AN INVESTIGATION INTO THE THEORY AND PRACTICE OF FORMATIVE ASSESSMENT IN KEY STAGE 3 GEOGRAPHY

by

Yonca Tiknaz

Thesis submitted to School of Education
University of Leicester, for the degree of
Doctor of Education (EdD)
December, 2003
ABSTRACT

Formative assessment has been identified in the School Improvement literature as an intervention which could significantly improve teaching and learning in the classroom and raise standards of student attainment for students of all ages (Black and William, 1998a). In order for this to happen, there is a need for an adequate conceptualisation of the process of formative assessment which seeks to identify the key components and their interrelationships.

The Literature Review of this thesis demonstrates that the current understanding of formative assessment is focused on three key components, namely, assessment tasks, teacher feedback and pupil self-assessment. The research studies as well as the major reviews in the area, mainly focused on teacher feedback, and pupil self and peer-assessment. The component which is given less emphasis in the literature, is the nature of assessment tasks and the identification of criteria for devising it to facilitate formative assessment.

The research for this thesis mainly investigates how geography Key Stage 3 teachers understand and use formative assessment in their practices. Specific emphasis is given to understanding the criteria that teachers employ when they devise tasks. To deepen the understanding of formative assessment, two selected classrooms were observed and pupils’ views were gained relating to assessment issues. Research took the form of 12 in-depth interviews with heads of geography departments, 14 hours of recorded classroom observations and 20 Key Stage 3 pupil interviews.

The research indicates that teachers’ assessment task design was mainly informed by the geography level descriptions but their incorporation into practice was problematic due to the difficulties of interpretation. Target setting appeared to be powerful, and was the commonest strategy for communicating to pupils what they needed to do to improve their learning and thereby facilitate formative assessment. Teachers and pupils acknowledged the benefits of pupil self and peer-assessment. Teachers had varied ways of facilitating pupil self-assessment, however, pupil-peer assessment was quite rare.

The predominant challenge for teachers appeared to be the planning for progression in pupils’ learning, which required a conceptual understanding of the nature of progression in Geography KS 3 and the ability to interpret level descriptions. The majority of teachers felt that they needed external support to facilitate progression more effectively.
ACKNOWLEDGEMENTS

I would like to express my gratitude to my supervisors Alan Sutton and Dr. Marlene Morrison for their ongoing professional and personal support for the completion of this thesis. Their valuable advice, encouragement and inspiration is greatly appreciated.

I would like to thank to the Turkish Ministry of Education (MEB) which gave me the opportunity to study in the UK through the funding of my studies.

I also wish to express my special thanks to the twelve geography teachers and twenty pupils who shared their views and made this research possible. Without their willingness to participate this study would have not been possible.

I would like to thank Jonathan Westgate, other academic and administrative staff at the School of Educating for their support over the years.

Special thanks to my sister for her support in the last stages of my studies.

Particularly, I would like to thank to Peter Crew who supported me over the years endlessly.

Yonca Tiknaz
CHAPTER 1: INTRODUCTION ................................................................. 1

SECTION 1: A CHANGE OF EMPHASIS TO ACCOMMODATE ASSESSMENT FOR LEARNING............................................... 2
  Conceptualising Diagnostic and Formative Assessment ..................................... 4
  The Components of Formative Assessment ..................................................... 6
  Assessment Tasks ......................................................................................... 6
  Pupil Self-Assessment .................................................................................. 7
  The Role of Teachers in Working with Pupils in Formative Assessment .......... 7

SECTION 2: ASSESSMENT IN THE NATIONAL CURRICULUM........................................ 8
  The Recommendations of the Task Group on Assessment and Testing (TGAT) ................................................................. 8
  Post TGAT ................................................................................................. 10

SECTION 3: THE AIMS OF THE PRESENT RESEARCH........................................ 11
  The Aims of Teacher Interviews .................................................................. 11
  The Aims of Classroom Observations .......................................................... 12
  The Aims of Pupil Interviews ...................................................................... 13

SECTION 4: THE SIGNIFICANCE OF THIS THESIS........................................... 14
  The Format of the Thesis ........................................................................... 15

CHAPTER 2: LITERATURE REVIEW........................................................................ 17

SECTION 1: ALTERNATIVE MOVEMENTS IN ASSESSMENT FOR LEARNING................. 17
  Dynamic Assessment .................................................................................. 18
  Performance Assessment ......................................................................... 20
  Authentic Assessment ................................................................................ 21
  Conceptualising Formative Assessment ...................................................... 22

