the teachers' judgements of pupil abilities related to the extended aims of primary education when account was taken of the teaching style experienced and pupil achievement on the tests of basic skills and performance on the structured activities. In the case of the analysis of pupil type effects on teachers' assessments using teaching styles and basic skills as covariates, the attention-seekers and quiet collaborators were over-estimated relative to their measured performance, whilst solitary workers and intermittent workers were under-estimated. The significant difference between these types of pupils appeared to be the context in which the interaction with the teacher took place. Attention-seekers and quiet collaborators tended to have more interaction with the teacher as an individual or as part of a group whilst solitary workers and intermittent workers tended to interact with the teacher as part of the class. This suggested that where the teacher has a significant number of interactions with individuals then these pupils tended to be over-estimated relative to their achievement on the basic skills.



The bias in teachers' assessments of different pupil types was no longer statistically significant when adjusted by teaching style with the structured activities scores as covariate although the same pattern of over- and under-estimation remained. This was consistent with the explanation that this bias in teachers' assessments was associated with the classroom behaviour of pupils rather than patterns of teacher-pupil interaction. The classroom behaviour of pupils was, to some extent, a function of teaching style experienced, since the proportions of different pupil types was found to vary from teacher to teacher, however, this evidence would suggest that this was not such a significant factor in influencing the nature of the bias in teachers' assessments when compared with the apparent effects of the number and quality of pupil-teacher interactions.

The relationship of teaching style to the bias in teachers' judgements of pupils was investigated in a similar manner. In this case the data was grouped by teaching style, and pupil type and achievement on the basic skills or performance on the structured activities were the covariates in the analysis. Statistically significant differences were obtained in both cases. These results showed that the six teaching styles could be grouped into those which tended to over-estimate pupil performance in relation to the extended aims of primary education investigated. These were the class enquirers, group instructors and infrequent changers and those which under-estimated pupil performance: rotating changers, habitual changers and individual monitors.

The former group were shown to be the most successful styles when progress on the modified tests of basic skills was calculated and they also placed more emphasis on these areas of the curriculum (Galton and Simon (eds), 1980). The second group of teachers were not as successful in terms of progress on the basic skills, and placed more emphasis on individualisation, independent study and enquiry in terms of class activities organised for such areas; for example, project and topic work. These teachers arguably valued those aims of the curriculum associated with independent study and enquiry or discovery learning more highly but it appeared that they judged pupil performance more harshly. This may well have been a function of their awareness of the complexity of the skills, abilities and attitudes required to perform effectively in relation to these extended aims which resulted in the under-estimation of performance in this area.

This over-estimation by the more successful teachers of pupils' abilities to carry out independent work in relation to their achievement on the basic skills and performance on



the structured activities and under-estimation of pupils' abilities by those teachers who were less successful overall also suggested some interesting questions in the context of previous research into the impact of teacher expectations on pupil performance. It has been noted previously that those teachers who over-estimated pupil abilities were the most successful teachers in terms of progress on the basic skills. It could be suggested that these teachers somehow communicated these expectations to their pupils which made some contribution to their success in achieving better than average progress in the basic skills.

This hypothesis was not investigated further within this research but has some interesting implications. If teachers do not make valid assessments of pupils as appeared to be the case here perhaps it is more beneficial for the pupils in their classes if they assume that performance is better than would be expected rather than under-estimate the performance of pupils as did the less successful teaching styles. This can only be considered a tentative suggestion since a number of other significant features associated with successful teaching styles have previously been described which accounted for their success. These have been fully discussed elsewhere (Galton and Simon (eds), 1980, pp.181-199) but in summary related to aspects of teacher-pupil interaction such as the frequency and quality of verbal feedback, task statements and questions. However, this does not preclude that the over-estimation of pupils by teachers with more successful teaching tactics was mediated by these types of interaction, particularly the quantity and quality of the feedback provided.

In the second year of the study the aims related to "the application of mathematical techniques to everyday situations" and "conveying meaning clearly and accurately through speech" were investigated. The results of this part of the study have been presented in detail in chapters 4 and 5. It should be remembered that this sample was made up fewer teachers some of whom had participated in the previous year of the study whilst others had not. The most significant finding relating to this data was the fact that teachers appeared to be able to make more valid assessments in relation to "everyday mathematics" than they had made in the previous year. The results suggested that particular pupil types were not differentially assessed by the teachers.

However, teaching style differences remained when pupil type and basic skills scores or structured activity scores were used as covariates. Two groups of teaching styles emerged which corresponded to those in the first year of the study. When



achievement on the basic skills was used as covariate together with pupil type, the group of teachers who over-estimated pupil achievement included individual monitors, class enquirers and group instructors whilst those who under-estimated pupil achievement included group instructors and infrequent changers. When the structured activity scores were used with pupil type as covariates then the group of teachers who over-estimated pupil performance included individual monitors, infrequent and other changers whilst class enquirers and group instructors under-estimated pupil performance. This analysis lent further support to the interpretation of the results presented for the first year of the study.

The next stage of the research was to investigate the differences if any in the performance of pupils on the various components of the structured activities. The results of this part of the research have been detailed in full in Chapter 5 and the following discussion explores the implications of these findings in relation to pupil characteristics such as age, sex, social class and type as well as the impact of experiencing different teaching styles. It should be remembered that this data represented a limited sample of pupil attainment in these areas since it was not possible to repeat the structured activities in the second year of the study as had been hoped in order to provide some measure of pupil progress in these areas. However, the data available on basic skills achievement allowed for the analysis of pupil performance to take account on differences in basic skills when making comparisons between pupils on groupings such as age, sex and social class so that more valid comparisons could be made between pupils than would have been otherwise possible.

Pupil performance on each of the factors identified in the analysis of the structured activities has been detailed in full previously. The following discussion presents the most significant findings from this study. Pupil performance on the structured activities was not shown to differ significantly on any of the pupil dimensions investigated once performance on the basic skills was accounted for except in the case of pupil age and sex for the factor of "questioning" and pupil age for the factor of "sequencing". In the first case older pupils were more able to ask questions and girls scored more highly than boys. In the second case a significant difference in pupil performance occurred between the ages of nine plus and ten plus.

The difference in performance between boys and girls on questioning was significant at the 5% level only so these results must be treated with caution. However, if



there is a real difference in the performance of boys and girls in respect to this aspect of the structured activities then this would appear to contradict the traditional stereotype of girls in the classroom if they are able to ask more relevant and open questions than boys. It should be noted, however, that the task did not involve an oral response since the pupils were expected to write down the questions. This may well have tended to inhibit the boys but allow the girls to express their questions in a manner which was more comfortable for them and less likely to attract attention.

The differences in performance which were associated with age have been accounted for in terms of developmental processes which it can be argued are more likely to occur in relation to these aspects of learning than perhaps with basic skills achievement. These differences may also have reflected teaching style differences if it was found that particular styles of teaching were more common with the younger pupils than with the older pupils.

The relative merits of different teaching styles were also discussed and it was found that the most successful styles overall for pupil performance on the study skills were the class enquirers and group instructors, followed by the individual monitors, habitual and infrequent changers, and finally the least successful style that of the rotating changers. A similar analysis was carried out in relation to the basic skills scores obtained during the first year of the study (Galton and Simon (eds), 1980, p.187). The order of most successful style to least successful style was found to be different in that the position of the individual monitors was much improved mainly as a result of the performance of the pupils on the originality component of the creativity structured activity.

One of the main findings in this phase of the research related to the invalidity of teachers' assessments on the particular criteria listed on the checklists of abilities for each of the aims under study. However, it was also clear that the teachers made a fairly accurate global assessment of pupil performance on some general academic ability which was more closely related to the qualities assessed by the Richmond tests of basic skills. Given that teachers' assessments have some impact on the quality of learning experiences provided to the pupils then arguably there is a need to develop appropriate strategies to facilitate the accurate assessment of pupil performance on the basis of a pupil's observed classroom behaviour. The opportunity to develop an in-service strand designed to achieve this end arose when I was lecturing on a full-time in-service B.Ed. Honours degree programme. The next phase of the research thus involved the evaluation



of this in-service course in relation to teachers' professional learning with particular reference to developing skills in the assessment of pupil learning.

In documenting the processes and outcomes of teachers' professional learning during the first term of the course it became apparent that the nature of the previous experience of individual teachers who came to the course was a critical factor in their ability to come to terms with the necessity to change their practice. If the teachers were inexperienced then their main concern was to develop a sense of confidence in their competency as a classroom teacher not to embrace the notion of "extended professional" as advocated by Stenhouse (1975). Those more experienced teachers who were placed in a relatively familiar context, either their own school or working within a similar subject area or age range, were found to benefit most from the course. They were beginning to develop appropriate skills of assessment and self-evaluation at the end of the first term after a relatively short period of three months on the course. Other experienced teachers found themselves in unfamiliar situations working with a different age or ability range of pupils. They also found it more difficult to adjust to the demands of self-evaluation since they focused on developing their knowledge of teaching materials rather than observing and assessing pupils or engaging in self-evaluation.

Another issue arising in the course was the threat which these teachers clearly felt from the "system" and their initial desire to reject any attempt at promoting self-evaluation which they perceived as undermining their professional autonomy. This has been discussed at length in Chapter 6 but it is suggested here that there are significant implications for providers of in-service courses. Teachers do not apparently view in-service in the benign way of the past, it must be seen to be relevant and also must offer the teacher a way of maintaining if not improving their status and sense of professionalism within the school context. The constraints and pressures felt by individuals are great in the current educational climate and the temptation for teachers is, to put in the words of one teacher, "*keep your head down ... don't rock the boat*". Such an attitude is unlikely to be conducive to teachers opening up to change by examining their own practice since this will clearly be seen as an attempt to undermine their professional autonomy. Thus careful consideration must be given to both the affective and cognitive dimensions of the course as well as the content and provision of school-based opportunities.

With regard to my own professional learning the most significant dimension of the

self-evaluation was the realisation that I often articulated in discussion with the teachers particular issues, questions or strategies for solving problems which were equally applicable to my own situation. However, it was only when the opportunity to reflect on what I had actually said was presented that I came to realise the appropriateness of the learning I was suggesting to the teachers for myself. The message of the self-evaluation was quite clearly: *"Teacher teach Thyself"*.



## APPENDICES

Page

1.	Questionnaire items from the survey conducted by Ashton <i>et al.</i> , 1975, Aims of primary education: a study of teachers' opinions.	1
2.	Pupil Record observation instrument	7
3.	Teacher Record observation instrument	11
4.	Item numbers of Richmond Tests (France and Fraser, 1975) used in the Oracle shortened version of the test.	
4.1	Reading test.	13
4.2	Mathematics test.	14
4.3	Language test.	15
5.	Aims of primary education listed as of "utmost" or "major" importance in the Ashton <i>et al.</i> survey (1975) presented to teachers.	16
6.	Letter and materials sent to teachers regarding involvement in the Oracle research programme and the development of the teacher-based assessments.	
6.1	Letter	17
6.2	Instructions and booklets for completion.	18
7.	Teacher development groups: discussion documents.	
7.1	Listening with concentration and understanding	21
7.2	Inventiveness and creativity	24
7.3	Acquiring information other than by reading	28
7.4	Everyday mathematics	32
7.5	Conveying meaning clearly and accurately through speech	36
8.	Checklists of abilities: final form	
8.1	Checklist of abilities for listening with concentration and understanding.	39

8.2	Checklist of abilities for inventiveness and creativity	40
8.3	Checklist of abilities for acquiring information.	41
8.4	Checklist of abilities for everyday mathematics.	43
8.5	Checklist of abilities for conveying meaning.	44
9.	Schedule for the development and application of the teacher-based assessments in Oracle schools.	45
10.	Structured activities: final versions.	
10.1	Structured activities for listening with concentration and understanding.	46
10.2	Structured activities for inventiveness and creativity	51
10.3	Structured activities for acquiring information.	52
10.4	Structured activities for everyday mathematics.	55
11.	Scoring schemes for structured activities:	
11.1	Structured activities for listening with concentration and understanding.	64
11.2	Structured activities for inventiveness and creativity	65
11.3	Structured activities for acquiring information.	66
11.4	Structured activities for everyday mathematics.	67
12.	Tape prepared for teachers for the assessment of the child's ability to convey meaning clearly and accurately through speech.	68
13.	Notes to application groups of teachers on teacher-based assessments.	69
14.	In-service course materials for "Assessment in primary schools"	80
15.1	A student's guide to the full-time B.Ed. (Honours) In-service degree programme offered at Worcester College of Higher Education.	90
15.2	Summary of programme structure and course submission of the full-time B.Ed. (Honours) In-service degree programme offered at Worcester College of Higher Education.	93



15.3	Syllabus for the full-time B.Ed. (Honours) In-service degree programme offered at Worcester College of Higher Education.	98
15.4	Teaching arrangements and assessment criteria B.Ed. (Honours) In-service degree programme offered at Worcester College of Higher Education.	104
15.5	Information to schools and teachers on the conduct of school-based assignments within the full-time B.Ed. (Honours) In-service degree programme offered at Worcester College of Higher Education.	106
16.	Information to teachers on the conduct of the professional project within the full-time B.Ed. (Honours) In-service degree programme offered at Worcester College of Higher Education.	108
17.	Summary of responses made by application groups of teachers to the teacher-based assessment questionnaire.	
17.1	Listening with concentration and understanding	112
17.2	Creativity and inventiveness	114
17.3	Acquiring information other than by reading	117
17.4	Everyday mathematics	120
17.5	Conveying meaning clearly and accurately through speech	124
18.	Pen portraits of course members derived from teacher self reports.	125
19.	The in-service course: a chronological summary of the course events commentary and selected tape extracts.	132
20.	Summary of assessed work and teacher self reports.	159

**APPENDIX 1      Questionnaire items from the survey conducted by Ashton *et al.*, (1975)**  
***Aims of primary education: a study of teachers' opinions.***

1.    The child should be able to communicate his feelings through some art forms; e.g. painting, music, drama, movement.
2.    The child should have an understanding of how his body works.
3.    The child should know how to acquire information other than by reading; e.g. by asking questions, by experimenting, from watching television.
4.    The child should know how to use mathematical techniques in his every day life; for instance, estimating distances, classifying objects, using money.
5.    The child should know the correct spelling of a basic general vocabulary.
6.    The child should have some knowledge of the Bible and Christian beliefs.
7.    The child should be an individual, developing in his own way.
8.    The child should be able to write legibly and know how to present his work attractively.
9.    The child should be developing the ability to make reasoned judgements and choices, based on the interpretation and evaluation of relevant information.
10.   The child should be developing awareness of the spiritual aspects of prayer and worship.
11.   The child should find enjoyment in some purposeful leisure time interests and activities both on his own and with others.
12.   The child should be able to read fluently and accurately at a minimum reading age of eleven.
13.   The child should find enjoyment in a variety of aspects of school work and gain



satisfaction from his own achievements.

14. The child should have a wide general (not subject based) knowledge of times and places beyond his immediate experience.
15. The child should know how to observe carefully, accurately and with sensitivity.
16. The child should have sufficient knowledge and skill to be able to engage in simple music making; e.g. singing, percussion, home made instruments.
17. The child should be industrious, persistent and conscientious.
18. The child should be generally obedient to parents, teachers and all reasonable authority.
19. The child should try to behave in accordance with the ideals of his own religion whether or not this is Christian.
20. The child should know how to behave appropriately in a variety of situations; e.g. talking to visitors, going on outings, answering the telephone.
21. The child should know how to convey his meaning clearly and accurately through speech for a variety of purposes; e.g. description, explanation, narration.
22. The child should be developing the capacity to form a considered opinion and to act upon it even if this means rejecting conventional thought and behaviour.
23. The child should know how to apply the basic principles of health, hygiene and safety.
24. The child should be beginning to acquire a set of moral values on which to base his own behaviour; e.g. honesty, sincerity, personal responsibility.
25. The child should be careful with and respectful of both his own and other people's property.
26. The child should be able to read with understanding material appropriate to his age

group and interests.

27. The child should be kind and considerate; he should, e.g. be willing to give personal help to younger or new children, to consider the elderly, the disabled.
28. The child should know how to engage in discussion; e.g. he should be able to talk about his own and others' opinions in a reasonable way.
29. The child should know how to compute in the four arithmetic rules using his knowledge of, for instance, number, multiplication tables and different units of measurement.
30. The child should be a good mixer; he should be able to make easy social contacts with other children and adults in work and play situations.
31. The child should know those moral values, relating to people and property, which are shared by the majority of members of the society.
32. The child should know how to speak in a clear and fluent manner appropriate to different situations; e.g. informal occasions with children and adults, formal occasions.
33. The child should have precise and economic body control for all ordinary physical activities including the handling of tools and equipment.
34. The child should know how to write clear and meaningful English appropriate to different formal purposes; e.g. factual reports, letters, descriptive accounts.
35. The child should have ordered subject knowledge in, e.g. history, geography.
36. The child should have some understanding of modern technological developments, e.g. space travel, tele-communications, automation.
37. The child should be developing a critical and discriminating attitude towards his experiences; e.g. of the mass media.
38. The child should be developing the ability to control his behaviour and his



emotions.

39. The child should know how to write interestingly and with sensitivity.
40. The child should be self confident; he should have a sense of personal adequacy and be able to cope with his environment at an appropriate level.
41. The child should know the basic facts of sex and reproduction.
42. The child should know the basic grammatical rules of written English.
43. The child should know how to behave with courtesy and good manners both in and out of school.
44. The child should be able to maintain lasting relationships with a few close friends.
45. The child should be adaptable to changing circumstances and flexible in outlook.
46. The child should be developing a personal appreciation of beauty in some of its forms both natural and artistic.
47. The child should be enthusiastic and eager to put his best into all activities.
48. The child should be beginning to understand aesthetic experiences and should be able to talk about them; for example, looking at pictures and sculptures, listening to poetry and plays.
49. The child should be able to play a musical instrument such as a recorder, violin, guitar.
50. The child should be able to listen with concentration and understanding.
51. The child should have a general knowledge of his local environment in some of the following aspects, historical, geographical, natural, economic, social.
52. The child should be beginning to understand his own emotions.

53. The child should be able to swim.
54. The child should know the appropriate techniques of some arts and crafts; for example, how to use paint, clay.
55. The child should have a wide vocabulary.
56. The child should be developing the skills of acquiring knowledge and information from written material; for example, summarising, taking notes accurately, the use of libraries.
57. The child should be developing the ability to plan independent work and organise his own time.
58. The child should be happy, cheerful and well balanced.
59. The child should know what to do in emergencies; for example, fire, sickness, accident.
60. The child should be developing his inventiveness and creativity in some fields; for example, painting, music, mechanical things, poetry, movement.
61. The child should have some knowledge of the beliefs of the major world religions other than Christianity.
62. The child should know some simple scientific experimental procedures and some basic scientific concepts; e.g. properties of materials, the nature and significance of changes in living things.
63. The child should know how to play a variety of games; e.g. football, skittle-ball, rounders.
64. The child should be beginning to feel community responsibility; e.g. he should be loyal to groups such as class and school of which he is a member and, where possible, the wider community, and willing to accept the responsibilities which membership implies.



65. The child should have a questioning attitude towards his environment.
66. The child should be developing tolerances; respecting and appreciating others, their feelings, views and capabilities.
67. The child should try to behave in accordance with the ideals of the Christian religion.
68. The child should be able to conduct a simple conversation in a foreign language.
69. The child should be beginning to realise that he can play an important part in his own development by e.g. recognising his strengths and limitations and setting his own goals accordingly.
70. The child should be able to listen to and enjoy a range of music, e.g. pop, folk, classical.
71. The child should know how to think and solve problems mathematically using the appropriate basic concepts of, e.g. the number system and place value, shape, spatial relationships, sets, symmetry and the appropriate language.
72. The child should have a range of movement and gymnastic skills.

## APPENDIX 2    The Pupil Record

Category	Item	Brief definition of item
<b>Coding the pupil-adult categories</b>		
<b>1</b> <i>Target's role</i>	INIT	Target attempts to become focus of attention (not focus at previous signal)
	STAR	Target is focus of attention
	PART	Target in audience (no child is focus)
	LSWT	Target in audience (another child is focus)
<b>2</b> <i>Interacting adult</i>	TCHR	Target interacts with teacher
	OBSR	Target interacts with observer
	OTHER	Target interacts with any other adult such as the head or secretary
<b>3</b> <i>Adult's interaction</i>	TK WK	Adult interacts about task work (task content or supervision)
	ROUTINE	Adult interacts about routine matter (classroom management and control)
	POS	Adult reacts positively to task work (praises)
	NEG	Adult reacts negatively to behaviour etc. (criticizes)
	IGN	Adult ignores attempted initiation
<b>4</b> <i>Adult's communication setting</i>	IND ATT	Adult gives private individual attention to target pupil
	GROUP	Adult gives private attention to target's group
	CLASS	Adult interacts with whole class
	OTHER	Adult gives private attention to another child or group or does not interact



## **Coding the pupil-pupil categories**

<b>5</b> <i>Target's role</i>	<b>BGNS</b>	Target successfully begins a new contact
	<b>COOP</b>	Target co-operates by responding to an initiation
	<b>TRIES</b>	Target unsuccessfully tries to initiate
	<b>IGN</b>	Target ignores attempted initiation
	<b>SUST</b>	Target sustains interaction
<b>6</b> <i>Mode of interaction</i>	<b>MTL</b>	Non-verbal, mediated solely by materials
	<b>CNTC</b>	Non-verbal, mediated by physical contact or gesture (with or without materials)
	<b>VRB</b>	Verbal (with or without materials, physical contact or gesture)
<b>7a</b> <i>Task of other pupil(s)</i>	<b>STK</b>	Same as target's task
	<b>DTK</b>	Different to target's task
<b>7b</b> <i>Sex and number of other pupil(s)</i>	<b>SS</b>	Target interacts privately with one pupil of same sex
	<b>OS</b>	Target interacts privately with one pupil of opposite sex
	<b>SEV SS</b>	Target interacts publicly with two or more pupils, same sex as target
	<b>SEV OS</b>	Target interacts publicly with two or more pupils, of whom one at least is of the opposite sex to the target
<b>7c</b> <i>Base of other pupil(s)</i>	<b>OWN BS</b>	From target's own base
	<b>OTH BS</b>	From another base

## Coding the activity and location categories

8 <i>Target's activity</i>	COOP TK	Fully involved and co-operating on approved task work (e.g. reading)
	COOP R	Fully involved and co-operating on approved routine work (e.g. sharpening a pencil)
	DSTR	Non-involved and totally distracted from all work
	DSTR OBSR	Non-involved and totally distracted from all work by the observer
	DSRP	Non-involved and aggressively disrupting work of other pupil(s)
	HPLY	Non-involved and engaging in horseplay with other pupil(s)
	WAIT TCHR	Waiting to interact with the teacher
	CODS	Partially co-operating and partially distracted from approved work
	INT TCHR	Interested in teacher's activity or private interaction with other pupil(s)
	INT PUP	Interested in the work of other pupil(s)
	WOA	Working on an alternative activity which is not approved work
	RIS	Not coded because the target is responding to internal stimuli
	NOT OBS	Not coded because the target is not observed for some reason
	NOT LIST	Not coded because the target's activity is not listed
9 <i>Target's location</i>	P IN	Target in base
	P OUT	Target out of base but not mobile
	P MOB	Target out of base and mobile
	P OUT RM	Target out of room



10 <i>Teacher activity and location</i>	T PRES	Teacher present with target through interaction or physical proximity
	T ELSE	Teacher privately interacting elsewhere with other pupil(s) or visitor
	T MNTR	Teacher not interacting but monitoring classroom activities
	T HSKP	Teacher not interacting but housekeeping
	T OUT RM	Teacher out of room

Reproduced from *Inside the Primary Classroom, pp.12-13*

**APPENDIX 3   The Teacher Record**

---

Questions	Task	Q1   recalling facts Q2   offering ideas, solutions (closed) Q3   offering ideas, solutions (open)
	Task supervision	Q4   referring to task supervision
	Routine	Q5   referring to routine matters
Statements	Task	S1   of facts S2   of ideas, problems
	Task supervision	S3   telling child what to do S4   praising work or effort S5   feedback on work or effort
	Routine	S6   providing information, directions S7   providing feedback S8   of critical control S9   of small talk
	'Silent' interaction, i.e., interaction other than by question or statement	Gesturing Showing Marking Waiting Story Reading
	Silence *	Not observed Not coded
	No interaction between teacher and any pupil in the class	Adult interaction Visiting pupil Not interacting Out of room



Audience      Class, group or individuals

Composition    Identification of pupils involved

Activity        E.g., Creative writing, practical maths etc.

---

\* While it was recognized that the term 'Silence' was in some instances a misnomer, its use for everyday purposes was preferred to the cumbersome term *silence or interaction other than by question or statement*.

Reproduced from *Inside the primary classroom*, p. 17.

APPENDIX 4

Richmond Tests items used in the long and short forms of the ORACLE tests of basic skills

Appendix 4.1

Reading tests

ORACLE reading tests

Richmond Tests of Basic Skills

R	SR	V	R
1	1	1	
2	2	4	
3	3	8	
4	4	15	
5	5	23	
6	6	31	
7	7	39	
8	8	46	
9		54	
10		62	
11		69	
12		77	
13		85	
14	9		6
15	10		7
16	11		8
17	12		9
18	13		10
19	14		11
20	15		37
21	16		38
22	17		39
23	19		40
24	19		41
25	20		42
26			132
27			133
28			134
29			135
30			136



Appendix 4.2      Mathematics tests

*ORACLE maths tests*

*Richmond Tests of Basic Skills*

*M*

*SM*

*M-1*

*M-2*

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

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
  
  
  
  
  
  
  
  
11  
12  
13  
14  
15  
16  
17

1  
3  
6  
8  
18  
23  
30  
36  
39  
47  
53  
56  
64  
73  
77  
84  
90  
94

1  
2  
18  
19  
23  
32  
36  
43  
49  
53  
59  
62

Appendix 4.3      Language tests

ORACLE language skills		Richmond Tests of Basic Skills			
L	SL	L-1	L-2	L-3	L-4
1	1	1			
2	2	12			
3	3	22			
4	4	31			
5	5	43			
6		52			
7		62			
8		73			
9		84			
10	6		1		
11	7		14		
12	8		27		
13	9		41		
14			55		
15			64		
16			80		
17	10			1	
18	11			12	
19	12			22	
20	13			32	
21	14			47	
22				53	
23				67	
24				79	
25	15				1
26	16				14
27	17				27
28	18				39
29					53
30					62



**APPENDIX 5      Aims of primary education listed as of "utmost" or "major" importance in the Ashton et al. survey (1975) presented to the development group of teachers.<sup>1</sup>**

Rank	Aim	Description
2	26	Read with understanding
4	12	Read fluently and accurately
10	21	Convey meaning clearly through speech
12	3	Acquire information other than by reading
13	50	Listen with concentration and understanding
15	4	Everyday maths
15	55	Wide vocabulary
20	29	Arithmetic - 4 rules
23	9	Make reasoned judgements and choices
24	65	Questioning attitude
25	8	Write legibly
30	60	Inventiveness and creativity
32	57	Plan independent work
33	5	Correct spelling
34	71	Modern maths

---

<sup>1</sup> Taken from: Ashton *et al.* (1975) *The aims of primary education: a study of teachers' opinions*. London: Macmillan Education.

## **APPENDIX 6.1 Letter sent to development group of teachers**

**October 1975**

**Dear**

### **THE NATURE OF CLASSROOM LEARNING IN THE PRIMARY SCHOOL**

**You have kindly offered to cooperate with us on one aspect of this research programme; that is, to attempt to increase the range and objectivity of teacher-based assessments. My colleagues and I hope with your help, to develop alternative assessment procedures covering a wider range of abilities and skills than are usually assessed, and to try to get away from straight 'pencil and paper' testing.**

**I hope that the guidelines and examples included here will help you in this exercise. The booklet contains the following sections:**

- 1. The aims, in full, to be considered**
- 2. Procedure**
- 3. Next step of the exercise**
- 4. Sheets provided for notes**
- 5. Example of a completed sheet**

**Thank you, in anticipation, for all your help and cooperation.**

**Yours sincerely,**

**Anne Jasman**



## **APPENDIX 6.2 Instructions and booklets for completion.**

### **AIMS FOR CONSIDERATION**

The following aims have been compiled by teachers, and they are considered by them to be important.

1. The child should be able to listen with concentration and understanding.
2. The child should know how to convey meaning clearly and accurately through speech for a variety of purposes, e.g. description, explanation, narration.
3. The child should develop his inventiveness and creativity in some field(s), e.g. painting, music, mechanical things, poetry, movement.
4. The child should know how to acquire information other than by reading, e.g. by asking questions, by experimenting, from watching television, etc.
5. The child should know how to compute in the four arithmetic rules using his knowledge of, for instance, number, multiplication tables and different units of measurement. The child should know how to use mathematical techniques in his everyday life; for instance, estimating distances, classifying objects, using money.

From this list of aims, you have been allocated one ( ) for particular consideration.

### **PROCEDURE**

There are three principle steps in the procedure:

1. To think what the aim means to you, making a note of the abilities and behaviours you would expect in a child whom you considered to have performed well in the area of the aim. You may also find it useful to note what the aim does not mean to you, and abilities you would not expect the child to show.

2. To look at what happens in the daily events of your classroom or school, noting activities where children have shown some evidence of learning, or achievement in the area of your aim. It would help to illustrate your comments with examples in which you judge the performance in the area of the aim to be satisfactory or unsatisfactory.
3. To record some activities you think would encourage the child to exercise his abilities in the area of the aim.

When you make these notes you may find it more convenient to record events as they happen, or perhaps at the end of the day make a brief summary on the sheets provided.

I have included one example of how the exercise could be done, but we should, of course, be happy with any alternative approach you prefer to use. You may find it helps to discuss points with your colleagues if several of you are working on the same aim. If, however, there are any problems you wish to discuss with me, please do not hesitate to contact me at the School of Education (Leicester 24211).

I would appreciate it if you could make these notes in the fortnight beginning ..... and during that time I will contact you to arrange a convenient time to visit the school and talk to you.

### **NEXT STEP OF THE EXERCISE**

When I have collected all the information and collated the points you have made under each aim, we will meet to discuss the possibility of drawing up a list of criteria for each aim.



---

**AIM:**

---

The aim consists of:

---

**ABILITIES, CONCEPTS, ATTITUDES ASSOCIATED WITH THIS AIM:**

---

The child should be able to .....

---

**EXAMPLES OF ACTIVITIES WHERE THESE MAY BE SEEN:**

---

## **APPENDIX 7.1: Teacher development group: discussion document for listening with concentration and understanding.**

**AIM: Children should be able to listen with concentration and understanding**

This discussion paper is based on the record booklets which you all completed for me. The format of the booklet has been used for this, and the discussion paper is presented under the following headings:

1. Aim, towards a definition.
2. Abilities, concepts, attitudes associated with the aim.
3. Examples of activities where these may be observed.
4. Examples of satisfactory and unsatisfactory performance in the area of the aim.

### **1. AIM, TOWARDS A DEFINITION**

There were many different interpretations of the aim, and these will be considered in two ways:

- a) What the children listen to.
- b) How they show that they have listened, concentrated on, and understood what has been said.

#### *1a. What the children listen to.*

It is obvious that children will be listening or engaging in conversation, discussions, etc., for a great part of the school day. However, there were particular occasions when you considered it very important for the children to be listening with concentration and understanding. These were as follows:

- i) Stories in the classroom, assembly, and in structured learning situations, e.g. Concept 7-9 and SRA listening skills.
- ii) Factual information - when the teacher talks to groups or the class, visitors come into the school, on visits to museums, etc.
- iii) Instructions - to perform specific tasks in games, P.E., etc. and dealing with routine classroom supervision.



- iv) Conversations - listening to the teacher, and each other, in small groups.
- v) Music/sounds - listening to records, tapes, etc.

*1b. How the children show that they have listened, concentrated, and understood what has been said.*

In considering this aspect of the aim, two different ways of approaching this arose from your records, i.e. looking at what the children are doing when they are listening, and then testing their recall of what had been said in different ways to see if they could, firstly, remember it, and, secondly, had understood it. I have discussed these two aspects together with reference to:

- i) Listening
  - ii) Concentration
  - iii) Understanding
- i) **Listening:** Children can appear to be listening by sitting quietly and appearing attentive, and by recalling information at the time, or after a short period of time.
  - ii) **Concentration:** Children can show concentration by listening for a longer period of time, not being distracted or fidgeting, and by recalling more detailed information, e.g., the order of events in a story.
  - iii) **Understanding:** Children can show understanding by using the information, acting upon instructions, and by asking relevant questions.

## **2. ABILITIES, CONCEPTS, ATTITUDES ASSOCIATED WITH THE AIM**

The abilities and attitudes expressed here are concerned with, firstly, listening and, secondly, showing that they have listened. These behaviours are listed below: following instructions, recalling verbatim, recording verbatim, recalling in their own words, recording in their own words, applying what they have listened to, extending what they have listened to and explaining what they have listened to.

## **3. EXAMPLES OF ACTIVITIES WHERE THESE MAY BE OBSERVED**

There are many occasions when children can be observed with respect to this aim. However, areas emerged from your reports where it seemed particularly important. These are now listed: P.E. lessons, handwork, craft, art lessons, TV programmes,

discussions, debates, stories, school assemblies, lessons where the teacher is talking for periods of time, Concept 7-9, S.R.A. "Listening Skills", music, sounds, on records or tape and memory recall games

#### **4. EXAMPLES OF SATISFACTORY AND UNSATISFACTORY PERFORMANCE**

Many good examples of satisfactory and unsatisfactory performances were given; some of these I have included here:

- a) The children were asked to listen to a BBC sound effects record, and guess the source of the noise.  
"Generally, the children listened carefully, and their descriptions, often vivid and imaginative...and I was surprised by the accuracy of their responses."  
"... inability of several children to listen carefully for a period exceeding just a few minutes."
- b) Children watched a TV programme.  
"A few of the older children had listened carefully and gleaned much information ... which the majority of children failed to notice or remember."  
"Where it had been necessary to pick up the information through listening, recall was disappointing. Information, however, gathered through visual happenings was markedly better."
- c) In a class debate for and against fireworks, two children "argued their cases clearly... they showed examples were they had listened to other children's contributions."

Many questions arise from the points which you have made, and I hope we as a group can begin to answer them. When we look at ways of assessing children's abilities, we could consider three aspects:

- i) Are there different levels of concentration and understanding which can be distinguished?
- ii) Are some 'listening situations' more appropriate than others?
- iii) How far should the assessment procedure be controlled; should we try and assess the children in their day-to-day classroom situation, or whilst the children perform a specific task, such as SRA listening skills, Concept 7-9?



## **APPENDIX 7.2 Teacher development group: discussion document for inventiveness and creativity.**

**AIM:** The child should develop his inventiveness and creativity in some field(s), e.g. painting, music, mechanical things, poetry, movement, etc.

This discussion paper is based on the record booklets which you completed for me. I have tried to clarify, from your examples, some of the criteria you use to distinguish between satisfactory and unsatisfactory performance in this area. The discussion paper is presented under the following headings:

1. Examples of activities where inventiveness and creativity may be observed.
2. Examples of satisfactory and unsatisfactory performance.
3. Possible assessment techniques.

### **1. EXAMPLES OF ACTIVITIES WHERE INVENTIVENESS AND CREATIVITY MAY BE OBSERVED.**

Children can be inventive or creative in many different contexts, e.g. maths lessons, science, discussions, etc. However, the areas which you considered most appropriate for observing children's inventiveness and creativity are given below:

- a) Art and craft
- b) Music, movement and drama
- c) Creative writing
- d) Other situations, e.g. maths and science in special cases.

### **2. EXAMPLES OF SATISFACTORY AND UNSATISFACTORY PERFORMANCE**

I have selected examples from your notes to illustrate the first two areas given previously. After each example I have tried to outline the main criterion for judging a piece of work as satisfactory. I have not, however, considered creative writing or particular examples related to other areas of the curriculum, as these presented special problems for assessment which I will discuss in more detail in a separate paper.

*a) Art and craft*

- i) A child who had watched a group of children modelling ships in school, had gone home and made a model ship from folded card, cocktail sticks, and stood it in a sea of 'card' waves.

(This child has copied an idea, but has been inventive in his choice and use of materials.)

- ii) A group of children were making models from 'junk' available in the classroom. They varied in their responses. "Most children had immediate ideas of their own ... two children copied other children's ideas ... two hadn't a clue what to make, and I had to give them ideas."

(The criterion here is clearly whether the children have their own ideas or not.)

- iii) Using charcoal and kitchen paper, children were drawing pictures on the theme of Hallowe'en. "One child drew a large shapeless monster trying to be friendly with a terrified human being ... other children only drew the usual stereotyped witches and cauldrons."

(A child has taken someone else's idea and expressed it in an original way, i.e. differently to the majority of the class but within the context of the theme.)

*b) Music, movement and drama*

- i) In drama, the children were asked to make a mime with a newspaper. "One child rolled it up and smoked it ... another made it into a baton and walked like a Sergeant Major, another child read it and said he was reading a comic."

(Two points seem to be important here: firstly, two children were clearly inventive in their mime, finding different uses for the newspaper; secondly, one child made a more appropriate mime considering the nature of the material, i.e. the newspaper.)



- ii) In a movement class, the children sat and listened to some music and were then free to move as they liked to it. "Many children seemed involved in the music, and unaware of the presence of others. After a discussion, I learnt that most of them had invented a personal drama for themselves ... Some children, however, simply listened to the music without thinking about it in any depth, and their 'movement' consisted of charging around and around the hall, or giggling with friends."

(The important point is, again, whether the children had their own ideas or not.)

### *Possible Assessment Techniques*

From the examples, a pattern appears to emerge which will, I hope, help us in assessing children's creativity and inventiveness. I have summarised this in the form of a table, which I have given below.

Context \ Child's response	The pupil has to be told what to do, or copies another's idea, or follows directions closely. 1	The pupil extends or adapts ideas which are given to him, doesn't think of his own 2	The pupil uses an idea of his own or produces a 'different' interpretation of a theme within its context. 3
1. Topic/subject/area	1/1	1/2	1/3
a Pupil's choice	a	a	a John
b Teacher set	b Tom	b	b
2. Choice of materials	2/1	2/2	2/3
a Pupil's choice	a	a	a John
b Teacher provided	b Tom	b	b
3. Application or techniques or use of materials	3/1	3/2	3/3
a Pupil adapts or extends normal use	a	a	a John
b Pupil applies or uses techniques or material in a conventional way	b Tom	b	b

I have included two examples to show how this may be used:

- i) *John* - he has decided on a project idea for himself, used many sources of information, and collected a wide variety of materials which are presented in an original way. He would, therefore, be ticked in boxes 1.3a, 2.3a, 3.3a.
- ii) *Tom* - he has to be given a topic for a project, is told what to do, uses only materials and books provided by the teacher, which he copies. He would, therefore, be ticked in boxes 1.1b, 2.1b and 3.1b.

These examples might be said to show that John exhibits a greater degree of inventiveness and creativity than Tom. Your examples are more realistic as they fall between these two extremes.

Is it possible to fit your examples into the grid? I would be very grateful if you could try this to see if we reach the same decisions. I have chosen two examples:

- a) Art and craft (i), for which I ticked 1.2a, 2.3a, 3.3a for the child who made the model ship at home.
- b) Music, movement and drama (i), for which I ticked 1.3b, 2.3b, 3.3a, for the child who made the newspaper into a baton and walked like a Sergeant Major.

I hope we can agree where each example fits within the grid. The next question that arises is whether this will help us to see if some children show a greater degree of creativity and inventiveness than others.

I hope that you will continue to help me by giving me any criticisms or comments on this technique, when I come to discuss this with you in the near future.



## **APPENDIX 7.3 Teacher development group: discussion document for acquiring information other than by reading.**

**AIM:** Children should be able to acquire information other than by reading

This discussion paper is based on the record booklets which you all completed for me. The format of the booklet has been used for this, and the discussion paper is presented under the following headings:

1. Aim, towards a definition
2. Abilities, concepts, attitudes, associated with the aim.
3. Examples of activities where these may be observed.
4. Examples of satisfactory or unsatisfactory performance in the area of the aim.

### **1. AIM, TOWARDS A DEFINITION**

There were many different interpretations of the aim, and these will be considered in three ways:

- a) How children acquire information
- b) Where children acquire information
- c) What information children acquire

#### ***1a How children acquire information.***

There were two main aspects which seemed to help you in thinking about "acquiring":

- i) 'Active' acquiring, e.g., asking questions.
- ii) 'Passive' acquiring, e.g. listening to a talk.

How children acquire information can be summarised as follows:

**LOOKING OR WATCHING** A child can observe any situation and acquire information. A child can look at and listen to a variety of stimuli, e.g. T.V., tape slide sequences, teacher showing how to do something.

<b>LISTENING</b>	A child can listen to talks, lectures, radio programmes, records. A child can actively seek information by asking questions and listening to the answers.
<b>ACTIVELY ACQUIRING</b>	A child can experiment, form hypotheses, and test them.

1b *Where children acquire information*

Many examples of different sources of information were given, but I think these can be broadly classified as follows:

- i) 'Factual' or 'pure' information, e.g., facts, data, evidence, etc.
- ii) 'practical' or 'applied' information, e.g., manipulative skills and techniques.

2. ABILITIES, CONCEPTS, ATTITUDES ASSOCIATED WITH THE AIM

A wide range of abilities and attitudes were expressed as important in the area of the aim. In the list below I have selected the key words from your descriptions of children's behaviours.

2a Behaviours (cognitive)

Observing	Recalling	Extending
Questioning	Talking fluently	Experimenting
Listening	Ordering information	Recording
Evaluating	Discussing	Reasoning
Examining	Expressing ideas	Deducing
Comparing	Continuing	Developing hypotheses

2b Behaviours (affective)

Curiosity	Eagerness	Concentration
Persistence	Excitement	Interest



### 3. EXAMPLES OF ACTIVITIES WHERE THESE MAY BE OBSERVED

Many activities were listed, for both in and out of the classroom. These are given below:

- a) TV, radio, tapes, records, films, slides, pictures, theatre
- b) Drama, movement, games, P.E.
- c) Art and craft
- d) Music, singing.
- e) Visits outside, to places of interest, nature walks, etc.
- f) Visitors to the classroom
- g) Discussions, class talks
- h) Practical experiments
- i) Assembly
- j) Examination of objects brought into the classroom

### 4. EXAMPLES OF SATISFACTORY AND UNSATISFACTORY BEHAVIOUR

Many good examples of satisfactory performance were given. It was clear that when you considered a performance to be satisfactory some of the behaviours listed in 2a and 2b had been shown by the children, e.g., "All the children were extremely *interested* at first, and many spiders' webs were *observed* and *compared descriptively*.... and the children were keen to write about their observations when we returned to the classroom."

In other examples where two children's work was compared, the 'quality' of these behaviours was mentioned, e.g., the children were asked to sketch the roofs of different houses, and consider the roof lines and roofing material: "accurate *observation* of line, perspective, indication of varying textures, proportion, pattern of tile laying .. less accurate *observation* - less appreciation of horizontal and vertical, pattern, proportion, etc." This gives rise to the question of whether one child can 'draw' better than the other, or whether one child 'sees' more than the other.

There are therefore many questions which I hope we as a group can begin to answer. When we look at ways of assessing children's abilities in this area, we could consider three aspects:

- i) Are there different levels of information acquired; are some kinds of information more important than others?
- ii) Are some ways of acquiring information more appropriate than others?
- iii) How far should be assessment procedure be controlled; should we try and assess the children in their day-to-day classroom situation, or whilst the children perform a specific task?



**APPENDIX 7.4 Teacher development group: discussion document  
for everyday mathematics**

- AIM:** The child should know how to:
- a. compute the four arithmetic rules, using his knowledge of number, multiplication tables and different units of measurement,
  - b. use mathematical techniques in his everyday life, i.e. estimating distances, classifying objects, using money.

This discussion paper is based on the record booklets which you completed for me. Parts (a) and (b) of the aim are considered separately.

*a) The child should know how to compute the four arithmetic rules.*

**1. AIM: ASSOCIATED ABILITIES, CONCEPTS AND ATTITUDES.**

From your notes I was able to clarify the term 'compute' with respect to the context, i.e., whether the question was written or verbal, and the different levels of competence shown by children in the way in which they answered questions involving the four rules. I have summarised your ideas in the form of a grid, given below:

The child:	Verbal context	Written context
knows symbols and terminology		
indicates that he knows what to do		
performs the operation correctly (with aids)		
performs the operation correctly (without aids)		
recognises the appropriate rule (variety of contexts)		

## **2. EXAMPLES OF ACTIVITIES WHERE THESE MAY BE OBSERVED**

This scheme is not dependent on the nature of the mathematics work in progress. It can be applied to a variety of situations, e.g., simple calculations using units only, or more complicated sums involving an understanding of place value, decimals, fractions, different measurement units, etc.