SECTION 2: FORMATIVE ASSESSMENT AND LEARNING THEORIES.................. 23
  Behavioural Learning Theory and Formative Assessment ......................... 24
  Constructivist Learning and Formative Assessment .................................... 25
  Formative Assessment as a Socio-Cultural Process .................................... 27

SECTION 3: KEY COMPONENTS OF FORMATIVE ASSESSMENT.......................... 29
  Teacher Feedback ...................................................................................... 29
  Pupil Self-Assessment .............................................................................. 31
  Assessment Task Design ......................................................................... 34
  Task Design Criteria for Alternative Approaches to Assessment ................. 35
    Task Design for Authentic Assessment .................................................... 36
    Task Design for Performance Assessment ............................................... 40
  Summary of Alternative Approaches to Assessment for Learning .............. 41

SECTION 4: FORMATIVE ASSESSMENT IN PRACTICE........................................ 44
  Review of Characteristics of Effective Teacher Feedback ......................... 45
    Specificity of Feedback: General or Focused ........................................... 45
    Timing of Feedback ................................................................................. 46
    Focus of Feedback: Task Referenced or Self-Referenced ......................... 46
  Research Studies on Pupil Self-Assessment ............................................. 47
    Quantitative Studies of Pupil Self-Assessment ........................................ 47
    Qualitative Studies of Pupil Self-Assessment .......................................... 50
  Research Studies on the Use of Specific Tools of Pupil Self-Assessment ........ 51
    Recording and Profiling .......................................................................... 51
    Portfolios and Processfolios ................................................................. 53
ABBREVIATIONS

ACCAC: Qualifications Curriculum and Assessment Authority for Wales
ARF: Assessment Reform Group
CBA: Curriculum Based Assessment
CBM: Curriculum Based Measurement
FE: Further Education
GA: Geographical Association
GCSE: General Certificate of Secondary Education
GNC: Geography National Curriculum
ICT: Information and Communications Technology
IQ: Intelligence Quotient
LEA: Local Educational Authority
OBS.: Observation
PGCE: Post Graduate Certificate in Education
POS: Programmes of Study
QCA: Qualifications and Curriculum Authority
SAT: Standard Assessment Task
SCAA: School Curriculum and Assessment Authority
SOAs: Standards of Attainment
T: Teacher
TA: Teacher Assessment
TGAT: Task Group on Assessment and Testing
TRF: Tropical Rain Forest
UK: United Kingdom
US: United States
ZPD: Zone of Proximal Development
CHAPTER 1: INTRODUCTION

Formative assessment or assessment for learning, as it has been recently called in the United Kingdom (UK), has been identified in the School Improvement literature as an intervention which could significantly improve teaching and learning in the classroom and raise standards of attainment for students of all ages (Black and Wiliam, 1998a). In order for this to happen, there is a need for an adequate conceptualisation of the process of formative assessment which seeks to identify the key components and their relationship with each other. Given the relative absence of classroom research on formative assessment and certainly in Geography, there is an urgent need to identify what formative assessment actually looks like in classrooms. This research focuses on how Key Stage 3 geography teachers understand and use formative assessment in their classrooms.

This chapter begins by charting the increasing interest in assessment for learning which has been associated with calls for a shift in emphasis away from summative assessment. This is followed by an attempt to conceptualise formative assessment by constructing a model which identifies the main components of the process of formative assessment and their relationships with each other. Each of the main components is then considered briefly with reference to the research literature. (A more detailed examination of each of these components will follow in the Literature Review chapter.) These components of the formative assessment process are assessment tasks, self-assessment by pupils on their performances, peer-assessment by pupils and the role of teachers in providing feedback and feedforward to pupils.

The next section provides essential contextual evidence on the Government’s policy on assessment in the National Curriculum in the UK in England and Wales. It shows how strong support for formative assessment (assessment for learning) by the Task Group on Assessment and Testing (TGAT), including claims that ‘formative and summative assessment were potentially compatible,’ was overshadowed by the Government’s determination to place much more emphasis on the summative assessment of learning as a way of driving up standards of
attainment. The chapter concludes with an explanation of reasons for choosing the research topic, and provides the main research objectives.

SECTION 1: A CHANGE OF EMPHASIS TO ACCOMMODATE ASSESSMENT FOR LEARNING

Summative assessment which can be characterised as the assessment of learning, refers to the use of assessment tasks, often externally devised tests, at the end of modules or key stages which serve a number of purposes, most importantly accountability and selection. The latter is a 'high stakes' purpose because it has an important bearing on pupils' future prospects for employment and entry into further and higher education. In this process of selection, individuals are assigned to particular institutions or educational programmes on the basis of their measured competence (Resnick and Resnick, 1992).