*b. The child should know how to use mathematical techniques in his everyday life.*

### **1. THE AIM AND ASSOCIATED ABILITIES**

I found two problems when considering this area. Firstly, there was a very wide range of topics mentioned in your notes, and secondly, each topic requires different combinations of abilities. I have tried to clarify this area by dividing the topics into two groups which seem to require similar combinations of abilities and listing these in order of difficulty. These can be divided into three major categories:

- i) those which show that the child knows what to do,
- ii) those in which the child actually does the operation,
- iii) those where the technique is applied in a more abstract situation.

These are shown in the grid overleaf.

These procedures do not include particular mathematical techniques which can be applied in a wide variety of everyday situations, e.g., finding the average, percentage, proportion, drawing to scale, and making graphical representations of information. I propose to consider these separately at a later stage as these present particular problems.

## **2 EXAMPLES OF ACTIVITIES WHERE THEY MAY BE OBSERVED**

Hopefully, this scheme is not dependent on the nature of the maths curriculum, and can be applied to a wide variety of situations, which is an important consideration when the range of mathematics courses available to schools is so great.



---

Applying techniques to topics involving: one dimension:	two dimensions:
money, time, length,	volume, speed, rate
weight, temperature.	force, costings.

---

The child:

1a knows symbol, terminology  
and notation,

1b indicates that s/he knows the  
appropriate method for  
everyday applications,

---

2a measures accurately, using the  
appropriate apparatus where  
necessary,

2b makes reasonable estimates,

2c distinguishes between  
relative differences (more than  
and less than),

2d converts into appropriate units of  
measurement,

---

3 applies the appropriate technique in an unusual  
situation.

---

### 3 SATISFACTORY AND UNSATISFACTORY PERFORMANCE

The most satisfactory performance would be when a child showed all the abilities listed in an appropriate part of the curriculum. For example, children were asked to work out the thickness of a sheet of paper. Some of them measured a stack of paper and then shared by the number of sheets. In this case, boxes 1b, 2a, 2d and 3 would be ticked as the children had shown that they knew the appropriate technique, measured accurately, converted to the appropriate units of measurement and applied the appropriate technique in an unusual situation. One child, however, said that it was 1mm, the smallest unit on the ruler, here box 1b only would be ticked, as the child only knew the appropriate technique, i.e., measuring, but was unable to apply

an appropriate technique. Clearly this is a less satisfactory performance than the first example.

I would be very grateful for your continuing help in the development of these procedures. I hope that you will be able to use the grid in a particular area, e.g., the use of money, length, weight, volumes, etc. If you find that the criteria within the grid do indicate differences in levels of performance, the next stage would be for me to visit you and discuss further developments.



## **APPENDIX 7.5     Teacher development group: discussion documents for conveying meaning.**

**AIM:**    The child should know how to convey meaning clearly and accurately through speech for a variety of purposes, e.g., description, explanation, narration.

This discussion paper is based on the record booklets which you completed for me. It is presented under the following headings:

- 1    The aim and associated abilities
- 2    Examples of activities where these may be observed
- 3    Examples of satisfactory and unsatisfactory performance

### **1    THE AIM AND ASSOCIATED ABILITIES**

From your notes it was clear that other modes of speech and aspects of voice control and articulation were considered important, apart from those listed in this aim. These can be broadly classified into three groups:

- a)    Abilities associated with voice control and articulation, i.e., the mechanics of speaking.
- b)    Areas where the child is repeating or conveying information from another source, e.g., conveying messages, reading from a book, etc.
- c)    Areas where the child conveys meaning, expressing himself in his own words, e.g., describing, narrating, explaining, instructing.

I feel that (c) is likely to be the most difficult area, and that we should consider this first. We can then turn to (a) and (b) at a later time.

From your examples and notes several abilities emerged which you considered important in the areas of describing, narrating, explaining, instructing. These are summarised in the grid overleaf:

---

The child:

Describing/narrating/explaining/instructing

---

1 uses appropriate vocabulary,

---

2 translates other sources into his own words,

---

3 distinguishes between relevant and irrelevant information for a particular purpose or context,

---

4 orders information logically and in sequence,

---

5 conveys meaning so that the purpose or context can be understood

---

## 2 EXAMPLES OF ACTIVITIES WHERE THESE MAY BE OBSERVED

Clearly, many different classroom situations lend themselves to assessing a child's ability to convey meaning clearly and accurately through speech, e.g., describing objects, situations, teaching a skill, telling a well-known story, telling the story briefly of a film, explaining the rules of a game. One obvious problem in any assessment procedure of verbal communication is that the teacher must hear the child speaking, and this may require setting up one-to-one situations for the assessment, or taping the child and making the assessment later. I feel that we could usefully explore this problem more fully in our discussions.



### 3 EXAMPLES OF SATISFACTORY AND UNSATISFACTORY PERFORMANCE

From your examples of different kinds of performance it is possible to illustrate the use of the list of abilities developed in the grid. For example, one child filled in gaps in an explanation of how to play snap, i.e., the object of the game, so clearly 5 would be ticked, as he had conveyed meaning so that the purpose of the game could be understood. Another child tried to explain how to play Ludo in four sentences. She may have used appropriate vocabulary, translated from other sources into her own words, but clearly had not distinguished between relevant and irrelevant information for the purpose of explaining the game.

I hope that in our discussions we can consider in more depth the practicalities of different assessment procedures and any problems of definition which may have arisen.

**Checklist of abilities for *listening with concentration and understanding.***

PLEASE COMPLETE THE ASSESSMENT FOR EACH ABILITY BEFORE MOVING ON TO THE NEXT.



APPENDIX 8.2

Checklist of abilities for *inventiveness and creativity*

CHECKLIST FOR ASSESSING INVENTIVENESS AND CREATIVITY

CLASS LIST	ORIGINALITY	APPROPRIATENESS

NOTES: 1 Complete the assessments using the following three point scales:

	ORIGINALITY		APPROPRIATENESS
A	Usually produces original work.	A	All the elements are highly appropriate to the context.
B	Sometimes produces original work.	B	While appropriate, some elements do not fit the context.
C	Rarely produces original work.	C	There is no relation between the context and the presented work.

2 Please complete the assessment for 'originality' for all the children before moving to the assessment of 'appropriateness'

### Checklist of abilities for *acquiring information*.

# CHECKLIST OF ABILITIES FOR ASSESSING A CHILD'S ABILITY TO ACQUIRE INFORM- ATION OTHER THAN BY READING

**The child is able to:**

1a. observe in a limited way

b. make his own independent observations

**ask relevant questions**

d. use appropriate language and terminology and appear to understand their meaning

2a. report in an appropriate way  
e.g. orally, by writing, model making, pictorially etc.

**Child's Name**

C. . . B. . . A

C. . B. . A

**C. . B. . A**

C. B. A.

2  
C B A

PLEASE COMPLETE THE ASSESSMENT FOR EACH ABILITY BEFORE MOVING ON TO THE NEXT.



### Checklist of abilities for *acquiring information*.

Please complete the assessment for each ability before moving on to the next.

## Checklist of abilities for *everyday mathematics*.

The child is able to:

state what mathematical processes are involved, using correct terminology.

state if any additional information is needed before a problem can be solved.

carry through the cooperation and represent his resolution to others.

arrive at an accurate result.

**Child's Name**

C. B. A

C. B. A.

C. H. A.

C. B. A.

THE NOTES FOR TEACHERS PROVIDED EXPLAIN THE PROCEDURE FOR COMPLETING THIS CHECKLIST.  
PLEASE COMPLETE THE ASSESSMENT FOR EACH ABILITY BEFORE MOVING ON TO THE NEXT.



# APPENDIX 8.5

## Checklist of abilities for *conveying meaning*.

<p>CHECKLIST OF ABILITIES FOR ASSESSING A CHILD'S ABILITY TO CONVEY MEANING CLEARLY AND ACCURATELY THROUGH SPEECH</p>								
<p>TEACHER:</p>	<p>The child should be able to:</p>	<p>1 vocalise audibly and articulately</p>	<p>2 use language and vocabulary appropriate to the situation or context</p>	<p>3 talk fluently without undue pausing, hesitation or repetition</p>	<p>4 maintain logical connections when describing objects, relating events, and during discussions</p>	<p>5 respond with reasonable speed and accuracy to questions and other contributions</p>	<p>6 modify and develop his own responses</p>	<p>7 link his own experiences to the present situation</p>
NAME								
EXAMPLE 1								
EXAMPLE 2								
EXAMPLE 3								
EXAMPLE 4								
EXAMPLE 5								

PLEASE COMPLETE THE ASSESSMENT FOR EACH ABILITY BEFORE MOVING ONTO THE NEXT

## **APPENDIX 9**

### **Schedule for the development and implementation of the teacher-based assessments in Oracle schools**

**1975**

**Autumn term**

Initial contact with teachers in the area by letter.

**October**

Meeting held for all interested teachers

Teacher development groups established, letter sent to teachers requesting completion of booklet (Appendix 6.1)

**November**

Teachers meetings commenced

**1976**

**Spring term**

Teacher development groups for each of the five aims of primary education meet at regular intervals to draw up checklists of abilities and develop structured activities.

**Summer term**

Reliability and validity pilot studies completed.

**Autumn term**

Introduction of "listening with concentration and understanding and "creativity and inventiveness" teacher-based assessment materials into Oracle sample schools.

Development work on "everyday mathematics" and "conveying meaning clearly and accurately through speech" continues.

**1977**

**Spring term**

Introduction of "acquiring information other than by reading" teacher-based assessment materials into Oracle sample schools.

**Summer term**

Finalisation of "everyday mathematics" and "conveying meaning clearly and accurately through speech" teacher-based assessment materials.

**Autumn term**

Introduction of "everyday mathematics" and teacher-based assessment materials into Oracle sample schools.

**1978**

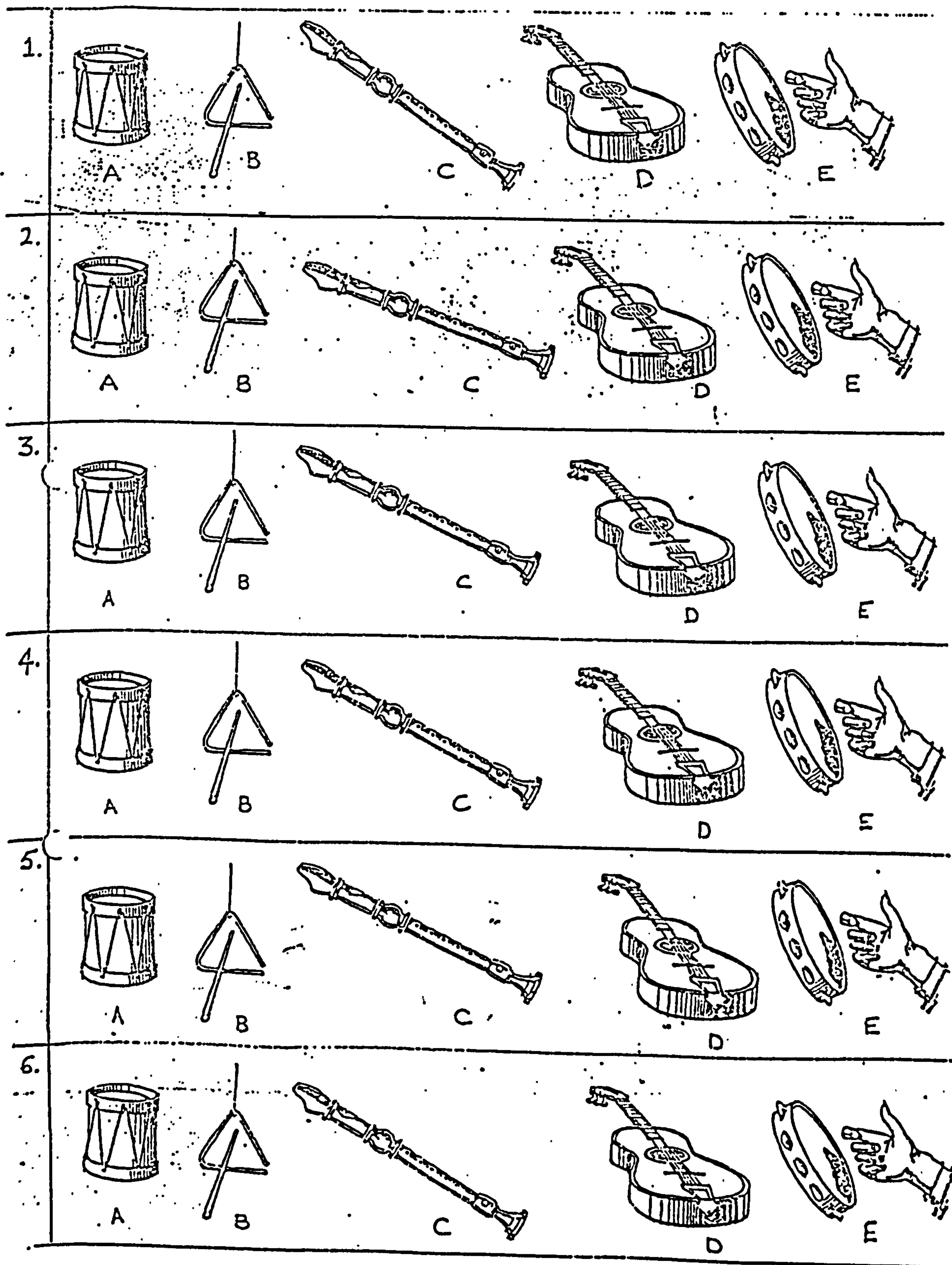
**Spring term**

Introduction of "conveying meaning clearly and accurately through speech" teacher-based assessment materials into Oracle sample schools.

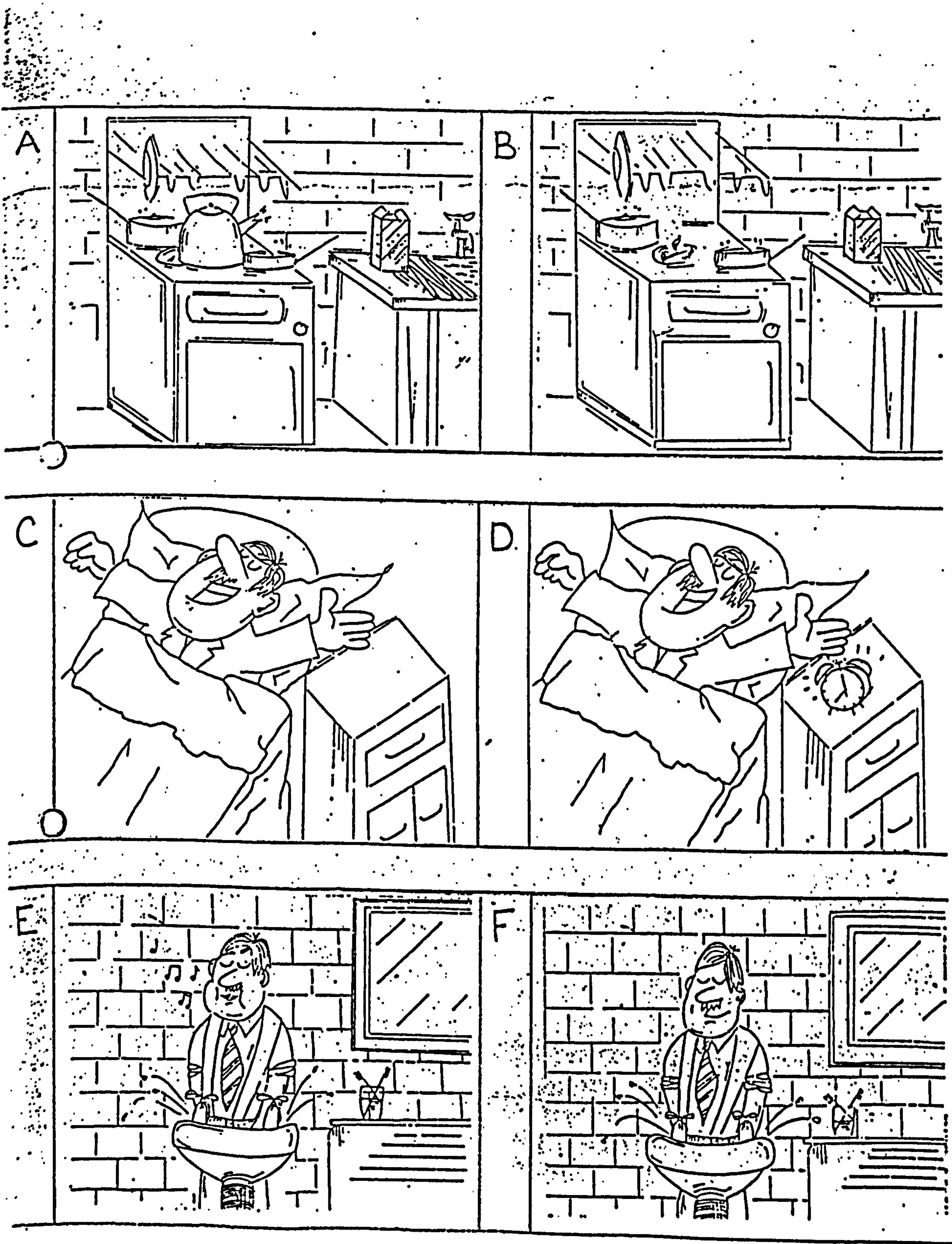


# APPENDIX 10.1

## Structured activities for listening with concentration and understanding.

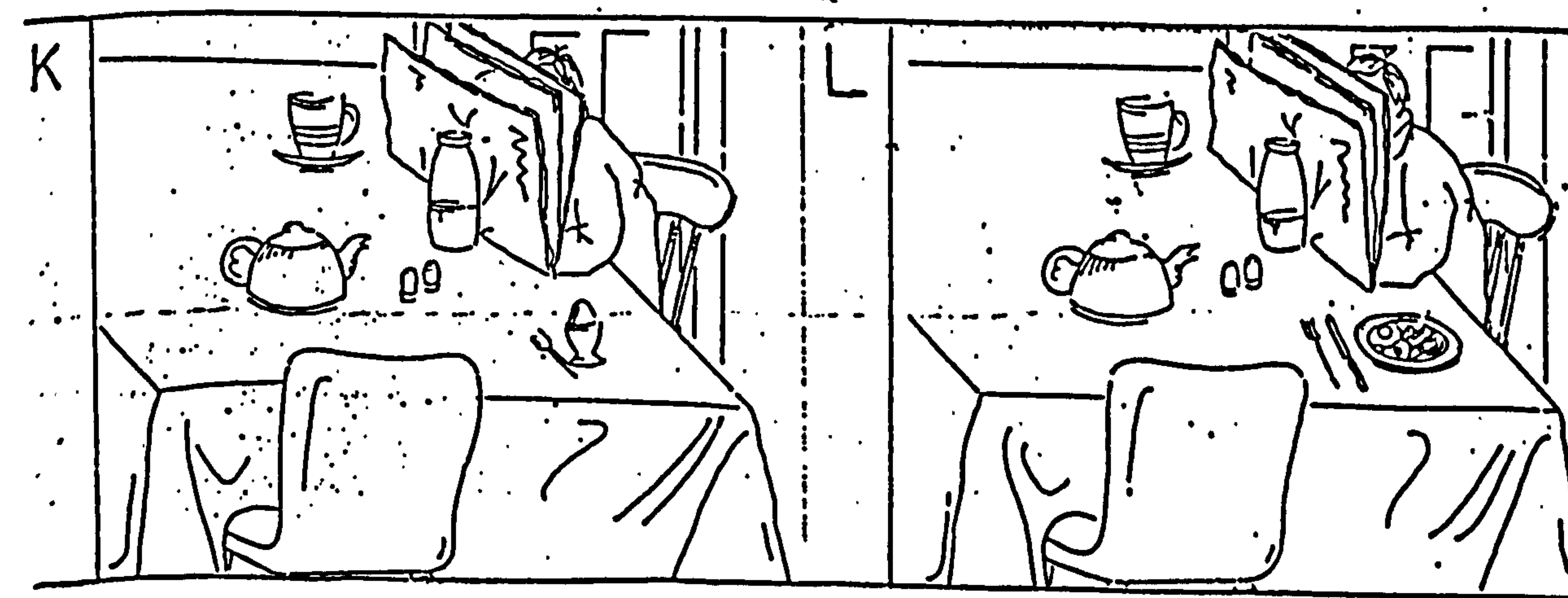
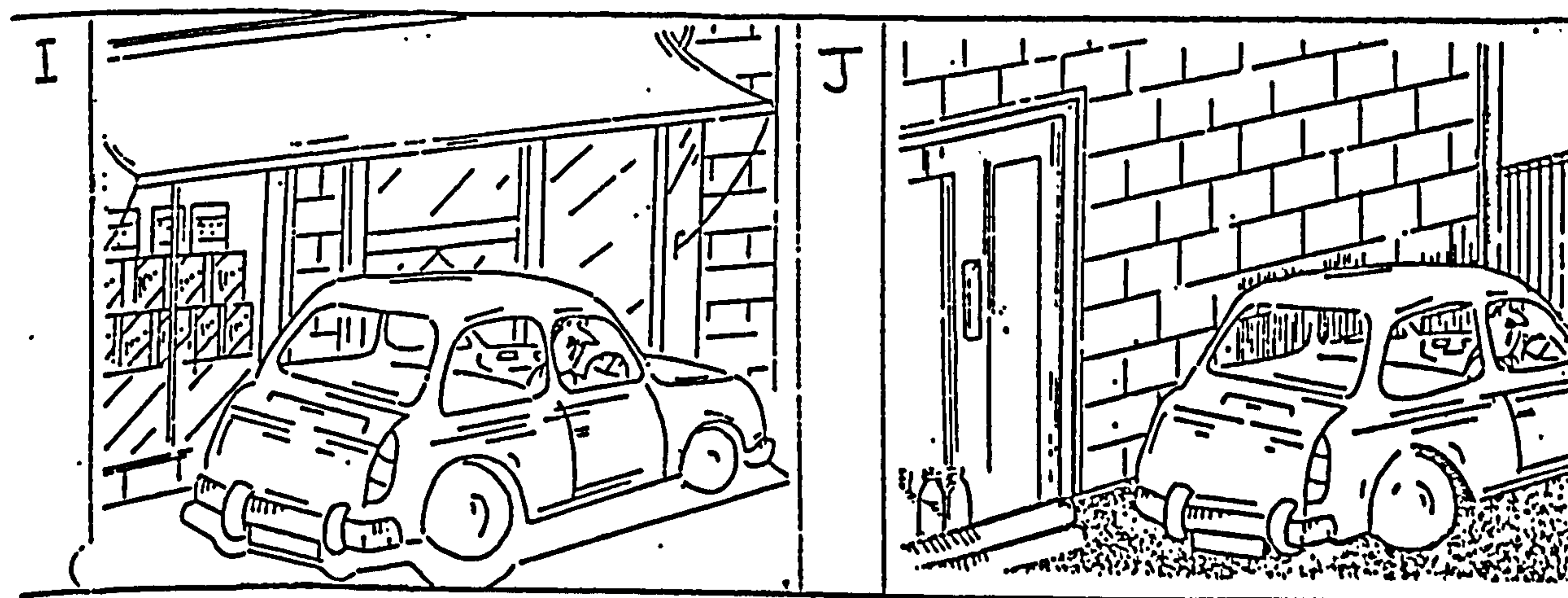
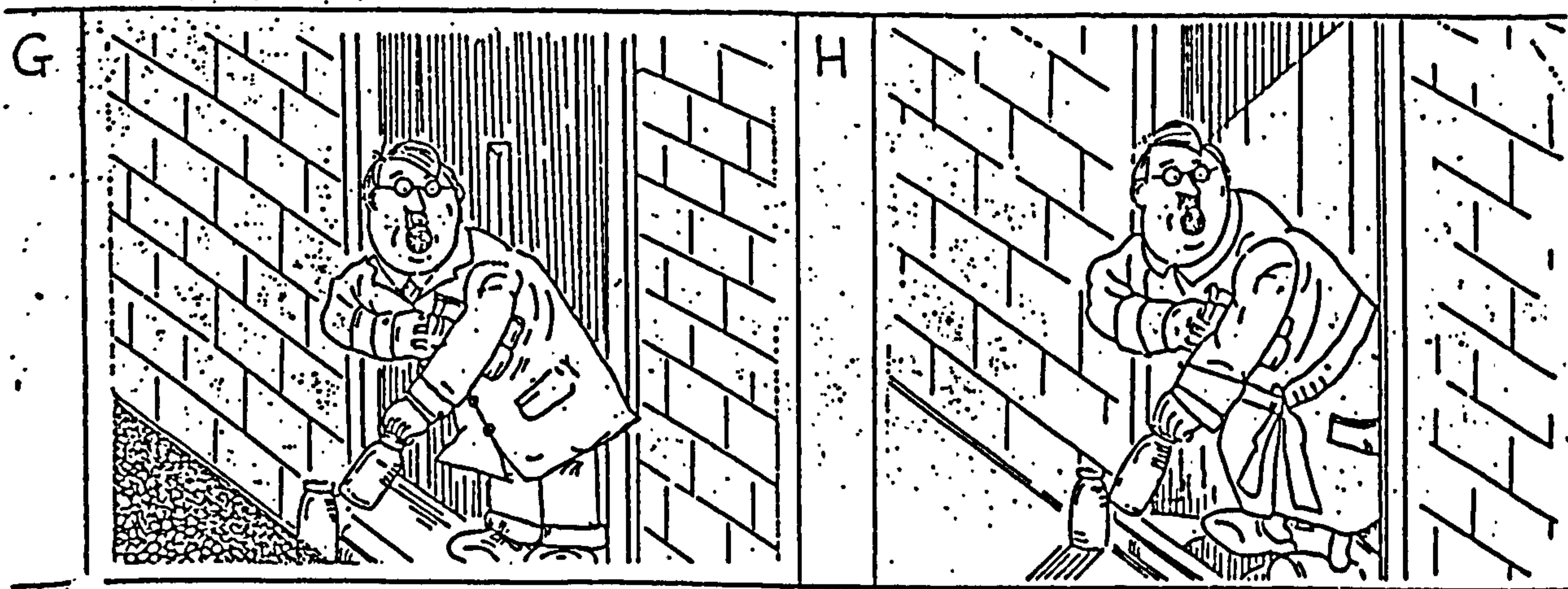


APPENDIX 10.1      Structured activities for listening with concentration  
and understanding.

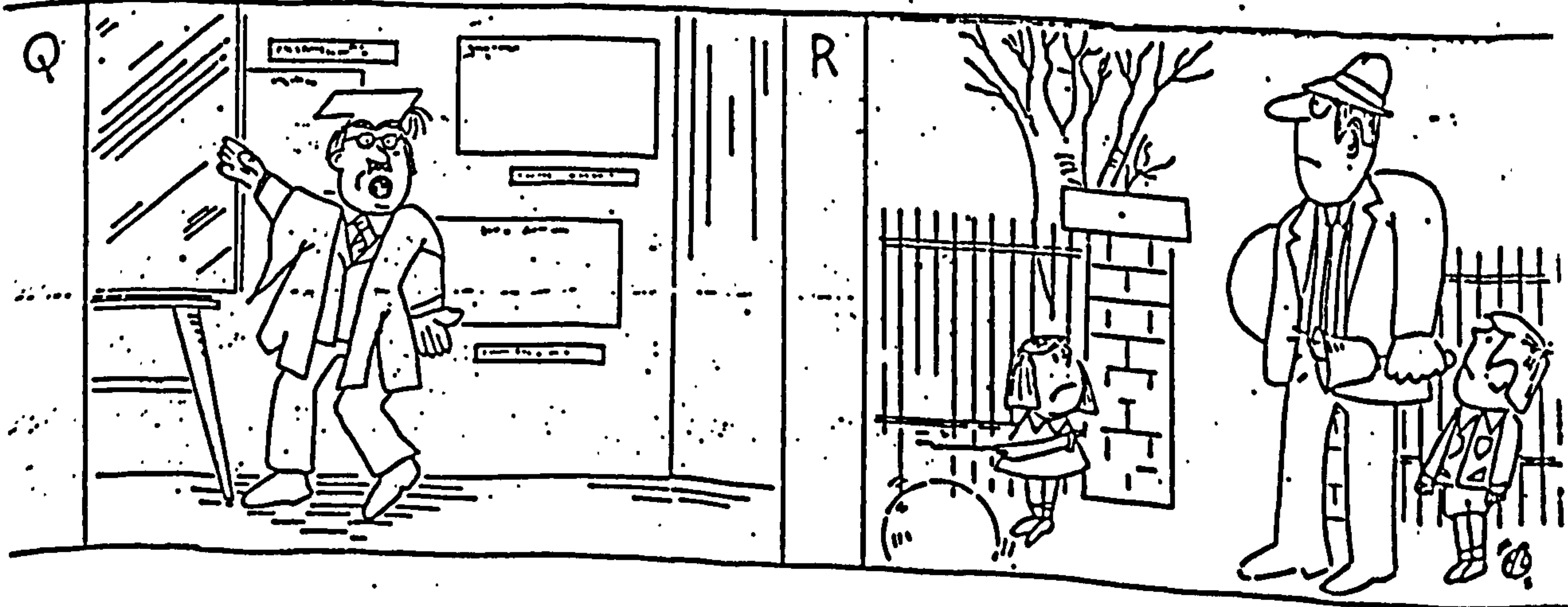
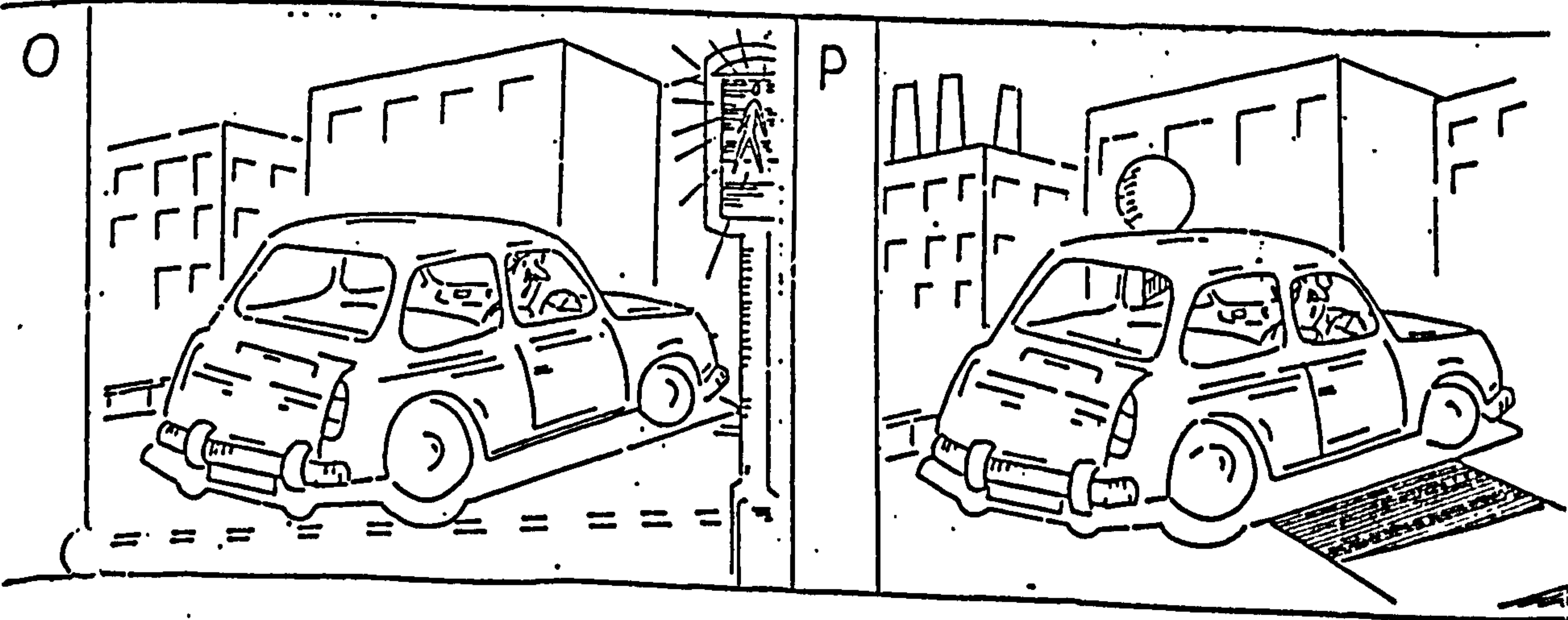
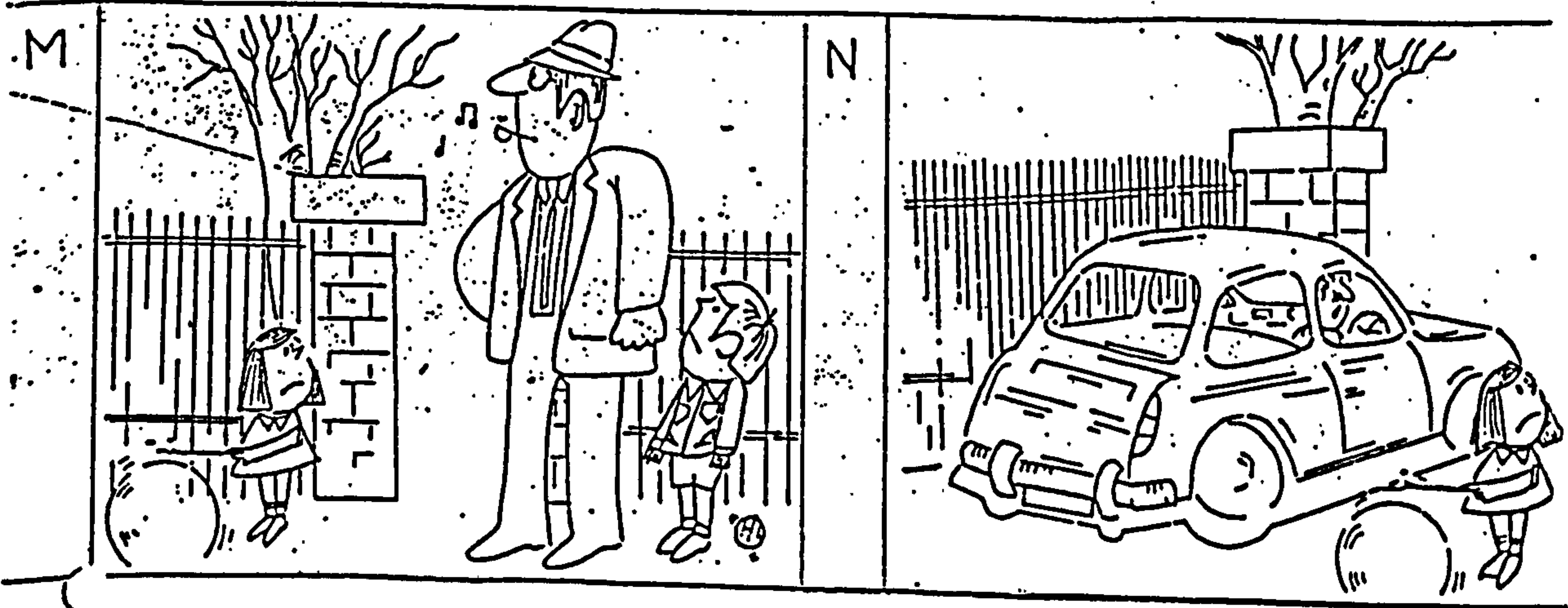




APPENDIX 10.1      Structured activities for listening with concentration  
and understanding.



APPENDIX 10.1      Structured activities for listening with concentration and understanding.





APPENDIX 10.1      Structured activities for listening with concentration and understanding.

Write the letters of the cartoon pictures you chose in the right order here:

- 1

2

3

4

5
- 6

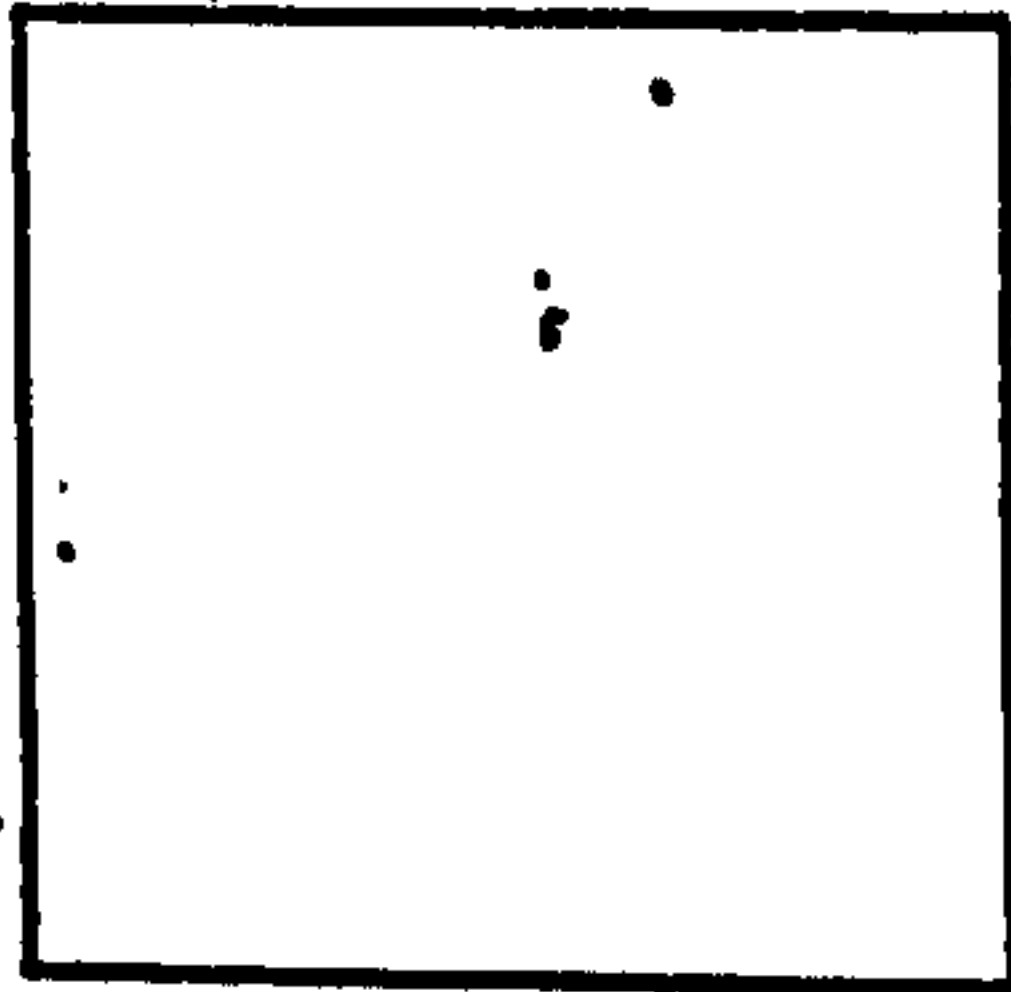
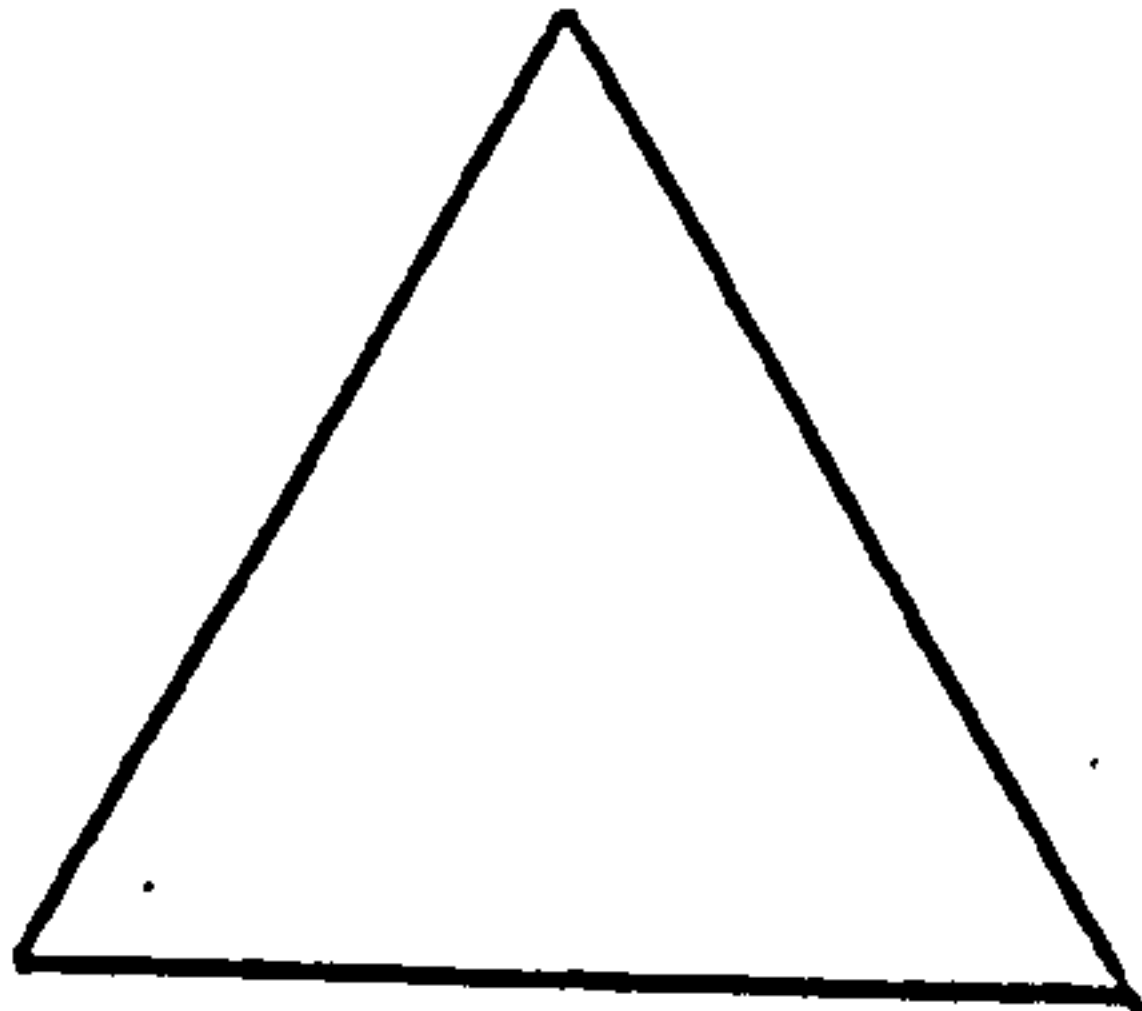
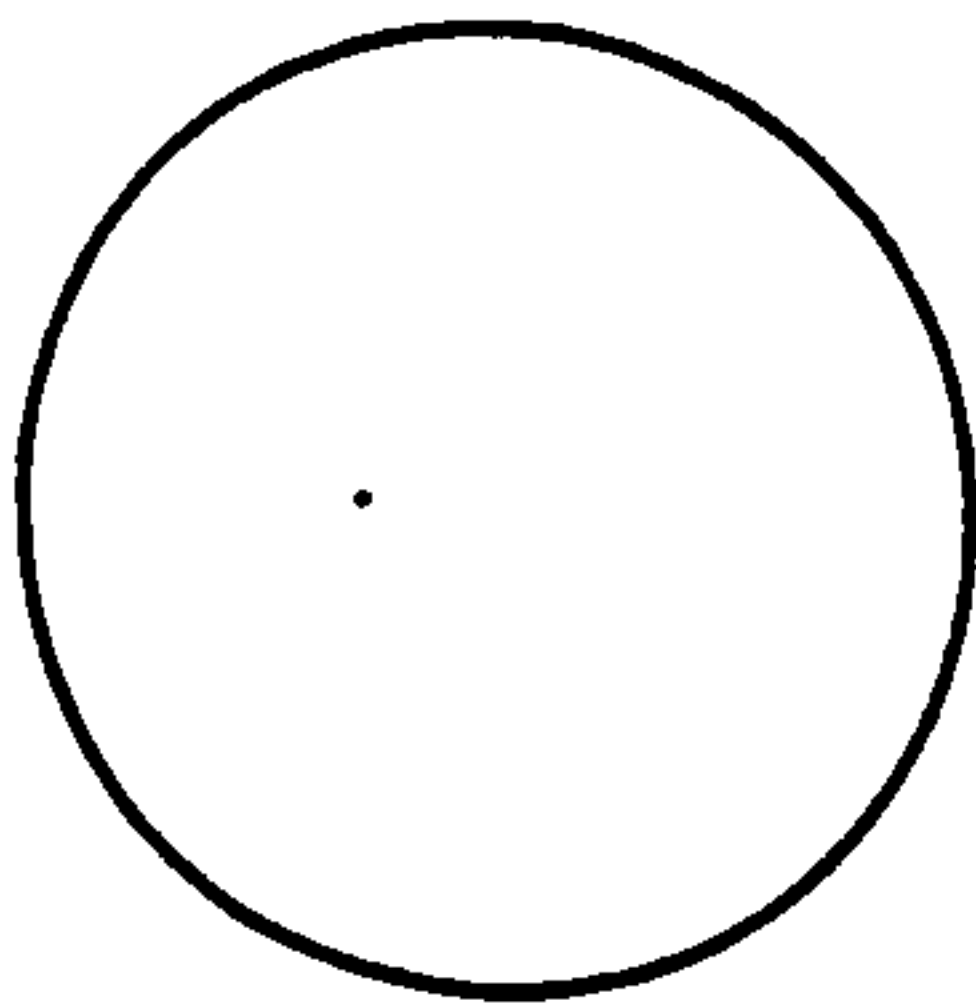
7