Assessment information is also used for purposes of accountability. The origins of the accountability functions of assessment are strongly rooted in the mass educational systems and have a concern for providing information to monitor school systems as a whole (Resnick and Resnick, 1992). Historically, as schooling became a mass undertaking, a need arose for the public to know how well this vast system was performing (ibid). This allows comparisons between institutions and also enables checks to be made on whether a particular institution is meeting its academic targets in relation to others.

Similar to selection / certification functions, the audience who uses assessment results for gauging accountability are positioned at some distance from the day to day educational practice. Further, the end results of this assessment do not reflect in detail on the performances of individual students. Rather, the focus in accountability is on how target 'groups' of students, considered as a population experiencing a particular program or instruction, are performing (Resnick and Resnick, 1992; Scott, 1994).

These two functions of assessment, accountability and selection have tended to dominate the educational systems of countries in the developed world for several
years. However, Gipps (1994) perceives that a major change is underway with assessment undergoing a paradigm shift to a broader model of assessment, from a testing and examination culture to an assessment culture. Some academics, such as Harlen et al. (1992), Gipps (1994), Murphy and Broadfoot (1996) have been critical of the excessive emphasis on summative assessment and have claimed it is unreliable, invalid and has ‘back-wash effects’, a term coined by Biggs (1993) providing negative feedback on teaching and learning, (teaching to the test) and on the curriculum and teachers’ professionalism. Wiliam (2003) has been particularly critical of the use of external tests for the core subjects in the National Curriculum, claiming that they are invalid and unreliable. Increasingly, these academics have called for much more emphasis to be placed on improving classroom based assessment to improve the quality of teaching and learning as a way of raising standards of attainment.

The term formative assessment is relatively new and the understanding of the term has been developing through the 1990s. Burgeoning levels of understanding has not being helped by the fact that other terms have been used particularly in the United States of America (USA) which implied similar meanings. These terms include classroom evaluation, curriculum based assessment (CBA) and curriculum based measurement (CBM). Development work on formative assessment has been preceded and paralleled by other movements in the area of assessment for learning, notably dynamic, performance, and authentic assessment. Each of these has different emphases and all have implications for assessment for learning in classrooms. The research literature on each of these traditions is extensive. Resnick and Resnick (1992), McNamara (1997), Stiggins (1994) and Meyer (1994), Broadfoot (1995) have written on performance assessment which is concerned with making direct assessment of pupils’ progress in the classroom. In dynamic assessment, Minick (1987), Lidz (1995) and Losardo and Syverson (2001) draw on the work of Vygotsky to emphasise the importance of interactions between the teacher and the pupils and the pupils themselves in the assessment process. In authentic assessment, Archbald and Newman (1988), Zessoules and Gardner (1991), Gardner (1992) and Eisner (1993) stress the importance of linking assessment to real world contexts to give relevance to the learning and prepare young people for adult life.
Interest in formative assessment in the UK was given a major boost by the publication of an extensive research survey on formative assessment by Black and Wiliam (1998a). This drew on a large number of quantitative and qualitative studies and confirmed the potential of formative assessment to produce significant and often substantial learning gains, in particular for the lower attaining pupils more than the rest. The effect size of using formative assessment as an intervention is claimed to be significantly high. However, Black and Wiliam (1998b) have identified a number of significant weaknesses in classroom assessment practice which need to be tackled if the potential of formative assessment is to be realised. They claim that some teachers are insufficiently aware of the nature of formative assessment and what they need to improve. Three main categories of weaknesses are identified: those which relate to the quality of assessment tasks on which the process of feedback is based; those which relate to the grading function as part of the process of feedback; those which relate to a failure to actively involve pupils in their own assessment in order to help them become increasingly more responsible for their own learning (Black and Wiliam, 1998b).

**Conceptualising Diagnostic and Formative Assessment**

The thesis defines formative assessment as a process which consists of a sequence of three interrelated components. The starting point is the perception by the teacher or the pupil of the existence of a learning gap between current and desired level of knowledge, understanding or skill (Black and Wiliam 1998:20). This is achieved by teachers devising *assessment tasks* to reveal the current learning evidence of pupils’ learning. Then, the teacher interprets and communicates this learning gap to pupil thorough *feedback* and *feedforward*. Here, pupils are also actively take responsibility for generating this information through *self and peer-assessment*. A key to this definition is that “assessment becomes formative when the evidence is actually used to adopt the teaching to meet learning needs (Black, 2003:3).

Research literature on formative assessment is relatively recent. Sadler (1989) was an early advocate of the use of formative assessment in classrooms to assist learning and made a significant contribution to its definition and conceptualisation.
along with Ramaprasad (1983) and more recently Black and Wiliam (1998a). Sadler described the purpose of formative assessment:

"Formative assessment is concerned with how judgements about the quality of student responses (performance, pieces, or works) can be used to shape and improve the student's competence by short-circuiting the randomness and inefficiency of trial and error learning" (Sadler, 1989:120).

Appropriate assessment practices can be used to diagnose the performance characteristics of pupils (diagnostic function) and to use this information to encourage the future development of the student (formative function).

To aid the conceptualisation of formative assessment a model is used which demonstrates the relationships between the component parts (Figure 1 below).

Formative assessment refers to the feeding back of information by the teacher to the pupil about his or her achievement on an assessment task to help him / her to improve, information which it is intended should be acted on by the pupil. This feedback information is intended to close the gap between the present level of learning (knowledge, understanding and skill development) of the pupil and the desired level. The teacher's responsibility is to recognise, interpret and translate ideas about the nature of the gap to the pupil (Black and Wiliam, 1998a). This feedback is only truly formative if it is used to alter the magnitude of the gap (Ramaprasad, 1983). If the information about the gap is just recorded and summarised, then the action cannot be formative (Sadler, 1989). Feedback can be written or oral; written feedback can be in the form of marks or grades and/or
comments. Sadler (1989) and Black and Wiliam (1998a) also claim that pupils have a key role in assessing their own progress which requires that they are aware of the gap between their present level of learning and the desired level through a process of self-assessment; pupils must also be prepared to act to close the learning gap.

The Components of Formative Assessment

Assessment Tasks

The process of formative assessment is enacted through a series of assessment tasks. A review of assessment literature suggests that assessment tasks have a paramount effect in shaping the quality of diagnostic and formative assessment (Bachor and Anderson (1994), Bell and Cowie (2001), Crooks (2001) and Black (1993). Formative assessment takes place when teachers devise assessment tasks in order to understand pupils’ strengths and weaknesses as a way of moving them forward in their learning (Black and Wiliam, 1998a). Black and Wiliam claim that there are weaknesses in the quality of assessment tasks on which the quality of feedback is based. They promote rote and superficial learning despite their intention to develop deep learning. Although the assessment tasks are crucial for the effective implementation of formative assessment, relatively little empirical research in the classroom has been carried out to investigate the design of assessment tasks to aid conceptualisation and improvement of practice in the classroom.

Research in this area has focussed on criteria for devising and appraising assessment tasks for authentic assessment (Eisner, 1993), on the accessibility of assessment tasks for different pupils, Murphy et al. (1995), on the effectiveness of different tasks formats in revealing the depth of pupils’ understanding (Larcher and Dumas-Carre, 1987; Caygill et al., 2001; Eley and Cygill, 2001). There are a few published studies of diagnostic and formative assessment in geography classrooms that refer directly to the nature of assessment tasks to support formative assessment (Leat and McGrane, 2000). Publications by the Geographical Association (GA) provide examples of planning for assessment for the short, medium and long term.
for Key Stage 3 (Hopkin, 2000; Howes, 2000b; Howes, 2000c; Keith and Flenders, 2000) as well as how level descriptions are used (Butt et al., 1995). One study by Hamson and Sutton (2000) identifies some criteria for the planning of a series of assessment tasks in Key Stage 3 Geography. Another, by Wood and Sutton (2002) examines the use of assessment tasks in a decision making exercise to make an initial diagnostic assessment which is then used for formative purposes in the teaching of A level Geography.

**Pupil Self-Assessment**

The role of pupils in their assessment process is a key component of formative assessment (Daws and Singh, 1996; Black and Wiliam, 1998a). This involves pupils reflecting upon their learning by evaluating their strengths and weaknesses (Klenowski, 1995). In order to assess their own work, pupils need to develop an understanding of what represents a quality performance for a specific task (Sadler, 1989). Self and peer-assessment during which other pupils’ performance is assessed, are essential components of formative assessment (Freeman and Lewis, 1998).