8

9

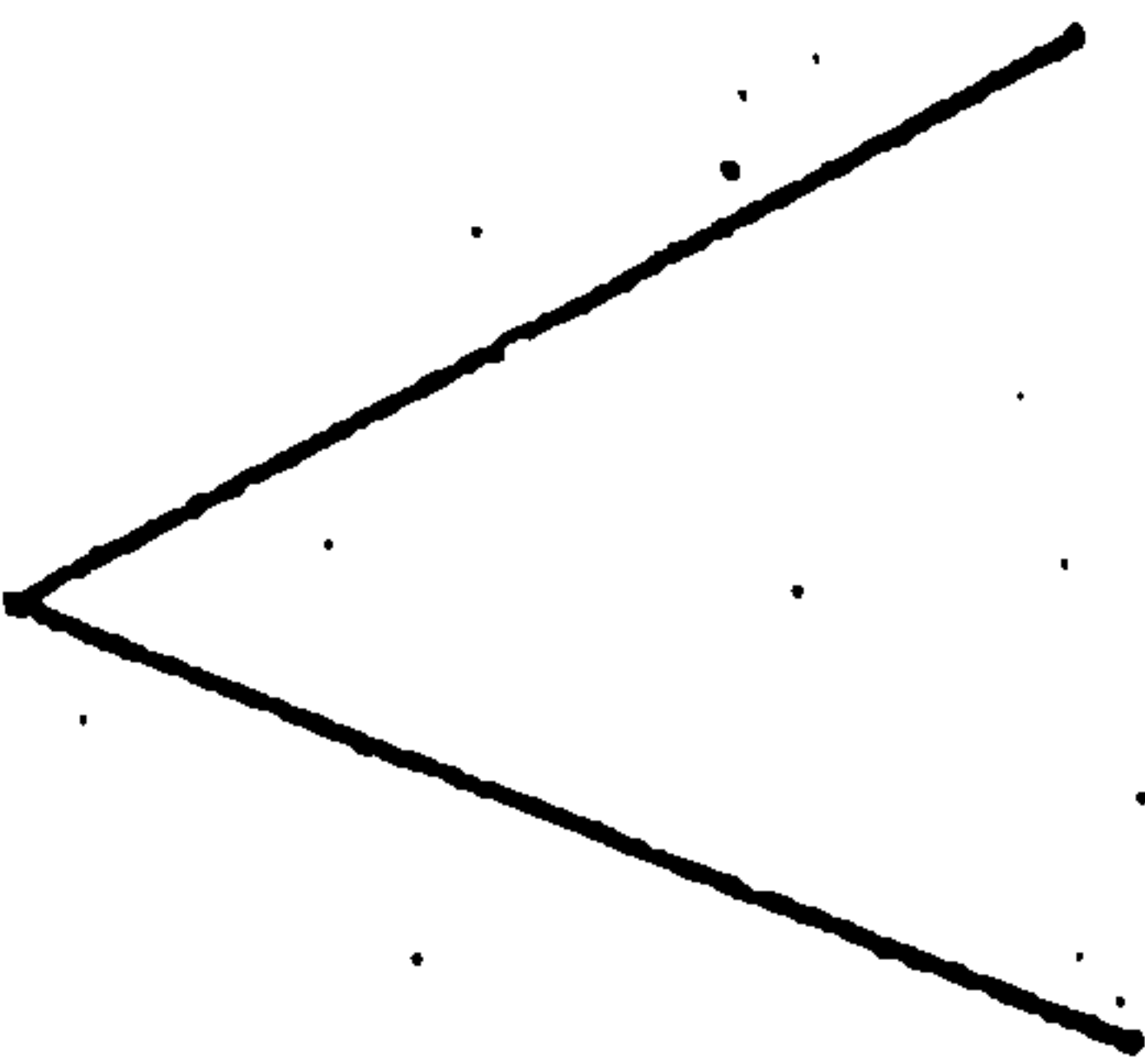
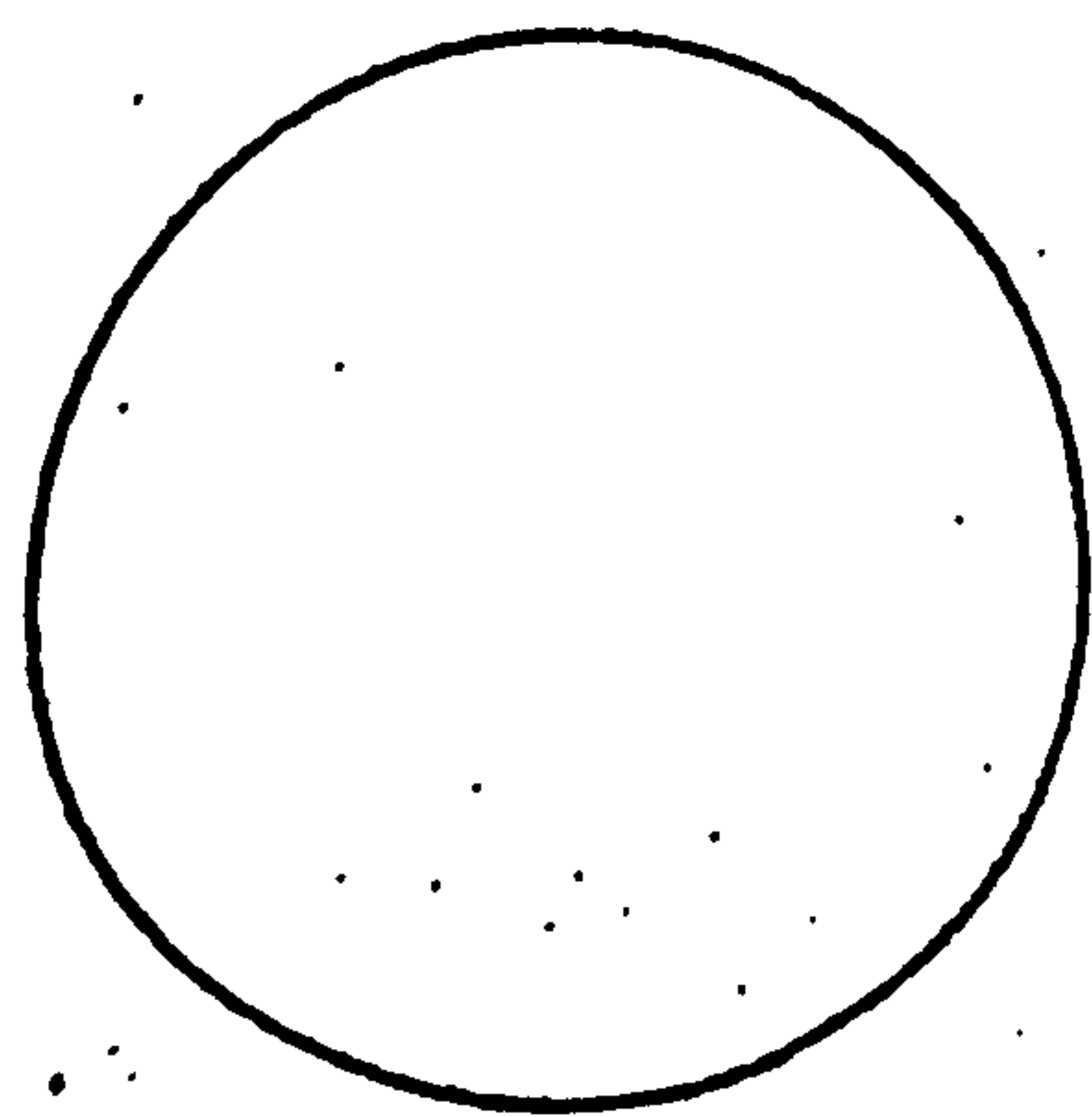
10

Following a set of instructions:



A	B	C	D	E	F	G	H	
I	J	K	L	M	N	O	P	Q
R	S	T	U	V	W	X	Y	Z

**APPENDIX 10.2      Structured activities for inventiveness and creativity**





### APPENDIX 10.3      Structured activities for acquiring information.

Here is a story about Suk who lived a long time ago. Listen carefully

The early morning air was cold, the wind which had been so wild was dying down. Suk welcomed the dawn. He had been on guard since his father had gone back to their hut to get a few hours sleep. These were anxious days and nights in the camp, watching and waiting to see if the enemy tribe was going to attack. It was not long since they had been living in the clearing in the wood in the valley. There it was sheltered and more friendly. With the danger of being attacked, Suk's tribe had moved to the top of a nearby hill. They had built themselves huts, collected all their animals together and were maintaining an hour to hour look-out. The guards, on a clear day, would be able to see if the enemy was approaching from ten miles away.

No-one was happy. The women were frightened, and grumbled because they had to climb down the hill to get their water; the children were quarrelsome because they were hungry and bored, even the animals seemed miserable for they had soon eaten all the grass. Perhaps they would be raided today. Suk heard the camp stirring. The young children were crying and shouting and the men were collecting their weapons in case they were needed. How hungry he suddenly felt. He wished the next guard would come and he could go and warm himself beside the fire in his hut.

As he looked at the low lying country he noticed curls of smoke rising into the sky. They came from a wooded part. He well remembered this wood, he had first learnt to track animals there. As the smoke continued to rise Suk realised what this meant. There had been no fires in the wood yesterday. The enemy must have moved closer to his tribe in the dark. As the guard on duty it was his job to raise the alarm. He forgot about his hunger and his frozen fingers and raced towards the dwellings to tell his news.

Now look carefully at the plan of the encampment.

(When you are satisfied that the children have had sufficient time to look carefully at the picture ask them to either turn over the sheets of paper or hand them back to you so they cannot see the picture).

## APPENDIX 10.3      Structured activities for acquiring information.

Here are some questions.

(Please allow enough time for the children to write or draw their responses, if they have any difficulty with spelling you may help them, or write their answer down for them).

1. What time of day was it?
2. What animals were kept?
3. How far away could the enemy be seen?
4. Why were they living on top of the hill?
5. How many entrances were there?
6. Where did Suk learn to track animals?
7. What did Suk see that made him think that the enemy was nearby?
8. Did the children have plenty of things to do?
9. What kind of house did Suk live in?
10. Write down what questions you would like to ask Suk about the way of life of his village? You can write down as many as you like, but tick the five questions which you think would give you most information.

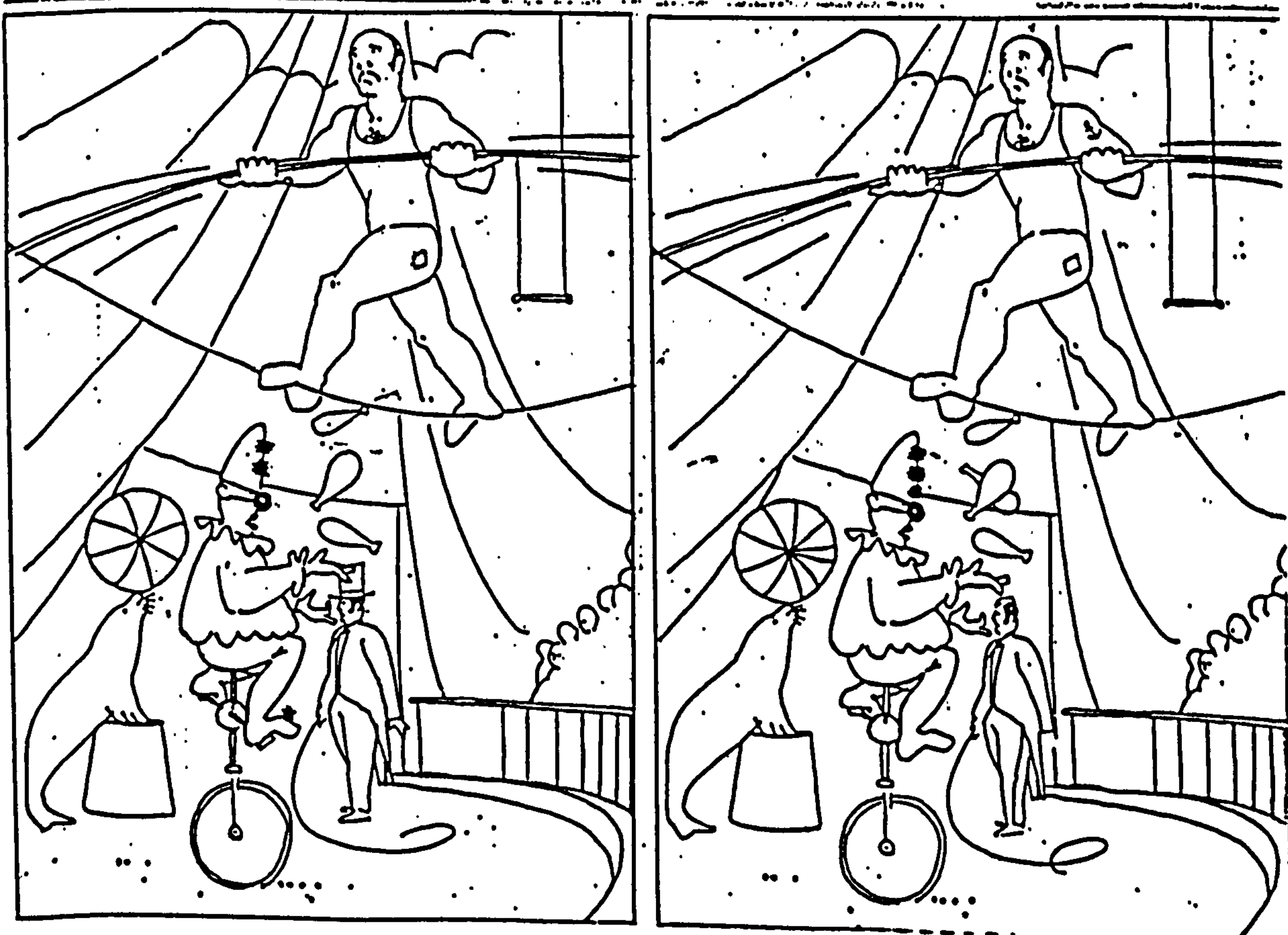
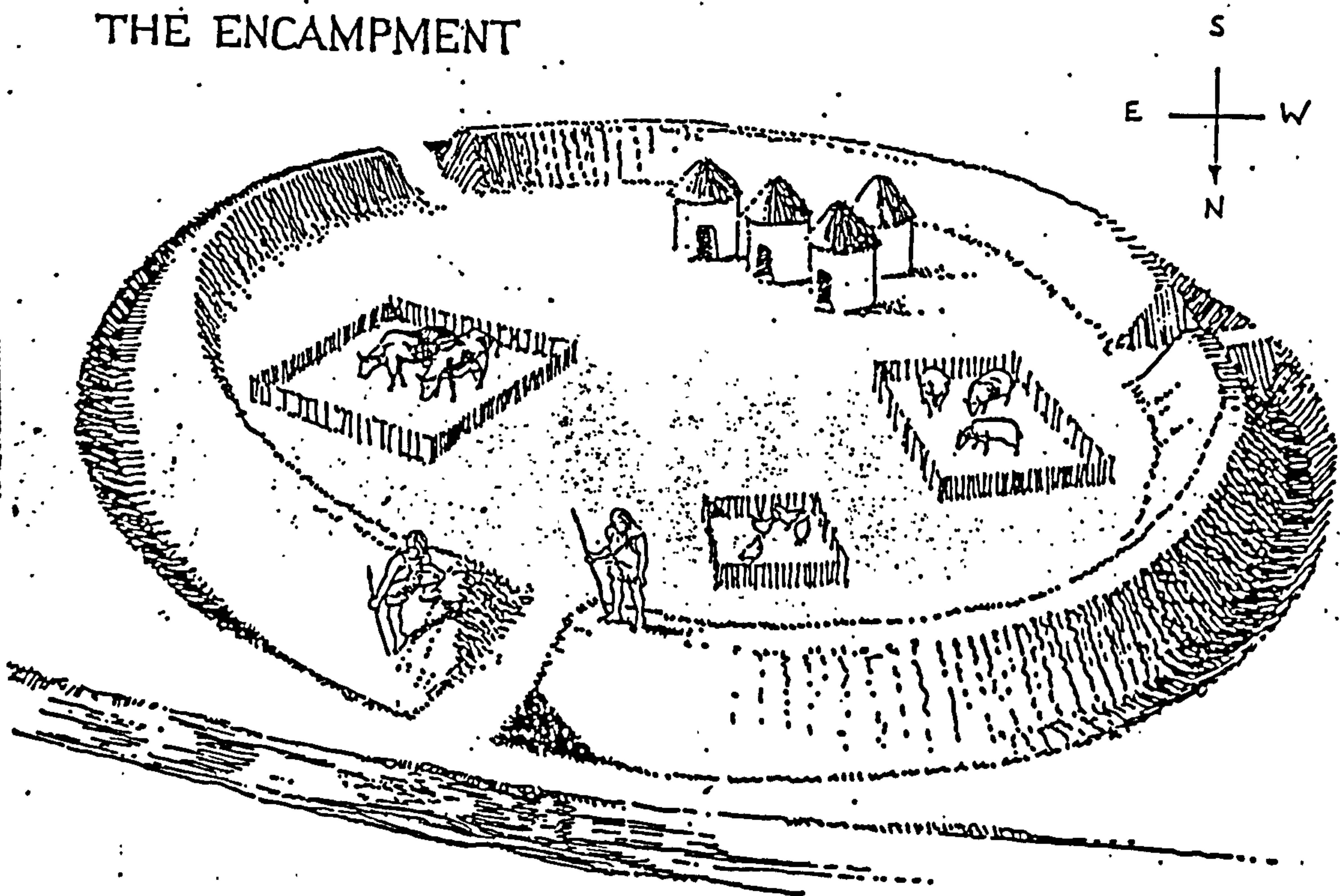
Story written by Mrs Wright  
Mayflower Junior School, Leicester.

NB Instructions are underlined. These may be repeated or rephrased at your discretion.



APPENDIX 10.3      Structured activities for acquiring information.

THE ENCAMPMENT



APPENDIX 10.4      Structured activities for everyday mathematics.

Here are some sums for you to do

1. Add these numbers together:

$\begin{array}{r} 9 \\ 6 \end{array}$	$\begin{array}{r} 24 \\ 51 \end{array}$	$\begin{array}{r} 38 \\ 67 \end{array}$	$\begin{array}{r} 528 \\ 430 \end{array}$	$\begin{array}{r} 109 \\ 697 \end{array}$
---------------------------------------	---	---	---	---

2. Subtract these numbers:

$\begin{array}{r} 7 \\ 2 \end{array}$	$\begin{array}{r} 69 \\ 45 \end{array}$	$\begin{array}{r} 80 \\ 13 \end{array}$	$\begin{array}{r} 479 \\ 136 \end{array}$	$\begin{array}{r} 580 \\ 274 \end{array}$
---------------------------------------	---	---	---	---

3. Multiply these numbers:

$\begin{array}{r} 8 \\ 4 \end{array}$	$\begin{array}{r} 31 \\ 6 \end{array}$	$\begin{array}{r} 75 \\ 9 \end{array}$	$\begin{array}{r} 243 \\ 8 \end{array}$	$\begin{array}{r} 576 \\ 19 \end{array}$
---------------------------------------	--	--	---	--

4. Divide these numbers:

$7 \overline{)21}$	$6 \overline{)282}$	$9 \overline{)8487}$	$12 \overline{)1440}$
--------------------	---------------------	----------------------	-----------------------

$21 \overline{)882}$

Please leave  
for your  
teacher

	1	2	3	4
a				
b				
c				
d				
e				

Here are some sums and the answers but the signs  
are missing. Write the correct sign (+, -, x, ÷)  
for each sum in the box.

5.  $9 \square 6 = 15$

9.  $8 \square 2 = 6$

6.  $9 \square 3 = 3$

10.  $5 \square 3 = 15$

7.  $7 \square 2 = 14$

11.  $7 \square 6 = 13$

8.  $28 \square 2 = 14$

12.  $13 \square 9 = 4$

a  
b  
c  
d  
e



APPENDIX 10.4      Structured activities for everyday mathematics.

Here are some impossible sums. Not even a genius could work out the answer because some information is missing. Read each question carefully. Decide what else you need to know before you could work out the problem. When you have decided, write this down in the space provided.

Please leave for your teacher

Here is an example:

A frog falls down a well. He climbs up 2 metres during the day and slips down 1 metre at night. How long does it take him to climb out of the well?

Missing information: The height or depth of the well.

1 Each of my hens lays 1 egg a day. How many eggs will I get in a week?

.....

2 I have noticed that my plane tree grows 10 cm in a year. How tall will it be in five years time?

.....

3 I bought 1 lb of apples for 30p and 1 lb of pears for 25p. How much change will I have?

.....

4 Mr Brown left Bradfield at 9.00 am for Nottley. His average speed was 45 mph. What time did he arrive?

.....

5 In two years time Mr Jones will be exactly twice as old as his son. How old is his son now?

.....

Here are some problems. For each question you must match one of the sums to the problem. When you have decided which sum would give you the right answer to the problem, ring the letter beside it.

Here is an example: John has 4 sweets. Mary has 2 sweets. How many more sweets has John?  
A.  $4 + 2$   
B.  $4 \div 2$   
C.  $4 - 2$   
D.  $2 - 4$

C is the right answer. Now draw a ring around the letter C.

PROBLEMS

6 Peter bought 4 stamps for 9p and 4 stamps for 7 p. How much money did he spend?  
A.  $4 \times 9 - 4 \times 7$   
B.  $9 \times 4 + 7 \times 4$   
C.  $9 + 4 \times 7 \times 4$   
D.  $4 \times 9 \times 4 \times 7$

a  
b  
c  
d  
e

a  
b  
c  
d  
e

APPENDIX 10.4      Structured activities for everyday mathematics.

- 7 Mrs Smith had 60p to share between her four children. How much did each child get?

8 John lost 2 of his 10 marbles to Mark and another 3 to Gary. How many marbles did John have left?

9 Mr and Mrs Street bought a cooker which costs £92. They paid a deposit of £20 and paid off the balance in 6 equal monthly payments. How much is each monthly payment?

10 A gardener grew twelve tomato plants. He shared them between four friends. Each plant produced ten tomatoes. How many tomatoes did each friend have?

A.  $60 \times 4$   
B.  $60 \div 4$   
C.  $4 \div 60$   
D.  $60 - 4$

A.  $2 + 10 - 3$   
B.  $2 - 10 + 3$   
C.  $10 - 2 + 3$   
D.  $10 - 2 - 3$

A.  $(92 - 20) \div 6$   
B.  $92 - 20 \div 6$   
C.  $(92 + 20) \div 6$   
D.  $92 - 20 \times 6$

A.  $4 + 10 + 12$   
B.  $12 \div 4 \times 10$   
C.  $10 \times 4 \div 12$   
D.  $12 + 4 \times 10$
- Please leave for your teacher

For the next questions you must write down the sum you would use to work out the answer. There is no need to work out the answer just write down the sum.

Here is an example:

In a block of flats there are 6 floors. On each floor there are three flats. How many flats were in the block?

The sum:  $6 \times 3 = \dots$   $3 \times 6 = \dots$

- 11 John bought 4 apples to take to his Granny. His mother gave him 2 more to take. How many apples did he take?

The sum: .....

- 12 Mr Walter has used 430 units of electricity. The Electricity Board charge 2.5p per unit. How much did the electricity cost?

The sum: .....

- 13 Mr Potter planted 10 bulbs; but only 7 grew. Each bulb produced 3 flowers. How many flowers did he have?

The sum: .....

- 14 When she called into a garage, Mrs Jones bought 2 cans of oil at 50p a can, and 4 gallons of petrol at 85p a gallon. How much did she spend?

The sum: .....

- 15 Andy has 25p pocket money and he spent 20p yesterday. He earned another 10p for cleaning the windows. How much has he now?

The sum: .....



APPENDIX 10.4      Structured activities for everyday mathematics.

Here are some problems for you to try and answer.  
If you have any difficulties please ask your teacher for help.

Please leave  
for your  
teacher

1. The No 10 bus leaves the station at these times:

8.10      17.40      11.00      9.00      13.00      15.00  
20.00      17.30      17.00      8.30      18.00      8.40

The journey to the end of the line takes 20 minutes.  
Make a timetable for Mr William. in the table below:

Departure Time	Arrival Time

Mr William catches the No 10 bus home every night. He gets off at the last stop and takes another 5 minutes to walk home. He likes to be home in time to see a television programme at 6.00 pm. Find out which is the latest bus he could catch to be home in time to see the programme.

Answer: .....

- a
- b
- c
- d i
- d ii
- d iii
- e

- a
- b
- c
- d
- e

APPENDIX 10.4      Structured activities for everyday mathematics.

2. John found an old piece of paper telling him how to find some buried treasure, but there was no map. Make a map for John from these instructions.

"When you step off the boat at the landing stage you take 4 paces forward in a northerly direction. Walk 20 paces east along the Coast Road. You are now standing next to a coconut tree. Face north and walk 5 paces forwards, a large white rock is in front of you. Walk 20 paces west and start digging for the treasure." Draw your map here:

Please leave  
for your  
teacher

a

b

c

d i

d ii

d iii

d iv

e

When you have drawn your map, work out how far John would have to carry the treasure to put it in the boat at the landing stage.

Answer: .....

In which direction did John walk towards the boat from the treasure spot?

Answer .....

a

b

c

d

e



APPENDIX 10.4      Structured activities for everyday mathematics.

3.      At the Park Primary School, the headteacher likes the school week to be split up like this. The chart shows how much time should be spent on each subject:

Subject	Hours	Subject	Hours	Subject	Hours
English	5	Science	1	Art	1
Maths	5	Swimming	1	Music	1
Topic	2	P.E.	1	T.V.	$\frac{1}{2}$
Clubs	1	Games	1	Singing	$\frac{1}{2}$

A teacher makes this timetable for her class:

Time	Monday	Tuesday	Wednesday	Thursday	Friday
9.00	Register	Register	Register	Register	Register
9.30	Assembly	Assembly	Assembly	Assembly	Assembly
	Maths	P.E.	English	Maths	T.V.
10.30					Singing
	B	R	E	A	K
11.00	Swimming	Maths	Topic	English	Maths
12.00					
	L	U	N	C	H
1.30	Topic	English	Clubs	Art	Music
2.30					
3.30	English	Science	Clubs	Topic	Games

Has the teacher managed to fit in the right amount of time for each subject? Beside each subject write more, less or correct.

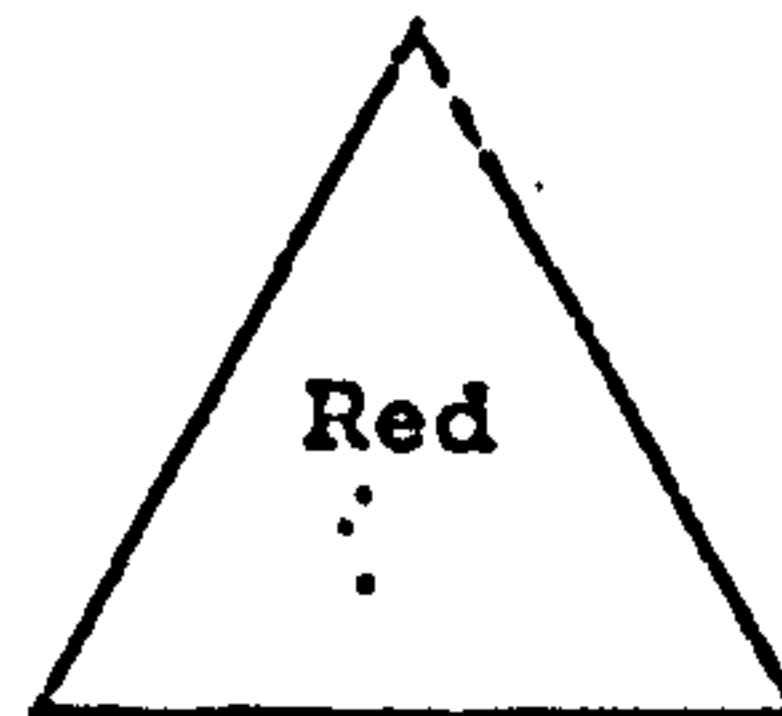
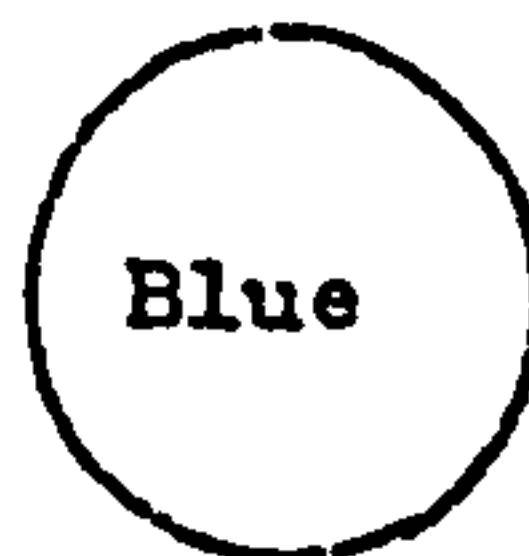
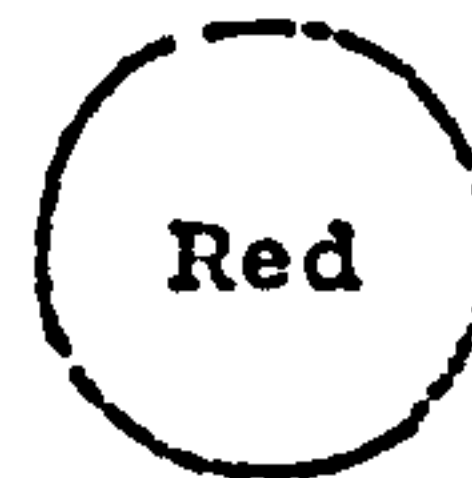
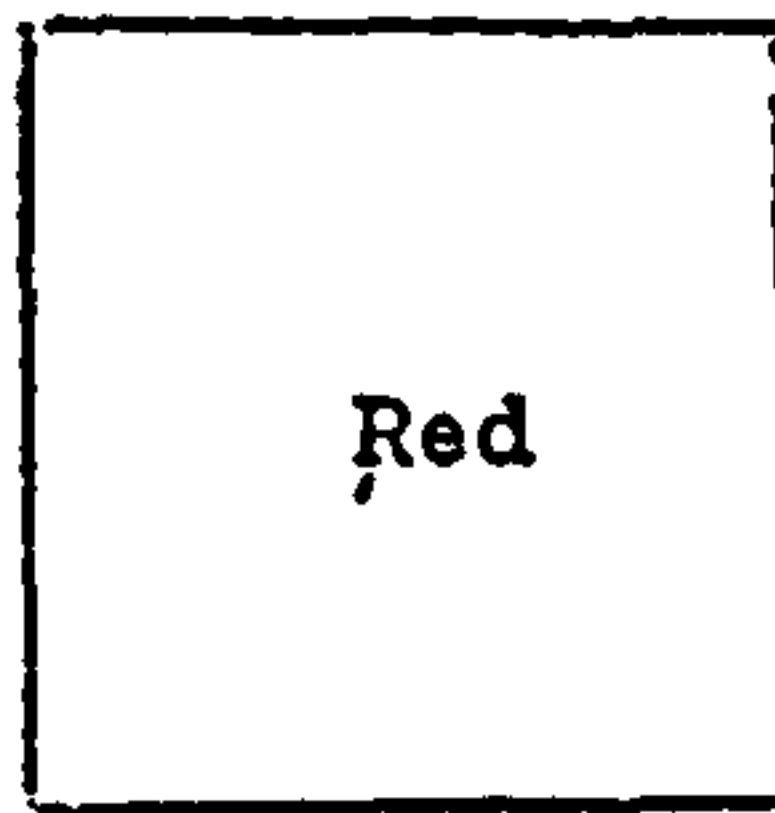
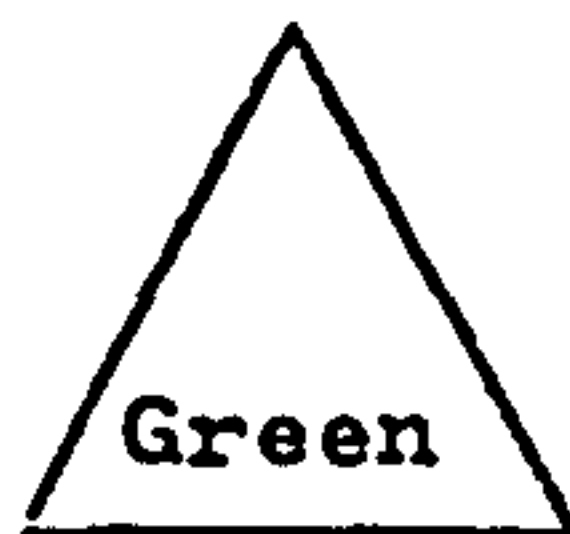
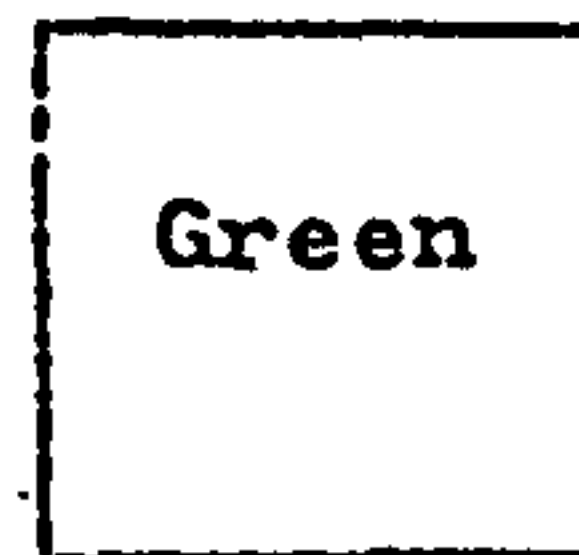
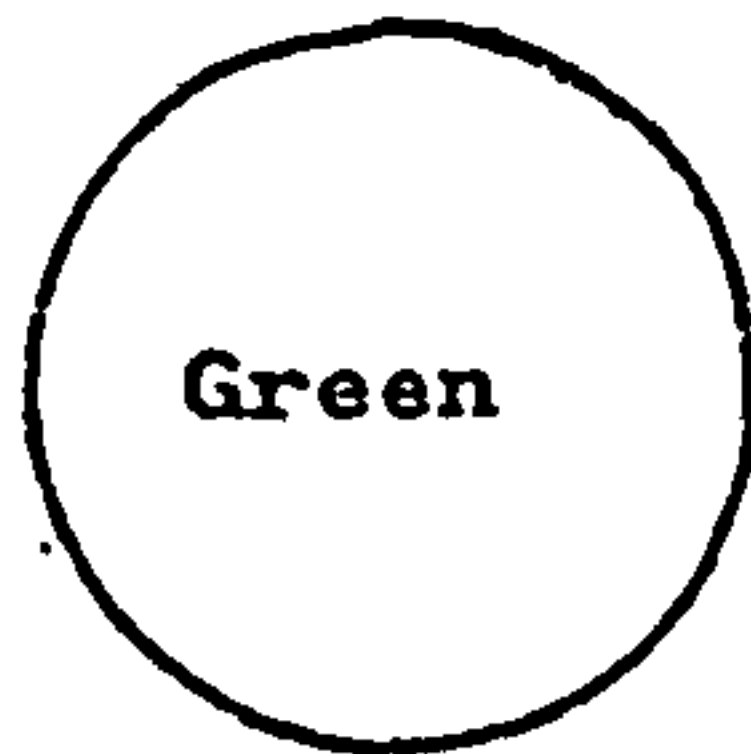
English	.....	Science	.....	Art	.....
Maths	.....	Swimming	.....	Music	.....
Topic	.....	P.E.	.....	T.V.	.....
Clubs	.....	Games	.....	Singing	.....

Please leave  
for your  
teacher

a  
b  
c  
d  
e

APPENDIX 10.4      Structured activities for everyday mathematics.

4. You can sort these shapes in several different  
ay



Write down the different ways here:

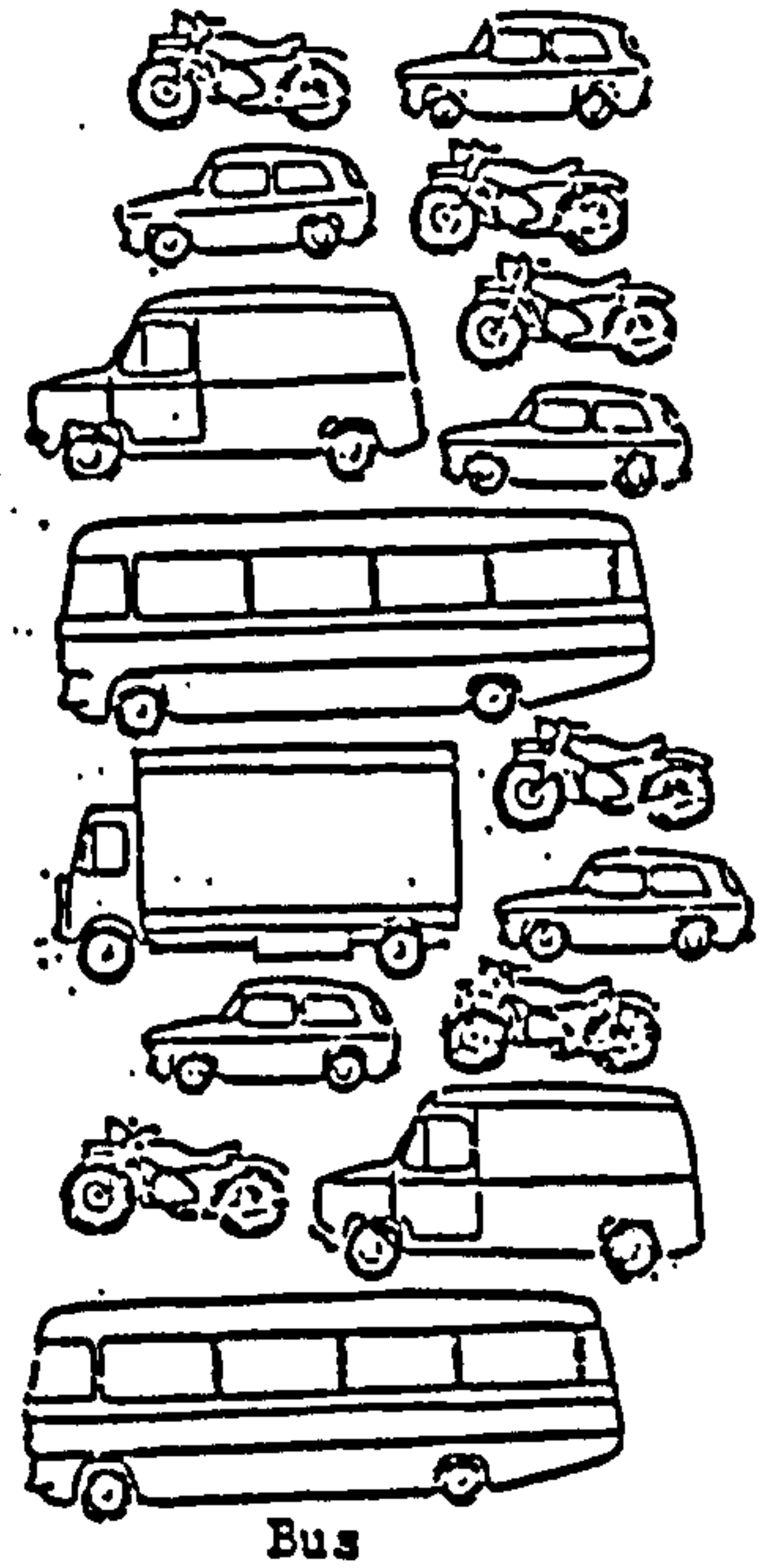
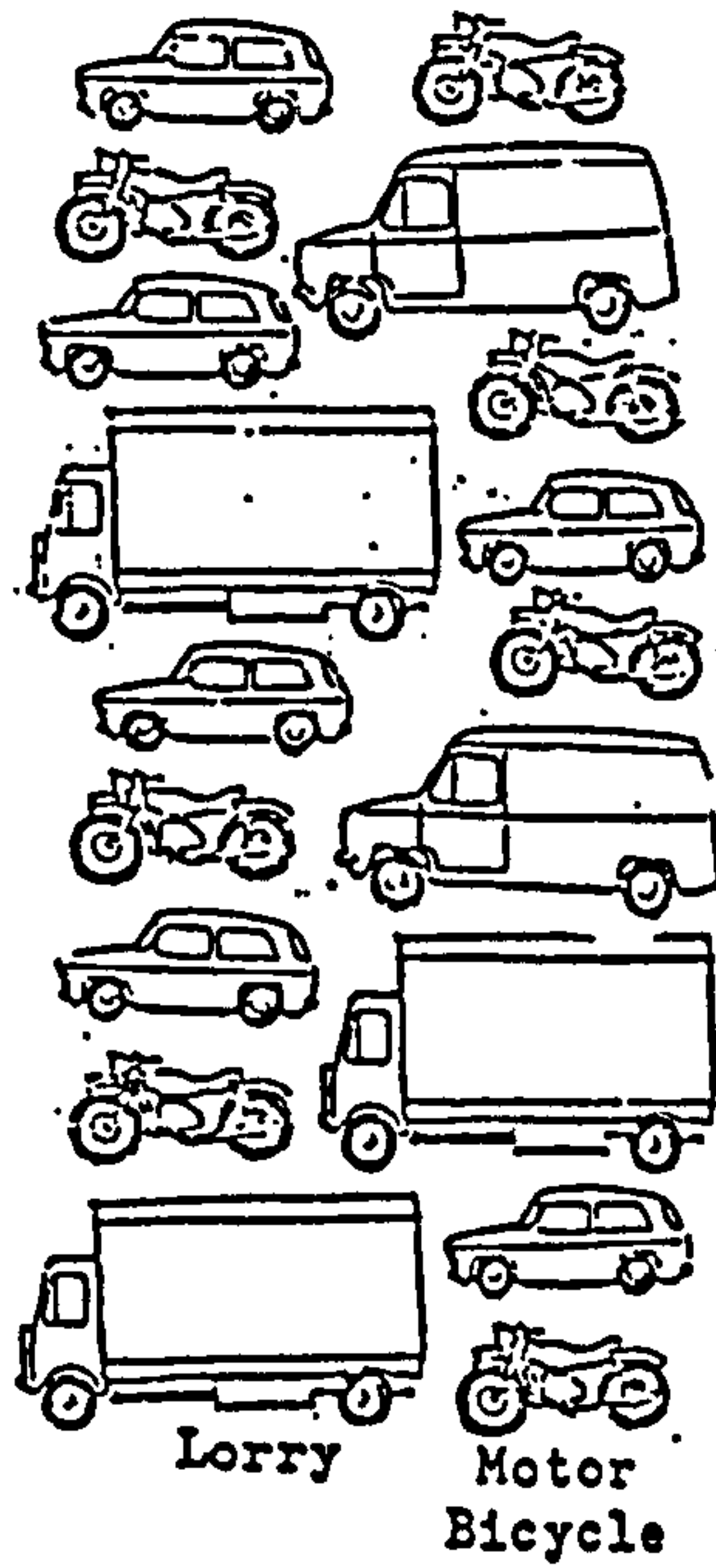
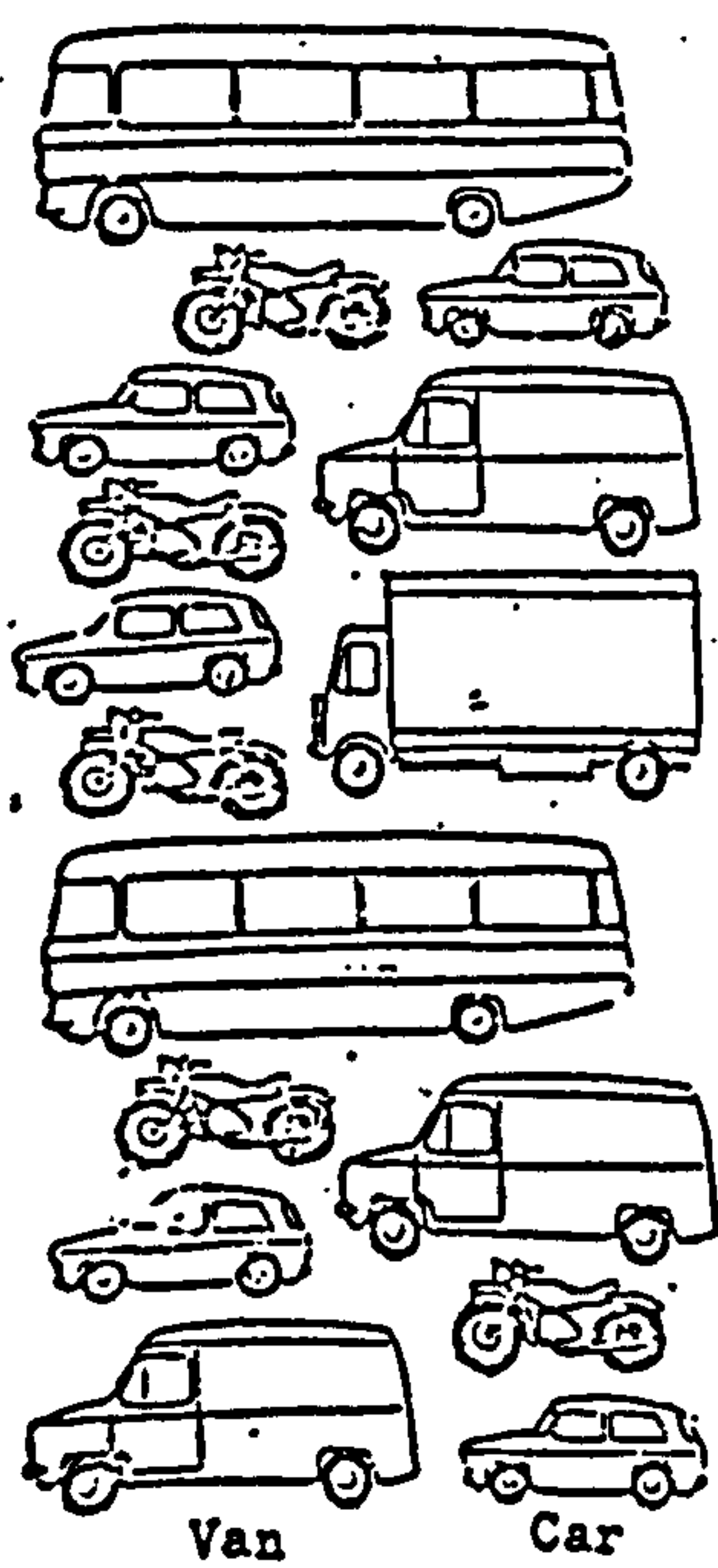
Please leave  
for your  
teacher