**The Role of Teachers in Working with Pupils in Formative Assessment**

Teachers have a key role in preparing their pupils to undertake informal assessment tasks in classrooms. Several research studies on scaffolding, mediated learning and expert tutoring (Feuerstein et al., 1979; Brown and Ferrara, 1985; Tharp and Gallimore, 1991) in dynamic assessment have referred to the importance of scaffolding by the teacher to enable pupils to complete an assessment task. Some have identified the importance of supporting pupils to complete assessment tasks by steadily releasing teacher control (Tharp and Gallimore, 1991). Wood and Wood (1996) have drawn attention to the need for teachers to use ‘contingency teaching’ which refers to the use of an appropriate level of support to complete the task. Bliss et al. (1996) have drawn attention to the problems created when there is an absence of conditions to assist scaffolding. When pupils have completed tasks, teacher feedback should be specific and include suggestions on how pupils can improve their learning. The effectiveness
of feedback depends on its planning and implementation (Black and Wiliam (1998a). Target setting is also an important process of formative assessment as a way to support pupils’ learning by indicating to them the next steps, for which they should aim.

SECTION 2: ASSESSMENT IN THE NATIONAL CURRICULUM

The Recommendations of the Task Group on Assessment and Testing (TGAT)

The National Curriculum in England and Wales was primarily a mechanism for raising standards of pupil attainment. This was to be achieved by the creation of centrally devised Programmes of Study (POS) and Standards of Attainment (SOAs) for the core and foundation subjects of which Geography was one when the National Curriculum was first introduced in the late 1980s. The responsibility for planning the arrangements for assessment fell to the Task Group on Assessment and Testing (TGAT) which was chaired by Professor Paul Black.

The recommendations of the TGAT report (1988) were based on a view that assessment for learning (the diagnostic and formative assessment) should play a key role in the National Curriculum. Assessment was perceived to be an integral part of the planning of teaching and learning.

"The assessment process itself should not determine what is to be taught and learned. It should be the servant, not the master of the curriculum. Yet it should not simply be a bolt on – addition at the end." TGAT Report, 1988 para. 4

The Report, (paragraph 5) also identified four main guiding principles for the design of the assessment provision:

- The assessment results should give direct information about pupils’ achievement in relation to the objectives: they should be ‘criterion referenced’.

- The results should provide a basis for decisions about pupils’ learning needs: they should be ‘formative’.
The scales or grades should be capable of comparisons across classes and schools, if teachers, pupils and parents are to share a common language and common standards, so assessment should be ‘calibrated’ or ‘moderated’.

The ways in which criteria and scales are set up and used should relate to expected routes of educational development, giving some continuity to pupils’ assessment at different ages: the assessment should relate to progression.

In a separate paragraph, the report came out strongly in favour of formative assessment and introduced the terms “feedback and feedforward”.

“Promoting children’s learning is a principal aim of schools. Assessment lies at the heart of this process... It should be an integral part of the educational process, continuously providing both feedback and feedforward. It needs therefore to be incorporated systematically into teaching practices and strategies”. TGAT Report, 1988, para. 4

TGAT (paragraph 25) also recognised the potential conflict between the formative and summative functions of assessment but suggested that the two functions could be reconciled. It suggested that a complete picture of pupil performance could be achieved through a multi-purpose system and that several purposes could be served by

“combining in various ways the findings of assessments designed primarily for a different purpose. It is possible to build up a comprehensive picture of the overall achievements of a pupil by aggregating, in a structured way, the separate results of a set of assessments designed to serve formative purposes”. TGAT Report, 1988, para. 25.

This statement provided a steer for schools on a possible direction they could follow in planning their arrangements for assessment. However, as Black and Wiliam (1998a) were to lament, practical guidance from government agencies on the implementation of formative assessment was lacking. Until the recent introduction of the National Strategy for Key Stage 3, little government support has been given to improving the quality of diagnostic and formative assessment and expenditure has been minimal in comparison to that on summative assessment (Daugherty, 1995).
The assessment arrangements which eventually emerged departed from the TGAT recommendations in the emphasis which was placed on summative assessment for the purposes of accountability. This was hardly surprising, given the Government’s interest in the publication of school assessment data in league tables to drive up standards and the general distrust of the reliability of teacher assessment. Pupils were to take external tests at the ages of 7, 11, 14 and 16 in the core subjects and teachers were to carry out their own teacher assessment (TA) in the foundation subjects which included Geography. In the first version of the National Curriculum (1988), pupils’ achievement was assessed on the basis of a series of attainment targets and Statements of Attainment on a 10 level scale between the ages of 5 to 16. In Geography, there were initially five attainment targets and 183 separate statements of attainment in Key Stage 3.

Teacher opposition to the National Curriculum developed quickly, partly as a result of the arrangements for assessment. A particular problem was the SOAs which were deemed to be unmanageable. Not only were there an excessive number but they were content driven and varied in their detail which made the devising of assessment tasks and the recording of achievement difficult.