a  
b  
c  
d  
e

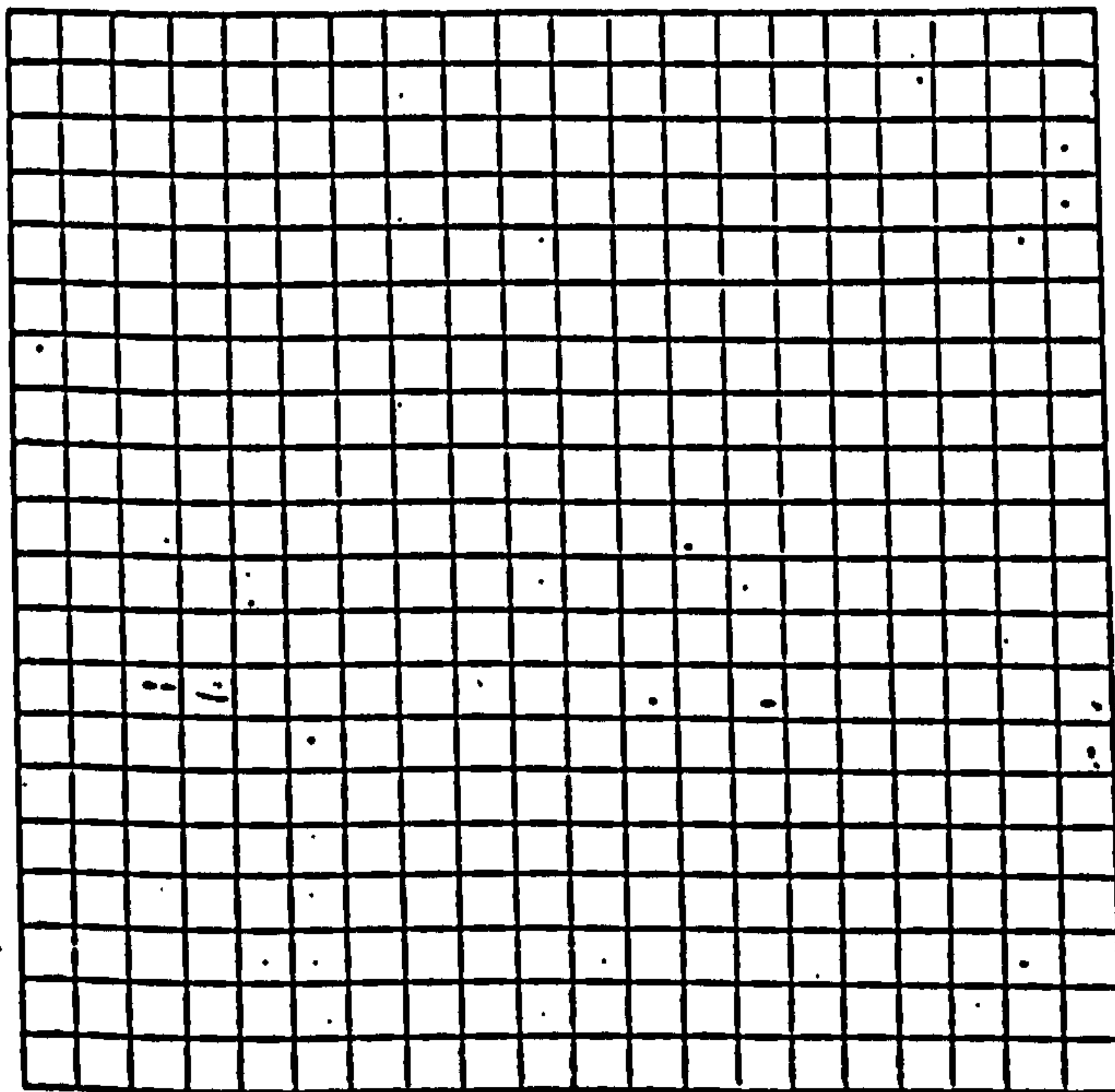


## APPENDIX 10.4      Structured activities for everyday mathematics.

5. A man did a survey of the vehicles in a car park. He tried to show what was there by drawing the picture below.



In the space below make a block graph to show what there was in the car park.



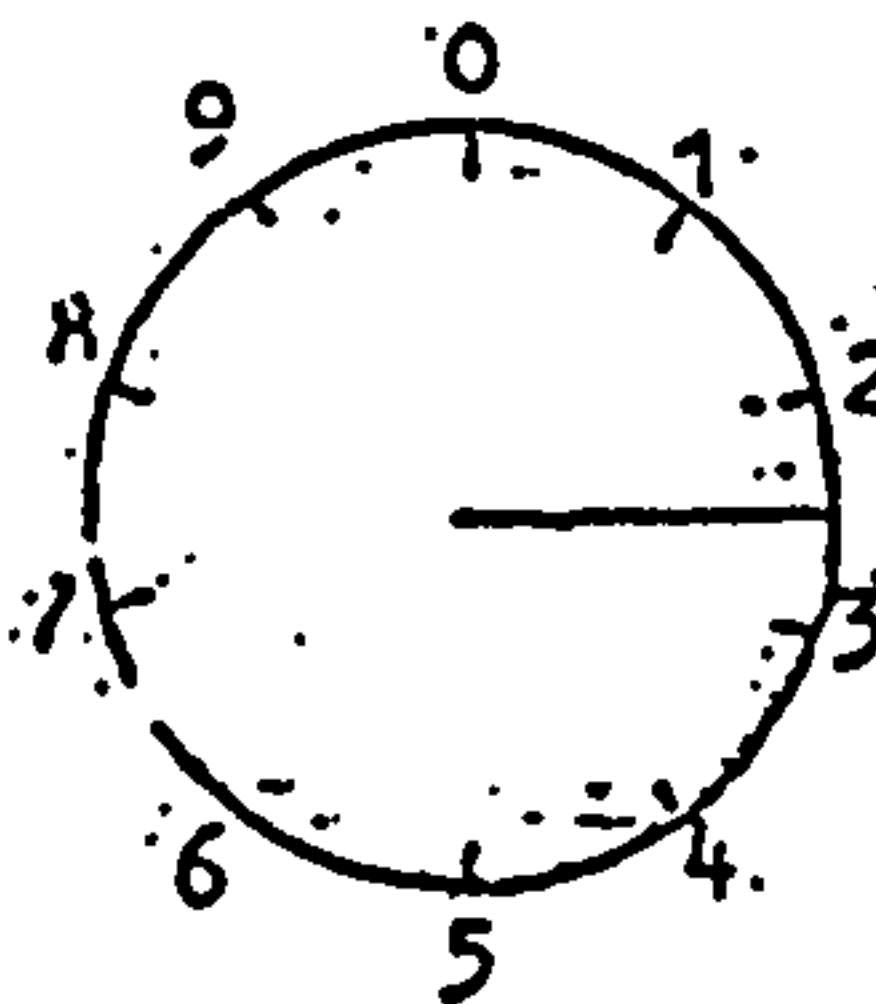
Please leave for  
your teacher

a  
b  
c  
d 1  
d 11  
d 111  
e

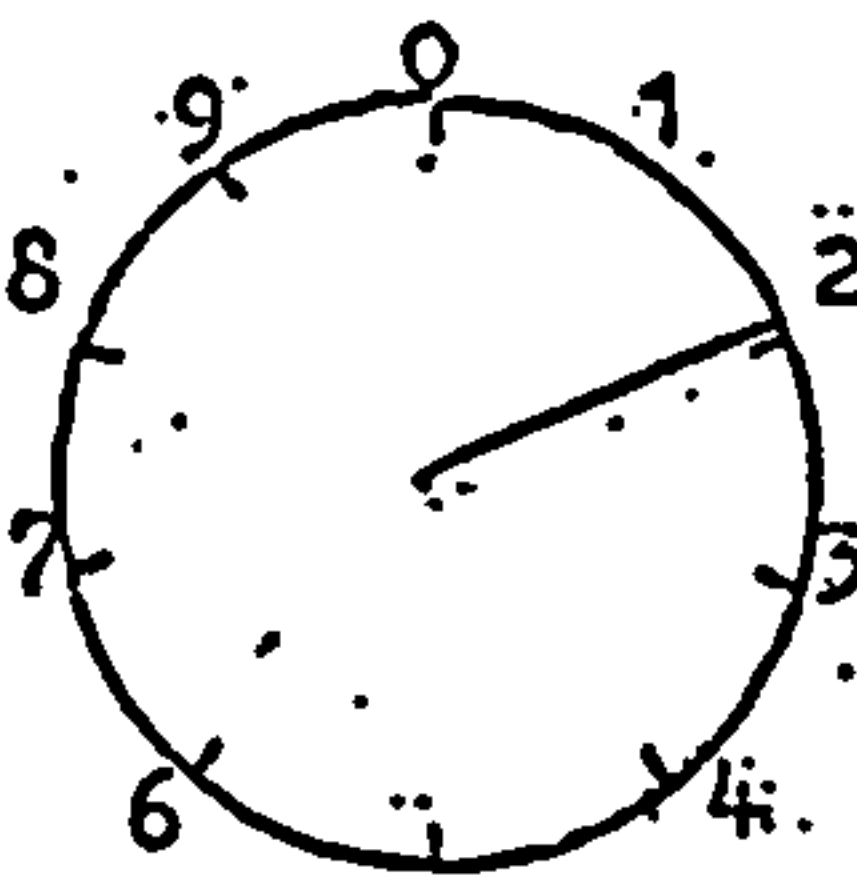
APPENDIX 10.4    Structured activities for everyday mathematics.

6. The gas man came to read the meter. He had to write down the reading from the dials. This is what he saw:

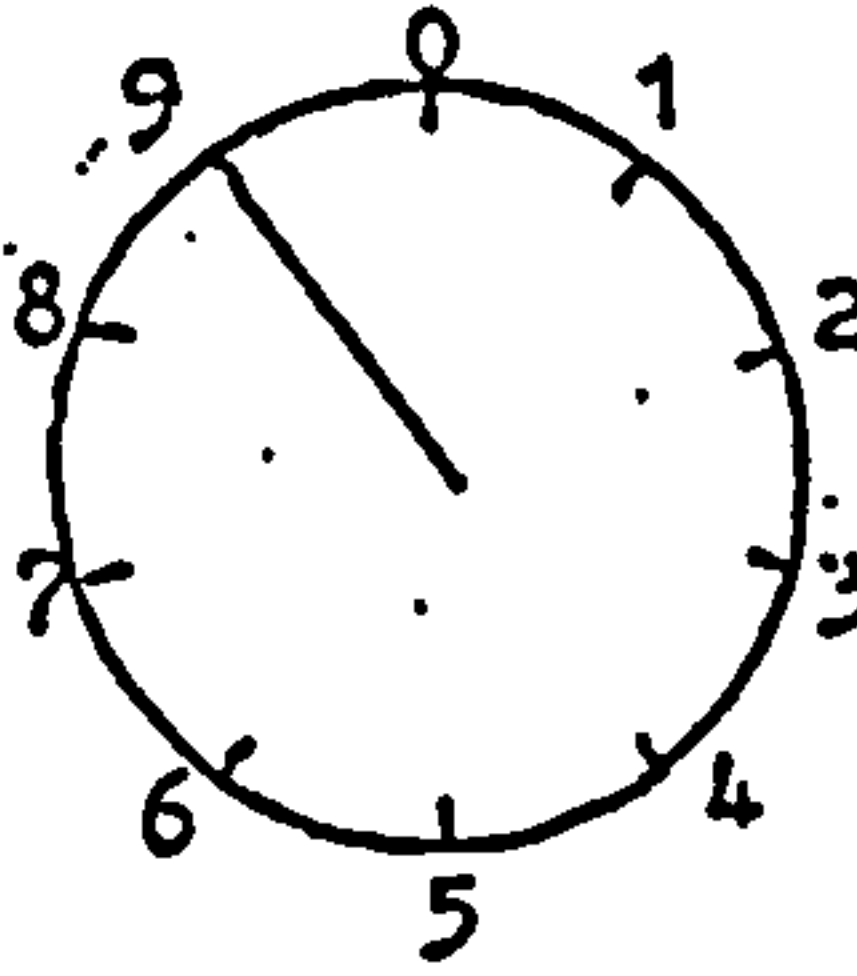
Please leave  
for your  
teacher.



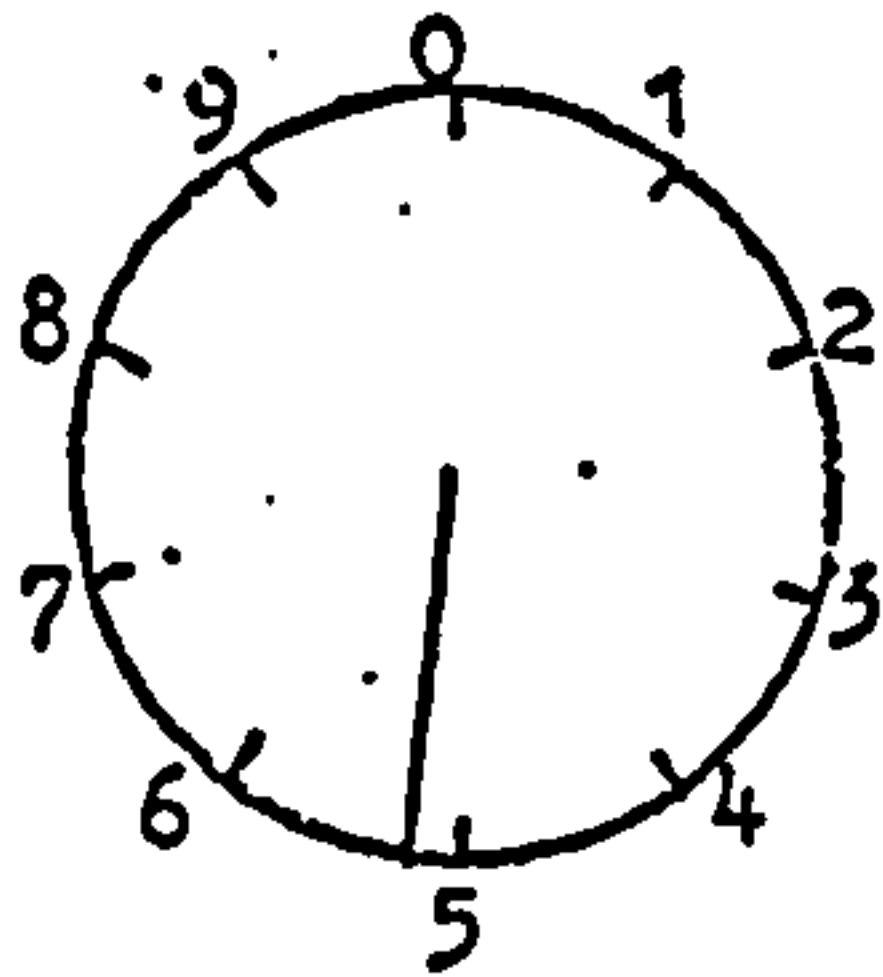
10,000 units



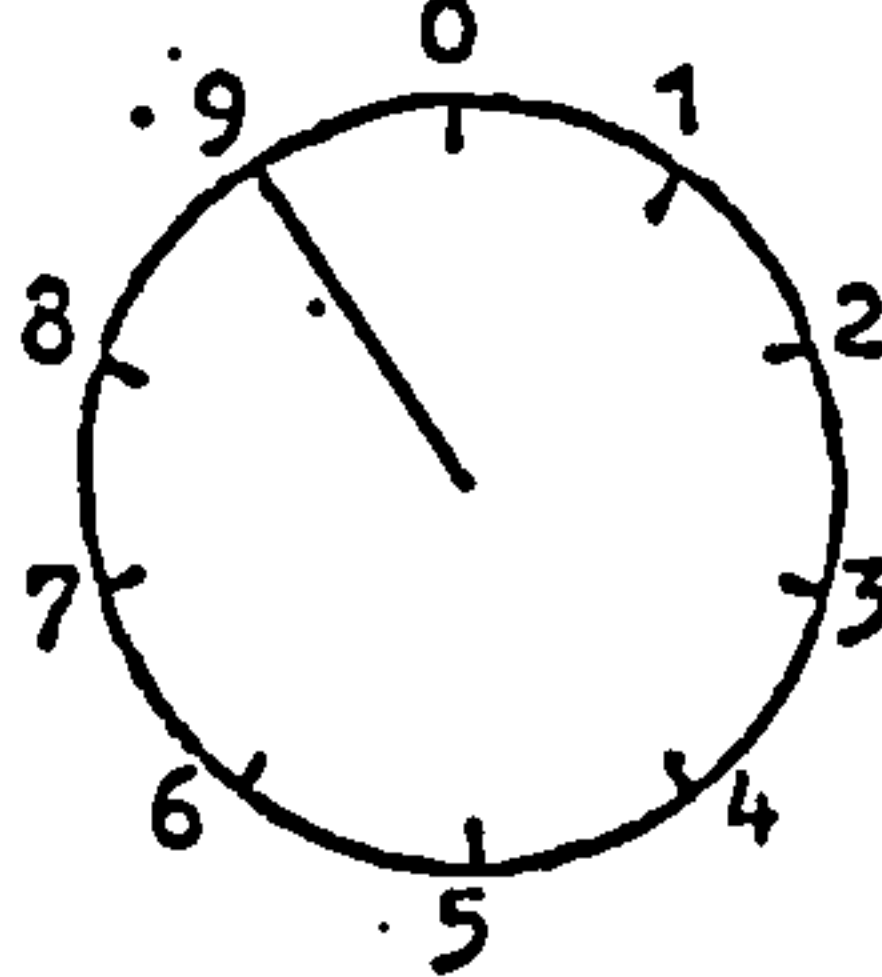
100 units



1 unit



1,000 units



10 units

- a
- b
- c
- d i
- d-ii
- i iii
- d.iv
- d v
- e

Write down the reading he took  
in these boxes

--	--	--	--	--

When the gas man called three  
months ago the meter read

2	4	1	9	9
---	---	---	---	---

How many units of gas have been used? .....

What will the meter read when one more unit of gas has  
been used?

.....



APPENDIX 11.1

Scoring scheme for structured activities for listening with concentration and understanding.

Scoring system	Listening with concentration and understanding	
	Each task scored according to correctness of each response	
Maximum score possible	Following instructions	13
	Sequencing	10

APPENDIX 11.2

Scoring scheme for structured activities for inventiveness and creativity.

Scoring system	Aspects of inventiveness and creativity
	Originality scored according to proportion of pupils with drawing of similar content
	Appropriateness scored by independent ratings of judges
Maximum score possible	Originality4
	Appropriateness3



APPENDIX 11.3

Scoring scheme for structured activities for acquiring information.

Scoring system	Acquiring information other than by reading	
	Comprehension exercises scored according to number of correct responses	
	Formulating questions scored according to	
	(i) number of questions	
	(ii) number of themes dealt with	
	(iii) number of 'open' questions	
Maximum score possible	Comprehension	12
	Formulating questions	no limit

APPENDIX 11.4
Scoring scheme for structured activities for everyday mathematics.

Section	No.	Description	Score	Max Score
A	1	Addition	5 sums in each case	20
	2	Subtraction	1 mark for each correct	
	3	Multiplication	answer	
	4	Division		
	5-12	Fill in missing sign in an equation e.g. $5 \quad 7 = 12$	1 mark for each correct answer	8
B	1-5	Impossible sums	1 mark for each correct answer	5
	6-10	Matching written problem to number sum	1 mark for each correct answer	5
	11-15	Write number sum to match written problem	1 mark for each correct answer	5
C	1	Bus timetable		10
	2	Drawing a map		13
	3	Class timetable		12
	4	Classification		6
	5	Drawing a block graph		10
	6	Meter reading		10

The following key was provided for each of the questions. Teachers indicated the degree of help offered to the pupils by ticking the appropriate letter in the help column.

The teacher:

- a. read the instructions.
- b. rephrased the instructions using different terms.
- c. gave other illustrative examples.
- d. explained the process of solving the problem.
- e. worked out the solution for the pupil.

Marks were deducted from pupil scores dependent on the degree of help provided for particular questions.



**APPENDIX 12      Tape prepared for the moderation of teachers'  
assessments of the child's ability to convey meaning  
clearly and accurately through speech.**

## **APPENDIX 13            Notes to application groups of teachers on the teacher-based assessments.**

### **NOTES TO TEACHERS**

Now that the research programme has begun we would like to start the teacher-based assessment which you have agreed to help us with. In order to help you plan these around your normal class activities more details of the timetable and the procedure involved are given below:

#### **Provisional time-table**

##### **Autumn Term**

- 1 The assessment of the child's ability to listen with concentration and understanding.
- 2 The assessment of some aspects of the child's inventiveness and creativity.

##### **Spring Term**

- 1 The assessment of the child's ability to compute in the four arithmetic rules using his knowledge of number, multiplication tables and different units of measurement.
- 2 The assessment of the child's ability to be able to acquire information other than by reading e.g. by asking questions, by experimenting, by watching television.

##### **Summer Term**

- 1 The assessment of the child's ability to use mathematical techniques in his/her everyday life, i.e. by estimating distances, classifying objects, using money.
- 2 The assessment of the child's ability to convey meaning clearly and accurately through speech for a variety of purposes: description, explanation, narration.



## The aims of the teacher-based assessment

The assessment procedures are being developed by groups of teachers to monitor pupils' progress in the areas listed previously. Such progress cannot be assessed by standardised testing procedures. We, therefore, need your help and co-operation, in making these assessment.

## Assessment procedures

The assessment of the child's performance is based on a checklist of abilities. These act as a criteria for judging the child's *level* of performance. We appreciate that it may not be easy to use the checklist in everyday class situations and the assessment is, therefore, in two parts.

- 1 *Preliminary assessment:* We would like you to try out these checklists of abilities making use of your previous knowledge of the children and by observing their behaviour in the specified area during a given week. *You will not have to set up any special activities for the children in order to do this.*
- 2 *Structured activities:* These have been developed by teachers and relate to the specific abilities listed as criteria on the checklist. We hope that these activities will make it easier for you to identify the level of performance of particular children. At the end of the exercise you will therefore have the opportunity to revise your preliminary assessment should you wish to do so.

What we are asking you to do:

*Week beginning 1st November*

You will be given a classlist and the checklist of abilities. We shall then ask you to rate the child's level of performance in *inventiveness and creativity* on a three point scale as outlined.

### *Week beginning 8th November*

- i Present the structured activities for *inventiveness and creativity* to the children in your class. As with the course work it would be better if this were done as an ordinary piece of classwork, rather than as a test. There is, therefore, no need for all the children to do it at the same time, if they would normally be doing such exercises individually or in groups.
- ii Assess each child's level of performance in the area *listening with concentration and understanding* during this same week. You will be asked to make your preliminary assessments using the procedure outlined in the following section.

### *c Week beginning 15th November*

Present the structured activities for *listening with concentration and understanding* to the children in your class. Again it would be better if these activities were given as an ordinary piece of classwork rather than a test.

Similar procedures will operate for successive terms for the other areas.

## **Evaluation**

When these exercises are completed you will then be asked for a comment on their suitability using a structured report form. Among the points of interest will be your estimate of the appropriateness of the activities for your children's age range, the reaction of the children and whether it caused you to revise your estimate of some children's performance.

## **Feedback**

At the end of the third week the observer or a member of the research team will collect the materials and report forms from you. You will receive a progress report on the first two assessments early in the spring term.



## Inventiveness and creativity

The assessment of a child's inventiveness and creativity is obviously very difficult. The techniques which have been devised are based on ideas developed by the teacher's group. The first problem is the definition of inventiveness. This has been done using two criteria. First the *originality* i.e. whether a child or many children produce a particular response, and second, the *appropriateness* i.e. whether the response fits the situation provided.

The originality of a piece of work could thus be assessed on the basis of the proportion of children producing similar pieces of work. For example, suppose you have set your class some art work on Guy Fawke's night. Most of the children's pictures would be likely to include a bonfire, and although each picture may be different in its quality, its content would be similar. However, one or two children in your class might have made very interesting pictures of the rockets exploding in the sky, while several other children had no idea of what to draw and either copied their neighbour or asked you what to do. There could be several levels of inventiveness operating here. Some children appear to have generated an unique idea, others have produced similar pieces of work based upon the same idea while a third group appear to have no ideas of their own.

When completing your preliminary assessment of the child's inventiveness and creativity. We would like you to make use of this idea of the originality of the piece of work. On the checklist of abilities there is a column of originality. In this column a child who you judge *generally* produces original work, as illustrated in the above example, would be rated as A. A child who *sometimes* produces original work would be given a B, and the child who *rarely* produces original work would be given a C.

The assessment of the appropriateness of a child's piece of work is based upon the suitable use of materials or the execution of ideas in relation to the themes involved. Thus it is the context of the piece of work that is important. For example, a child who paints a picture of a bonfire, but includes a large sun would have produced a more original piece of work, but it might not be thought appropriate to the theme 'Bonfire Night'. In assessing the appropriateness of a particular piece of creative work, we should like you to try to use a three point scale as follows:

- A All the elements in the child's work are highly appropriate to the context and nothing could be said not to fit that context.
- B While appropriate, there are included elements which do not fit the context.
- C There is no relation between the context and the presented work.

You will be given a report form on which there is a class list and six columns. The first one is for your preliminary assessment of the 'originality' of the child's work; the second is provided for any revisions you may wish to make. The fourth column is for your assessment of the 'appropriateness' of a child's piece of work and fifth column for any revisions you may wish to make. These assessments should be made in the light of your knowledge of the children and we would ask you to complete this using the letter A, B, or C as appropriate using the above criteria. Please complete the section on originality for all children before going on to consider the appropriateness.

### Structured activities

We have also devised a simple structured activities for the children which can be assessed in terms of originality and appropriateness. This activity is called '*Make a Picture*'. This is a sheet of paper with a circle and a V-shape printed on it. The children are asked to make these into a picture. An example is included at the end of the booklet.

When you give out the picture to the children, we would ask you to say "Here is a piece of paper with a circle and a V-shape. I would like you to make these two shapes into any picture you want. If you want to colour them you may do so." It is best to give the picture out so that the circle and V-shape are on the left and right hand side of the paper respectively.

As the children complete these tasks you may see evidence which leads you to want to revise your preliminary assessment, if so, please fill in your revised assessment in the appropriate columns on the report form.



## **Listening with concentration and understanding.**

The groups of teachers have identified three levels of ability for listening with concentration and understanding. These are given below.

The child should be able to:

**LEVEL 1** Listen and repeat or recall some content.

**LEVEL 2** Repeat or recall the order or sequence of the content heard.

**LEVEL 3** Apply what has been heard and understood.

For example, when you read a story to the children, one child is always able to recall the sequence and order of what has been heard, answering any questions concerning the story. In this case the child would be assessed as having achieved levels 1 and 2, that is s/he can repeat some content in the correct sequence. Often children are given instructions to carry out some task, their ability to do this successfully is an indication that they have achieved level 3. On the report form are nine columns, for LEVEL 1 there are two columns, the first to be filled in with your preliminary assessments, the second for any revision you may wish to make. Columns 4 and 5 are for the completion of your preliminary assessment and any revisions you may wish to make for LEVEL 2, and LEVEL 3 is completed in the same way in columns 7 and 8. Again we would ask you to make a preliminary assessment using the following three point scale on the report form provided.

- A** The child always achieves this level of performance
- B** The child generally achieves this level of performance.
- C** The child rarely achieves this level of performance.

Please complete LEVEL 1 for each child before moving to LEVEL 2, similarly complete LEVEL 2 assessments before moving to LEVEL 3.

## **Structured activities**

A problem here was to find some way of assessing different levels of performance without requiring written evidence from the children, as this is clearly unfair on those who find writing difficult. We are, therefore, making use of a series of picture cartoons and taped sounds. The children listen to the tape and follow the instructions so that no reading or written answers are required apart from the instance of writing their name. There are three parts to the structured activity. These may be given one after the other or at different times.

**PART 1** Identification of musical sounds

**PART 2** Listening to a sound story and choosing pictures to represent it.

**PART 3** Carrying out a set of instructions.

For these structured activities, it is necessary for a tape recorder or cassette recorders to be available for playing the tapes. For convenience we are using cassettes. However, if you would prefer to use a 'reel to reel' tape recorder we can provide you with tape for this. The tapes and associated materials will be delivered to the school for the week beginning the 15th November and will be collected at the end of the week by the observers. Again we hope that if at all possible these structured activities will be given as a classroom activity rather than under test conditions.

**N.B.** If there are any children in your class who have known hearing difficulties you can indicate this on the report form at the bottom of the page in the space provided.

I hope that this has helped to clarify the first term's work on the teacher-based assessment. If there are any problems or queries please do not hesitate to contact us about them. We fully realize that the assessment procedures may involve problems of organisation, but we hope that you and the children will find these activities interesting and worthwhile, and this will in some measure compensate for the effort. So far our experience has shown this to be true. Thank you very much for your help and co-operation.



## ***Summary of Teacher-based Assessment Programme***

### **Week**

<b>1st - 5th November</b>	<b>Preliminary Assessments of the child's ability to show inventiveness and creativity in some field.</b>
<b>8th - 12th November</b>	<b>Structured activity for inventiveness and creativity <i>Make a picture.</i></b>  <b>Preliminary assessment of the child's ability to listen with concentration and understanding.</b>
<b>15th - 20th November</b>	<b>Structured activities in listening skills</b>  <b>a) Identification of musical sounds.</b>  <b>b) Listening to a sound story and choosing pictures to represent it.</b>  <b>c) Carrying out a set of instructions.</b>
<b>Spring Term</b>	<b>Progress report</b>

**All enquiries regarding the teacher-based assessments should be made to:-**

**Anne Jasman, Research Associate, SSRC Programme,  
School of Education, Leicester University, 21 University Road, LEICESTER.  
Tel.No. Leicester (0533) 24211 Extn. 28.**

## **Teachers' report form**

In planning future assessments, the teachers' groups would find it helpful to know how you felt about the materials which have been developed. It would help them if you used the following questions as a framework for making your comments. If there is anything further you would like to add we hope you will do so.

1. Did you have any problems fitting the procedures into your usual class activities?
2. Were the materials appropriate to the age range in your class?
3. Were there any parts of the procedure which the children found difficult to understand?



## Teachers' report form

4. How did children with special learning difficulties cope with the materials?
5. Did any of the children surprise you by their level of performance in these activities?
6. Did you find the checklist clear enough to use for assessing during the structured activities?
7. Did the checklist help you to assess children's abilities during the structured activities?

## Teachers' report form

8. Have you any further suggestions for activities which might be used for assessing levels of performance in this area?
9. Do you think that the criteria are appropriate to the area being assessed?  
Were there other abilities which could have been included?
10. Could these materials be used in a teaching context?



## **APPENDIX 14    Assessment in the primary school: in-service course materials**

### *What do we mean by assessment?*

1. judge the quality of performance
2. involves measuring to provide information on attainment (with any measurement it is desirable to have some ways of ensuring the measuring device doesn't change and that it is appropriate to what is being measured).

### *Why do we assess?*

1. to assess current performance in order to make predictions and selections e.g. 11+
2. to assess present stages of development, ability or skills, in order to match this to future learning situations.
3. to compare groups of children within a school or to national standards in order to evaluate the success of a particular curriculum content, style of management etc.
4. to monitor the experiences and progress of individual pupils in order to provide information to other teachers, parents and institutions.
5. to diagnose particular learning difficulties in order to provide appropriate remedial action.

### *Who do we assess for?*

1. Pupils
2. Other teachers and ourselves
3. Headteacher
4. Parents
5. Other schools
6. LEA/Advisors/HMI/DES

### *How do we assess?*

1. By using standardised testing procedures e.g. reading, comprehension, mathematics, listening, etc.
2. by assessing classwork e.g. rate of working, stage in a programme of work, quality of products - pictures, stories, projects etc.
3. by questioning, talking and listening to pupils
4. by observing children as they engage in activities, which gives insight into their skills, abilities and attitudes.

### *How well do various procedures fulfil the different purposes outlined previously?*



## *Second session: summary sheet*

This session considered two main questions:

1. What content, skills, abilities, attitudes etc., might we wish to assess?
2. How do we record assessments?

As these two questions are interdependent they were considered together.

We considered procedures on the following continuum.

### *Information for the classroom*

Assessments are more detailed, individual, diagnostic designed to help match teaching and learning/monitoring progress over short time periods. Less likely to be used to make comparisons with other classes / groups of children.

### *Information for other institutions*

Assessments are less detailed, cumulative, to provide overview of attainment rather than progress. Collected over a long time span 1/2 year to 1 year. Should be able to compare these with those of others, i.e. understood by other teachers, schools, etc.

### *School record card*

This should include: health and physical development, personality, academic attainment, social background, emotional and psychological problems. Guidelines from Record keeping in the progressive primary school (Rance, 1971).

### *Examples of techniques of recording*

Breakdown of reading skills, language and mathematics content and skills from Recording individual progress: schooling in the middle years (Foster, 1971) and Surrey record card (Dean, 1976)

Some of the areas listed under language and mathematics are readily amendable to diagnostic or standardised assessment procedures. However, others are not; for example, science, topic and project work, oral communication skills, art and craft. Techniques are available for some areas e.g. Match and mismatch (Harlen, 1978), for monitoring progress in science, Listening to children talking: a guide to the appraisal of the child's use of language (Tough, 1976) and Learning through listening (Wilkinson *et. al.*, 1976). It was to help fill some of these gaps in assessment that the work was begun on developing teacher-based assessment procedures at Leicester.



### *Third session - Summary sheet*

#### **1. Description of methods of assessing achievement (from Harlen, 1978)**

- a) Presentation:    Film/tape/slide to a group or individual  
                         Individual problem (oral)  
                         Paper  
                         Normal work
- b) Response:        Oral, written (closed/open), action
- c) Categorization:  Objective, i.e. right or wrong answer  
                         Criteria applied on the spot  
                         Criteria applied to a record
- d) Interpretation:  Norm-referenced  
                         Criterion-referenced  
                         Pupil referenced

**Thus Progress in learning science checklists would be described as follows:-**

**Presentation: normal work ; Response: action ; Categorization: criteria applied on the spot; Interpretation: criterion-referenced.**

#### **2. Purposes of assessment (from Harlen, 1978)**

- a) Purposes affecting individual pupils; e.g. diagnosis, matching, recording, progress, feedback.
- b) Purposes affecting individual pupils through school; e.g. selection, streaming, grouping, passing of information to other teachers.
- c) Purposes affecting individual pupils beyond school; e.g. information to parents, employers, F.E. and H.E.
- d) Purposes of accountability of the school e.g. information for governors, parents, L.E.A., and others within the community.
- e) Purposes of national monitoring; e.g. comparing the performance of groups using different curricula, monitoring trends in national performance.

#### **3. Development of teacher-based assessment procedures**

- a) Teachers were invited to general meeting and given information on the research programme.

- b) Interested teachers were asked to opt for one of the areas given on the list taken from **Aims of primary education: a study of teachers' opinions (Ashton *et al.*, 1975)**
- c) Teachers groups formed to consider the following areas:
  - i. The child should be able to listen with concentration and understanding.
  - ii. The child should develop some inventiveness and creativity in some fields.
  - iii. The child should know how to acquire information other than by reading.
  - iv. The child should know how to convey meaning clearly and accurately through speech for a variety of purposes.
  - v. The child should know how to compute in the four arithmetic rules and use mathematical techniques in his everyday life.
- d) Each teacher made notes under the following headings:
  - i. What does the aim mean to you?
  - ii. What abilities, concepts and attitudes are associated with the aim.
  - iii. What types of classroom activities have occurred where these abilities might be observed.
- e) These notes formed the basis of a discussion paper which was developed through several meetings into a checklist of abilities.
- f) Checklist used by teachers and myself, and the following problems encountered:
  - i. too many children
  - ii. too much activity not related to the aim or skills on the checklist.
  - iii. too much time taken to make assessment.
- g) Preparation of structured activities to match abilities on the checklist.
- h) Trial and revision of checklists and structured activities by teachers' groups into present form (see Appendices 8, 10 and 11)



#### *Fourth session: Summary sheet*

Discussion on the teacher-based assessment materials and other procedures.

The following points were made in this discussion:

- 1 *Relevance* - it is impossible to produce any materials which are directly relevant to everyone's curriculum. With the TBA (teacher -based assessment procedures) it is between standardised testing (which doesn't relate to individual curriculum content) and individual teachers' assessments (which are directly related to curriculum content). With any assessment there must be a compromise between making it relevant to your own situation and being able to communicate this to others.
- 2 *Development* - in the process of setting up aims for oneself and thus designing appropriate assessment procedures, there is the advantage of becoming aware of the categories you wish to assess, and then being able to apply these to everyday situations.
- 3 *Novelty* - although TBA activities resemble everyday class situations clearly there is some 'novelty' in using them and a possibility of biasing in the assessment of children's performances. This is a problem with any procedure other than normal class work; although performance is variable here too.
- 4 *Definition of abilities* - it is possible by discussion within a staff group to reach some sort of consensus on the meaning of criteria in terms of what children do. However, difficulties are bound to arise when new staff join the school who have not been party to the development process.
- 5 *'Fiddling' test results* - this is a problem particularly when the teachers feel that the head or others may evaluate their performance adversely on the basis of their pupils test scores.
- 6 *Structure* - as soon as you begin to assess then the teaching of that topic will improve. It forces you to sit down and think about what you are trying to do and one's teaching improves.

- 7 *Observation/opportunity* - it helps to sit and watch if at all possible, however there is no guarantee that you will be able to see every child when they show evidence of attainment in one of the abilities listed. A framework for observing is useful, since it helps a teacher to focus on particular abilities, one at a time.
- 8 *Assessment of project work* - the detail was rather daunting at first but would be useful once the criteria were familiar. Concern was expressed at how project work can be used as a time filler, the checklist might prompt teachers to think more critically about their use of project work. Often only the product is assessed, but individual children may just copy from books, it is important to assess how children go about doing projects.
- 9 *Overlap* - different aims were not mutually exclusive, and there is overlap between areas. This may be useful initially to focus on a limited range of abilities and gradually build up towards the assessment of project work. The next stage is clearly to refine the checklists, drawing out key elements and match these more closely to everyday situations.
- 10 *Expectations* - Some teachers were surprised with children's performances on the structured activities; some less able did better than expected, some more able did worse than expected.



### *Fifth session: summary sheet*

Details of assessment procedures used in a sample of schools in Leicestershire and Warwickshire were discussed. Two groups produced outlines for assessment materials in areas which they were interested in. Details are given below:

#### **1. Mathematics - Infant Level**

Practical - need to know what maths skills they have mastered, need to list concepts. Pre-number concepts. Grading - size (bigger/smaller taller/shorter). Sorting - size, shape, colour. One to one correspondence. Two to two correspondence (twoness of two). Counting up to 10 - comes out of sorting. Writing to 10, recognition of symbols. mapping addition/subtraction of sets/figures to 10

Assessment by observation, asking, giving tasks that show they understand. Graded as follows: always correct, frequently correct or occasionally correct.

#### **2. Topic work (Junior)**

What should the children do with a book when they've obtained it in topic work?

1. Know what they're looking for - i.e. "I'm looking for information about ...."
2. Survey book\* or section of book, if book is on varied subjects to find out what its about.

\* Subject-based book (as opposed to encyclopaedia).

#### *Assessment*

1. Know what they're looking for. Ask the child "what are you looking in that book for?" The child may be at one of two stages:
  - a) Surveying the general area e.g. "I'm finding out about giraffes" (more appropriate to lower juniors).

- b) Finding the answers to specific questions e.g. "I'm finding out what giraffes eat" (more appropriate to upper juniors).
2. Survey book to find out what its about. Ask the child "What else does the book tell you about?"

The second group decided to continue their involvement by meeting a regular intervals to develop and refine this checklist (see Jasman and Ashby, 1980).



## **APPENDIX 15.1**

### **B.Ed (HONOURS) IN-SERVICE PROGRAMME FULL-TIME A Student's Guide to the B.Ed. (Honours) In-Service Programme**

*What are the aims of the B.Ed. (Hons.) In-Service degree?*

The degree will help you to consider your professional practice in a rigorous and systematic way, relating your studies to actual school and classroom situations. It has two aspects, firstly it seeks to develop the notions of "being a professional" and "thinking professionally" through consideration of the teacher as a decision maker. Secondly, it seeks to inform and apply this analysis to a specific area of professional concern: either *Remedial Teaching* or the *Teaching of Reading* according to your choice of study.

The programme will develop your abilities to analyse professional situations, to critically examine educational information and proposals, to make well founded and reasoned recommendations for professional action and will enable you to evaluate professional activities and results.

*What areas of study are covered in the degree?*

There are two courses: Core Studies and Pathway Studies. In your Core Studies you will examine the complex and specific problems encountered by teachers in schools. Core Studies will consider such topics as the handling of educational ideas, the nature of educational theory and problem solving, critical approaches to education and the relevance of theory to practice.