In order to solve the problem, in April 1993, the Secretary of State for Education invited Sir Ron Dearing to undertake a review of the National Curriculum. The School Curriculum and Assessment Authority (1994) (SCAA) noted that several assessment issues needed attention:

- The future of the ten level scale;
- The simplification of the testing arrangements;
- The administration of the National Curriculum and the tests.

In the resulting report, the existing 10 level scale was retained. The number of standard national tests was restricted to the core subjects only. In Geography, in Key Stages 1, 2 and 3, the derivation of a final level for individual pupils was left as the responsibility of individual teachers. But the most important change was the
replacement of the Statements of Attainment (SOAs) by Level Descriptions which were to be used for best fit judgements of pupils’ ability at the end of a Key Stage. These are composite statements of attainment for each of 8 levels and an exceptional level which define progression in pupils’ learning between the ages of 5 and 14. In the latest version of the National Curriculum in Geography 2000, the 4 progression strands are: enquiry and skills, patterns and processes, places and environmental education, and sustainable development. The level descriptions which encapsulate these progression strands are not age specific. They provide a framework for planning for progression and differentiation in pupils’ achievements.

The latest statutory requirement in National Curriculum assessment has been for schools to set quantitative targets. As from 1998, schools were obliged to set targets for 14 and 16 year olds. From June 1997, schools were obliged to report the result of teacher led assessment. This data is publicised and became one of the strongest measures of judging schools’ success in the form of league tables. League tables led to quantitative target setting and both became a major tool for the Government to monitor the improvements in the standards of education.

Having described the different functions of assessment as well as the diverse requirements that it is supposed to meet, it becomes important to understand how assessment for learning (formative assessment) is understood and implemented by teachers within the current National Curriculum.

**SECTION 3: THE AIMS OF PRESENT RESEARCH**

The main aim of this study is to investigate how geography teachers in Key Stage 3 report on the planning and implementation of formative assessment in their classroom.

**Teacher Interviews**

The following sub questions are developed to deepen the understanding of their perceived practices of formative assessment.
1. How do teachers use a series of assessment tasks across the key stage?
   - What criteria do teachers use to devise assessment tasks?
   - Do teachers use the SCAA exemplar assessment tasks? Do they adapt them?
   - How frequently do teachers set assessment tasks over the key stage?
   - How do teachers differentiate the tasks?
   - Do teachers allocate time for pupil self and peer assessment when they plan for assessment tasks?

2. How do teachers use the results of the assessment tasks to inform pupils of their progress?
   - How do teachers make pupils aware of the standards by which they are being judged i.e. the use of level descriptions?
   - Do teachers use marks or grades and or general comments related to the level descriptions to feed back to their pupils?
   - Do teachers set pupils targets for improvement of their learning? What is the nature of these targets? Are they related to the level descriptions?

3. How do teachers use performance data from the results of the assessment tasks to decide on a final summative level for each pupil at the end of Key Stage 3?
   - How do teachers aggregate qualitative and quantitative data of pupils’ progress to make summative judgements level descriptions?

Classroom Observations

A subsidiary aim of this research is to describe formative assessment practices through observing one selected teacher over a period of time. The main research questions addressed by classroom observations are:

1. How does the teacher elicit information about pupils’ learning?
   - What strategies does he employ to support formative assessment?
   - What factors influence the elicitation of assessment evidence of pupils’ learning?
2. How does the teacher use feedback to communicate the weaknesses and strengths of pupils’ learning?
   - What strategies are potentially positive towards achieving formative assessment?

3. How does the teacher communicate the level descriptions to pupils?
   - Does the teacher mention the level descriptions?
   - At what stage and in what context are they mentioned in lessons?

4. How is self and peer-assessment used?
   - What tools are used to facilitate pupil self and peer-assessment?
   - What is the nature of the assessment criteria that are employed for pupil self and peer-assessment?

Pupil Interviews

A further aim of this study is to explore pupils’ views and explanations on a number of assessment issues which are mainly informed by the classroom observations. In order to achieve this aim, a series of subsidiary questions were identified:

1. What do pupils think about the design of the assessment tasks?
   - What criteria do pupils think make ‘good’ assessment tasks?
   - What type of assessment tasks do pupils prefer?