In your Pathway Studies you will have the opportunity to apply the techniques, skills and insights gained from the Core Studies to the practical problems of either Teaching Reading, or Remedial Teaching. At the same time you will undertake a careful analysis of the theories and information associated with your chosen focus from such disciplines as Psychology, Sociology and Linguistics. Course studies are related throughout to the practice of schools and you will be attached to a school for the duration of your course.

A third element of your studies will be the carrying out of a Professional Project. In this you will have tutorial guidance in selecting a practical problem in school and bringing your studies to bear on its analysis and solution over a period of time, and with particular reference to your own teaching situation.

*At what level are the degree studies?*

The award is an honours degree and will demand a critical and rigorous approach. This is linked throughout to an awareness of the essentially practical nature of teaching, but the concern will be to examine practice in a detailed and systematic way. The emphasis is on developing the power to recognise the nature and extent of professional problems and to generate solutions and recommendations based on a thorough understanding of principles.

*How is the programme arranged?*

You will be required to register during the Summer School preceding your year of full-time study and during the Summer Term and you will be expected to carry out preliminary reading. The course will begin with the Summer School (during the last week of July and the first week of August) which is a key feature of the B.Ed. In-Service programme. During this time you will work intensively on professional studies and have opportunity for discussion and consultation with tutors and other students, both full and part-time.

The Autumn Term will be occupied with Core Studies and there will be no specialisation of interests at this stage other than in a school placement. This early work will be assessed in order to ensure that the course is appropriate to your interests and that your achievement is appropriate for your progress to honours. These grades will not be a part of the final honours assessment. An assignment during the Christmas vacation will lead you into the Pathway Studies (either Teaching Reading or Remedial Teaching) which will occupy the whole of the Spring Term and a part of the Summer Term. The Summer Term will reintroduce Core Studies designed to help you to put your theory into practice through the development of a Professional Project. Your work in these two terms will be assessed by course work assignments and an oral examination, both of which contribute one third towards your final honours grade.



The third element of assessment is the Professional Project. You will be required to identify a project area during the Easter vacation, to refine and develop this into a clear proposal during the Summer Term and to undertake the reading and practical work involved during the Summer vacation and Autumn Term following your full-time study year. This timing has been chosen so that the pay-off from the course will actually be in your own school and not in relation to some artificial situation which we might set up at the College. Tutorial help will be available during the Summer vacation and the Autumn Term when a day school and a weekend school will be arranged to provide tutorials and guidance. There will also be a further day school at the end of the Christmas vacation for the discussion of preliminary drafts of the project report. The report should be submitted by January 31st. for a summer graduation.

## **APPENDIX 15.2**

### **WORCESTER COLLEGE OF HIGHER EDUCATION**

**Faculty of Education and Teaching Studies**

**Board of Studies for B.Ed. In-Service (CNAA)**

**Proposal to the Council for National Academic Awards for validation within the approved Degree of B.Ed. (Honours) In-Service of a Full Time Mode.**

#### *Rationale*

The College has considerable experience of in-service work with teachers embracing professional, education and subject studies. The present proposal recognises an area of professional need for experienced teachers and brings to bear the College's resources and previous experience in meeting this by extending the approved B.Ed. (Honours) In-Service programme to include a full time mode of study.

Three considerations predominate in the proposal: a coherent approach to professional studies, the application of theoretical insights from academic studies to the solution of professional problems and the testing of these solutions in real situations.

A core of professional studies is especially designed to meet the needs of qualified, experienced teachers and has a continuing reference to the professional experience and situation of all students. Appropriate theoretical considerations and analyses are brought to bear on characteristic issues and problems of professional practice. Students will be required to complete a professional project in which course work will be focussed on a field problem in their own school.

Experience with existing in-service courses and consultations with servicing teachers suggests two areas as the focus of studies in the programme - Teaching Reading and Remedial Teaching. The courses consider problems raised by these areas of professional concern and seek practical solutions by drawing upon the insights from relevant theoretical studies. The emphasis throughout is upon the application of theory to practice and a rigorous testing of practice in real situations.



As the programme will be followed by full time study students may be distant from their normal professional situations. In order to ensure that students' studies are continuously tested against the criteria of practical relevance, School Placements will be arranged so that each student has an ongoing contact with a school in the locality of the College. Placements will take account of the student's own professional background and focus of study.

In order to make maximum use of vacation study there will be a reading programme in preparation for a Summer School prior to the full time study. Each student will be allocated a tutor who will be available for tutorial guidance within the programme.

Through their studies and assignments in the programme students will develop and evidence powers of critical analysis and synthesis of both practical issues and relevant theory. These will be brought together in systematic study in which the student will demonstrate the practical professional application of the programme aims and will evaluate the outcomes of planned action in specific settings with an awareness of their wider context. These activities will be pursued in a scholarly and rigorous fashion, employing procedures of enquiry and development appropriate to the achievement of honours level study.

## **Proposal**

### ***1. General Principles***

The degree is a professional degree for experienced qualified teachers. The Board of Studies believes that central to a B.Ed. degree, which is to be acceptable as an in-service qualification, must be thorough and rigorous professional studies, closely linked to theory and knowledge relevant to the particular student's professional situation.

The programme of study relates directly to the actual or intended professional situation of the students, involving Professional Studies directly linked to work in schools. Each student will also make a systematic analysis of an aspect of a professional situation as a substantial part of the programme.

## **2. *Practical Considerations***

The professional situations of practising teachers are such that it is necessary to provide for different areas of interest and expertise. In designing the degree structure the Board of Studies has provided courses of study suitable for a variety of professional needs. Each student has a measure of choice, within the programme, in relation to particular interests and professional goals.

## **3. *Aims***

The principle aim of the B.Ed. (Honours) In-Service degree is to influence and improve professional practice. It seeks to achieve this by a concentration on the professional behaviour of the teacher, by examining the notions of professionalism and accountability and applying this analysis rigorously to selected situations in simulation and then to real situations through school placement exercises. A basic premise of this approach is that practical decisions must be informed by the best theoretical insights available and their outcomes must be objectively evaluated against recognised criteria.

The programme seeks to put students in touch with both "educational research" and "research on education", as distinguished by Elliott, and to enable them to evaluate published reports. It also aims to give them the necessary skills for undertaking educational research as part of the professional approach to the task of teaching.

## **4. *Structure***

- Programme:** the set of courses leading to the degree of B.Ed. (Honours) In-Service.
- Course:** a self-contained scheme of work forming a substantial and separately assessed part of the programme.
- Core studies:** that part of the programme which is followed by all students.
- Pathway studies:** that part of the programme in which students exercise choice according to their professional interests.



## 5 *Modular Structure*

- 5.1 The academic year is divided into three terms, each providing ten full weeks of study time. Administrative procedures are arranged outside this basic teaching term. During the summer vacation prior to full time study a residential school of two weeks is provided.
- 5.2 The programme is arranged in termly elements, each preceded by vacation study. The Summer School will introduce students to the College, its broad approach to the degree and to professional studies. It will enable students to rediscover academic study habits and submit work for scrutiny and advice. An assignment will be undertaken in anticipation of the work of the Autumn Term.

Core studies in the Autumn Term will involve exploration of professional knowledge and skill demanding rigour, independence of thought and self-criticism. The emphasis will be on the consolidation and development of previous professional education and experience in the context of an honours degree programme.

Pathway studies in the Spring Term will consider content and approach appropriate for honours level, drawing upon core studies and on professional experience. The work will require students to be aware of current developments in the area and will examine practical issues in the light of theoretical perspectives, involving the use of primary source material wherever appropriate.

The Summer Term will take as its focus the application of theory to the solution of practical problems. Students' time will be equally divided between further Pathway Studies, a re-introduction of Core Studies focussing on the theme 'Theory into Practice' and individual work leading to a Professional Project.

## 6 *General objectives*

The student will:

- 6.1 become competent in independent study and in handling ideas in the context of educational theory;

- 6.2 acquire insight into the nature of educational theory and its applications;
- 6.3 demonstrate ability to discuss orally and in writing issues and problems in the process of education;
- 6.4 be able to distinguish and identify evaluative and empirical assumptions and assertions concerning the process of education;
- 6.5 show an appreciation of the relevance of education theory to practical issues in the process of education;
- 6.6 understand and be able to apply theoretical insights to educational problems;
- 6.7 be able to suggest and justify solutions to problems arising in their own experience by analysis and systematic investigation of the complex factors involved.

## ***7. Examinations***

All major submissions of work - the Course File, the School Placement File and the Professional Project will be examined viva voce. The examinations will work to a published timetable and each candidate will discuss the work submitted with internal and external examiners for a period of 45-60 minutes. There are no written examinations.



## **APPENDIX 15.3**

### **INDEX TO UNITS**

The course elements listed below are presented in full on the following pages and appear in the same sequence in which they will be encountered by students.

#### ***SUMMER SCHOOL***

Thinking Professionally

#### ***CORE STUDIES***

Learning Difficulties and Reading:  
Analysis of Practice

#### ***PATHWAY STUDIES***

Either Reading Pathway  
or Remedial Pathway

#### ***CORE STUDIES***

Theory into Practice

## **B.Ed. (Honours) In-Service CORE STUDIES FULL TIME**

### **Autumn Term**

**Title:** Learning difficulties and Reading: Analysis of practice

#### *Objectives*

The students will:

- 1 understand the complexity of skills involved in the reading process;
- 2 understand the sources of handicaps and learning difficulties;
- 3 know and understand techniques for the recognition of learning difficulties and special needs;
- 4 describe a wide variety of approaches to the development of the language arts;
- 5 devise learning experiences and choose teaching strategies to meet the group and individual needs of children;
- 6 evaluate children's progress in reading, writing, talking, listening, number and study skills;
- 7 show awareness of current research in the teaching of the language arts and in teaching children with learning difficulties or special needs;
- 8 understand and be able to apply theoretical insights to educational problems;
- 9 be able to suggest and justify solutions to problems arising in their own experience by analysis and systematic investigation of the complex factors involved.

#### *Syllabus*

- 1 Educational data and sources; retrieving and organising source material in relation to educational problems.
- 2 The concept of education; nature of educational theory; value judgements and empirical data; purposes, process and decisions; analysis of selected concepts and decisions.
- 3 Education theory and decision making in the classroom, in the school and in society - the application of theoretical insights to educational problem solving.
- 4 Language development; nature/nurture and language; questions of compensation, modification, intervention; pre-reading programmes; reading, writing, talking, listening in the classroom; current research interests.



- 5 Teaching reading; methods and approaches, research evidence; children's literature; systematic skills development; using context; silent reading; purposeful reading - group methods; critical and comprehension skills; study skills; reading for pleasure.
- 6 The Bullock Report; sources of learning handicaps; individual difficulties and disabilities; adverse environmental conditions; backgrounds of disadvantaged children; diagnostic record keeping; the 'Warnock Report'.
- 7 Assessment of reading and number ability - formal/informal, norm/criterion reference; uses and abuses of assessment; identification and diagnosis of difficulties - physical, intellectual, emotional, social, cultural, ethnic; specific reading disability.
- 8 Working with older children and adults; English as a second language.

### *Teaching Methods and Time Allocation*

Lectures	40 hours	
Seminars	40 "	
Workshops	30 "	
Tutorials	10 "	
Private Study	180 "	
School Placement	60 "	Total 360 hours

Practical work in schools will involve specified assignments agreed between Tutor, Student and School.

### *Assessment Procedure*

A Course File: the file will contain:

- 1 Evaluations of reading programmes, books and materials for 'average','remedial' and 'gifted' pupils.
- 2 Reviews of key books and journal articles relating to language development, the language arts and teaching exceptional children.
- 3 Notes on a range of tests/assessments.
- 4 Reviews of methods relating to the development of primary, intermediate, higher order and study skills.
- 5 Review of learning handicaps and treatments.
- 6 Review of methods relating to the development of number competencies.
- 7 Notes of workshops relevant to professional situation.

## **B Seminar paper:**

Students will present a seminar paper on a selected topic. Topics will be suggested but students may offer their own choice for consideration. Seminar papers will be presented orally to the group in the first instance. The final written version should take account of the discussion arising from the initial paper.

## **C School assignment file:**

Students will be assigned to an appropriate school which they will visit regularly throughout the course. This contact will be the focus of school assignments.

- 1 The student will apply a range of formal/informal assessments of reading and number to a range of children and evaluate the exercise with reference to relevant material from test manuals, other reading, the specific needs of the school and the comparative needs of his own professional situation.
- 2 The student will produce a learning programme together with a record keeping scheme suitable for pupils identified in collaboration with the school as having 'special needs'. The programme will have regard to group and individual needs.

### *Weighting*

A	Course file	35%
B	Seminar	25%
C	School assignment	40%

### *Tutors Responsible*

D.S. Dennis, J.I.Ellis, H.Haines, A.Jasman.



## **B.Ed. (Honours) In-Service CORE STUDIES FULL TIME**

### **Summer Term**

*Title*      Theory into Practice

### *Objectives*

The student will:

- 1 analyse and evaluate a complex problem in education from the perspectives of their previous studies;
- 2 evaluate the relative contributions of theories about education to a given problem;
- 3 understand the nature of educational theory as a synthesis of contributory perspectives;
- 4 understand the nature of educational research and its difficulties and specifically the demands of small scale research in classrooms;
- 5 develop a small scale research project in relation to an identified problem area in their own schools.

### *Syllabus*

- 1 The analysis of educational issues from alternative perspectives; descriptive and evaluative approaches.
- 2 Curriculum integration: nature of knowledge; sociology of knowledge; teacher's and pupil's models of reality; cognitive power and child development; the content of the curriculum; integration of content and objectives.
- 3 The management of learning; disciplines of knowledge; integration of space; integration of time; group dynamics and social integration; differentiation of activities and materials; evaluative criteria.
- 4 Education and schooling; implications of integration; social and individual purposes; value judgements in education; the functions of the school in society; the role of theory.
- 5 Educational enquiry; research methods - survey - experiment, sampling, variables, data analysis; applicability to classroom research; the teacher as researcher; reading research reports.

- 6 The teacher as change agent: classroom evaluation and analysis; development of individual projects.
- 7 Teacher as change agent: the curriculum and organisation of the school.

*Teaching Methods and time allocation*

Lectures	10 hours	
Seminars	10 hours	
Group Work	10 hours	
Tutorials	10 hours	
Private study	80 hours	Total 120 hours

*Assessment*

These studies will immediately precede and lead into the work for the Professional Project (Appendix 16) and will be assessed as part of the project, through the skill and understanding shown by the student in developing the theoretical underpinning of the practical work undertaken.



## **APPENDIX 15.4**

### **Teaching arrangements and assessment criteria**

#### **I *Teaching Arrangements***

1. Lectures will provide a framework of understanding and cover the main areas of the syllabus.
2. Workshops will allow students to evaluate published materials, to discuss materials with tutors and fellow students, to produce their own materials to meet identified individual needs of their pupils.
3. Seminars will allow students to scrutinize each others appraisals of key issues and allow interaction between students concerned with pupils having differing needs. They will put students in touch with a wider range of sources than could be encompassed individually.
4. Tutorials for individuals or small groups will allow maximum specific guidance for students in relation and interpretation of relevant texts.

#### **II *Assessment Criteria***

- 1 A good seminar paper will reveal:
  - a) an awareness of relevant key issues.
  - b) an ability to select and synthesise details from a range of relevant sources.
  - c) an ability to present a methodical, coherent account of the problem under discussion.
  - d) an ability to respond appropriately to matters raised in the group discussion.
  - e) an ability to reflect the group discussion in the final written presentation of the paper.

Seminar papers should be between 2,000 and 3,000 words long.

## **2 A good file will reveal:**

- a) the student's methodical, coherent and systematic approach.**
- b) comprehensive discussion of the relevant areas with significant and relevant insight.**
- c) precise but adequate justifications for decisions taken.**
- d) interpretation of findings with appropriate selection of significant points.**
- e) comprehensive evaluation including synthesis of lecture material and relevant reading.**
- f) evidence throughout of critical analysis rather than descriptive reporting.**
- g) appropriate clear style and presentation making a coherent file.**

**The file should be approximately 7,000 words in length.**



## APPENDIX 15.5

### Guidelines to students on in-service courses: carrying out school based assignments

- 1 The title and outline of the assignment should be agreed by your tutor and by the head teacher of the school.
- 2 The report form should be completed by you to give details of the outline and must be countersigned by your tutor and head teacher *before commencing the assignment*.
- 3 Should it not be possible to carry out the assignment as originally planned by yourself and the tutor you should amend the assignment in consultation with the head teacher and have the amendment confirmed by the tutor *before commencing the assignment*.
- 4 The report form should be handed to the head teacher on the commencement of your assignment. It will be used by him to provide a report to the College on the completion of the work. Please ensure that you have entered at the foot of the report form the date on which the report should be returned to the College.
- 5 Whilst the assignments are designed to relate directly to the actual professional work in which students are involved, it is inevitable that from time to time the responsibilities of teaching in other areas will come into conflict with assignment work. It is of course essential that precedence is given to professional responsibility to the pupils and the assignment report should be annotated appropriately. *Involvement in project or assignment work should never be a reason for failure to carry out other professional work.*
- 6 The customary procedures to ensure professional confidentiality must be observed; in summary these are:

#### 6.1 School documents

No school document or copies should be included in an assignment report without the explicit permission of the head teacher for them to be used in this

way. Should be documents contain any identifying names, these should be erased or occluded.

## **6.2 *Individual pupil records***

No data abstracted from individual pupil records should be included in your reports without the explicit permission of the head teacher. Any such data should be identified by code names and the identity of the pupil must not be revealed in the documentation of the assignment.

## **6.3 *Identification of school***

A code name should be used throughout for the identification of the school.

7. At the conclusion of your assignment the head teacher must have adequate opportunity to look over the completed work before its submission.
8. The completed assignment must be submitted by the due date in the manner and binding/covers indicated by the course tutor.



## **APPENDIX 16**

### **THE PROFESSIONAL PROJECT**

#### **1 Introduction**

In order to proceed to the degree of B.Ed. (Honours) In-Service, a student will be required to carry out a Professional Project, approved by the Board of Studies and carrying a credit rating equivalent to course work or examination grades. The project will show evidence of sustained individual work over an extended period of time.

#### **2 Aims and objectives**

##### **2.1 *General aims***

The aim of the Professional Project is to provide students with the opportunity to develop autonomously a project of their own choice, producing work of some depth in an appropriate area of professional practice and analysis.

##### **2.2 *Objectives***

Through the project students will demonstrate the following abilities:

###### **2.2.1 Analysis**

The analysis of a professional issue and the recognition of material relevant to associated practice, decisions and judgements; the analysis of previous work in the field, including an awareness of its point of view, bias or line of thought.

###### **2.2.2 Synthesis**

The synthesis of knowledge of previous work in the field, and its incorporation and development in the project; the organisation and presentation of any new data, evidence and opinions collected in the course of the project.

### **2.2.3 Evaluation**

The awareness of, and the ability to apply, procedures and standards appropriate to the issues examined in the project.

### **2.2.4 Communication**

The effective communication of the material of the project in accordance with criteria relevant to the procedures, studies and practice involved.

## **3 *Acceptability of topics***

Acceptable topics will allow an approach in which the student will systematically examine a professional issue in an actual setting. This may involve the development, carrying out an analysis of professional work by the student, or in the detailed analysis and evaluation of the professional practice and procedures of others. The project will be characterised by a reasoned and careful consideration of practice in the light of appropriate evidence and theory, and will include appropriate evaluation procedures.

## **4 *Presentation***

As students will be free to work on a wide variety of issues and professional settings there will be considerable variety in the manner of presentation, including the submission of practical work. In such cases it is expected that written evidence of relevant analysis, synthesis and evaluation will be presented.

## **5 *Procedure: registration***

- 5.1** By the end of the Spring Term, candidates, in consultation with the advisory tutor, will submit a title and outline proposal for their professional project to the programme committee for approval.
- 5.2** Not later than the first Friday of the Summer Term each candidate will submit the definitive title and a brief outline of his project to the committee which shall report these to the Director of Studies of In-Service B.Ed. The Director of Studies will then publish a consolidated list of students showing the approved



titles and fields of their projects together with the names of the supervising tutors. These lists shall be made available to the Board of Examiners.

5.3 Completed projects, including all practical and written material, shall be submitted not earlier than January 1st of the year following the taught course and normally not later than August 31st of that year.

5.4 Projects shall be submitted, as directed, between January 1st and 31st of the year in which the student wishes to be finally assessed.

5.5 Each student shall have a final tutorial with his supervising tutor to discuss the work as finally submitted.

## **6**     *Procedure: assessment*

6.1 The panel of examiners shall appoint an examiner in accordance with the directions of the Board of Examiners.

6.2 Supervisors of projects will submit a report on the project work of the student to the examiners, and shall make an initial assessment of the grade of the completed study, in accordance with the directions of the Board.

6.3 The examiner or examiners responsible for each study shall recommend the grade to be awarded to the Board in accordance with the directions.

6.4 The assessment procedure will include a viva voce examination.

## **7**     *Supervision of Professional Projects*

7.1 The B.Ed. In-Service Board shall nominate supervisory tutors for those projects which have been given approval.

7.2 Supervisors will provide regular tutorial guidance in accordance with the directions of the B.Ed. In-Service Board.

7.3 Students will attend for tutorials and make submissions of work in progress as required by supervising tutors in accordance with the directions of the B.Ed.

### **In-Service Board.**

- 7.4 In nominating supervisors the B.Ed. In-Service Board shall have regard to the following criteria:**
  - 7.4.1 Tutors will have recent experience as teachers and examiners on an Honours B.Ed. programme. Normally they will be required to possess a higher degree. They will be required to maintain professional contact with schools.**
  - 7.4.2 In addition, a supervisor shall himself have had appropriate research, or other relevant experience.**
- 7.5 The student may also be assigned to an associate supervisor, who will assist the supervisor.**
  - 7.5.1 An associate supervisor will have experience of supervising and assessing dissertations on an Honours B.Ed. programme.**
  - 7.5.2 An associate supervisor will normally fulfil a specified specialist function e.g. supervision of the use of computer facilities for data analysis.**
- 7.6 Where approved projects involve particular skills in experimental design or statistical analysis the Board will arrange specific programmes of tutorials - either for individuals or small groups - to ensure that suitable levels of competence are attained.**



## APPENDIX 17.1      Summary of responses made by application groups of teachers to the teacher-based assessment questionnaire for listening with concentration and understanding

### Comments on teachers reports

#### *How do the materials fit in with usual class activities?*

Similar difficulties of timing and organisation arose with this exercise as with the activities for *inventiveness and creativity*. A special difficulty was the need for quiet in order to hear the tape. This was a particular problem in open plan situations especially when carpenters were putting up notice boards at the time.

#### *How appropriate are the materials to this age and ability range?*

Most teachers found the materials appropriate to the age and ability ranges in their classes. Inevitably with children from 8+ - 11+ some parts of the exercise were thought too easy or too difficult for some of the children. The 'musical instrument' section presented few problems, and most children coped well with 'following a set of instructions'. However, the 'sound story and sequencing' was difficult. Although the idea was considered appropriate the actual tape could be considerably improved. The main criticisms were summed up by one teacher's comments: *"the sound story was not explained clearly enough, there were no breaks between each set of pictures and it was not clear which set of sounds corresponded to which pictures"*. Criticism was made of the technical production of the tape; for example, parts of the sound story were hard to distinguish, or the tape was too quiet to use in the classroom. Some teachers said their children had little difficulty with the exercise: *"it ties up with S.R.A. laboratory procedures"*, others overcame some of the problems by repeating or replaying parts of the tape when it was not clear enough to be heard.

The responses of children with special learning difficulties were varied. Teachers commented that more precise instructions for noting choice would have helped and that these should be given more slowly as these children tended to panic if they missed a set of instructions. However, one teacher commented that they coped *"fairly well, as long as they listened"*, and another was particularly struck by the capability of his remedial group who coped as well, if not better than, the more able children in his class. From the comments it was clear that surprises occurred in both directions. Some more able

children found difficulty in following instructions and some less able children did better than expected.

### *The checklist and criteria for assessment*

About half the teachers who responded thought the checklist was clear enough to assess class activities. Most difficulties were caused through indicating the level of performance by A, B and C. The procedure has now been changed to a five-point scale which should prove easier to use. The abilities listed were considered too general. The majority of teachers did not find the checklist particularly useful in assessing the structured activities. Criticisms made included: the three-point scale was insufficient for accurate assessment, there was insufficient time to assess every child at the time of the test, the abilities listed on the checklist were too broad to be used and no precise instructions were given about the nature of assessment during the activities.

### *Other ideas and the use of structured activities for teaching*

Other activities mentioned for assessing listening ability included; the use of SRA materials, listening to a short story or non-musical sounds. A number of comments were made here on the inadequacy of the structured materials. For example, one teacher noted that some children rapidly became fatigued and lost concentration if instructions continued for a long time. It worried some teachers that we attempted no assessment of long term memory. One teacher commented that whereas in real life situations information is picked up against a background of other stimuli, some relevant and some irrelevant, in the exercise the children were cued in unnaturally. Other abilities mentioned which could have been included on the checklist were; the time span of concentration and an understanding of content.

Most teachers thought the materials could be used in a teaching context. However, it was generally agreed that the quality of the tape should be improved and each exercise should incorporate a range of levels of difficulty.



## **APPENDIX 17.2 Summary of responses made by application groups of teachers to the teacher-based assessment questionnaire for creativity and inventiveness**

### **Comments from report forms**

#### *How do the materials fit in with usual class activities?*

Two types of problems were met in trying to fit in the procedure with usual class activities. First, some teachers found it necessary to organise the activities on a class basis to prevent copying. Where group activities or an integrated day were operated some teachers found it difficult to incorporate them into their routine. One teacher commented that more original work was produced when no attempt was made to prevent copying. Second, in some schools the short time allowed for the activities posed particular problems; for example, when visits had been previously arranged, when teachers had special responsibilities and other classes to teach, and where schools were undergoing reorganisation.

#### *How appropriate are the materials to this age and ability range?*

Most teachers thought that the materials were appropriate to the age and ability range in their class. The main reservation concerned the 'speech bubble'; for example, the space allowed for writing was too small and children with below average reading ages or writing difficulties found the exercise difficult. Most teachers felt that the children could express themselves better in pictures. Children with special learning difficulties seemed to require more explanation and more time to finish. Some also showed some hesitation in starting and tended to copy. However, at the other extreme teachers noted that most coped surprisingly well, possibly because the materials were pictorial.

In general children had few problems in understanding the procedure, although some were unsure with the 'Circle and V' whether to redraw the shapes to make their picture, incorporate both shapes into one picture or treat them independently. Some children wanted suggestions of what they could do.

Two-thirds of the teachers found the children surprised them by their level of performance. Not surprisingly it was the children who were thought by their teachers to lack originality, and children from remedial groups who most often exceeded their teachers' expectations in this exercise.

### *The checklist and criteria for assessment*

The clarity of the checklist posed most problems in this exercise. Concern arose that the concepts of originality and appropriateness would be different for individual teachers. The levels of performance indicated by A, B and C were restrictive and did not cope with the wide range of performances exhibited within a class. Teachers also found it difficult to make assessments when they did not take one class for the whole day. This arose particularly in open-plan situations. The checklist helped some of the teachers to focus on children's abilities in this area, whilst others found it restrictive. There was insufficient time during the exercise to observe every child, and teachers suspected that children made an immediate shallow response. One teacher felt a more accurate assessment could be made at the end of the year.

A number of teachers felt that both criteria were relevant to the area being assessed given that the understanding of the terms may differ from teacher to teacher. One teacher was concerned that no researcher working in this field had given an adequate definition of criteria for creativity. However, the relevance of the activities for assessing creativity was questioned. For example, some children who are not particularly talented in art or creative writing do have a flair for pottery, craft work or 'verbal' creativity - drama and story telling. Some may not have original ideas in content but may present the ideas in an original way.

A recurring criticism was that one specific activity did not provide enough basis for assessment. This is true but the reason for asking you to assess individuals is to gain a more accurate picture of the children's real abilities. The structured activities provide you with the opportunity for up-grading a child should they show abilities which were not previously recognised. We would not down-grade your assessment on the basis of a one-off experience.



*Other ideas and the use of structured activities for teaching?*

Suggestions for other activities which might be used to aid assessment in this area included: Edward de Bono's dog exercising machine, model making with plasticine or card, observing children in P.E./Dance, creative writing based on a common stimulus, assessing a piece of art or craft work, listing uses of objects e.g. the different uses for a brick and perhaps a battery of picture making and caption tasks, and creative music making.

The responses to the possible use of these materials in a teaching context was varied. Ideas included: using the 'speech bubble' for teaching the use of speech marks, and developing the idea of direct speech, helping a teacher with a new class to grade work appropriately, helping to extend knowledge of children. One teacher commented that a child's response is sometimes almost diagnostic.

**APPENDIX 17.3     Summary of responses made by application groups of teachers to the teacher-based assessment questionnaire for acquiring information other than by reading**

**Comment from report forms**

*How do the materials fit in with the usual class activities?*

Only six of you had difficulty in fitting in the procedures with your usual class activities. This was as a result of vertical grouping within the school, which sometimes made it difficult to arrange for half of the class to complete the structured activities whilst the remainder proceeded on their usual class work.

*How appropriate are the materials to this age and ability range?*

Most teachers thought the materials were appropriate to the age and ability range in their class. Some reservations were expressed. A number of you thought that question 10, which required the children to think up appropriate questions, was a little too advanced for the majority of the age group. The children had difficulty in thinking of a number of questions, then writing these down and deciding which were of most importance.

Children with special learning difficulties, on the whole, coped reasonably well with the activities. Some of them found the comprehension a little beyond them, partly because instructions and content were too difficult and partly because they did not retain much of the information. However, most children coped well with the cartoons and only one teacher commented that they seemed to be satisfied with finding few differences. Some of these children attempted to draw their answers or sought help from the teacher or each other.

The procedure appeared to be fairly clear to most of the children although one teacher had difficulty in helping the slow ones with spelling while the test was on, as other children could have been disturbed or copied ideas. Another said that the children would have found writing the answers easier if they had been provided with lines and a re-cap on instructions was occasionally necessary. Children with special learning



difficulties may have found comprehension more difficult as the questions were spoken and they could not refer back to them once they had been asked.

There were some surprises in children's performances on these activities. A few of the slower workers did better than their teachers anticipated as they were able to complete each task, although one teacher said that many of the children did not put a great deal of effort into answering the questions about the passage, that the spelling was very poor and the children's own questions did not seem to be very imaginative.

### *The checklist and criteria for assessment*

In general you found the checklist clear enough to use for assessing class activities. Reservations were expressed however concerning level 3 as this appeared to be unclear in parts. About four or five teachers found the checklist either too complicated or too detailed with some repetition of areas. One teacher commented that it seemed to be set out so that you would grade the child on his previous school work, as was its intention. These opinions about the clarity of the checklist were reflected in the comments on whether it helped to assess the children's abilities during the structured activities. Several teachers did not find it of any help although some commented that the points in the checklist were valuable in directing their thoughts, making their own assessments of each child and in helping to quantify different aspects. Your comments seem to indicate that the criteria were appropriate to the area being assessed but a number of additional abilities were suggested which could be included. For example, expressive ability through oral speech and drama, acquiring information by logical deduction, and possibly a further assessment to determine retention could all be included. Other criteria mentioned which teachers felt were important but difficult to assess were qualities like persistence and the power of concentration for long periods of time. You also commented that it was difficult to make assessments without intensive observation of children working on appropriate activities and that the structured activities provided were inadequate for this purpose. However, if a more comprehensive series of activities had been provided they would have taken up far more of the children's time.

### *Other ideas and the use of structured activities for teaching*

One other suggestion was made for an activity which might be used for assessing in this area. This was the use of more practical tasks, for example, making simple electrical circuits. The majority of teachers felt the materials could be used in a teaching context but that they were limited. Several teacher commented that they were particularly relevant for use with observation work to train children to look and listen carefully and to take notice of things which happen around them.



**APPENDIX 17.4**      **Summary of responses made by application groups of teachers to the teacher-based assessment questionnaire for everyday mathematics**

**Comments from the report forms**

*How do the materials fit in with usual class activities?*

Two main problems were encountered. Firstly, the time needed to complete the structured activities was underestimated particularly Part C. It was difficult for some of you to complete the exercises as Christmas approached. Secondly, where you had organised the activities on a class basis, the amount of individual attention and supervision required presented some administrative problems. One teacher commented that "*.... in future I would fit the procedures into group activities during the normal integrated day.*" However, where the materials were used as part of maths work on a group or individual basis there were fewer administrative difficulties.

*How appropriate are the materials to this age and ability range?*

The specific points you made for the various parts of the exercise are detailed below:

*Part A*

Questions 1-4. Computation using the four arithmetic rules:

"Questions 3 & 4 (multiplication and division) were too hard."

"Children didn't recognise the symbol  $\frac{1}{2}$ ."

"Too easy to give any indication of the ability of the more able children"

Questions 5-12. Insertion of appropriate sign (+, -, x, ÷, :)

These were not difficult enough.

*Part B*

Comments on this section ranged from "*gave the most problems*" to "*fairly appropriate*".

#### Questions 1-5. Impossible sums

The children found this procedure difficult to understand, one of you commented that *"many children initially attempted to answer the questions"*. Another teacher, however, was surprised by the children's performance as the language was thought to be too subtle for some second language children, but this did not prove to be the case.

#### Questions 6-10. Matching sums to written problems

Careless reading presented difficulties to the children in understanding what was wanted.

#### Questions 11-15. Writing sums for written problems

The example worried some children, some insisting that *"it was 6 sets of 3 and not 3 sets of 6"*, this reversal was confusing. The language of question 9 caused some difficulty and several children asked about this.

#### *Part C*

This section was consistently thought the most advanced, needing a lot of preliminary work and preparation before the children could attempt some or all of the questions. The specific problems associated with each question are listed below:

1. Constructing a bus timetable. *"Some children thought the times given were departure or arrival e.g. 8.10 - 8.30."* Ordering the times presented difficulties in one class.
2. Map. The main difficulty was how to represent 'paces', squares would have helped. Other children were unsure whether to actually measure a pace and convert it into an accurate scale drawing.
3. Class timetable. This frustrated some children, one teacher commented that their school timetable was the other way round and another that it needed a lot of explanation.
4. Grouping shapes etc. The pictorial format baffled some children, the teacher thought they could have coped better with real materials e.g. logic blocks. One of you commented that very few children had any real idea of setting, and several that the children found this exercise difficult to understand.



5. Meter Reading. This question presented most problems. The comments included: *"the numbers were too big"*, *"10's and 1's units were slightly inaccurate"*, *"some children recorded units by reading left to right instead of numerical order"*, *"the units below the dials could be confused with the reading e.g. some children wrote 1111"*, finally two of you felt that it was totally outside the children's experience.

Some general points also emerged. Teachers of 3rd year children found the materials a little difficult for this age group. Most of you thought that it was, as might be expected, appropriate for the most able, reasonably attempted by the middle ability range, and too difficult for the less able. One of the main criticisms concerned the language of the questions. This presented particular difficulties both for you and the children. For example, one teacher commented that *"once the materials had been rephrased or explained to the children they were able to cope fairly comfortably"*, another that *"in almost every case there was difficulty on every sheet, whether it was specific words misunderstood or the entire problem misinterpreted or not understood at all. Most often an illustrative example sufficed to clear up the difficulties..."*. However, despite these criticisms most of you were relatively happy about most of the questions and on the whole felt the activities were appropriate.

Children with special learning difficulties generally found the materials difficult to cope with. This appeared to be a result of both poor comprehension and difficulty in applying mathematical techniques. This was partly overcome by reading and rephrasing the questions to them and giving examples. This was time-consuming and some children became frustrated and eventually switched off. However, several of you found that they coped reasonably well, and even if they performed poorly enjoyed doing it, particularly the treasure map.

There were fewer surprises in the level of the children's performance. Two of you commented that the children did better than expected particularly on the "impossible" sums. Several were surprised by the underachievement of "above average" children.

#### *The checklist and criteria for assessment.*

Most of you found the checklist clear enough to use for assessing class activities, although one teacher thought that children's attainment varied with different activities.

One criticism was that it was *"rather too general to cover so many different aspects."* Again most of you found the criteria appropriate and they could be used to assess children's abilities during the structured activities. One teacher commented however that *"as a lot of the material .... bore little relationship to the classwork... it was rather unfair to relate the children's assessment with their performance in the structured activities. If asked to assess how well each child would complete the proposed tests I would have graded them all slightly lower."*

#### *Other suggestions and use of structured activities in teaching context*

One suggestion was the use of different terminology in an activity i.e. symbols and words. Most of you felt that there was scope for using the materials in a teaching context but with some reservations e.g. using them later in the year when the children have mastered basic mathematical skills sufficiently well to tackle appropriate mathematical problems, and providing more examples and detailed preparation. A marking system was also requested.



**APPENDIX 17.5 Summary of responses made by application groups of teachers to the teacher-based assessment questionnaire for conveying meaning clearly and accurately through speech.**

**Comments from report form.**

**1. Did you find the criteria on the checklist clear enough to use?**

The majority of teachers found the criteria easy to apply with the exception of 6 and 7 on the checklist. Where this was the case, as one teacher commented "I didn't really understand what was wanted. In fact I have been unable to complete the column on the checklist," then the intention was not assessed.

**2. Did you find the tape helpful as a guide to completing the checklist?**

All teachers found the tape helpful with the exception of the two teachers who found criterion 7 difficult to assess.

**3. Did you find any problems in assessing the taped examples of children talking?**

Again all teachers found the examples easy to assess with the exception of criterion 7.

## **APPENDIX 18      Summary of teacher self reports.**

This appendix presents pen portraits of the nineteen teachers enrolled on the in-service course on the basis of their responses in the self reports. The teachers have been identified by an initial to ensure their anonymity. This initial has been used throughout the presentation and discussion of results in this thesis.

- 1 AB taught Art and Physical Education at private co-educational school (4 - 16) of some 200 pupils. He had very little experience of assessment within his school context and noted that he was not involved in decisions relating to assessment. He was only expected to comment on Art 'O' level entries and stated that his knowledge was minimal except at the general level of monitoring pupil performance in the classroom.
- 2 AL taught geography, health education and social studies to examination classes in a comprehensive school (11-16). She was a fourth year tutor and had some experience of writing pupil references. The school assessment policy involved testing the first year on entry with Richmond tests for setting. Effort and achievement assessments were made at the end of the second, third and fourth years. Examinations were given at the end of each year and school reports were based on these results. This teacher expressed cynicism and skepticism on the apparent reasons for keeping records. She was concerned over the degrees of fluidity and permanence of records, what they were recording and the usefulness of the information. She said she had nil knowledge of methods of assessment.
- 3 AM was second in the history department of an 11-16 comprehensive which had recently been formed as a result of secondary school reorganisation. She identified a number of problems associated specifically with the reorganisation such as the high proportion (60%) of less able pupils and travelling on a split site. She taught mostly C.S.E. classes and some health education and general studies. School policy on assessment included subject and form tutor's reports twice a year. Richmond Tests were given in years 1 and 2. Examination results were very important too and she noted that many pupils were entered who should not have been. Her comments on assessment indicated that oral assessment of the less able was often misleading and that there was little time to evaluate pupils' efforts since teaching in remedial classes was centred on containment of the pupils. She also asked for help in subject-based



assessments to solve this problem.

- 4 B was a scale 2 teacher of basic skills in an 11 - 18 comprehensive school of 900+ pupils. He taught mainly years 1 to 3 with some 4th and 5th year pupils. He also had responsibility as a tutor for the general pastoral care of third year pupils. The school assessment policy involved offering a range of external examinations such as 'A' level, 'O' level, C.S.E., 16+, C.E.E., City and Guilds and 365. First, second and fourth year pupils were examined internally twice a year and a report was sent home at this time. There were also half termly tests and continuous assessments. In the third year, fourth year option choices were determined on the basis of an internal examination and discussions between the Deputy head and Year tutor. The reports indicated the age of the pupil, percentage and position in the group for each examination, attitude and attainment. This report was being reappraised at that time. Assessment within the basic skills department involved the use of a number of standardised tests such as Daniels and Diack or Schonell. The tests being used were undergoing a reappraisal. Individual programmes were designed for the pupils with a reassessment every half term. This teacher did not provide information on his concerns regarding assessment.
- 5 C was head of rural studies in comprehensive upper school (14 -18) and was involved with development of a school leaving certificate. Within the school all pupils had an individual leaving certificate. Competition was not encouraged within the school and there were no rankings or examination grades. All pupils were entered for public examinations. The pupil profile on the leaving certificate was designed to reflect only positive comments and examination results were not included. One period a week was allocated to the tutor for the preparation of the report. This teacher expressed a number of concerns largely related to his involvement in the development of the leaving certificate. Within the school system he noted that there was always compromise and there was no time to sort out the practicalities because of time constraints. He felt that the positive nature of expected comments constrained the teacher, that they did not reflect pupil performance and were seen as a waste of time. He also identified the difficulty of deciding who the record was for and commented that class assessments did not feed into the profiles. On a personal level he noted that he did not take the feeder schools assessments into account, relying on his own judgements and that examinations confirm what you know.



- 6 D was employed as Director of Physical Education in a comprehensive school of some 1200 pupils. She also had responsibility for teaching a non-examination child development course to 4th and 5th year pupils (14 - 16 years old). School assessment policy included half termly records of achievement and effort for all pupils. Within physical education performance was recorded in relation the pupil's teaching group. The grades were coordinated by Heads of year and form tutors later discussed progress with the pupils. She considered that she had limited experience of assessment since she was in charge of physical education and viewed the assessment sheets she was required to complete as a necessary evil, since she valued the regular informal discussions which took place with pupils on their individual performance more. She also noted that it was important to monitor less able pupils and troublesome pupils to pinpoint their difficulties. In particular, she objected to a set pattern of grades for achievement since this seemed to devalue the grades and did not reflect the pupil's performance and achievements.
- 7 F was a teacher in a middle school (9 -13) of 600 pupils. She taught general subjects to a class of ten and eleven year olds in a mixed ability group. The school policy regarding assessment was varied. In the first two years GAP and cloze procedures were used to determine reading age. Richmond Tests were used throughout the school which F commented were often used to reinforce the teachers' opinions and that little was done to diagnose individual problems. Grading was on an A to E scale for achievement with the expectation that the grades awarded would conform to a normal distribution curve. There was no restriction on effort grades awarded. Records of achievement based on the Richmond Tests, I.Q. and reading age were referred to for the identification of over- and under-achievement. F identified the diagnosis of individual problems and provision of remediation as of concern. She noted that ultimately pupil assessments were based on teachers' observations of pupils supplemented by individual testing for clarification.
- 8 G was a home studies teacher in a 9 - 12 middle school of 300 pupils. She also was a third year class teacher in mathematics and english and taught french to the first years. The school assessment policy included the testing of reading ages twice and the administration of the N.F.E.R. mathematics test in the third year. She had devised her own test for home studies and graded on an A to D scale with qualifying comments. Her main concern was with the assessment of practical work. She kept brief notes on practical performance since she found it difficult to remember



individual pupil performance overall.

- 9 H was a teacher in a first school (5 - 9) of approximately 180 pupils. She was the advisory teacher for mathematics and taught the eight and nine years olds prior to transfer to the middle school. Assessment policy within the school was based on the local education authority record cards. There was little formal assessment before the final year when the Richmond Tests were given and scores derived for reading age and spelling. A to E grades were given for language, number and attitude and these were seen as being for middle school purposes. She expressed concern that she had little formal experience in assessment and that she needed a lot of help. Her particular areas of interest included: social development, number, reading and language.
- 10 J taught integrated humanities in the lower school in an 11 - 18 comprehensive school of 2000 pupils. In addition she had a 5th and 6th year class. School assessment policy involved the initial grouping of pupils on their junior school record which was amended after one week. As a class teacher she kept detailed records of course work to inform movement of pupils to other groups at the end of the year. Her comments on assessment stressed the importance of knowing pupils and the fact that the school tended to emphasise examination performance. She also found it difficult to accept the pupils lack of ability.
- 11 K was a teaching head working almost full time in the classroom in a rural school with three classes (5 - 11). The school's policy on assessment included the informal continuous assessment by class teachers and formal assessment of reading at least once a year. Mathematics testing took place in the third and fourth year juniors and a non-verbal I.Q. test in the fourth year which contributed to the High school record. He commented that he was beginning to feel that assessment and monitoring of pupil performance in the classroom should be considered more seriously. He considered that he was competent in maintaining records of achievement in reading and mathematics and individual pupil records on other curricular areas, personal needs and achievement. However, he considered that he needed help with the production of records of achievement in language within a broader context than reading.