2. How do pupils interpret their teacher’s initiations in the classroom?

3. What do pupils think about assessing themselves and their peers?

4. How do pupils interpret and justify levels that are given on a particular assessment task?
   - What do the levels mean to pupils?
SECTION 4: THE SIGNIFICANCE OF THIS THESIS

Examination of the research literature shows that, historically, understanding of the nature of formative assessment has been unclear. Indeed, the conceptualisation of formative assessment has been evolving progressively and this seems to have occurred in parallel with other assessment for learning movements (dynamic, performance and authentic assessment). Important work by Black and Wiliam (1998a) has helped to crystallise thinking on formative assessment and to understand its key components. However, it is clear that much classroom practice in the use of formative assessment contains major weaknesses as a result of teachers having an inadequate understanding of the process of formative assessment (ibid). The current research aims to contribute to the conceptualisation by revealing geography Key Stage 3 geography teachers’ understanding of formative assessment.

In this investigation, formative assessment is conceptualised on the basis of a theoretical model which involves three key components (assessment tasks, teacher feedback and pupil self- and peer-assessment). As the Literature Review will demonstrate, the current understanding of formative assessment practices is mainly informed by major reviews which emphasise how pupils’ progress is communicated through teacher feedback (Black and Wiliam, 1998a). Although some research studies have investigated the ways in which pupils’ performance is affected by the nature of assessment tasks (Crooks, 1988; 2001, Bachor and Anderson, 1994), relatively little attention has been devoted to understanding how the nature of assessment tasks affect the other components of formative assessment. This thesis contributes to the closure of this gap by exploring the criteria that geography Key Stage 3 teachers use in their task design and how this links to other components of formative assessment.

Insight into the assessment task design practices of geography Key Stage 3 teachers is valuable in a number of ways. Firstly, there are few publications available to guide geography teachers in their task design. Secondly, in contrast to core subjects, which are accompanied by external summative tests, assessment in geography in Key Stage 3 is entirely dependent upon teacher assessment. Given
the lack of research studies on how teachers understand formative assessment in general, an understanding of how geography teachers devise their assessment tasks could contribute to the conceptualisation of formative assessment.

Research studies have identified that formative assessment is grounded in the socio-cultural contexts within which they are constructed. This relationship has been investigated in terms of how classroom dynamics and culture affect the formative assessment process (Bell and Cowie, 2001; Filler 1993; 1995). Earlier, Edward and Mercer (1987) investigated interactions between pupils and teachers as a three step sequence (teacher initiation, pupil response and teacher feedback). Having accepted the significance of social context in explaining formative assessment, the current research aims to illustrate instances of formative assessment by classroom observations and also through drawing on the teacher’s and pupils’ reflections. Such insight could assist in the conceptualisation of formative assessment by adding real classroom examples to illustrate the process of formative assessment as well as the dilemmas involved.

This introduction has provided information on the Geography National Curriculum and has investigated the scope for formative assessment within the current eight-level system that it promotes. Although Butt (1995), Butt et al. (1995) and Lambert and Digby (1996) provided suggestions to help geography Key Stage 3 teachers interpret the level descriptions, very few published studies are available which investigate the implications of the level descriptions for the process of formative assessment. This information has the potential to enrich the current conceptualisation of formative assessment by revealing how the current national framework shapes teachers’ thinking of their formative assessment practices.

The Format of the Thesis

This thesis consists of five chapters each being divided into sub-sections to organise the information provided. The following summarises the outline of the each chapter.
Chapter 1 is the introductory chapter. It introduces the key concepts of formative assessment and provides brief information on the National Curriculum assessment arrangements. This chapter also outlines the aims and objectives of the study.

Chapter 2 reviews the relevant literature on formative assessment. This Chapter provides information on alternative approaches to assessment for learning and explores the ways in which these movements have informed the current thinking on formative assessment. The chapter continues with information on the components of formative assessment (assessment tasks, teacher feedback and pupil self and peer-assessment) before exploring the Geography National Curriculum as a context for formative assessment.

Chapter 3 focuses on the philosophical underpinnings of this research before justifying the research methods selected. It also provides information on the design of research instruments, data collection and analysis procedures before addressing the issues of validity and reliability.

Chapter 4 provides the findings and analysis of this thesis. This chapter consists of three parts which report findings and analysis of teacher interviews, classroom observations, and pupil interviews respectively. The part on teacher interviews explores teachers' views on a number of issues relating to planning and implementing formative assessment. The part on classroom observations presents research data on the implementation of formative assessment in two different Key Stage 3 classrooms. The pupil interview part provides pupils' views and explanations on a number of selected assessment issues which were mainly identified through classroom observations.