- 12 L was head of girls physical education in a comprehensive high school (13 - 18) of 900 pupils with responsibility as group tutor for collecting reports, updating confidential records and pinpointing areas of concern. Her comments on assessment related largely to her own practices within physical education. She noted that assessments were based on personal observation throughout the year, which weighed up individual pupils on lesson performance. Success and failure tended to be seen in the context of the "superstar" but she aimed to give some sense of success to all pupils. She sited lack of contact with formal methods of assessment as a weakness but experience as a strength. She was also concerned that quite competent pupils were overlooked, she saw them as having potential to influence others and a resource to be tapped.
- 13 M taught craft, design and technology in a comprehensive school (11 - 16) of some 1100 pupils. School assessment policy involved profile reports on 5th year pupils for school records and employer's references. These were based on confidential records and information from subject reports during the first four years of secondary schooling. These reports were administered by the heads of year and coordinated by the head of the special needs department. Achievement and effort grades were given three times a year for internal use and school reports prepared for parents. M felt confident of assessing pupils' practical ability in his own area. His concerns were largely with the injustice of norm referencing which resulted in some pupils never achieving an A grade because comparisons were made across whole year groups.
- 14 N was a remedial teacher in a Roman Catholic co-educational comprehensive school (11 - 16). She taught withdrawal groups and the bottom sets in mathematics and English. The school assessment policy included the use of Richmond Tests for setting pupils for the first three years, the identification of pupils with special needs by the use of such tests as Schonell, Daniels and Diack and Ravens. Reports were completed at the end of each term with grades 1 to 7 used in each set. This teacher considered it was unfair to give the same grades in different sets. She also felt that there should be proper facilities such as more staff, help from outside agencies, in-service within the school for helping teachers with pupils with learning difficulties and for helping under-achievers.



- 15 O was a science teacher at an 11 - 16 comprehensive school of 800 pupils where she was acting head of physics and taught 4th and 5th year groups half time. External examinations offered by the school included: 'O' level, C.S.E., 16+ in chemistry, T.V.E.I. Non-examination pupils left at Easter and she stated that there was no assessment policy for this group. This teacher commented that she assessed regularly each lesson, linking this to pupil learning. She believed that assessment should also reinforce the pupils in the belief that they could do the work.
- 16 P taught within the prison system at a remand centre for 16 - 21 year olds. He had responsibility for providing classes of an appropriate nature dependent on the charge, remedial tuition and general classes. The assessment of the young men sent to the centre involved an initial brief interview of a few minutes and local tests for screening remediation needs. The nature of his concerns were bounded by this context, such as the lack of time available for assessment, the uncertain conditions regarding the length of the stay within the institution and the "penal attitude". He identified the need to have a quick and accurate way of identifying individual problems and needs before the "student" was allocated to other jobs.
- 17 U was Deputy Head in a first school (4 - 9) with responsibility for the school library and clubs. There was very little assessment for the younger pupils. The teachers completed a record card with general details at the end of each year for reading and number work and included examples of the pupils' pictures and writing. Richmond Tests were also used for pupils aged seven and nine and there was informal assessment by the teachers. This teacher said that he did not feel happy "chatting" with pupils, seeing this as wasted time. He also felt some confusion over aims and the expectations of the Head, parents and the local education authority. He considered that reading records were vital and he was disappointed by the lack of cooperation from the staff. He thought also that folders of pupils' work proved very valuable.
- 18 V was a home economics teacher in an 11- 16 comprehensive school of 305 pupils. In addition to teaching this subject at 'O' level and C.S.E. to years 1 to 5, she taught history, geography, religious education and physical education to years 1 and 2. The school assessment policy involved termly internal examinations. The pupils were then moved from one class to another on the basis of their performance in the examination and their attitude to work or if there was a clash of personality. She

found difficulty in assessing an individual with respect to the group, in other words where the individual performance was good for that pupil but not necessarily in relation to the group as a whole. In this way she would grade carefully monitoring for improvement in an individual's performance. She thought it was important not to compare one pupil against another and to be seen as fair, backing up her assessments. She was worried that assessments could be affected by the mood of the teacher and preconceived ideas.

- 19 W worked in a junior high school (11 - 14) teaching mostly mathematics to lower ability pupils and some religious education. The school assessment policy involved personal pupil information being recorded by the form tutor. Half term assessments were made with grades for effort and achievement. These contributed to the pupil's final record of achievement. This teacher did not feel confident in constructing personal records even though she recognised their usefulness. She also thought that there should be a detailed breakdown of skills in mathematics to keep a check on when and how skills were acquired or not.



## **APPENDIX 19: The in-service course: a chronological summary of the course events, commentary and selected tape extracts.**

The following section describes the course which was implemented during the autumn term of 1984 by reference to significant details of the tape transcripts, reflections from the diary which I kept and, where appropriate, a brief description of content covered to contextualise the reported discourse. The events are reported in chronological sequence and have been used as the basis for the results presented in Chapter 6.

### ***Session 1: 27th September 1984***

I considered it was important to enlist the support and cooperation of the teachers participating in the course before embarking on monitoring events during the workshop sessions. Since the evaluation proposal was not prepared for this session and this was my first opportunity to meet the teachers, I decided not to record it on audiotape. It was not possible, therefore, to check for the match between my intentions and actions in this case.

The session was planned to introduce the concept of appraisal as described by Tough (1977). This content was reviewed on 4th October in the following way.

*"If you remember last week we talked about the concept of appraisal and we looked at the way in which you could write about a pupil at various times, we looked at the ways you could make that a little bit more systematic by dividing it into talk with other pupils, the nature of that talk, general attitude and behaviour at various times and the final example I gave you which was the checklist approach, where particular categories were identified and ticks made against particular times. So those were three methods of recording information about a pupil, in this case devised to make an appraisal of their use of language."*

*"I did not give you the purposes of language that you would be looking for with this particular method. I did that intentionally because what I was interested in getting across to you was the difficulties there are in actually recording information about the pupils based on observation and the different summaries or conclusions that people reach using different kinds of methods because they have different perceptions, different ways of responding to individual children, that they come to the situation with as a result of previous experience ... own ideas, knowledge, expectations and so on."*

This extract supported my interpretation of events that the planned activity was introduced and formed the substantive part of the session. However, there was no opportunity to further validate this since alternative sources of data were not available. This extract also illustrated one of the basic premises which I incorporated into my



espoused theory which was to emphasise the need for accurate observation and appropriate methods of recording so that teachers could make more valid and reliable assessments of pupils. This central consideration is reiterated throughout the sessions which follow.

### *Session 2: 4th October 1984*

In the original course outline it was planned that the teachers should develop their observation skills through the use of the Pupil Record as an example of a systematic observation schedule (see Galton, 1978b). However, it was necessary to start the session with a handout relating to the evaluation project since I wanted to ensure the cooperation of the teachers from the outset. Also, in order to match their needs more closely it was necessary to collect some preliminary information on their previous experience, understandings of assessment and assessment practices within the schools where they had previously worked. Information was requested in the following areas:

- a) an account of the teachers' previous school experience,
- b) a general description of assessment practices within the school and those which the teachers particularly used,
- c) a personal statement identifying areas of strength and weakness,
- d) identification of important aspects of assessing pupils' school performance.

Previous experience in working with both the development groups of teachers and the application groups of teachers during the Oracle research programme had suggested that using such written communications and requests alone was often an unsatisfactory way of gaining teacher cooperation. In addition, teachers were often lacking in confidence and unwilling to ask questions or seek clarification if asked to do something called "research".

Therefore, I anticipated that the teachers might need time to discuss the project and the opportunity to clarify their uncertainties. By using course time for the completion of the self reports I was able to discuss any points of concern during the session. The audiotape transcript of this session suggested that setting aside this time was necessary. This may have been due, in part, to lack of clarity in the written statement presented to the teachers but a number of issues did arise as the teachers completed their self reports. I asked if there were any questions about



*"what my intentions are, any problems you might have with this particular proposal?"*

One teacher [B] responded with

*"This is presumably a two way thing you would help us as we help you."*

I took the opportunity at this point to spend some time explaining and clarifying my intentions. The following extract is quoted verbatim from the tape transcript.

*"I would hope so because what I am doing is evaluating what I provide for you in terms of the materials and the work to see if there is any impact on the work and the way your ideas, your perceptions of pupil achievement, pupil performance (develop) and also if it makes any difference in the way you operate within the classroom. That's what I'm interested in looking for - those are my areas of concern. Clearly it has an impact on you because I have to be prepared for the course, I have taught it before - but I'm going to be monitoring my own inputs closely. I haven't put that on the end but obviously my commitment is to actually make a record of all the course provision, the handouts I prepare, the nature of the contact we have during the taught element, finally perhaps videotaping it and using the transcripts from the tapes in order to get a feel for what the course was like. My purpose is to try to match your experience of the course with my intentions, my aims, what I did - and whether my intentions, my aims were actually fulfilled by the way I delivered the programme for you and whether you perceived it in the way I intended."*

[Inaudible teacher contribution]

*"I am doing, in a sense, what we are expecting you to do in terms of monitoring your own intentions, the way in which you deliver and whether in fact it was received in the way it was intended."*

This clarification was intended to alert the teachers to my espoused theories and theories-in-use related to self-evaluation, offering what I proposed to do as a model of this approach. I then identified the use of course time for the completion of the self reports within the session as a problem and suggested that the teachers might have to finish their self reports later. Since I had made no allowance in planning the in-service course for time to introduce the evaluation project or complete the self reports, this caused me some initial conflict in relation to my dual role as lecturer and as researcher. In the former role I had a responsibility to provide the teachers with the appropriate skills and knowledge as described in the course outline. In my role as researcher it was necessary to set aside some time to explain the nature of my research and elicit the support and cooperation of the teachers. I also needed time for the teachers to provide the additional information which would enable me to investigate their perceptions of their needs and the course in progress. A diary entry also supported this view since I stated that *"I felt concern for the use of course time in conducting this research"*.



It was an unexpected realisation that I was concerned about the amount of time it took for the teachers to provide this additional information on their understandings of assessment, previous experiences and areas of concern. My espoused theory would have suggested that this initial data gathering exercise was valuable since I believed in the importance of observation, diagnosis and assessment of learners in order to match their needs more closely. Whereas in practice I was more concerned with meeting the course requirements as outlined in the programme summary (Appendix 15). This provided some insight into the conflicts which might also be experienced by teachers in using informal methods of assessment and engaging in classroom-based or school-based evaluation.

Other questions which I raised during the introduction of the evaluation included the issue of confidentiality. Since there was no discussion of this I concluded that the teachers were not concerned about confidentiality. However, as the teachers completed the self reports it was apparent that some wanted to provide the information they thought that I wanted rather than express their understandings and concerns as they saw them. The following dialogue illustrates teacher L's uncertainty about the kinds of activity which counted as monitoring and assessment.

*AJ "...I don't want to constrain the response that you give me, because I want it to be the way you feel rather than the way you think I want you to write - the way you interpret is very important."*

*L "...with some assessments practices - obvious forms - exams, gradings on reports, C.S.E.'s etc, do you want us to go into confidential records as group tutors?"*

*AJ "Do you want to?"*

*L "As a P.E. teacher you may find that that's the only thing I know anything about - nothing to do with formal assessment in that sense."*

*AJ "O.K. - you've answered your own question."*

This was immediately followed by a conversation with the teachers which highlighted that they themselves were learners and that my expectations and actions were providing an uncomfortable experience of what it felt like to be a pupil in their own classes.

*AM "...doing it in sentences,"*

*AJ "... whatever you might want to give me."*

*AM "... now we know how the children feel, isn't it? Are we doing what she wants us to do?"*



*AJ "Which is not what I'm asking you to do."*

*AM "I know but that's what children must feel."*

[Audiotape recorder was switched off for short period of time to change cassette.]

*AJ "I don't want to tell you what to write or how much to write. I want you to respond in the way that you feel is appropriate to address the particular points that I've raised. I've not phrased them as questions either just made them fairly brief areas - I've left them fairly open-ended intentionally."*

*H "Well I've finished does that mean I haven't done it right - nobody else has."*

A number of issues were emerging here such as a concern for "right answers", "is this what is expected of me" and "are we giving you what you want". This supported my previous experiences which suggested that the teachers lacked confidence in expressing their own ideas, perceptions and anxieties in the written form but were prepared to talk about these issues informally. I also expected at this early stage of the course that the teachers would need reassurance that they could cope with the demands of further study and achieve success within the course. The teachers proceeded to finish their self reports until some 45 minutes had been spent on this activity.

At this stage I introduced the activity on the Pupil Record observation schedule as originally time-tabled. The preamble to this activity was a review of the previous workshop session which is documented under September 27th. The Pupil Record was then introduced as an example of a particular type of systematic observation instrument. A number of points were mentioned relating to this type of observation schedule; these included the length of training, time sampling procedures, the use of distinct mutually exclusive categories to describe pupil behaviour, the low inference nature of the instrument and finally the differences between observation and inference.

The teachers were particularly concerned with the problems of introducing an observer into the classroom and the effect that this might have on the teacher and pupils who were being observed. This led to a discussion of the differences between participant observation such as I was using within the evaluation and non-participant observation such as the Pupil Record which was used within the Oracle research programme. In addition, the discussion introduced some of the strategies used by observers to minimise the amount of interference in classroom events. These concerns suggested that the teachers were viewing observation as something which was done to them (and their



pupils) rather than a technique which they could use to increase their knowledge and understanding of their pupils' and their own learning.

The remainder of the session was then spent on the introduction, explanation and illustration of the various categories which comprised the Pupil Record (Appendix 3). In particular, the categories under pupil activity were explored. At the close of the session I suggested to the group that they should try out these and other methods of recording and categorising their observations of pupils as outlined by Tough (1979) on placement. The response to engaging in this activity was unexpectedly negative but this is explored in detail in the account of the next session.

### ***Session 3: 11th October 1984***

The original intention for this session was to introduce the teachers to techniques of informal assessment. The checklists developed within the Schools Council project, *Progress in Learning Science* (Harlen *et al.*, 1977a) were selected as an example of such methods of assessment. However, this was not covered within the session as planned. Two factors contributed to the changes in this session. First, the teachers revised their self reports since I considered it important to provide them with an opportunity to read, correct and add material to the summaries which I had made. The time constraint identified during the previous week was thus further exacerbated by this activity. Second, the loss of teaching time resulting from the introduction of the evaluation and completion of self reports meant that the majority of the Pupil Record categories were still to be covered. My diary also reflected this concern since I noted on 16th October the following comments.

*"My separation of the roles of lecturer and researcher may well parallel the teachers' problems associated with the use of observation and assessment procedures based on the informal model. For example, how valid is this use of time, how much certainty can I place on the results, can't I go with my own judgements, why do I need to check them out? Does this help my professional development."*

The teachers spent the first ten minutes of this session correcting their self reports. I then continued the session with a review of the work on the Pupil Record which I had introduced briefly the previous week. The time constraint was again identified in the transcript of the workshop since I stated that the previous week was *"fairly rushed"*. The main points to arise during this review was the discussion of time sampling as a way of selecting observations to record, given that it was impossible to record all aspects of the



classroom and pupil behaviour. After reviewing the pupil activity categories the teachers then coded written descriptions of pupil behaviour using all the Pupil Record categories. As this activity was in progress, I took the opportunity to reflect back to the teachers my interpretation of their concerns regarding the uncertainty of what was expected of them as learners within the course.

*"Somebody commented at the beginning of last session - now we know how the pupils feel - one of the things I think is quite important is that we should be able to learn to reflect on our own learning and use that to inform the way in which you plan and structure your teaching - you are learners in this situation, you're coming into this situation probably not having been involved at that end - if you like - for a period of time and you may find that the adjustment is quite traumatic and painful - having to get used to a new situation - that's one of the things I think you ought to think about ."*

I selected this extract since it illustrates my espoused theories well, in particular, the importance of taking into account the impact of such a course on the feelings of the teachers in line with the recommendations noted previously (Argyris and Schon, 1976; Day, 1980). However, the shift from we to you through the passage was an indication that I still had to come to terms as provider of the course with the trauma of self-evaluation which I was identifying for the teachers. My theories-in-use were not as congruent with my espoused theories as I would have liked to believe. This continued in the following way emphasising the importance of processes as well as content.

*"Keep in mind you can learn a lot from your own experience of being on the course, as a course member, as a learner, as much as the course work we do."*

The next part of the session involved the review of the teachers' codings of pupil behaviour. The teachers raised a number of points with respect to this type of category system. The first concern was that nothing was being recorded about the quality of the interactions. The second concern was with checking agreement between observers within the context of the Oracle research programme. Reliability checks used in the research were described. Further consideration was then given to training procedures and the differences between the use of such a system in the classroom by teachers and when used by researchers. This was another indication that the teachers were not understanding the course content as I had intended. They did not see the relevance of such methods of recording pupil behaviour for their own practice.

The teachers were also concerned about the difference between pupil intentions and observable behaviour. This arose during discussion of the disruptive category on the



Pupil Record. As one teacher put it *"if it's in their mind to be disruptive then they are disruptive"*. At this point I emphasised that the category system was based on observable behaviour and not on the observer inferring pupil intentions. The elaboration of this point was in terms of such category systems providing a more balanced view of pupil behaviour by focusing on the routine events in the classroom rather than only seeing significant events. It was important for me to emphasise this since one aspect of my espoused theory was to help teachers observe more effectively, since I believed that this would result in them then being able to make more valid assessments of the pupils' classroom performance. The following extract illustrates this position.

*"I understand what you are saying to me - classroom life is more complicated than this - this is not going to give me information about particular pupils that I think I need but what it might do is focus your attention on patterns of behaviour. For example, I've known people who are surprised at the amount of disruption exhibited by a pupil they thought of as being cooperative - they'd worked out a very good way of concealing their non-cooperation so there's that dimension. On the other hand there are other pupils that we think of as being very distracted or disruptive or aggressive, we suddenly find their behaviour isn't mirrored by the observations. There might be one instance in ten minutes, the rest of the time they may be cooperating on their task. All it's designed to do is balance the immediacy of certain events which kind of are lifted into high relief because they're very powerful events. It's saying 'yes these events do occur but what's happening the rest of the time?'."*

However, as this discussion continued, it became clear that the teachers were seeing this type of observation schedule as a research tool used by other people not as a technique which might be of value to them in the classroom. The focus of the discussion shifted back to a consideration of the value of such research in general. This was particularly true for those teachers from secondary schools who did not see the value of this work to their own classroom practice. This was expressed quite clearly by one of teachers (C) who said *"I don't see how this, interesting though it may be, is going to help me in the future."*

The following justification of the course content illustrates my concerns and espoused theories regarding the value of the activities I had planned.

*"How I expect it might help you ... two things ... I don't agree that a secondary school teacher does not have time to make these records because I think, not in the way it's used here, but by devising a checklist of behaviours that you want to observe. Pupils are not sitting listening to you all the time. You monitor what pupils do when you're doing practical work, when they are writing, answering questions, doing various types of activity. Therefore, the first reason for giving you this kind of way of looking, and that is all it is, is to highlight the problems of observing what pupils are actually doing, actually categorising the types of*



*behaviour to be confident about the judgements you are making about pupils. We all observe all the time..."*

The idea of categorising observations as a means of making the task of recording pupil behaviour and performance more manageable was then discussed in response to the teachers concerns for creating the time to do this. There were a number of significant contributions at this stage regarding the feasibility of using such methods which continued the debate as to their value for practising teachers. These are now detailed in full together with my responses since I consider this was one of the more significant events throughout the course.

*O "How can you observe when..."*

*AM "... remedial kids jumping around with the possibility of .. just can not do it."*

*? "Only time I could do it was when they cooperated."*

*J "It depends on how you see your role as a teacher if you see yourself there as an observer or getting stuck in and if they don't understand helping the individual children especially low ability kids. I feel personally that this type of tick list ... example of social sciences ... individuals are all individual. I don't think you can classify them in little boxes, much better being in with kids for a year getting to know them ... things that happen at playtime, things that they tell you, the work they've done, things about home. I think being involved with them you get much more out than sitting down ... the kids there ... and you ticking little boxes."*

*AJ "O.K. I'm not suggesting that you should operate as an observer in the classroom first of all. I agree with a lot of what you are saying from a personal standpoint but I don't see that that should stop me from providing you with alternative ways of looking and that is what I'm trying to do. What you are saying is 'I get to know my kids, how do you get to know them?' ... [inaudible response] ... so what you are saying to me is that your observations are going to be more reliable than the information I'm going to collect in this way."*

*R "I don't agree with that from experience .."*

*G "... different teachers have different things get on their nerves so there is no objective way of saying what the [pupil] was actually doing with all the emotion cut away."*

*AJ "Do you think it is important that one should try to work towards some ... closer approximation to what that child is really about?"*

*C "What this is to me [is] more help monitoring my own performance than the kids' performance because what they are doing depends on what I'm doing if they are distracted, that is not their fault that is my fault."*

This teacher was more alert to the underlying assumptions in my espoused theory than



I had expected. He was aware that through the assessment of pupils he would become more aware of his own theories-in-use and thus would have to face the probable lack of congruence between these and his espoused theories. The following extract illustrates how I attempted to explore this perspective with the teachers.

AJ *"Is the teacher or the pupil the centre of the classroom? That seems to me to be the difference in perspective, it's saying to me 'do I respond to the pupils or do pupils respond to me?' In other words .. what I do determines what the pupils do and ... here we have a difference which is saying what pupils do determines the work that I do."*

[Inaudible responses]

AJ *"What I'm saying is one polarises when one feels threatened into one position or the other which is to say either I'm responding to the kids, it's really important to be responsive to their particular needs at particular times and treat them as individuals and it doesn't matter what I go in with, my main concern is the pupil and the other perspective is one which is saying I have this body of knowledge, sets of skills which I must communicate if it doesn't come across, it's my fault not the kids' fault. .. [inaudible] .. one tends to polarise. I sense a lot of resistance to the idea that one should look at what the pupils do and make a record of that at the time or as soon afterwards as you can ."*

C *"... creating own problems by categorising."*

? *"... differences in judgement ... that does happen suppose you did something with ... you might be able to persuade PTA."*

AJ *"That might be a reasonable way of reaching a closer approximation of what you want , what is happening? "*

? *"The assumption here is that we are all the same ... [inaudible] ... personality problem ... poles apart in judgement, observation may stop argument."*

AJ *"It may actually demonstrate that the child behaves differently in two situations and observations are accurate. It may actually demonstrate something different and you [were] both wrong in the sense that the pattern of behaviour is not as marked in either direction or it may help to persuade a person who is not perhaps seeing certain types of behaviour. We all have our own specific way of looking at what's happening. We all have our own set of glasses, pigeonholes and categories."*

*"What I'm trying to do .. I'm not sure how successfully at the moment, we'll wait and see, is to suggest that we can talk about those categories and argue their relative merits but we can't do that whilst they're in here. [I pointed to my head.] We can only do that if we have some way of bringing them out. What is it I'm looking for, what is a good pupil, what criteria am I using to assess pupils, what do I base my judgements on and that's the focus. Now this system, I've no great desire for you to go into [it] to use it as it stands. It's not a series of hoops for you to jump through. If you would find it helpful in certain situations such as the one L. described and you have the opportunity to do some observation of pupils then use it. If it doesn't work for you and you don't want to use it justify why not. That's the position ... no conversion,*



*[I'm] not trying to sell one system. I happen to have used it for a number of years and there are other systems which I have which you can also have a look at."*

This discussion proceeded during the changeover of the tape and the following transcript illustrates the continuation of the debate concerning the value of such methods to teachers. This supported an interpretation of events which suggested that the teachers were rejecting the use of more systematic methods of looking at classrooms. It highlighted the threat to their own espoused theories and theories-in-use just as my transcription and analysis of each workshop tape resulted in self-evaluation and ultimately the restructuring of the course and the learning experiences for the teachers.

*AJ "If the data is available for somebody to inspect then they can reach their own judgements about the quality of the interpretations and that's something that you should be able to do by the end of this course. Reach your own judgements about pieces of research and as soon as you start reading around areas of interest and start being critical about what other people have written the better, because I happen to have written an article doesn't mean that it's gospel .... no good belly-aching about it, saying people can use statistics to prove anything they like. You can interpret those statistics and make sensible criticisms of the uses to which they've been put. A lot of the problem is the use that is made of them after they've been presented because what statistical evidence does is simply give some idea of how confident you can be about a particular statement that has been made. Other people take it away and use it to justify a particular value position. Now that's something quite different, that's not something the researcher is doing that is something which is being done within the political arena and the same evidence can be used to support two quite contrary value positions which the researcher has got no control over, ever. I'm not asking you to be researchers I'm asking you to observe and to look at what is actually happening and to use that information in order to reflect on your own practice ."*

Again the question of congruence between my espoused theories as described to the teachers and my own theories-in-use was raised. I appeared on reflection to be stating my own position regarding choice and personal judgement but allowing very little opportunity within the discussion for the teachers to make significant contributions to the debate. This discussion continued with a reference by me to the group searching for some understanding of how effectiveness in teaching could be measured. At this point it became apparent why the teachers' responses to the task of observation and assessment of pupils was largely negative.

The following transcript clearly illustrates the feelings of the teachers regarding the political and economic climate of the time where funds were restricted and there was little valuation of the worth of teachers by the public or politicians. In retrospect these teachers



may well have been reflecting the build up of frustration and anger concerning the lack of resources, financial reward and loss of professional status which culminated in the protracted teachers' strike during 1985 in the United Kingdom.

*AJ "They want to get an answer, because ultimately an effective teacher depends on what values you've got."*

*AM "But they've already done that. In .... in the reorganisation they decided on who was effective and who wasn't. Some got promoted but some went down an awful long way ... [inaudible] ... that's exactly it they didn't know what was effective."*

*[General laughter]*

*B "[You] see their role as being a twenty four hour social worker, twenty four hours a day, seven days a week."*

*AJ "No that would be unrealistic, I know what you're saying I mean I'm parodying but the nature of the professional role of a teacher is changing."*

*B "Undoubtedly, but what I'm saying is .. whoever reads what we say and what we think about or makes effective use of the kinds of observations we have .. [inaudible] .. teachers come out with ideas .. bolshei radical should be booted out."*

One of the teachers went on to describe her experiences with one local education authority which had recently undergone reorganisation from grammar and secondary modern schools to the comprehensive system. The following transcript illustrates how much apprehension had been produced for this teacher by insensitive handling during the change over. Examples of problems experienced by other members of the group followed so that their feelings of frustration were openly expressed and this appeared to be a necessary catharsis for the teachers concerned.

*AM " I wasn't even interviewed. I was obviously so dynamic he remembered my interview, gave me a job and I'd never even been in the school before and that shows you how ineffective they are ... shows how ineffective I am, if I'm so ineffective they didn't even notice if I had an interview or not. What good is it me saying what teachers are effective. I know if I'm effective or not, I think this is useful for seeing how effective I am but beyond me and my classroom..."*

*B "Are you measuring the pupils? Are you measuring the fact they all get eight 'O' levels that's great that makes somebody who gets the pupils through ... I teach within the school system and that to me is the important crucial point we're talking about modifying the system.. working with this within our school [reference to the professional project]. We're going back to this in a system that doesn't really work and we can all see, we all know when we think about it but who the hell.... I try within my own little small way through professional associations then I get branded as a trouble maker and a stirrer."*



*'Oh! he's off on his union again!' I really resent that because what I'm doing is to benefit my colleagues and ultimately the pupils. People are defending the status quo. What we're in great danger of doing this morning is defending a system .. they're part of the system anyway it's in their interest to defend their ... because it's the system that put them there in the first place so if you change the system .. [inaudible] .. so what, no good comparing my school to your school, because there's a totally different set of teachers and pupils, syllabi and exams ... so totally different that it's almost impossible to compare two places.'*

? *"... reading the results"*

B *"that's the interpretation people, so called experts, who interpret educational research statistics, with the local rag that I happen to read occasionally, they just can't get anything right. I know because when I was a trade union member I used to be misquoted. Totally opposite to what I actually said but they'd interpreted it wrongly ... so until we get down off our ... and look at ways of interpreting things together ... [laughter] ... O.K. teacher's pet ... we've got a system here of observing pupils or teachers that we all can research to some degree."*

We then discussed what constituted a successful teacher in terms of pupil learning. The group recognised that they gave different priorities to different aims dependent on a number of factors. It was at this point that I introduced the notion of mismatch between aims and intentions and classroom practice. However, the teachers resumed the previous discussion focusing on external accountability. They addressed the issues related to the assessment of pupil achievement across the curriculum not just success in public examinations. Finally the discussion drew to a close on a rather pessimistic note as illustrated by the following transcript.

AM *"There's no future in teaching. The majority of teachers know that they are not upwardly mobile, they know they are going to spend the next twenty years doing the same job. The majority have got to learn to live with that. They are going to be in that job for a long time."*

B *"People do what they see as being their job, keep the head down, don't rock the boat. I'll be alright."*

AM *"We're completely governed by the system. People feel they're paying off their mortgage, I don't agree with that. I'm really there because I really basically like teaching but the trouble is the system squashes you so much you end up being cynical."*

The most significant feature of this part of the discussion was that both teachers (AM and B) were prepared to embark on self-evaluation in the hope of improving their classroom practice even though they expressed anger and concern about the way external accountability had influenced them in the past. It was at this point that the teachers began



to understand my intentions and espoused theories and reflect these in their contributions to the discussion that followed. This incident also helped me to see that although I often dominated the discussion with statements of my own position the climate in the group was such that such expressions of personal feeling, conflict, anger and finally some resolution were also possible. Given my concerns for creating an atmosphere of mutual trust and respect this discussion was very important for the development of the group as a whole. My concluding remarks indicated that there was a need for me to go away and consider ways of changing the course to meet both the affective and cognitive needs that had been expressed by the teachers.

*AJ "I think one of the things we've managed to do this morning is to air a lot of the hidden agenda that was around and coming up in previous discussions. That's great because it gives me a lot more insight into where each of you individually is coming from and your concerns. For that reason I suspect I might well modify the way in which I'm going to approach the next few weeks so can you leave that with me and not hold me to my outline just yet."*

These concerns were further reflected in my diary entries during the following week. This was written immediately after the session before transcribing the audiotape.

*"I seemed to be fighting a hidden agenda initially which finally came out. Issues I perceived were first accountability and the publication of exam results. Do 'they' know what they are looking for. Second, we know our pupils too well, therefore we don't need any system for organising our thinking. Third, the teacher's professional role is changing, too many demands on time, how do you cope? Fourth, resources are not available to do the job."*

These concerns were equally applicable to my own professional development where severe time constraints were operating as the college was undergoing considerable change as a result of government policy and the full-time in-service B.Ed. (Hons.) programme was being evaluated for revalidation by the Council for National Academic Awards. As a result of the discussion during this session and my reflections on events, the course were reorganised to take account of the concerns expressed by the teachers and reduce my own concerns regarding the time constraint and conflict between my roles as lecturer and researcher.

My intentions in modifying the course were to raise a number of questions centred on the issue of effective teaching such as "how do you know you are effective" and "what measures are there of effectiveness"? I also intended to direct the following discussions to look at the purposes of assessment and provide more information on the range of methods available.



#### *Session 4: 18th October 1984*

This session began with an introduction which outlined for the teachers the decisions I had made during the intervening week and the changes to the course for the rest of the term. This introduction is given in full since it describes the nature of the changes I proposed, my feelings about the previous sessions and the reasons I offered to the teachers for these changes.

*AJ "As a result of the discussion we had last week I've reviewed the programme that I had intended we should follow. I think what you're addressing are certain important issues relating to, firstly, your role in assessment as a teacher, what part you should play in that and secondly, what I became aware of as I was going over what we were talking about last session, was a comparable problem to the one that was facing me in talking about and working with you in doing my research. So I was reflecting on the fact that I was experiencing much the same kind of conflict as I saw you experiencing in terms of assessing pupils and your role as a teacher.*

*Now my justification for researching in this area is that I believe it makes a difference to the way that I teach or lecture. In other words the way I programme what I'm doing, my expectations and so on are affected by reflecting on what happens in these sessions. Could I just suggest to you that perhaps you might be aware of that kind of conflict in me and be aware of that conflict within yourselves ... it may be that you don't actually agree that formalising and making explicit, in the kind of way I was suggesting you do, in the assessment of pupils would actually make a difference. You may actually believe it will actually confirm what you already know. I suppose I'm asking you to suspend your judgement of that until you've actually done it, to give it a try.*

*In the sense that I'm finding I'm having to review and reappraise what I'm doing in the light of things which come up during the sessions I'm suggesting that you might like to do that too and suspend your disbelief and have a try in the light of that reflection on my part. I do understand what it feels like, the conflict, should I really be doing this or should I discontinue. I believe I ought to be doing it. I've reworked the plan of the term, if you'd like to have a look at these and see if there are any comments you would like to make."*

This statement was a turning point for me and supported the suggestion that the identification of the learners' needs through the close observation and reflection on events could provide a more appropriate way of confronting any discrepancies between my own espoused theories and theories-in-use and that this might also apply for the teachers' professional learning.

The major change in programming related to the organisation of groups and workshop



sessions to consider a number of issues related to methods of assessment linked to a consideration of purposes, including accountability. However, before we could proceed with the new programme, it was necessary to complete the work already started two weeks previously on the Pupil Record. This took up the majority of the session. The teachers were working in groups and discussing the categories on the worksheet. During this time I circulated among the groups helping with the coding of examples where appropriate. When the teachers had completed the worksheets the examples were reviewed and instructions given for operating this in the placement schools. This review reiterated that there was more than one way of collecting information and that the teachers were to use their judgement in observing, assessing, selecting and recording information on pupils for the case study presentations at the end of term. Some time was also spent clarifying the nature of the course assessment through the school and course files (see Appendix 15).

The next activity was related to the nature of assessment. The teachers were directed to consider a series of questions. These were elaborated as follows within the session.

AJ *"What do you understand by assessment? What are its characteristics? What makes an assessment an assessment? What are the purposes you would be required to assess for in school? There are quite a number you are assessing for. How might you assess for these different purposes? This workshop will draw on your knowledge of school assessment procedures but it should not be framed by your knowledge of particular assessment procedures. How do you want to organise yourselves? Can I suggest you work on the first part of the workshop this session and we will have feedback next session."*

This session continued with the construction of line-labelled concept maps to illustrate their current understandings of assessment and discuss the questions on the worksheet provided. The teachers worked in groups and generated a number of key words they associated with assessment. The following list gives all the key words which were used: profiling, assessment, measurement, judgement, quantification, formative, summative, appraisal, observation, monitoring, examination, continuous, final, testing, formal, informal, norm-referenced, criterion-referenced, ratings, diagnosis, checklists, interpretation, records, screening, reports, coursework, evaluation, internal, external, process, product, quantity and quality.

The instructions given to the teachers required that they should group and organise these key words in such a way as to clarify their meaning in relation to assessment and link them with labelled lines to show the nature of the relationship, if any, between them.



In the next session I intended to review with the teachers the use of the Pupil Record in the placement schools, consider the line-labelled concept maps and obtain feedback on the discussions within the teacher groups.

*Session 5: 25th October 1984*

During the first thirty minutes of this session the use of the Pupil Record in the placement schools was reviewed and the line-labelled concept maps were collected. The teachers also asked for clarification regarding the concept maps. The teachers discussed the problems they found in completing the concept maps.

*AJ "Can I just ask if there were any problems you encountered with those?"*

*AM "Understanding some of the words, if you didn't understand them you could leave some of them off, O.K.?"*

*D "I found it very complicated, I did a simple one first and I did a more complicated one and in the end I went back to the very simple one for me. I got lost with all the words and how you could in fact link them up. [I] abandoned a lot of them."*

*L "There's a similarity with a lot of words, I don't know if you work all the time with the terminology you know, its specific meaning for, just off the top of your head seem very similar."*

*J "The main trouble is that most of the words should be used and yet having room to actually put them in and relate them ..."*

This transcript indicated that the teachers, at this stage of the course had not clearly worked out their understanding or conceptualisation of the term assessment. This was expected since many of the teachers had expressed unfamiliarity with the field in their self reports. However, there were some indications here of the ways in which the teachers tried to make sense of the large number of terms which had been generated. One response involved simplifying the task by reducing the number of words included in the concept map. Another response was to group similar ideas to clarify understandings of terms.

The teachers then broke up into groups and continued with the discussion of the questions presented during the previous workshop which were designed to help the teachers link and make sense of the key works they had identified. For some reason the teachers were reluctant to have these discussions taped so in order to facilitate the establishment of the working groups I did not press them to have the content recorded. I then began interviews with each individual teacher during the workshop and tutorial



sessions.

At this stage of the course I was collecting and analysing three different types of data, which included the concept maps produced by the teachers, information from the teacher interviews as well as the evidence of the group processes and content covered within the college-based part of the course. An initial appraisal of the line-labelled concept maps suggested that the teachers found some difficulty in conceptualising assessment at this stage of the course, although particular organising concepts were emerging. For example, the idea of formal and informal methods of assessment, evaluation based on a range of assessment methods as an end point and different purposes of assessment were identified in some of the line-labelled concept maps. The evidence from the teacher interviews suggested that they were also experiencing some difficulty in coming to grips with the school placement requirement to apply their skills and knowledge in the practical context of the classroom.

These developing understandings of the teachers provided valuable feedback as to the particular misconceptions and possible directions for course discussions and workshops. The next stage for the teachers was to develop their understanding of assessment. I proposed, therefore, to consider different conceptualisations and compare these with my own content structure diagram during the following session. In addition, more emphasis on the application of the college-based workshops to the school placement was necessary.

#### *Session 6: 1st November 1984*

During this session the construction of the line-labelled concept maps was reviewed. I presented my content structure diagram to the group as an example of how I saw the content of the course and explained how I had constructed this diagram. I then asked for questions or comments on the content structure diagram. None were forthcoming, but the teachers did write down my conceptualisation of the course for reference.

As the teachers were completing this I explained the relationship between observation, interpretation and evaluation with reference to different assessment methods. This was directly related to the issues of the validity and reliability of assessment procedures which I intended to discuss during this session in order to help the teachers in the application of different assessment methods to the school placements. The following definition was given "*the validity of a test can only be judged in relation to its purpose.*" My elaboration of the concept of validity continued with a consideration of construct validity. The



comments that ensued indicated that the teachers were surprised by my statement that different tests which were designed to measure the same aspect of pupil performance or achievement could yield different results. This led to a discussion of error in educational measurement and assessment.

This was not my intention at the outset of the session but clearly the need was being expressed by the teachers for clarification of the terms used in the discussion of formal assessment procedures, particularly in interpreting standardised test scores. The teachers pursued this discussion with reference to why we assess and test pupils which were considerations directly related to the questions that had been raised in the previous workshop session. In addition, some of the teachers had raised the issue of norm-referenced tests in the interviews. This was clearly an issue for the teachers and I interpreted this as a response to the school file requirement to use, review and evaluate different methods of assessment as well as a concern about assessment in general.

Different purposes of assessments were identified in the ensuing discussion. Distinctions were made between norm-referenced and criterion-referenced assessments and the following extract illustrates the concerns that these teachers expressed about the self-fulfilling prophecy and the impact of their expectations on the pupils. My interpretation of this was that the teachers were beginning to articulate the relationship between their own teaching and the pupils' learning and questioning the match between their expectations and the abilities of the pupils they worked with as a result of their recent experiences on school placement.

*AJ "In order to make a judgement about quality you have to compare it with something ... in order to make sense of it, in order to know where to go. We all have standards implicitly. We all have expectations of what children should be able to do at particular ages, we may not agree about that, but we do have that underlying notion about what children should be able to do at particular stages."*

*D "When we draw up schemes of work we're thinking of what we hope them to be capable of, aren't we? Our areas may be different but our expectations come through, what we expect the children to be able to do in that year."*

*AJ "Whatever you plan to teach has embodied in it expectations of what the pupils should be able to cope with ..."*

*AM "...which is compared with what you've taught before ..."*

*D "We say this year's better than that year, that's a marvellous year they've reached that standard now ..."*



*J     "I don't think that's always a good thing. If you set standards children only work to those standards and depending on your teaching experiences your standards might be too low so you might be pulling the children down by doing this sort of thing."*

In the following extract various purposes of assessment, which were previously identified, were linked with the practical classroom situation where the teachers were confronting the problems of collecting, recording and using assessment data in their case studies in preparation for planning for individual needs. This aspect of my espoused theory is illustrated by the following statements.

*AJ    "... What are appropriate levels of expectation? What are appropriate goals? My feeling is you can't actually work out what an appropriate goal is until you've got some picture of what the child is doing at the moment. Now you may be wrong."*

*AJ    "The issue you are raising is an important one, which is, why are we assessing in the first place? So this notion of why we're actually doing it , is very important in understanding and making judgements about the quality of the assessment procedure you're using."*

The session continued with a consideration of different methods of assessment and their appropriateness for different purposes. In particular, the developments with graded tests and the problems associated with them were considered. This led to a discussion of the implicit value system underlying selection within the education system. This discussion was very wide ranging and drew out a number of different perspectives regarding the relative merits of the selective system and comprehensive schooling. This was an issue which was close to this group of teachers as was suggested in previous discussions regarding accountability.

My interpretation of this concern was in the context of the effects of secondary reorganisation on one third of the teachers in the group. As noted previously some of the teachers were on the course because they were given secondment to upgrade their qualifications as a result of the redeployment of staff by particular local education authorities. Although this part of the workshop was not planned, I believe the teachers found it valuable to discuss the issue of selection. This also enabled me to introduce the notion of predictive validity to the group with reference to a problem arising from the teachers' concerns.

I then reviewed the two different forms of validity: construct and predictive which had arisen in these discussions. The concept of content validity was then introduced and we



considered the validity of the 'O' level examinations. The teachers expressed concern about the role of memory in such examinations and developed their ideas about continuous assessments and the practical difficulties associated with marking and operating Mode 3 syllabi and examinations. In the ensuing discussions the teachers identified the changing nature of their professional role in terms of the lack of time to reflect, develop and evaluate the curriculum as they would wish. These concerns were also important to my self-evaluation since I was still finding such factors a constraint within the evaluation.

- C *"... they haven't got time to change. If you've got to change all your system, you need so much more time to assess what you're doing and put it into practice ... most teachers haven't got that time."*
- ? *"It is possible to do. I've done it, my school is changing at a phenomenal rate and the examining system is changing at an enormous rate ... continuous assessment has been a major part of all the exams we've done in school for the past eight /ten years. It's becoming increasingly important in other subjects, isn't it? Other subjects are being asked to be taught in a more practical way."*
- D *"Over the last ten years we get less and less time allowed in schools."*
- AJ *" and more and more is expected of you, of the teaching profession, which is why you're here, one of the reasons you're here: to have time out, to reappraise, to look again, to have time to think about what you're doing and why you're doing it. There is already evidence that you are beginning to think about what you're doing and why you're doing it."*
- H *"That's definitely why I came."*
- D *"You're not given it otherwise. You can't redirect your career or anything. You're never given time to sort it out, to think about it ..."*

At this stage the teachers appeared to be coming to terms with the opportunities that the course offered as well as reflecting on the often unrealistic demands which were being made on them in school (and perhaps in the course as well). I then reviewed the session in the following way.

- AJ *"There's some important issues that have come up so far. I showed you my conceptualisation, my structure diagram which tried to relate the various bits and pieces together, that's the way I see it. We talked a bit about the problems of making judgements of quality, that's really where most of our discussion has centred today. It's about making judgements of quality, what criteria you use to judge quality. I think a lot of the discussion we've been having centres on the fact that some of us value different things or rather than valuing different things, we have different priorities. We have a different set of priorities in our aims, in what we think we should be doing. So that when somebody says 'this is a bad teacher' it may actually be a statement which says that 'I don't like the values of that person'. If someone says 'that's an ineffective teacher' it may be*



*they don't like what they're doing but they may be quite good at doing it. I think that's something that's quite important when we think about evaluating anything is making clear the criteria we're using to make the judgement because then people can argue with you about it. Up to that point it's not a debate, not a communication ... if the criteria are explicit you can argue about it. Let's have a look at the issues that arose out of the second workshop: choosing the mode of assessment and devising assessment procedures."*

The teachers had not had an opportunity to discuss this part of the assessment workshop materials previously so I explored this area with them and gained feedback from some of teachers who had used the Pupil Record on school placement the previous day. They were surprised by how easy it was to observe and record using this type of systematic observation.

The remainder of the session involved a consideration of the Richmond tests of basic skills (France and Fraser, 1975). I asked the teachers to go through the items on the tests and label each item with the skill they thought it tested using the skills classification provided with the test. I suggested that they might work in pairs and discuss the match between the listed skills and items. No interviews were conducted as I circulated among the teachers discussing the workshop activities.

#### ***Session 7: 6th November 1984***

The start of this session as planned entailed a review and comparison of the skills classifications of the Richmond tests as described in the manuals and those arrived at by the teachers in their discussions the previous week. This was designed to explore the criterion-related and content validity of the test. The teachers were surprised by the disagreements between their understandings of the criteria and the way in which the test constructors had described each particular item. I identified the problem in the following way.

*AJ "I think it is a very real difficulty in actually trying to make explicit the nature of the learning you are assessing ... it applies whether it is a standardised test of this kind or if it is an assessment procedure you've developed for use within you own teaching, diagnostic procedures actually describing various criteria which you are going to use to judge pupils' performances."*

*L "It's quite useful to have them because you can reject them, if you've got something to work on in the first place."*

*AJ " ... It does raise some important issues about how much value such a procedure has for diagnosis. If you don't actually think these items are testing what they say they're testing then how can you plan a programme on the basis of the*



*skills breakdown derived from the Richmond tests."*

The lecture continued with a consideration of the reliability issues associated with standardised tests of various kinds using an item analysis of the Richmond tests as an example. The general discussion that followed reviewed the issues of reliability and validity of different forms of assessment. I then gave the teachers copies of an article (Jasman, 1981) which summarised the issues which we had covered in the previous workshops. The discussion then moved to a consideration of pupil profiling as I intended. This activity was designed to focus the teachers' attention on the wider applications and purposes of assessment, particularly at the school level. In addition, I hoped that it would provide some useful insights which would help the teachers select and organise the information they were collecting on pupils for the case studies since this need had been expressed during the teacher interviews. These intentions are illustrated in the following extract of the tape transcript.

AJ *"How do we describe pupils overall in order to get some picture of their various abilities. We've all talked about what criteria we use to make a judgement of the pupil. So in order to open up the exercise away from the classroom level, to think about the overall achievement and experience of pupils, I'm going to suggest that you work in groups of your own choosing, to draw up profile categories appropriate to either pupils leaving school or one of the other transfer points prior to leaving school. So if you are a secondary school teacher you might want to work with a primary orientated teacher to determine what kinds of information would be appropriate to transfer. One or two of you might be interested at looking at transfer into the system. I want you to be aware of the implications of the four issues I've already identified: the purpose, validity, reliability and practicality of that particular profile. The idea is to try and support the school files where you have been asked to produce a profile of a particular pupil, to alert you to the kinds of things you might be looking for. Can you keep in mind what instructions or arrangements might be necessary to help teachers complete such a profile, given that you know what some of the problems are ... a pupil profile for a particular purpose which you will define.*

The teacher groups then worked on producing their profiles during this and the following session whilst I continued to conduct individual interviews and tutorials with the teachers during the workshop time. This arrangement went some way towards alleviating my concerns about encroaching on the teaching time allocated in conducting the interviews and providing for individual needs. This organisational strategy of grouping the teachers for discussion and workshop activities was one which I found particularly valuable in helping to maximise the available time for the research and reduce the conflict I experienced between my lecturing and researching roles.



### *Session 8: 13th November 1984*

This session was devoted to the workshop activities on profiling begun in the previous session and conducting interviews with some of the teachers. The teacher discussions during the workshop were not recorded on audiotape. However, the substantive content of their discussions was reflected in the review carried out in the next session.

In the diary entry made after this session I noted that there appeared to be some difference in the way the secondary and primary teachers viewed assessment. The secondary teachers were more confident in talking about formal methods of assessment such as public examinations and testing for grouping, whilst the primary teachers were more familiar with informal methods of assessment such as checklists and the use of anecdotal comments on record cards. Interestingly those teachers who worked in middle schools seemed to have more knowledge of assessment than either group. This observation has also been reflected in the teacher self reports and interviews.

### *Session 9: 20th November 1984*

I started the review of this session with a reminder to the teachers that the pupil profiles which they were developing should provide them with some help towards the work they were doing on school placement in preparing for the case study presentations and suggested that we could review their progress at the end of the session. The content of this session was to examine the relationship between teaching styles, pupil behaviour and pupil performance. I introduced this session in the following way.

*AJ "My intention is to give you some information relating to each of those areas and to look at how they interrelate so that you get a better feel for the whole interaction of the teaching process and what strategies may or may not work for different areas of the curriculum."*

I then presented my conceptualisation of the teaching-learning process. I proceeded to go through each of the phases represented with reference to the activities the teachers had been engaged in during the course. I saw this as a way of reviewing the college-based activities and the relationship of these to the school placement. The teachers spent some considerable time making a record of this content structure diagram. They indicated that it was useful to them in making sense of the work they had completed that term. I then talked with individual teachers clarifying points which had arisen during the interviews



and tutorials. During these discussions it was clear that the teachers were experiencing considerable anxiety as the submission dates for the course and school files approached, as well as the presentations of their pupil case studies. These were the first pieces of graded and formally assessed work which they had to successfully complete to progress to the next stage of the course. The session continued with a discussion of the problems and possible strategies for dealing with these. The following extract illustrates the way in which the teachers were coming to terms with the complexity of the relationships in integrating theory and practice.

*AJ "You've got a lot of information and you don't know what to do with it. How to organise it? How to make sense of it? Perhaps I could address you to some questions: why have you collected that information?"*

*D "... because you told us to."*

*AJ "Have you any other purpose for collecting it? "*

*D "... to find out if they can do it."*

*AJ " Can you express that in appropriate language which will help you to fit it into what we've been talking about this morning."*

*D "What the pupils actually gained from all the teaching-learning processes."*

*AJ "To assess what they've gained from their previous experience. Why are you doing it though?"*

*D "So you know what to go on to?"*

*AJ "Right! Now that's why you're collecting the information - to make a diagnosis of their present position, what they've learnt previously, where they're at now so that you can make decisions about where they go to next."*

*"It enables you to explore the range of techniques that are available by focusing on one pupil in particular but also to gain a general assessment of the remaining pupils you are working with in the group. Clearly, each of you have focused on different areas of the curriculum, because you are working with different ages who are at different stages of development, where there are different priorities of what you think they ought to know or you feel you need to know about in order to plan their learning. It would be inappropriate for me to give a list of headings for organising the information that you have collected because that should arise from the work that you've been doing in relation to J's and H's work. You could have global categories like pupil behaviour, which might be fed by the observations you've been doing with the pupils either the diary, using the Pupil Record or some other type of observation you've made whilst you've been doing something with them."*

*"So that's one area you need to think about what information you have relating to that pupil's behaviour in the context of the classroom. You may want to focus on dimensions of their language, another global area ... you can begin to make some sense of the information you have relating to their oral ability, their*



*reading and written skills. We've also asked you to have a look at number, at numeracy. Another area that might be appropriate to think about is the kind of social relationships that develop. I'm sure you can think of other areas you've looked at in terms of what pupils do and how much they've learnt. I think it's up to you to devise the categories, the organising concepts."*

The session proceeded with further clarification of the nature of the school and course files. The teachers also talked about the feelings they associated with being assessed by different tutors and their uncertainty about our expectations. They did not, however, press me as they had done in an earlier session to provide detailed guidelines for completing the assessed work. I believe it was important to leave the method of presentation open so that the teachers could work through the assignments in a way that was appropriate for them as individuals. In order to clarify this further I outlined what I considered characterised a "good school file".

*AJ "What categorises a 'good school file' is your ability to identify a central thread for yourselves. It's one where you have begun to make an analysis of the information, begun to synthesise it in a way that makes sense to you and that you have the ability to evaluate it and make a realistic judgement and appraisal of what you've done. It's not how much data you've collected, which would be worthless unless you used it in a worthwhile way to inform what you do ... it's where you try to make sense of your own experience, your own way of working, your own decisions in terms of making assessments and judgements about what pupils can do and helps you to make decisions about what to teach next."*

In reflecting on this statement my own understanding of the self-evaluation process and the difficulties of representing this to others were highlighted. This was an unexpected insight which has greatly facilitated the writing process for me and enhanced my own professional development. The parallels between the experiences of the teachers on the course and my own experience in conducting this evaluation were again reinforced by the process of self-evaluation. In the remaining thirty minutes of the session I introduced the Oracle research programme to the teachers as an example of a study that investigated the complex interrelationships between teaching style, pupil type and pupil performance. The discussion of this research was continued in the following session and the teachers presented their case studies in the last two sessions of the course. I did not, as I had planned, review the pupil profile data that the teachers had prepared during this session because of the time constraint.

#### *Session 10: 27th November 1984*

At the beginning of the session I gave out examples of the teacher-based assessment



materials to the teachers as a number of them had discussed the use of these in previous interviews and during tutorials. The distribution of the teacher-based assessments caused some concern to those teachers who had not had the opportunity to use the materials on their school placements. I interpreted this anxiety as a concern that they might be penalised for "failing" to take this opportunity even if the material was inappropriate to their needs. In the following discussions I reassured them and reviewed again the criteria that were used to assess the school file (see Appendix 15). A number of concerns emerged which again related to the way in which the school files would be assessed. One teacher raised the issue of not having any data, stating that she did not have test results. This developed into a consideration of the value of different types of data in the files and the organisation of this material. As I noted in discussion

*AJ "It's about being critical of the way in which you collect and interpret and reach conclusions about an individual pupil's performance but it's within the context of the classroom, the professional situation."*

The implication of this discussion was that the teachers were still coming to terms with the changes which we were expecting of them within the course. I then returned to consider teaching styles and their relationship with different pupil types and pupil performance as found within the Oracle research programme. I spent some time explaining and clarifying the types of analysis used within the study and in particular the notion of statistical significance. This obviously concerned the teachers and this occupied much of the remainder of the session which became a lecture on statistical procedures and the meaning of terms such as standard error, mean, standard deviation and norms.

The judgements made by the teachers of the Oracle study indicated that they still had a jaundiced view of educational research, as one teacher asked "*what did this actually achieve?*" Judging from the discussions during this session this may have been because the content did not relate directly to their immediate concern of presenting case studies and the submission of their school and course files. However, I justified the inclusion of this emphasis at the end of the course as a means integrating different perspectives of classroom events. In this way I intended to give some understanding of the complexity of the interactions between teaching style, pupil type, progress and performance which would provide a basis for developing their thinking during Part 2 of the programme.

#### *Sessions 11 and 12: 4th and 11th December 1984*

During this session the teachers presented their case studies of individual pupils.



## **APPENDIX 20: Summary of assessed work and teacher self reports.**

### ***Teacher AB***

This teacher was absent for the oral presentation of the pupil case study. The school file was descriptive and information presented sketchy. However, the following extracts illustrate how this teacher saw his problems on the school placement.

*"I have very little experience of classroom teaching and certainly no experience of testing, formal or otherwise."*

*"I felt that given longer I could have made far more progress with my group. My lack of experience with testing and my initial lack of understanding of what was expected of me made me feel rather diffident with regard to the task in hand. At times I thought that I knew what I was trying to achieve but lacked the confidence to do it. It was not until the latter part of the term that I really felt I was getting things together. My main concern is that due to my late start in getting organised the depth of my studies does not reflect my understanding."*

This outlined his concern with lack of experience and confidence in his abilities which were previously noted in the interview. The particular needs of this teacher stemmed from his lack of classroom experience within the state system. He had a lot more learning to do at the basic level of competent teaching than most of the other teachers and did not perform particularly well at this stage of the course.

This teacher made no reference to the course work on assessment in the course file. He did not complete an overview in the file and clearly had experienced difficulties in coping with the level of the work expected within the course. This was not a surprise given the comments he had expressed during the interview and in the school file.

### ***Teacher AL***

This teacher presented a very detailed case study of a boy in her class. This showed how she had come to understand different aspects of his behaviour and abilities over time. However, there was no resolution in practical terms of how best to structure the learning for the future. A lot of information was presented but each new item of data sent the teacher on a different tack. She recognised this herself as indicated in the following extract from the school file.

*"I was also nearing the end of my placement experience when I became dissatisfied with the rate at which I was gleaning information. My thinking needs to be far*



*more specific. Many of the things which I have spotted in the children are still waiting to be tackled."*

This was one of the few teachers to use the teacher-based assessment materials systematically and she found them very useful for the boys she was working with in a slow learner unit. Her comments were positive and showed that they had helped her to structure pupil learning, effectively blending number and language through small group work using the everyday mathematics materials.

In her evaluation of the school placement this teacher noted how previously her work in secondary schools was determined by the needs of the examination board and how this experience of teaching in a slow learner unit had worked for her. As she commented

*"I have not enjoyed teaching so much for years! "*

However, she also recognised the changes she was undergoing as a result of her experiences.

*"I did not learn the value of diagnostic teaching until I was into the second half of the term. As a result of an invigorating term's work and thought in this area, my school placement work enabled me to begin to feel capable, competent - a professional. However, there is still much to do: I still need to work to find fast, accurate methods of record keeping and storage. Part of this lies in the materials which I would choose for a classroom but a major part needs to be my own regime. However, I now feel that I am on my way to asking the right questions."*

In the course file this teacher wrote as follows about her perceptions of the course and what she had learnt in the process, which provides an effective summary of the course intentions.

*"This course seems designed to make the teacher a confident educational practitioner. The teacher is seen as a competent detector of the intellectual, social, emotional and motor skills of pupils and as an efficient treater of these needs. In order to detect and treat, the right questions need to be asked. A learning environment needs to be created which will point up the children's needs and a curriculum needs to be organised in such a way that the child can be treated as an integrated whole."*

### **Teacher AM**

This teacher was disadvantaged when compared with the rest of the group as she began her school placement some five weeks before the start of the autumn term. However, she was continuing to work in her own school and, therefore, did not have the



problems associated with adjusting to an unfamiliar situation. Her case study was detailed and focused on behavioural problems exhibited by a fourteen year old boy, since this was to her the cause of his learning difficulties. The analysis was detailed and effective in that some progress was made by the boy in his learning. The following extract which was taken from her evaluation of the school placement illustrates her lack of skill at the start of the course and the way in which she describes her own professional development.

*"The course has given me time to begin to consider the various ways I could start to improve my teaching. I have come to give value to the part that knowledge can play in the professional work of every teacher. As a result of this knowledge and values I would now set myself pedagogic aims, which would have been impossible prior to coming on this course."*

In this teacher's course file reference was made to a number of features on the school placement in an attempt to integrate theory and practice. The overview of the course to the end of term 1 considered the way in which this teacher's perceptions of the child with special learning difficulties changed through exposure to the course material. Her attitude at the beginning of the course was *"to be understanding and anxious to effect change"*. By the end of the course she expressed the following views.

*"This course puts me in the position of developing strategies and therefore dealing effectively with those who have special educational needs."*

This teacher, despite her initial antagonism to the course which resulted from her recent experience of redeployment, espoused the theories of self-evaluation within her school and course work and there was evidence that this had affected her practice.

### ***Teacher B***

Teacher B presented a concise account of the testing programme he had carried out in conducting the case study of a seven year old boy. He described language and number abilities as well as details of attitudes to adults. However, the case study did not make recommendations for future needs and was largely descriptive.

This teacher also presented an overview of both course and school files which represented an attempt to synthesise both the practical and theoretical inputs of the course. However, this was not successful since it did not reflect personal learning. It was a more academic piece of writing which did not meet some of the assessment criteria. Overall I



concluded that this teacher still had not made the transition to reflecting on his own and the pupils' experiences and learning in order to inform his decision-making in the classroom.

### ***Teacher C***

The case study presentation was long and descriptive. There was a detailed description of the pupil based on observations, the school's records and assessments but little to indicate how the teacher intended to use this information to meet the pupil's individual needs.

The school file was similar to the case study in that this teacher presented much detailed descriptive information relating to the problems of the pupils but he did not reflect the impact, if any, on his own approach to the pupils. The only comment which appeared to indicate this was a reference to the difficulty he experienced in pacing work for slow learners. At this stage of the course this teacher was still unable to come to terms with the implicit and explicit demands of the course which required that he should reflect on his own experiences as well as describe and analyse those of the pupils.

This teacher experienced considerable difficulty adjusting to the demands of the course. He was aware of the need to evaluate his own performance and yet did not use the theoretical insights which he had been offered in the lectures and workshops in the course file. He made reference to the lectures and workshops but did not move from description to analysis, synthesis and evaluation of the material.

### ***Teacher D***

In the school file there was evidence which indicated that this teacher had a well balanced diagnostic teaching experience. She gave many examples of personal initiatives, application of ideas from the course and the development of her own ideas in relation to assessment. She also made recommendations based on her findings and interpretation of the data she had collected on individual pupils. She had clearly implemented many of the ideas that had been raised in the course in the school placement since she made reference to the literature, developed a line of argument in the evaluation and drew conclusions from the case study. This teacher produced a thorough course file covering all four strands of the course. In the overview this teacher identified her own personal professional development in the following ways:



*"It has made me aware of how varied a child's response is to a given task; how wide the range is of children's knowledge because of different school background, different academic ability, interest shown by the child's parents, help given at home, emotional problems due to our expectations or non-acceptance of educational problems, physical and mental handicaps."*

*"I have become aware of how much the remedial child needs approval and to succeed. It is up to us as teachers to help them succeed independently as far as possible in the world outside school."*

This indicated that she had made some progress towards extending her professional role through the school placement and college-based course.

### **Teacher G**

In her account of her personal professional development this teacher identified that the placement with younger children as *"a totally new experience"*. Her immediate concern was to establish her authority and credibility in the classroom and she used the time available on the school placement both to observe the teacher and the pupils. Her adjustment to an infant classroom took time but she adapted well to the integrated day and later identified some of the assumptions she held about teaching and learning. As she noted,

*"I don't think I was consciously aware of all the different activities that make up a cookery lesson before doing it with five year olds."*

*"I think previously I might have rushed in to help where they obviously learned more by doing it for themselves."*

*"I have always taught with the premise that you should assume the child knows nothing until you have found out just exactly where they are in a given area."*

*"I learnt to take one step at a time, go back and do it again and again before taking another step. I have C to thank for that with her difficulties in number work."*

*"I have learnt not to judge children on other people's opinions. It is all too easy to fit a child into a stereotype created by someone else."*

*"Above all I learnt to appreciate the difficult task facing teachers of infant classes, where every child really is a separate individual working at his own pace and at his own level."*

In her course file this teacher reflected a similar approach and developed a pupil profile together with two other teachers. She appeared not to be threatened by working with younger children and maintained an open, reflective attitude to all the learning she did on



the course.

### ***Teacher H***

This teacher experienced considerable difficulty in coming to terms with using a range of assessment procedures in constructing the case study. She expressed considerable unease at not finding suitable "tests" to assess young children and found it hard to tolerate the differences in perception between herself, the classroom teacher and the supply teacher who regularly worked in that class. However, once she had come to terms with these factors she gained in confidence in her own ability to assess pupils and diagnose learning difficulties.

In the school file this teacher identified two particular advantages which were: first, the opportunity to work with a small group of children and, second, that these children were younger than those she would normally have taught. However, these advantages were seen in terms of making the job easier in the first instance and, second, providing a new experience for her without the particular stresses involved in teaching a whole class. Personal professional development was not commented on.

The teacher attempted to synthesise her learning within the course file but did not provide an effective overview on her own professional development. The discussion was mainly in terms of material covered and showed little extension beyond what had been covered in the course workshops, seminars and lectures.

### ***Teacher J***

The school file contained detailed work outlines and assessments of six pupils. In her evaluation, this teacher compared her school placement with previous teaching experiences. Her comments related to the need for more careful observation of pupils to recognise the variations in pupil behaviour and the need to provide for individuals within the class context. However, the evaluation tended to describe what was needed for all teachers rather than reflect the impact of the school placement on her own teaching. She tended also to describe her experience in general terms with little reference to significant events which may have led to her conclusions.

This teacher attempted in the course file to *"draw connections between the different strands of the course"*. Her overview tended to describe the content rather than identify

any underlying themes or concepts. However, there was a level of understanding about the processes in the course which was reflected in the following comment.

*"The overall aim of each strand in terms of method, encourages us to look at everything we approach be it reading material, test material, educational models or research articles from an evaluative point of view, looking critically at the articles involved and critically analysing them. Attempting to make us question accepted ideas and material in terms of what we actually require."*

Although this teacher wrote about increased awareness as a result of looking at various methods of assessment she did not reflect this in the way she approached the school placement or by including critical analyses, evaluations or other evidence that she had applied these ideas in practice.

#### **Teacher K**

In his evaluation of the school placement and his own learning this teacher noted the following as "learning about myself".

*"... teaching must be tailored very carefully to individual needs ... stereotyping is always a danger. I think I was guilty of this with K."*

He had come from a small village school and initially found adjusting his style to the placement school difficult. However, he showed evidence in his evaluation of the school placement that he was able to contextualise his own concerns and consider the implications of his own experience in relation to the system as a whole. Thus he noted that

*"Surely, we cannot be satisfied with a situation in which the curriculum is not being delivered to ensure that each individual child gets an education suited to his age, ability and aptitude."*

*"If a child is not apparently learning we should examine our teaching methods critically."*

In relation to his learning about assessment he commented that his experience and thinking had been *"broadened by the placement"*. As noted in the self report this teacher's previous experience of assessment was the administration of norm-referenced standardised tests. The course had given him access to *"a range of diagnostic tests, kits and tools available"*. He also reflected that he had undergone some change in his perceptions of the value of different methods of assessment.



*"As far as observation techniques are concerned I initially received these with some scepticism: thinking that as a class teacher I was always observing individuals sufficiently. However, having done some observation on placement I can see that such very individual attention can reveal aspects of a child that 'day to day' class routine can hide."*

The course file showed that the teacher was effectively synthesising different strands in the course. His overview indicated how he had internalised the material he had encountered. He analysed the assessment strand in the following way:

<i>"Difficulty:</i>	<i>What are children's needs?</i>
<i>Understanding:</i>	<i>Appraisal, pupil behaviour, informal assessment, the reliability and validity of tests.</i>
<i>Teacher skills:</i>	<i>Pupil and teacher behaviour, style, methods of recording, pupil profiles and case studies."</i>

Overall this teacher displayed an open minded approach to the school experience and college-based course and a willingness to try out different approaches and judge them on their merits in practice.

### *Teacher L*

This teacher had used a wide range of both formal and informal assessment procedures in her case studies but the course file did not directly reflect any of the assessment strand. This teacher did, however, make a considerable effort in producing her school file to reflect many of the aspects of the assessment strand. Her evaluation of the school placement indicated that she had learnt a great deal, *"understanding the task better when I finished than when I started."* Her previous experience was as a high school physical education teacher and she expressed some concern about her lack of experience of the classroom.

She was confronted with the problem that *"time is an enemy in the classroom"* since she was trying to provide individual attention in a large middle school class. However, a number of methods helped her in the processes of coming to know the pupils. In particular, she made use of a diary despite an initial reluctance to go into the school and start *"formal testing"*. This teacher's observations of the placement illustrate that the hidden curriculum of the course was perhaps not as hidden as I might at first have expected.



*"Undoubtedly, knowing that your focus is assessment you tend to key into every opportunity to "find out". This is almost a hidden curriculum built into the placement - if it moves assess it! In fact, I think it was very valuable and sharpened up my perception of just what assessment entails."*

### **Teacher M**

The school file showed some evidence of reflection on his professional role. He commented on his choice of placement in a junior school in the following way

*"My background in education may seem to some extremely limited. I teach at the same secondary school I attended as a pupil ... I have often felt unable to base judgements on the wide sphere of differing educational establishments with which other colleagues have had experience. For this reason, I see this one year secondment as a chance not only to increase my knowledge in the specific areas of reading and special needs, but also as background experience on which I may draw in the future."*

He also echoed the problems he had mentioned in the tutorial about the ability to pitch work at an appropriate level and the difficulties he had found in adjusting to the junior school classroom. He also questioned why some children fail and what he could do to improve his own performance. In this way it could be said that he had begun to set his own agenda for professional development. He indicated also that his perceptions of the children changed throughout the placement and had developed a range of strategies both for assessment and curricular planning.

This teacher was largely descriptive in summarising the contents of his course file rather than providing an overview of his learning, although he did identify that he had

*"gained valuable practical experience in diagnostic testing when trying to classify the learning difficulties experienced by a group of five primary school children."*

He did not reflect the assessment strand in this file since he had seen this as more related to the school file.

### **Teacher N**

The case study presented by this teacher was detailed and covered a number of relevant aspects: for example, social skills, language and number. This teacher did not address the evaluation of her own personal professional development effectively in the school file. In the course file this teacher reviewed the course content reasonably well but



did not appear to have reached any conclusion or evaluation of the knowledge she had gained in the course. It was as if the lack of right answers had produced a sense of powerlessness to act or decide on one course of action over another. This teacher appeared to need the security of being sure that what she was doing was right and could not tolerate the possibility that what might be "right" today could well be inappropriate tomorrow.

### ***Teacher O***

This teacher had previously taught science to examination classes so that the adjustments she had to make in her teaching within the school placement were great. She was also relatively inexperienced as a teacher with little formal training. The insecurity she felt was reflected in the interviews and also in the case study, school and course files. She expressed considerable confusion regarding what was expected of her in terms of the course and school placement in relation to the assessed component. As she noted,

*"With Anne's piece on assessment, I could not comprehend what I was supposed to be doing. I am not sure if work on that area was supposed to be included at all."*

The only reference made to assessment in the course file referred to the content by repeating the lecture notes in a way which read as a list of rules or guidelines to follow. There was no analysis or evaluation of methods and this teacher was still working on developing a basic competency in the classroom rather than becoming an extended professional.

### ***Teacher P***

This teacher had come from the prison system and found working in a school situation less of an immediate challenge to him than he had expected. He commented that

*"I felt until well on towards the end of this term that having a group of only slightly below average pupils with no really significant problems being apparent, did not give me much to work on in comparison with others who were working with extreme cases. As the term progressed I began to see the class situation in a different context and the reasons for a particular type of behaviour began to emerge."*

In the course file this teacher attempted to synthesise the four strands of the course. This overview was limited and did not reflect much significant personal professional



learning. This was seen in terms of the adequacy of the assessed work rather than exploring the possible applications of course work to the school placement.

### ***Teacher U***

The case study presented by this teacher was based on a pupil selected by the teacher. However, the information collected was detailed and thorough if largely descriptive. In the school file this case study showed development of thinking and understanding on the part of the teacher, which resulted in changes of strategy and tactics with subsequent evaluation of these. This teacher was effectively implementing in practice the methods which were being advocated in the college-based course. This was reflected in the comments made by the teacher in his evaluation of the school placement.

*"At the beginning I based my assessment of the children on what they were producing for their own teacher, which reading scheme book they were on and which Scottish maths scheme book they were using. I found, however, that I needed some sort of test to help with the Pictogram programme. .. I had therefore to invent my own test."*

In the course file the teacher noted that it *"has helped me fit my ideas together and focused my thoughts"*. He expressed ideas in the overview which showed an awareness of the importance of aims and clarifying meaning in educational debate, without becoming dogmatic and inflexible. As he concluded,

*"It has been very interesting to see how well the strands have gone together so that new thoughts have been sparked off. The course file work has also fitted in well with the school file and they have complemented each other well. The whole exercise has had a very stimulating effect upon me and has helped me to take a fresh look at what I have been doing and rethink several of my ideas."*

In relation to assessment the teacher had clearly internalised many of the more important aspects of the work I had been doing with the group during the term. In particular he reflected his changing perceptions and understandings as the course developed by noting differences in his views as time progressed. He was operating at this level in both the school placement and in his understandings of the course.

### ***Teacher V***

This was evidence within the school file that this teacher had still to come to terms with the impact of the theoretical concepts and knowledge provided within the course in



the context of the classroom. Her overview indicated that the school placement had not been

*"as beneficial to me in relation to the course as it might have been for several reasons."*

Among these reasons were references to being used to cover absent teachers, the unfamiliarity of the situation and the difficulties associated with getting information from the class teacher about particular pupils. It would seem that this teacher, who was relatively young and inexperienced, was finding it difficult to adjust to her new professional role. She viewed the school placement as another teaching practice rather than an opportunity to explore a diagnostic teaching approach. The overview was otherwise largely descriptive and did not contain reflections on the teacher's own learning.

This teacher effectively organised the course file into the four strands of the course. Within the section on assessment this teacher reported the work done on the line-labelled concept maps, the purposes and modes of assessment and finally her case study. The presentation of this information was largely in the third person providing further indications that she had still not moved into a personal, reflective mode of approaching the work she had done.

### *Teacher W*

The written presentation of her case study indicated that she had conducted a number of tests and made some specific recommendations for the pupil although these were behaviourally orientated rather than exploring the individual pupil's needs. In the school file this teacher introduced the work she had completed in the following way.

*"My brief is to learn by doing. I am going into an infant classroom for the first time, to see how the course and my own previous experience will enable me to cope with assessing and working with 5 and 6 year olds."*

Her conclusion to the school experience indicated that to a large extent this teacher had achieved what she set out to do which was to cope with working with this age group. However, by dealing with such an unfamiliar age group she also had to come to terms with new teaching materials and methods, as did other teachers, which resulted in less reflection and analysis of her own teaching and evaluation of the pupils. The course file

produced by this teacher showed some evidence of synthesis and evaluation of the work she had completed. As she noted

*"During the term my own ideas and conceptions have been greatly extended and enriched."*

However there was less effective and coherent analysis of the course work than this teacher had achieved in her school placement.



## BIBLIOGRAPHY

- Acland, H. (1973) "Streaming in English primary schools", *British Journal of Educational Psychology*, Vol. 43, No.2, pp.151-160.
- Adelman, C. and Walker, R. (1974) "Open space - open classrooms", *Education 3-13*, Vol.2, No.2, pp.103-7.
- Advisory Committee on the Supply and Training of Teachers (1974) Department of Education and Science *In-service education and training: some considerations*. (Mimeo). London: Her Majesty's Stationery Office.
- Aitken, M., Bennett, S. and Hesketh, J. (1981) "Teaching styles and pupil progress: a re-analysis", *British Journal of Educational Psychology*, Vol.51, No.2, pp.170-186.
- Alexandar, R.J. (1984) *Primary teaching*. London: Holt, Rinehart and Winston.
- Allender, J. (1972) "New conceptions of the role of the teacher", in Silverman, M.L. (ed) *The psychology of open teaching and learning: an enquiry approach*. Boston, Mass: Little, Brown and Company. pp.230-238.
- Amabile, T. and Stubbs, M. (1982) *Psychological research in the classroom*. London: Pergammon Press.
- American Psychological Association (1974) *Standards for educational and psychological tests*. Prepared by a joint committee of the American Psychological Association, the American Educational Research Association and the National Council on Measurement in Education. Washington, D.C.: American Psychological Association.
- Anderson, R.G., Spiro, R. and J. and Montague, W.F.(eds) (1977) *Schooling and the acquisition of knowledge*. Hillsdale, N.J.: Erlbaum.
- Anthony, W.S. (1979) "Progressive learning theories: the evidence", in Bernbaum, G. (ed) *Schooling in Decline*. London: Macmillan.
- Anthony, W.S. (1982) "Research on progressive teaching", *British Journal of Educational Psychology*, Vol.52, pp.381-385.

- Argyris, C. (1953) *Executive leadership*. New York: Harper
- Argyris, C. (1970) *Intervention, theory and method*. New York: Addison-Wesley.
- Argyris, C. and Schon, D.A. (1976) *Theory in practice: increasing professional effectiveness*. London: Jossey-Bass.
- Armstrong, M. (1980) *Closely observed children: a diary of a primary classroom*. London: Writers and Readers in association with Chameleon Press.
- Ashton, P., Keen, P., Davies, F., and Holly, B. (1975) *The aims of primary education: a study of teachers' opinions*. London: Macmillan Education.
- Ashton, P., Henderson, E.S., Merritt, J.E. and Mortimer, D.J. (1982) *Teacher education in the classroom: initial and in-service*. London: Croom Helm.
- Auld, R. (1976) *The William Tyndale junior and infant schools: report of the public inquiry into the teaching, organisation and management of the William Tyndale junior and infant schools, Islington, London, N.1*. London: Inner London Education Authority.
- Badcock, E.H., Daniels, D.B., Islip, J., Razzell, A.G., and Ross, A.M. (1972) *Education in the middle years: the first report from the Schools Council Middle Years of schooling project*. London: Evans/Methuen Educational for Schools Council.
- Banks, O. and Finlayson, D. (1973) *Success and failure in the secondary school: an interdisciplinary approach to school achievement*. London: Methuen.
- Barker Lunn, J.C. (1970) *Streaming in the primary school: a longitudinal study of children in streamed and non-streamed junior schools*. Slough: National Foundation for Educational Research.
- Barnes, D. (1982) *Practical curriculum study*. London: Routledge and Kegan Paul.
- Barth, R.S. (1975) "Open education: assumptions about children, learning and knowledge" in Golby, M., Greenwold, J. and West R. (eds) *Curriculum Design*. London: Croom Helm. pp.58-81.



- Bassey, M. (1978) *Nine hundred primary school teachers*. Slough: National Foundation for Educational Research.
- Bealing, D. (1972) "The organisation of junior school classrooms", *Educational Research*, Vol. 14, pp.231-5.
- Becher, T., Eraut, M.R. and Knight, J. (1981) *Policies for educational accountability*. London: Heinemann.
- Becker, H. (1952) "Social class variations in teacher-pupil relationships", *Journal of Educational Sociology*, Vol.25, pp.451-65.
- Bell, A., Ziprusky, M. and Switzer, F. (1976) "Informal or open area education in relation to achievement and personality", *British Journal of Educational Psychology*, Vol.46, No.3, pp.235-43.
- Bell, A., Abrahamson, D., and Growse, R. (1977) "Achievement and self-reports of responsibility for achievement in informal (open-space) and traditional classrooms", *British Journal of Educational Psychology* . Vol.47, No.3, pp.258-67.
- Bell, S. (ed) (1977) *The organisation of open plan primary schools: report of a national course on organisation of open-plan primary schools*. Glasgow: Jordanhill College of Education.
- Bennett, S.N. and Jordan, J. (1975) "A typology of teaching styles in primary schools", *British Journal of Educational Psychology*, Vol.45, No.1, pp.20-28.
- Bennett, S.N. (1976) *Teaching styles and pupil progress*. London: Open Books.
- Bennett, S.N. (1978) "Recent research on teaching: a dream, a belief and a model", *British Journal of Educational Psychology*, Vol.48, pp.127-147.
- Bennett, S.N., Andreae, J., Hegarty, P. and Wade, B. (1980) *Open plan schools: teaching, curriculum and design*. Slough: National Foundation for Educational Research.

- Bennett, S.N., Desforbes, C., Cockburn, A. and Wilkinson, B. (1984) *The quality of pupil learning experiences*. London: Lawrence Erlbaum Associates.
- Bentley, C. and Marsden, D. (1983) *Guides to assessment in education: mathematics*. London: Macmillan Education.
- Bernbaum, G. (ed) (1979) *Schooling in decline*. London: Macmillan Education.
- Bernstein, B. (1970) "A critique of the concept of compensatory education", in Rubinstein, D. and Stoneman, C.(eds) *Education for democracy*. London: Penguin Education Special. pp.110-121.
- Black, H.D. (1980) *Collaborating with teachers - a researcher's perspective*. Paper presented at Teachers in Research Seminar. Edinburgh: Scottish Council for Research in Education.
- Black, H.D. and Broadfoot, P. (1982) *Keeping track of teaching in the modern classroom*. London: Routledge and Kegan Paul.
- Black, H.D. and Dockrell, W. (1981) *Diagnostic assessment in secondary schools*. Edinburgh: Scottish Council for Research in Education.
- Black, H.D. and Dockrell, W.B. (1984) *Criterion-referenced assessment in the classroom*. Edinburgh: Scottish Council for Research in Education.
- Blackie, J.E.H. (1967) *Inside the primary school*. London: Her Majesty's Stationery Office.
- Blackie, J.E.H. (1974) *Changing the primary school: an integrated approach*. London: Macmillan.
- Blenkin, G. and Kelly, A. (1981) *The primary curriculum*. London: Harper and Row.
- Bloom, B.S. (ed) (1956) *Taxonomy of educational objectives. Handbook 1: cognitive domain*. New York: Longman.



- Blyth, W.A. (1965) *English primary education: a sociological study*. Volumes 1 and 2. London: Routledge and Kegan Paul.
- Blyth, W.A. and Clayfield, R. (1982) "Primary education", in Cohen L., Thomas, J., and Manion, L. *Educational research and development (1970-1980)*. Slough: National Foundation for Educational Research/Nelson. pp.152-164.
- Bolam, R. (ed) (1982) *School-focused in-service training*. London: Heinemann Educational Books.
- Bowen, H.C. (1903) *Froebel and education by self activity*. London: Heineman.
- Boydell, D. (1974) "Teacher-pupil contact in junior classrooms", *British Journal of Educational Psychology*, Vol.44, pp.313-18.
- Boydell, D. (1975a) "Individual attention: the child's eyeview", *Education*, 3-13, Vol.3, April, pp.9-13.
- Boydell, D. (1975b) "Pupil behaviour in junior classrooms", *British Journal of Education Psychology*, Vol.45, No.2, pp.122-129.
- Boydell, D. (1979) *The primary teacher in action*. London: Open Books.
- Boyer, E.G., Simon, A. and Karafin, G. (eds) (1973) *Measures of maturation: an anthology of early childhood observation instruments*, Vol.II. Philadelphia: Research for Better Schools.
- Bradburn, E. (1976) *Margaret McMillan: framework and expansion of nursery education*. Surrey: Denholm House Press.
- Brandis, W. and Bernstein, B. (1974) *Selection and control: teachers' ratings of children in the infant school*. London: Routledge and Kegan Paul.
- Brandt, R. (1973) "Three weeks in British infant schools", in Hearn, B.D., Burdin, J. and Katz, L. (eds) *Current research and perspectives in open education*. Washington D.C.: EKNE.

- Brandt, R. (1975) "An observational portrait of a British infant school", in Spodek, B. and Walberg, H. (eds) *Studies in open education*. New York: Agathon Press. pp.101-25.
- Braun, C. (1976) "Teacher expectation; socio-psychological dynamics", *Review of Educational Research*, Vol.46, pp.185-213.
- Brearley M. *et al.* (1969) *Fundamentals in the first school*. Oxford: Blackwell.
- Brimer, E. (1969) *Bristol Achievement Tests*. London: Nelson.
- Broadfoot, P. *et al* (1981) *Politics and educational change; an international survey*. London: Croom Helm.
- Brophy, J.E. and Good, T.L. (1974) *Teacher-student relationships - causes and consequences*. New York: Holt, Rinehart and Winston.
- Brophy, J. (1983) "Research on the self-fulfilling prophecy and teacher expectations", *Journal of Educational Psychology*, Vol.85, No.6, pp.631-61.
- Brown, M.G. and Precious, N.G. (1970) *The integrated day in the primary school*. London: Ward Lock.
- Brown, R.L. (1969) "A study of children's behaviour in primary school classrooms", *Pedagogica Europaea*, Vol.5, pp.28-40.
- Bruner, J.S. (1966) *Towards a theory of instruction: essays*. Cambridge, Mass.: Belknap Press of Harvard University Press.
- Burgess, R. (ed) (1985) *Field methods in the study of education*. East Sussex: Falmer Press.
- Buros, O.K. (ed) (1938-1972) *Mental measurement: Yearbooks*. Highland Park, N.J.: Gryphon Press.
- Buros, O.K. (ed) (1974) *Tests in print .Volume II*. Highland Park, N.J.: Gryphon Press.



- Burt, C. (1945) 'The reliability of teachers' assessments of their pupils', *British Journal of Educational Psychology*, Vol.15, pp.80-92.
- Calderhead, J. (1984) *Teachers' classroom decision-making*. East Sussex: Holt, Rinehart and Winston.
- Callaghan, J. (1976) "Towards a national debate - The Prime Minister's Ruskin speech", *Education*, Vol.22, October, pp.332-333.
- Campbell, D.T. and Stanley, J. (1963) "Experimental and quasi-experimental design for Research and Teaching", in Gage, N.L. (ed) *Handbook of research on teaching*. Chicago: Rand McNally.
- Chambers, J.H. (1983) *The achievement of education: an examination of key concepts in educational practice*. London: Harper and Row.
- Chambers, P. and Powney, J. (1982) "School-based curriculum research", *British Educational Research Journal*, Vol.8, No.2, pp.133-139.
- Choppin, B. (1969) "An item bank using sample-free calibration", in Wood, R. and Skurnik, L.S. (eds) *Item banking*. Slough: National Foundation for Educational Research.
- Clifford, M. and Walster, E. (1973) "The effect of physical attractiveness on teacher expectations", *Sociology of Education*, Vol.76, pp.248-258.
- Clift, P., Weiner, G. and Wilson, E. (1981) *Record keeping in primary schools*. Schools Council Research Studies. London: Macmillan Education.
- Coard, B. (1971) *How the West Indian child is made ESN*. London: New Beacon.
- Cohen, L. and Deale, R.N. (1977) *Assessment by teachers in examinations at 16+*. Schools Council Examinations Bulletin 37. London: Evans/Methuen.
- Cohen, L. and Manion, L. (1980) *Research methods in education*. London: Croom Helm.

- Cohen, L., Thomas, J. and Manion, L. (1982) *Educational research and development in Britain (1970-1980)*. Slough: National Foundation for Educational Research/Nelson.
- Coleman, J., Coser, L. and Powell, W. (1966) *Equality of educational opportunity*. Washington: U.S. Government Printing Office.
- Cooper, H., Baron, R. and Lowe, C. (1975) "The importance of race and social class information in the formation of expectancies about academic performance", *Journal of Educational Psychology*, Vol.67, pp.312-319.
- Cooper, H. (1979) "Pygmalion grows up: a model for teacher expectation communication and performance influence", *Review of Educational Research*, Vol.49, pp.389-410.
- Cox, C.B. and Dyson, A.E. (eds) (1969a) *Fight for education. Black paper No.1*. London: The Critical Quarterly Society.
- Cox, C.B. and Dyson, A.E. (eds) (1969b) *The crisis in education. Black paper No.2*. London: The Critical Quarterly Society.
- Cox, C.B. and Dyson, A.E. (eds) (1970) *Goodbye Mr. Short. Black paper No.3*. London: The Critical Quarterly Society.
- Cox, C.B. and Dyson, A.E. (eds) (1975) *The fight for education. Black paper No.4*. London: Dent.
- Cox, C.B. and Dyson, A.E. (eds) (1977) *Black paper No.5*. London: Temple Smith.
- Craft, M., Raynor, J.F. and Cohen, L. (eds) (1967) *Linking home and school*. London: Longmans, Green & Co. Ltd.
- Croll, P. (1981) "Social class, pupil achievement and classroom interaction", in Simon, B. and Willcocks, J. (eds) *Research and practice in the primary classroom*. London: Routledge and Kegan Paul.



- Cronbach, L.J. (1970) *Essentials of psychological testing* (3rd Edition). New York: Harper and Row.
- Crossland, R.W. (1967) *Report of an individual study of the Nuffield Foundation Primary Science Project* (unpublished). Manchester: University of Manchester.
- Daniels, J.C. (1961) "The effects of streaming in the primary school II, a comparison of streamed and unstreamed schools", *British Journal of Educational Psychology*, Vol.31, pp.119-127.
- Daniels, J.C. and Diack, H. (1970) *The standard reading tests*. London: Chatto and Windus.
- Davidson, H.H. and Lang, G. (1960) "Children's perceptions of their teacher's feelings towards them, related to self-perception, school achievement and behaviour", *Journal of Experimental Education*, Vol. 29, pp.107-118.
- Davie, R., Butler, N. and Goldstein, H. (1972) *From birth to seven*. London: Longmans.
- Day, C. (1981) *Classroom-based in-service teacher education: the development and evaluation of a client-centred model*. University of Sussex Education Area Occasional Paper 9. Sussex: University of Sussex.
- Day, C. (1985) "Professional learning and researcher intervention: an action research perspective", *British Educational Research Journal*, Vol.11, No.2, pp.133-152.
- Dean, J. (1972) *Recording children's progress*. London: Macmillan Education.
- Dearden, R. F. (1968) *The philosophy of primary education*. London: Routledge and Kegan Paul.
- Dearden, R.F. (1976) *Problems in primary education*. London: Routledge and Kegan Paul.
- Department of Education and Science (1972) *The James Report : teacher education and training*. London: Her Majesty's Stationery Office.

- Department of Education and Science (1974) *Educational disadvantage and the educational needs of immigrants*. Cmnd 5720. London: Her Majesty's Stationery Office.
- Department of Education and Science (1975a) *Bullock Report: A language for life*. Report of Committee of Inquiry appointed by Secretary of State for Education and Science under Chairmanship of Sir Alan Bullock. London: Her Majesty's Stationery Office.
- Department of Education and Science (1975b) *Education: a framework for expansion*. Cmnd. 5174. London: Her Majesty's Stationery Office.
- Department of Education and Science (1977a) *Education in schools: a consultative document*, Cmnd. 6869. London: Her Majesty's Stationery Office.
- Department of Education and Science (1977b) *Taylor Report: a new partnership for our schools*. London: Her Majesty's Stationery Office.
- Department of Education and Science (1978a) *Primary education in England: A survey by H.M. Inspectors of Schools*. London: Her Majesty's Stationery Office.
- Department of Education and Science (1978b) *Making INSET work*. London: Her Majesty's Stationery Office.
- Department of Education and Science (1983a) *9-13 Middle Schools: an illustrative survey*. London: Her Majesty's Stationery Office.
- Department of Education and Science (1983b) *Teaching Quality: a white paper*. London: Her Majesty's Stationery Office.
- Dewey, J. (1900) *Educational principles*. The Elementary School Journal.
- Dewey, J. (1916) *Democracy and education*. New York: Collier-Macmillan (1966 edition).
- Dobby, J. and Duckworth, D. (1979) *Objective assessment by means of item banking*. Schools Council Examinations Bulletin 40. London: Evans/Methuen Educational.



- Doll, R.C. (1974) *Curriculum improvement: decision-making and process*. Boston: Allyn and Bacon.
- Donoughue, C. (ed) (1981) *In-service: the teacher and the school*. London: Kogan Page in association with Open University Press.
- Douglas, J.W.B. (1964) *The home and school: a study of ability and attainment in the primary school*. London: Macgibbon and Kee.
- Downey, M. (1977) *Interpersonal judgements in education*. London: Harper and Row.
- Dusek, J. and Joseph, G. (1983) "The bases of teacher expectancies: a meta-analysis", *Journal of Educational Psychology*, Vol. 85, pp.327-46.
- Duthie, J.H. (1970) *Primary school survey: a study of the teacher's day*. A report by Dr. J.H.Duthie [to the] Scottish Education Dept. Edinburgh: Her Majesty's Stationery Office
- Ebbutt, D. and Partington, D. (1982) "Self-monitoring by teachers: aftermath of a project" in Bolam, R. (ed) *School-focused in-service training*. London: Heinemann Educational Books.
- Education Act (1944) *White paper: educational reconstruction*. Cmnd. 6824. London: His Majesty's Stationery Office.
- Eggleston, J.F., Galton, M. and Jones, M. (1975) *Processes and products of science teaching*. London: Macmillan.
- Eggleston, J. and Kerr, J. (1969) *Studies in Assessment*. London: English Universities Press.
- Eggleston, J.F. and Lobel, R.F. (1969) "Assessing attainment in History", in Eggleston, J.F. and Kerr, J.F. (eds) *Studies of assessment*. London: The English Universities Press pp.77-120.

- Eggleston, J. and Newbould, C. (1969) "Assessing attainment in science at C.S.E. level", in Eggleston, J.F. and Kerr, J.F. (eds) *Studies in assessment*. London: The English Universities Press. pp.15-76.
- Eisner, E. (1979) *The educational imagination: on the design and evaluation of school programs*. New York: Macmillan Publishing Co.
- Elbaz, F. (1982) *Teacher thinking: a study of practical knowledge*. London: Croom Helm.
- Elliott, J. (1976) "Preparing teachers for classroom accountability", *Education for teaching*, No.100, pp. 49-71.
- Elliott, J. (1977) *Conceptualising relationships between research/evaluation procedures and in-service teacher education: a case study of a developing methodology*. Schools Council. Progress in learning science project. Dissemination study Interim Working Paper IX. Cambridge: Cambridge University Institute of Education.
- Elliott, J. (1978a) "Classroom accountability and the self-monitoring teacher", in Harlen, W. (ed) *Evaluation and the teachers' role*. London: Macmillan Education. pp.47-90.
- Elliott, J. (1978b) "What is action research in schools?", *Journal of Curriculum Studies*, Vol.104.
- Elliott, J. (1978c) "Classroom research; science or commonsense?", in McAleese, R., and Hamilton, D. (eds) *Understanding classroom life*. Slough: National Foundation for Educational Research.
- Elliott, J. (1979a) "The case for school self-evaluation", *Forum*, Vol.22, No.1 pp.23-25.
- Elliott, J. (1979b) "Implementing school-based action research: some hypotheses", *Cambridge Journal of education*, Vol.9, No.1, Lent, pp.55-71.
- Elliott, J. (1980a) "Who should monitor performance in schools?", in Sockett, H. (ed) *Accountability in the English educational system*. Kent: Hodder and Stoughton.



- Elliott, J. (1980b) "Implications of classroom research for professional development", in Hoyle, E. and Megarry, J. (eds) *Professional development of teachers: world year book of education*. London: Kogan Page.
- Elliott, J. (1983) "Self-evaluation: professional development and accountability", in Galton, M. and Moon, B. (eds) *Changing schools ... changing curriculum*. London: Harper and Row.
- Elliott, J. and Adelman, C. (1975) *Ford Teaching project*. East Anglia: University of East Anglia: Centre for Applied Research in Education.
- Elliott, J. and Adelman, C. (1976) *Innovation at the classroom level: a case study of the Ford Teaching Project*. Unit 28, E203 The Open University. Bletchley: The Open University Press.
- Enright, L. (1979) "Learning in my classroom", *Forum*, Vol.21, No.3, pp.78-81.
- Eraut, M. (1972) *Occasional paper 4: in-service education for innovation*. London: National Council for Educational Technology.
- Evans, J. (1975) "An activity analysis of U.S. traditional, U.S. open and British open classrooms", in Spodek, B. and Walburg, H.J. (eds) *Studies in open education*. New York: Agathon Press.
- Eyers, S. and Richmond, J. (eds) (1982) *On becoming our own experts*. Talk Workshop Group. London: The Inner London Education Authority English Centre.
- Fairbrother, R. (1975) "The reliability of teachers' judgements of the abilities being tested by multiple-choice items", *Educational Research*, Vol.17, No.3, pp.202-210.
- Ford, J. (1969) *Social class and the comprehensive school*. London: Routledge and Kegan Paul.
- Foster, J. (1971) *Creativity and the teacher*. London: Macmillan Education.

- Fowler, G. (1981) "The changing nature of educational politics in the 1970's", in Broadfoot *et al.*, *Politics and educational change: an international survey*. London: Croom Helm. pp.13-28.
- France, N. and Fraser, I. (1975) *Richmond tests of basic skills*. London: Nelson.
- Gage, N.L. (ed) (1963) *Handbook of research on teaching*. Chicago: Rand McNally.
- Galton, M. (1978a) "Accountability and the teacher: some moves and counter-moves", in Richards, C. (ed) *Power and the Curriculum: issues in curriculum studies*. Driffield: Nafferton Books pp.135-146.
- Galton, M. (1978b) *British mirrors: a collection of classroom observation systems*. Leicester: School of Education, Leicester University.
- Galton, M. (ed) (1980) *Curriculum change*. Leicester: Leicester University Press.
- Galton, M. and Moon, B. (eds) (1983) *Changing schools, changing curriculum*. London: Harper and Row.
- Galton, M. and Simon, B. (eds) (1980) *Progress and performance in the primary classroom*. London: Routledge and Kegan Paul.
- Galton, M., Simon, B. and Croll, P. (1980) *Inside the primary classroom*. London: Routledge and Kegan Paul.
- Galton, M. and Willcocks, J. (eds) (1983) *Moving from the primary classroom*. London: Routledge and Kegan Paul.
- Gannoway, H. (1976) "Making sense of school", in Stubbs, M. and Delamont, S. (eds) *Explorations in classroom observation*. London: Wiley and Sons.
- Gardner, D.E.M. and Cass, J.E. (1965) *The role of the teacher in the infant and nursery school*. Oxford: Pergamon Press.
- Gardner, D.E.M. (1966) *Experiment and tradition in primary schools*. London: Methuen.



- Garner, J. and Bing M. (1973) "Inequalities in pupil-teacher contacts further evidence and a typology", *British Journal of Educational Psychology*, Vol.43, pp.234-43.
- Garner, J. and Bing M. (1974) "The elusiveness of pygmalion and differences in teacher-pupil contacts", *Interchange*, Vol.4, pp.34-42.
- Ghaye, A.L. (1984) *Discovering pupil mindscapes: a participant observation study of children in the middle years of schooling*. Unpublished Ph.D. Thesis. University of London, Institute of Education.
- Gillham, C. and Hesse, K.A. (1970) *The Nottingham number tests: basic number concepts and skills age 9:1 to 11:0*. London: Hodder and Stoughton.
- Gillham, C. and Hesse, K.A. (1973) *The Leicester number tests: basic number concepts*. London: Hodder and Stoughton.
- Gillham, C. and Hesse, K.A. (1976) *Basic Number Screening*. London: Hodder and Stoughton.
- Ginsberg, M., Meyenn, R.J., Miller, H.D. and Ranceford-Hadley, C. (1977) *Teacher reaction to the Great Debate and to education cuts*. Paper presented to the British Education Research Association Conference. Nottingham.
- Golby, M., Crane, G. and Tann, N. (1981) "A strategy for professional development", *British Journal of In-service Education*, Vol.7, No.3, pp.153-157.
- Golby, M., Greenwald, J. and West, R. (eds) (1975) *Curriculum design*. London: Croom Helm.
- Goldstein, H. (1979) "Consequences of using the Rasch model for educational assessment", *British Educational Research Journal*, Vol.5, No.2, pp.211-220.
- Goldstein, H. (1980) "Dimensionality, bias, independence and measurement scale problems in latent trait test score models", *British Journal of Mathematical and Statistical Psychology*, Vol.33, pp.234-246.

- Goldstein, H. and Blinkhorn, S. (1977) "Monitoring educational standards - an inappropriate model", *Bulletin of the British Psychological Society*, Vol.30, pp.309-11.
- Goldstein, H. and Blinkhorn, S. (1982) "The Rasch model still does not fit", *British Educational Research Journal*, Vol.8, No.2, pp.167-170.
- Good, T. (1980) "Classroom expectations: teacher-pupil interactions", in McMillan J. (ed) *The social psychology of school learning*. New York: Academic Press.
- Good, T., Biddle, R. and Brophy, J. (1975) *Teachers make a difference*. New York: Holt, Rinehart and Winston.
- Good, T.L. and Brophy, J.E. (1974) "Changing teacher and student behaviour: an empirical investigation", *Journal of Educational Psychology*, Vol.66, No.3, pp.390-405.
- Goodacre, E. (1967) *Teachers and their pupil's home backgrounds*. Slough: National Foundation for Educational Research.
- Gower, D. (1981) "The chemistry of the classroom", in Nixon, J. (ed) *A teachers' guide to action research*. London: Grant McIntyre.
- Gowan, J.C., Demos, G.D. and Torrance, E.P. (eds) (1967) *Creativity: its educational implications*. New York: Wiley.
- Gray, J. and Satterley, D. (1981) "Formal or informal? A re-assessment of the British evidence", *British Journal of Educational Psychology*, Vol.51, pp.187-196.
- Green, T. (ed) (1971) *Educational planning in perspective: forecasting and policy-making*. London: IPC Business Press Ltd.
- Gross, N., Chiaquinta, J.B. and Bernstein, M. (1971) *Implementing organisational innovations*. London: Harper and Row.
- Grundy, S. (1981) *Three modes of action research*, Mimeograph. Victoria, Australia: Deakin University.



- Grundy, S. and Kemmis, S. (1981) "Social theory group dynamics and action research", in *The professional development of teachers through involvement in action research projects*. Victoria, Australia: Deakin University.
- Guildford, J.P. (1973) *Fundamental statistics in psychology and education*. Tokyo: McGraw-Hill.
- Habermas, J. (1971a) *Towards a rational society*. London: Heinemann.
- Habermas, J. (1971b) *Knowledge and human interests*. Boston, Mass: Beacon Press.
- Habermas, J. (1974) *Theory and practice*. London: Heinemann.
- Haddon, F.A. and Lytton, H. (1968) "Teaching approach and the development of divergent thinking abilities in primary schools", *British Journal of Educational Psychology*, Vol.38, No.2, pp. 171-180.
- Hadow Report (1926) *Education of the adolescent: report of the consultative committee*. London: His Majesty's Stationery Office.
- Hadow Report (1931) *Report of the consultative committee on the primary school*. London: His Majesty's Stationery Office.
- Hadow Report (1933) *Report of the consultative committee on infant and nursery schools*. London: His Majesty's Stationery Office.
- Hallworth, H.J. (1962) "A teacher's perception of his pupils", *Educational Review*, Vol. 14, pp.124-33.
- Halsey, A.H. (1972) *Educational priority*. London: Her Majesty's Stationery Office.
- Hambleton R.K., Swaminathan, H., Algina, J. and Coulson, D.B. (1978) "Criterion referenced testing and measurement: a review of technical issues and developments", *Review of Educational Research*, Vol.48, pp.1-47.
- Hamilton, D. (1976) *Curriculum evaluation*. London: Open Books.

- Hamilton, D. (1980) "Some contrasting assumptions about case study research and survey analysis", in Simons, H. (ed) *Towards a science of the singular*. East Anglia: University of East Anglia.
- Hammersley, M. and Atkinson, P. (1983) *Ethnography: principles in practice*. London: Tavistock Publications.
- Harari, H. and McDavid, J. (1973) "Naive stereotypes and teachers' expectations", *Journal of Education Psychology*, Vol 65, pp.222 - 225.
- Hargreaves, A. and Tickle, L. (eds) (1980) *Middle schools: origins, ideology and practice*. London: Harper and Row.
- Hargreaves, D. (1967) *Social relations in a secondary school*. London: Routledge and Kegan Paul.
- Hargreaves, D. (1972) *Interpersonal relations and education*. London: Routledge and Kegan Paul.
- Hargreaves, D.H. (1977) "The process of typification in classroom interaction: models and methods", *British Journal of Educational Psychology*, Vol.47, pp. 274-284.
- Hargreaves, D., Hester, M.K. and Mellor, F.J. (1975) *Deviance in classrooms*. London: Routledge and Kegan Paul.
- Harlen, W. (1967) *An evaluation of the development of scientific concepts in children taking part in the Oxford Primary Science Research Project (1966-1967)*. Bristol: Research Unit, University of Bristol, Institute of Education.
- Harlen, W. (1975) *Science 5-13: a formative evaluation*. London: Macmillan Education.
- Harlen, W. (1976a) "Change and development in education strategy", in Tawney, D. (ed) *Curriculum evaluation today: trends and implications*. London: Macmillan Education for Schools Council.
- Harlen, W. (1976b) "Towards a stronger teacher role in curriculum development", *British Educational Research Journal*, Vol 2, No. 2. pp.20-27.



- Harlen, W., Darwin, A. and Murphy, M (1977a) *Match and mismatch: raising questions*. Leaders' guides. In-service training materials for teachers produced by the Schools Council Progress in learning science project. Edinburgh: Oliver and Boyd.
- Harlen, W., Darwin, A. and Murphy, M. (1977b) *Match and Mismatch: finding answers*. In-service training materials for teachers produced by the Schools Council Progress in learning science project. Edinburgh: Oliver and Boyd.
- Harlen, W. (ed) (1978) *Evaluation and the teacher's role*. London: Macmillan Education.
- Harlen, W. (1979) "Accountability that is of benefit to schools", *Journal of Curriculum Studies*, Vol 11, pp.287-97.
- Harlen, W. (1983) *Guides to assessment in education: science*. London: Macmillan Education.
- Hartley, D. (1978) "Teachers' definitions of boys and girls - some consequences", *Research in Education*, Vol. 20, pp.23-35.
- Hartley, D. (1979) "Sex differences in classroom behaviour of infants: the views of teachers and pupils", *British Journal of Educational Psychology*, Vol 49, pp.188-93.
- Havelock, R.G. (1971) "The utilization of educational research and development", *British Journal of Educational Technology*, Vol. 2, No. 2, pp.84-98.
- Hearn B.D., Burdin, J. and Katz, L. (eds) (1973) *Current research and perspectives in open education*. Washington D.C: EKNE.
- Herbert, G. (1974) "Teachers' ratings of classroom behaviour factorial structure", *British Journal of Educational Psychology*, Vol. 44, pp.233-240.
- Hindley, C.E. (1962) "Social class influences on the development of ability in the first five years", in Stand, A.G. and Husen, T.(eds) *Child education*. Copenhagen: Munksgaard.

- Hilsum, S. and Cane, B. (1971) *The teacher's day*. Slough: National Foundation for Educational Research.
- Hirst, P. (1980) "The logic of curriculum development", in Galton, M. (ed) *Curriculum change*. Leicester: Leicester University Press.
- Holmes, E. (1911) *What is and what might be: a study of education in general and elementary education in particular*. London: Constable.
- Holsti, O.R. (1968) "Content analysis", in Linzey, G. and Aronson, E. (eds) *The Handbook of social psychology, Vol.2, Research methods*. New York: Addison-Wesley.
- Hoste, R. and Bloomfield, B. (1975) *Continuous assessment in the CSE: opinion and practice*. Schools Council Examinations Bulletin 31. London: Evans Methuen.
- House, E. (ed) (1973a) *School evaluation*. Berkeley, Cal: McCutchan.
- House, E.R. (1973b) *The price of productivity: who pays?* (mimeograph) Urbana, Ill: Centre for instructional research and curriculum evaluation. University of Illinois.
- Hoyle, E. (1973) "Strategies for curriculum change", in Watkins, R. (ed) *In-service training: structure and content*. London: Ward Lock Educational.
- Hoyle, E. and Megarry, J. (eds) (1980) *Professional development of teachers: world year book of education*. London: Kogan Page.
- Humanities Curriculum Project (1970) *The Humanities Curriculum project: an introduction*. London: Heinemann Educational Books.
- Hunter, M. (1980) *Advantages and disadvantages of collaborative research with particular reference to two studies: 'Remedial education in the primary school' and 'Investigating a technique for developing written language'*. Paper presented at Teachers in Research Seminar. Edinburgh: Scottish Council for Research in Education.



- Hurlyn, A. (1975) "Open plan schools and inquiry/discovery learning", *Cambridge Journal of Education*, Vol.5, No.2, pp.98-103.
- I.L.E.A. (1979) *Keeping the school under review*. Inner London Education Authority.
- Ingleby, J.D. and Cooper, E. (1974) "How teachers perceive first year school children: sex and ethnic differences", *Sociology*, Vol.8, pp.463-473.
- Isaacs, S. (1930) *Intellectual growth in young children*. London: Routledge and Kegan Paul.
- Isaacs, S. (1933) *Social development in young children: a study of beginnings*. London: Routledge and Kegan Paul.
- Jackson, B. (1964) *Streaming: an education system in miniature*. London: Routledge and Kegan Paul.
- Jackson, B. and Marsden, D. (1962) *Education and the working class*. London: Routledge and Kegan Paul.
- Jackson, D. (1981) "Food for thought", in Nixon, J. (ed) *A teacher's guide to action research*. London: Grant McIntyre.
- Jackson, P.W. and Lahaderne, H.M. (1967) "Inequalities in teacher-pupil contacts", *Psychology in schools*. Vol.4, pp.204-211.
- Jarman, C. (1977) "The organisation of open plan primary schools: the Oxfordshire experience", in Bell, S. (ed) *The organisation of open-plan primary schools: report of a national course on organisation of open-plan primary schools*. Glasgow: Jordanhill College of Education.
- Jasman, A. (1977) "Assessment in the primary school: filling in some gaps", *Forum*, Vol.19, No.3, pp.13-17.
- Jasman, A. (1979) "Developing teacher-based assessment techniques for the primary school", *British Educational Research Journal*, Vol.5, No.1, pp.125-134.

- Jasman, A. and Ashby, C. (1980) *Teachers in research: exploring the nature of collaborative research*. Paper presented at Teachers in research seminar. Edinburgh: Scottish Council for Research in Education.
- Jasman, A. (1981) "Teachers' assessments in classroom research", in Simon, B. and Willcocks, J. (eds) *Research and practice in the primary classroom*. London: Routledge and Kegan Paul. pp.144-160.
- Jeffrey, A. (1970) "Infants in an open plan school", in Mason, S. (ed) *In our experience*. London: Longman.
- Jones, J.A.G. (1979) *Industrial training services: an in-school approach to in-service training*. London: Industrial Science.
- Kallett, A. (1970) "Some recent changes in the primary school", in Mason, S. (ed) *In our experience*. London: Longman.
- Keddie, N. (ed) (1973) *Tinker, Tailor...* Harmondsworth: Penguin.
- Kellmer-Pringle, M.L., Butler, N.R. and Darrie, R.V. (1966) *11,000 seven year olds*. London: Longman.
- Kelly, G.A. (1955) *The psychology of personal constructs*, Vol.1. New York: Norton.
- Kelly, P.J. and Lister, R.E. (1969) "Assessing practical abilities in Nuffield A-level Biology", in Eggleston, J.F. and Kerr, J.F. (eds) *Studies in Assessment*. London: The English Universities Press. pp.121-152.
- Kelman, H. (1961) "Three processes of social influence", *Public Opinion Quarterly*, Vol. 25. N.J.: Princeton.
- Kemp, L. (1955) "Environmental and other characteristics determining attainment in primary schools", *British Journal of Educational Psychology*, Vol. 25, pp.67-77.
- Kerr, J.F. (ed) (1968) *Changing the curriculum*. London: University of London Press.



- Kimmins, C.W. and Rennie, B. (1932) *The triumph of the Dalton plan*. London: Nicholson and Watson.
- King, R.A. (1969) *Values and involvement in a grammar school*. London: Routledge and Kegan Paul.
- King, R.A. (1973) *School organisation and pupil involvement: a study of secondary schools*. London: Routledge and Kegan Paul.
- King, A. (1978) *All things bright and beautiful? A sociological study of infants classrooms*. Chichester: John Wiley and Sons Ltd.
- Kitwood, T. (1977) *Some sociological aspects of testing*, Conference proceedings: British Educational Research Association. Nottingham.
- Kogan, M. (1971) "English primary schools: a model of institutional innovation", in Green, T. (ed) *Educational planning in perspective: forecasting and policy-making*. London: IPC Business Press Ltd. pp.40 - 48.
- Kogan, M. (1975) *Educational policy-making: a study of interest groups and parliament*. London: George Allen and Unwin.
- Lacey, C. (1970) *Hightown Grammar: the school as a social system*. Manchester: Manchester University Press.
- Law, B. (1984) *Uses and abuses of profiling: a handbook*. London: Harper & Row.
- Levy, P. and Goldstein, H. (eds) (1984) *Tests in Education: a book of critical reviews*. London: Academic Press.
- Linzey, G. and Aronson, E. (eds) *The handbook of social psychology, Vol. 2, Research methods*. New York: Addison-Wesley.
- Lovell, K. (1963) "Informal versus formal education and reading attainments in junior schools", *Educational Research*, November, pp.71-76.

- Lynch, J. (1980) "Legitimation crises for the English middle school", in Hargreaves, D. and Tickle, L.(eds) *Middle Schools: origins, ideology and practice*. London: Harper and Row.
- McAleese, R. and Hamilton, R. (eds) (1978) *Understanding classroom life*. Slough: National Foundation for Educational Research.
- MacDonald, B. (1971) "The evaluation of the humanities curriculum project: a holistic approach", *Theory into Practice*, Vol. 10, pp.163-167.
- MacDonald, B. (1978) "Hard Times: educational accountability in England", *Educational analysis*, Vol. 1, No.1, pp.23-45.
- MacDonald, B. & Walker, R. (1976) *Changing the curriculum*. London: Open Books.
- McIntyre, D., Morrison, A. and Sutherland, J. (1966) "Social and educational variables relating to teachers' assessments of primary school pupils", *British Journal of Educational Psychology*, Vol.36, No.3, pp.272-9.
- McMillan, J. (ed) (1980) *The social psychology of school learning*. New York: Academic Press.
- McNiff, J. (1984) "Action research: a generative model for in-service support", *British Journal of In-service Education*, Vol. 10, No.8, pp.40-46.
- Mc V. Hunt, J. (1961) *Intelligence and experience*. New York: Ronald Press Company.
- Mager, R.F. (1962) *Preparing instructional objectives*. Belmont, Cal: Fearon Press.
- Mager, R.F. (1968) *Developing attitudes towards learning*. Belmont, Cal: Fearon Press.
- Marjoram, D. (1978) "The APU and assessment in the middle years", *Education 3-13*, Vol. 6, No.2, pp.31-36.
- Marsh, L.G. (1970) *Alongside the child in the primary school*. London: A. & C. Black



- Marton, F. and Saljo, R. (1976) "On qualitative differences in learning: I - outcome and process", *British Journal of Educational Psychology*, Vol.46, No.1. pp.4-11.
- Mason, S. (1964) *The Leicestershire experiment and plan*. London: Councils and Education Press.
- Mason, S. (ed)(1970) *In our experience*. London: Longman.
- Medley, D.M., Quirk, T.J., Schluck, C.G. and Ames, N.P. (1973) "The personal record of school experience", in Boyer, E.G., Simon, A. and Karafin, G. (eds) *Measures of maturation: an anthology of early childhood observation instruments*, Vol.II. Philadelphia: Research for Better Schools.
- Mehrens, W.A. and Lehmann, I.J. (1975) *Standardised tests in education*. New York: Holt, Rinehart and Winston.
- Millman, J. (1974) "Criterion referenced measurement", in Popham, W.J. (ed) *Evaluation in education: current applications*. Berkeley, Cal: McCutcheon.
- Ministry of Education Building Bulletin (1949) *New Primary schools*. London: His Majesty's Stationery Office.
- Ministry of Education (1952) *Annual report: education in 1952*. London: His Majesty's Stationery Office.
- Ministry of Education (1959) *Handbook on primary education*. London: Her Majesty's Stationery Office.
- Miskimmin, E. (1980) *Collaborative research: a happy accident*. Paper presented at the Teachers in Research seminar. Edinburgh: Scottish Council for Research in Education.
- Mitchell, J. (1970) "Freedom in the junior school", in Rubinstein, D. and Stoneman, C. (eds) *Education for democracy*. London: Penguin education special, pp.104-109.
- Monjan, S.V. and Glassner, S.N. (1979) *Critical issues in competency based education*. New York: Pergamon.

- Montessori, M. (1912) *The Montessori method*. London: Heinemann.
- Moran, P. (1971) "The integrated day", *Educational Research*, Vol.14, No.1, pp. 65-69.
- Morris, L. and Fitz-Gibbon, C.T. (1978) *How to measure achievement*. London: Sage.
- Morrison, A. and McIntyre, D. (1969) *Teachers and teaching*. Harmondsworth, Middlesex: Penguin Books.
- Morrison, A. and McIntyre, D. (1972) *Social psychology of teaching*. Middlesex: Penguin modern psychology readings.
- Murphy, R. (1984) *Educational assessment: changing the face and facing the change*. Conference proceedings. British Educational Research Association.
- Musgrove, F. and Taylor, P.H. (1969) *Society and the teacher's role*. London: Routledge and Kegan Paul.
- Nash, R. (1973) *Classrooms observed: the teacher's perception and the pupil's performance*. London: Routledge and Kegan Paul.
- Nash, R. (1976) *Teacher expectations and pupil performance*. London: Routledge and Kegan Paul.
- National Union of Teachers (18.4.69) *Teacher*.
- Neale, M.D. (1958) *Neale analysis of reading ability*. London: Macmillan Education.
- Newsom Report (1963) *Half our future: report of the Central Advisory Council for Education (England)*. London: Her Majesty's Stationery Office.
- Newson, J. and E. (1977) *Perspectives on school at seven years old*. London: Allen and Unwin.
- Nie, N.H. et al. (1975) *Statistical package for the Social Sciences*. 2nd edition. Berks: McGraw Hill.



- Nitko, A.J. (1980) "Distinguishing the many varieties of criterion referenced tests", *Review of Educational Research*, Vol.50, pp.461-485.
- Nixon, J. (ed) (1981) *A teacher's guide to action research*. London: Grant McIntyre.
- Ogilvie, E. (1973) *Gifted children in primary schools*. London: Macmillan Education.
- Parkhurst, H. (1922) *Education on the Dalton Plan*. New York: Bell.
- Parlett, M. and Hamilton, D. (1972) *Evaluation as illumination: a new approach to the study of innovative programmes*. Edinburgh: University of Edinburgh, Centre for Research in the Educational Sciences. Occasional paper No.9.
- Pask, G. (1976) "Styles and strategies of learning", *British Journal of Educational Psychology*, Vol.46, No.2.
- Persell, C. (1977) *Education and inequality: the roots and results of stratification in American schools*. New York: The Free Press.
- Piaget, J. and Inhelder, B. (1969) *The psychology of the child*. London: Routledge and Kegan Paul.
- Pigeon, D.A. (1970) *Expectation and pupil performance*. Slough: National Foundation for Educational Research.
- Plowden, B. (1967) *The Plowden Report. Children and their primary schools: a report of the Central Advisory Council for Education (England)*. Vol. 1: Report; Vol. 2: Research and surveys. London: Her Majesty's Stationery Office.
- Pluckrose, H. (1975) *Open school: open society*. London: Evans Bros.
- Poole, R. (1972) "Characteristics of the taxonomy of educational objectives: cognitive domain - a replication", *Psychology in schools*, Vol 9, pp.83-88.
- Popham, W.J. (ed) (1974) *Evaluation in education: current applications*. Berkeley, Cal: McCutcheon.

- Popham, W.J. (1978) *Criterion-referenced measurement*. Englewood Cliffs, N.J.: Prentice Hall.
- Popham, W.J. (1981) *Modern educational measurement*. Englewood Cliffs, N.J.: Prentice Hall.
- Popham, W.J. and Husek T.R. (1969) "Implications of criterion - referenced measurement", *Journal of Educational Measurement*, Vol. 6, No. 1, pp. 1-9.
- Popham, W.J. and Lindheim, E. (1980) "The practical side of CRT development", *NCME Measurement in Education Series*, Vol.10, No.4, Spring 1980.
- Posner, G. (1978) "Tools for curriculum research and development: potential contributions from cognitive science", *Curriculum Inquiry*, Vol.8, No.4, pp.311-340.
- Pumfrey, P. (1976) *Reading: tests and assessment techniques*. London: Hodder and Stoughton.
- Raban, B. (1983) *Guides to assessment in education: reading*. London: Macmillan Education.
- Rance, P. (1971) *Record keeping in the progressive primary school*. London: Ward Lock Educational.
- Rasch, C. (1960) *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danmarks Paedagogiske Institut.
- Rawls, J. (1971) *A theory of justice*. Oxford: Oxford University Press.
- Reason, P. and Rowan, J. (eds) (1981) *Human inquiry: a sourcebook of new paradigm research*. Chichester: Wiley.
- Reid, J. (1975) *Breakthrough in action: an independant evaluation of 'Breakthrough to literacy'*. London: Longman.



- Resnick, L.B. (1972) "Teacher behaviour in an informal British infant school", *The School Review*, Vol.81, pp. 63-83.
- Richards, C. (ed) (1978) *Power and the curriculum: issues in curriculum studies*. Duffield: Nafferton Books.
- Richards, C. (ed) (1982) *New directions in primary education*. London: Falmer Press.
- Ridgeway, L. (1977) *The task of the teacher in the primary school*. London: Ward Lock Educational.
- Rintoul, K.A.P. and Thorne K.P.C. (1975) *Open plan organisation in the primary school*. London: Ward Lock Education.
- Rochdale, Metropolitan Borough of (1978) *Record-keeping in primary schools*. Rochdale: Borough Education Department.
- Rogers, C.R. (1967) *On becoming a person: a therapist's view of psychotherapy*. London: Constable.
- Rogers, G. (1982) *A social psychology of schooling*. London: Routledge and Kegan Paul.
- Rosenthal, R. and Jacobsen, L. (1968) *Pygmalion in the classroom: teacher expectation and pupils' intellectual development*. New York: Holt Rinehart and Winston.
- Ross, A.M. (1960) *The education of childhood*. London: Harrap.
- Ross, A., Russell, A. and Badcock, E. (1975) *The curriculum in the middle years*. Schools Council Working paper No.55. London: Evans/Methuen Education.
- Rowan, J. (1976) *Ordinary ecstasy: humanistic psychology in action*. London: Routledge and Kegan Paul.
- Rowlands, S. (1982) "Progressive education: a reformulation from close observation of children", in Richards, C. (ed) *New directions in primary education*. London: Falmer Press.

- Rubinstein, D. and Simon, B. (1973) *The evolution of the comprehensive school, 1926-1972*. London: Routledge and Kegan Paul.
- Rubinstein, D. and Stoneman, C.(eds) (1970) *Education for democracy*. London: Penguin Education Special.
- Ruddock, J. (ed) (1982) *Teachers in partnership: four case studies of in-service partnership*. London: Longman for the Schools Council.
- Ruddock, J. and Kelly, P. (1970) *The dissemination of curriculum development*. Slough: National Foundation for Educational Research.
- Rumelhart, D.F. and Ortony, A. (1977) "The representation of knowledge in memory", in Anderson, R.G. *et al.* (ed) *Schooling and the acquisition of knowledge*. Hillsdale, N.J.: Erlbaum.
- Rusk, R.R. (1933) *A history of infant education*. London: University of London Press.
- Salford, City of (1977) *The primary school profile*. Salford: City Education Department.
- Salford, City of (1982) *Profile '82*. Salford: City Education Department.
- Satterley, D. (1981) *Assessment in schools*. Oxford: Basil Blackwell.
- Schonell, F.J. and Schonell, F.E. (1963) *Diagnostic and attainment testing*. London: Oliver and Boyd.
- Schools Council (1971a) *Working paper No.39: Social studies 8-13*. London: Evans/Methuen.
- Schools Council (1971b) *Working paper No.37: Physical education 8-13*. London: Evans/Methuen.
- Schools Council (1972a) *With objectives in mind: a guide to science 5-13*. London: MacDonald Educational for the Schools Council.



- Schools Council (1972b) Pamphlet 14 *Dissemination and in-service training*. Report of the working party on dissemination. London: Evans/Methuen.
- Schools Council (1973a) *Working paper No. 48: Environmental studies 5-13: the use of historical resources*. London: Evans/Methuen.
- Schools Council (1973b) *Concept 7-9 materials: Unit 1 - Listening with understanding; Unit 2 - Concept building; Unit 3 - Communication; Unit 4 - The dialect kit*. Leeds: E.J. Arnold.
- Schools Council (1974) *Children's growth through creative experience: art and craft education 8-13*. London: Van Nostrand Reinhold.
- Schools Council (1975) *Development of scientific and mathematical concepts in children between the ages of 7-11*. London: Evans/Methuen.
- Schools Council (1983) *Working paper No. 75: primary practice: a sequel to "The practical curriculum"*. London: Methuen Educational.
- Science Research Associates (1972) *Listening Laboratory Series*. Chicago: Science Research Associates.
- Scottish Council for Research in Education (1977) *Pupils in profile: making the most of teachers' knowledge of pupils*. Edinburgh: The Council.
- Seaborne, M. (1971) *Primary school design*. London: Routledge and Kegan Paul.
- Seaborne, M. and Lowe, R. (1977) *The English school: its architecture and organisation*. London: Routledge and Kegan Paul.
- Seefeldt, C. (1973) "Open spaces, closed learning?", *Educational Leadership*, Vol 30, No. 4, pp.355-7.
- Selleck, J. (1972) *English primary education and progressives 1914-1939*. London: Routledge and Kegan Paul.

- Sharp, P.R. and Green, A. (1975) *Education and social control: a study in progressive primary education*. London: Routledge and Kegan Paul.
- Silberman, C.E. (1970) *Crisis in the classroom: the remaking of American education*. New York: Random House.
- Silverman, M.L. (ed) (1972) *The psychology of open teaching and learning: an enquiry approach*. Boston, Mass: Little, Brown and Company.
- Simon, B. and Willcocks, J. (eds) (1981) *Research and practice in the primary classroom*. London: Routledge and Kegan Paul.
- Simons, H. (ed) (1980) *Towards a science of the singular*. East Anglia: University of East Anglia.
- Smetheram, D. (1980) *Pedagogues as perquisitors or to put it another way: teachers as researchers*. Paper presented at Teachers in research seminar. Edinburgh: Scottish Council for Research in Education.
- Society of Education Officers (1975) *Management in the education service: challenge and response*. London: Routledge and Kegan Paul.
- Sockett, H. (1976) "Teacher accountability", *Proceedings of the Philosophy of Education Society of Great Britain*, Vol.10. Oxford: Basil Blackwell.
- Sockett, H. (1980) *Accountability in the English educational system*. London: Hodder and Stoughton.
- Spodek, B. and Walberg, H.J. (eds) (1975) *Studies in open education*. New York: Agathan Press.
- Stabler, T. (1978) *Drama in primary schools*. Schools Council Drama 5-11 project. London: MacMillan Education.
- Stake, R. (1973) "Measuring what learners learn (with a special look at performance contracting)", in House, E. (ed) *School evaluation*. Berkeley, Cal: McCutchan. pp.193-223.



- Stand, A.G. and Husen, T. (eds) (1962) *Child education*. Copenhagen: Munksgaard.
- Steadman, S. (1976) "Techniques of evaluation", in Tawney, D. (ed) *Curriculum evaluation today*. London: Macmillan Education.
- Steadman, S., Parsons, C. and Salter, B. (1978) *Impact and take-up project: an inquiry into the impact and take-up of Schools Council funded activities*. London: Schools Council publications.
- Stenhouse, L. (1971a) "The humanities curriculum project: the rationale", *Theory into practice*, Vol. 10, pp.154-162.
- Stenhouse, L. (1971b) "Pupils into students", *Dialogue*, No. 5, pp. 10-12.
- Stenhouse, L. (1975) *An introduction to curriculum research and development*. London: Heinemann.
- Stenhouse, L. (1980a) *Curriculum research and development in action*. London: Heinemann.
- Stenhouse, L. (1980b) "The study of samples and the study of cases", *British Educational Research Journal*, Vol.6, No. 1, pp 1-6.
- Stephens, L.S. (1974) *The teacher's guide to open education*. New York: Holt, Rinehart and Winston.
- Stevenson, H.W., Parker, T., Wilkinson, A., Hegron, A. and Fish, E.F. (1976) "Predictive value of teachers' ratings of young children", *Journal of Educational Psychology*, Vol .68, pp.377-400.
- Stott, D.H. and Marsden, H.C. (1970) *Bristol social adjustment guides*. London: Hodden and Stoughton.
- Stott, D. (1978) *Helping children with learning difficulties: a diagnostic teaching approach*. London: Ward Lock Educational.

- Stubbs, M. and Delamont, S. (eds) (1976) *Explorations in classroom observation*. London: John Wiley.
- Sturges, D. (1970) "Happenings in a primary school", in Rubinstein, D. and Stoneman, C. (eds) *Education for democracy*. London: Penguin Education Special pp. 97-103.
- Sumner, R. and Robertson, T.S. (1977) *Criterion referenced measurement and criterion referenced tests: some published work reviewed*. Slough: National Foundation for Education Research.
- Sutcliffe, J. and Whitfield, R. (1976) "Decision-making in the classroom: an initial report", *Research Intelligence*, Vol.2, No.1, pp.14-19.
- Sylvester, D. (1974) *Robert Lowe and education*. London: Cambridge University Press.
- Taba, H. (1962) *Curriculum development: theory and practice*. New York: Harcourt, Brace and World.
- Tawney, D. (ed) (1976) *Curriculum evaluation today: trends and implications*. Schools Council Research Studies. London: Macmillan Education.
- Taylor, J. and Ingleby, T. (1974) *The Scope storybook*. Schools Council Project in English for immigrant children. London: Longmans for Schools Council.
- Taylor, M. (1976) "Teachers' perceptions of their pupils", *Research in Education*, No.16, pp. 25-35.
- Taylor, P.H. (1962) "Children's evaluations of the characteristics of a good teacher", *British Journal of Educational Psychology*, Vol.32, pp.258 - 266.
- Tempest, N.R. (1974) *Teaching clever children, 7 -11*. London: Routledge and Kegan Paul.
- Thorndike, R. and Hagen, E. (1969) *Measurement and evaluation in psychology and education*. London: Chapman and Hall.



- Torbett, W. (1981) "Why educational research has been so uneducational", in Reason, P and Rowan, J. (eds) *Human inquiry: a sourcebook of new paradigm research*. Chichester: Wiley.
- Torrance, E.P. (1967) "Non-test ways of identifying the creatively gifted", in Gowan, J.C. et al. (eds) *Creativity: its educational implications*. New York: Wiley.
- Tough, J. (1976) *Listening to children talking: a guide to the appraisal of children's use of language*. London: Ward Lock Educational for the Schools Council.
- Tough, J. (1977) *Talking and learning: a guide to fostering communication skills in nursery and infant schools*. London: Ward Lock Educational for Schools Council.
- Tough, J. (1979) *Talk for teaching and learning*. London: Ward Lock Educational.
- Tyler, R.W. (1949) *Basic principles of curriculum and instruction*. Chicago: University of Chicago Press.
- Vernon, P.E. (1957) *Secondary school selection*. Leicester: British Psychological Society.
- Verrier, R. (1981) *School-based in-service education: a case-study of school consultancy relationships*. Sussex: University of Sussex Education Area Occasional paper No.8.
- Vincent, D., Powney, J., Green, L. and Francis, J. (1977) *Review of tests and assessment techniques for reading and language*. Slough: National Foundation for Educational Research Nelson.
- Walburgh, J. and Thomas S.C. (1971) *Characteristics of open education: towards an operational definition*. Newton, Mass: Educational Development Centre.
- Walker, D.F. (1973) "What curriculum research?" *Journal of Curriculum Studies*, Vol.5, No 1, pp.58-72.
- Warburton, (1964) "Attainment and the school environment", in Wiseman, S. (ed) *Education and environment*. Manchester: Manchester University Press. pp.101-124.

- Watkins, R. (1972) *Curriculum development and teacher development - a suggestion*. Paper submitted to Schools Council working Party on dissemination. London: Schools Council Publication.
- Watkins, R. (1973) *In-service training: structure and content*. London: Ward Lock.
- West, C. and Anderson, T. (1976) "The question of preponderant causation in teacher expectancy research", *Review of Educational Research*, Vol. 46, pp.613-630.
- Whitbread, N. (1972) *The evolution of the nursery-infant school*. London: Routledge and Kegan Paul.
- Whitehead, D. (1980) *The dissemination of educational innovations*. London: Hodder and Stoughton.
- Whitehead, J. (1977) *The process of improving education within schools*. Conference proceedings British Educational Research Association, Nottingham.
- Whitehead, J. (1980) "In-service education: the knowledge base of educational theory", *British Journal of In-service Education*, Vol. 6, No.2.
- Wilkinson, A., Stratta, L. and Dudley, P. (1976) *Learning through Listening*. Schools Council Oracy Project Listening Comprehension Tapes. London: Macmillan Education.
- Willcocks, J. and Jasman, A. (1980) "Study skills and pupil performance", in Galton, M. and Simon, B. (eds) *Progress and performance in the primary classroom*. London: Routledge and Kegan Paul. pp.99-140.
- Wiseman, S. (1964) *Education and environment*. Manchester: University Press.
- Wiseman, S. (1967) "Education and environment", in Craft, M. *et al.*, *Linking home and school*. London: Longmans, Green & Co. Ltd. pp. 107-122.
- Wolfson, B.J. (1968) "Pupil and teacher roles in individualised instruction", *Elementary School Journal*, Vol.68, pp.357-366.



- Wood, R. and Naphali, W.A. (1975) "Assessment in the classroom: what do teachers look for?", *Educational Studies*, Vol. 1, No.3. pp 151-161.
- Wood, R. and Skurnik, L.S. (eds) (1969) *Item banking*. Slough: National Foundation of Educational Research.
- Woods, P. (ed) (1979) *Teacher strategies: explorations in the sociology of the school*. London: Croom Helm.
- Worrall, N. and Cowan, R. (1983) "Situational biases in teacher impressions", *British Journal of Educational Psychology*, Vol. 53, pp. 249-252.