Chapter 5 summarises the conclusions reached in this study. The chapter starts by discussing the main conclusions and then the limitations of this research study. A following sub-section is then dedicated to the implications of this research on policy and practice. Finally, there is a brief section on the directions for future research which stem from the current work.
CHAPTER 2: LITERATURE REVIEW

Introduction

Interest in formative assessment predates the Black and Wiliam (1998a) research. The difference is that whereas before, it was the concern of only a relatively small number of academic researchers, it is now being actively promoted in the UK by the Government as a way of improving teaching and learning in order to raise standards of pupil attainment. Teachers are being asked to review and develop their performance in this area. However, little empirical research on formative assessment in practice in classrooms has been carried out until recently and practical guidance has only recently become available with the launch of the National Strategy for Key Stage 3. Instead, the emphasis has been on summative assessment for the purpose of accountability to raise standards of pupils’ attainment.

Formative assessment first appeared in the research literature with the term formative evaluation by Scriven (1967) and Bloom et al. (1971) and its introduction was designed to improve the curriculum. Since then, successive attempts have been made to clarify the nature of formative assessment with major contributions by Sadler (1989), Ramaprasad (1983), Crooks (1988) and Natriello (1987) and more recently in the 1990s by Torrance (1993), Torrance and Pryor (1998), Bell and Cowie (2001), and Black and Wiliam (1998a). Preceding and in parallel with this research on formative assessment has been the promotion of other approaches to assessment for learning; these include dynamic, performance and authentic approaches to assessment. Research in these areas has carried on separately but the findings can assist in an improved conceptualisation of the nature of formative assessment.

This chapter consists of five sections. The chapter begins with a consideration of these alternative movements in assessment for learning, notably their origin, purpose and distinguishing characteristics. The second section attempts to conceptualise formative assessment by understanding the key components (teacher feedback, pupil self-assessment and assessment task design) with reference to theories of learning and theoretical perspectives on the alternative assessment for
learning movements. The third section examines research on the implementation of the separate components (teacher feedback and pupil self-assessment) of formative assessment before investigating the whole process of formative assessment in classrooms. The fourth section reviews published research on pupils’ views on a number of assessment related issues. The chapter concludes with a consideration of the arrangements for criterion referenced assessment in the National Curriculum in Key Stage 3 Geography which provides the context within which the process of formative assessment has to operate; in particular, it deals with the relationship between formative assessment and continuity, progression and differentiation.

SECTION 1: ALTERNATIVE MOVEMENTS IN ASSESSMENT FOR LEARNING

Dynamic Assessment

The motive for the development of dynamic assessment procedures was dissatisfaction with standardised tests which failed to provide information on how pupils learn (Lidz and Elliot, 2000). Its central purpose is to ascertain not only what pupils have already learned but their potential for further learning (Losardo and Syverson, 2001:119). Dynamic assessment is based on the work of Vygotsky and in particular his theory of the zone of proximal development (ZPD):

“If he is to fully evaluate the state of the child’s development, the psychologist must consider not only the actual level of development but the zone of proximal development” (Vygotsky, 1978, p.86)

“The difference between the level of solved tasks that can be performed with adult guidance and the level of independently solved tasks is the zone of proximal development” (Vygotsky, 1978:86)

These quotations suggest that social interaction is crucial to the promotion of learning; as teachers engage with the ZPD of the pupil, both become part of the learning process. In dynamic assessment, teachers interact directly with their pupils in the classroom (Lidz and Elliot, 2000:2). Teachers examine how pupils use their cognitive and emotional resources in response to social interactions, tasks and materials (Losardo and Syverson, 2001:120). They also examine pupils’ ability to
learn and change in response to the amount of scaffolding which is provided (Losardo and Syverson, 2001:120).

Two approaches to dynamic assessment have been identified, assessment of the ZPD and assessment in the ZPD. The former was designed to provide quantitative measures of a child’s learning potential (Budoff, 1974; Brown and French, 1979; Brown and Ferrara, 1985; Brown et al, 1992; Allal and Ducrey, 2000). In this approach, the assessment procedure consists of a pre-test/train/post-test sequence. The comparisons of performance before and after the training provides a measure of the change created by the assessment intervention process (Losardo and Syverson, 2001).

In contrast, Minick (1987:127) suggests that Vygotsky viewed the analysis of the ZPD not as a means of assessing a child’s learning potential or efficiency but as a means of gaining insights into the kinds of psychological processes that the child might be capable of in the next or proximal phase of development. In this second approach to dynamic assessment, assessment in the zone of proximal development, the primary concern is the in-depth diagnostic exploration of the child’s learning and appropriate intervention planning (Lidz, 1995). This intervention, in the form of scaffolding, is central to assist and guide pupils’ learning.

Tharp and Gallimore (1991) conceived scaffolding as consisting of four stages which (in ideal situations) represented a transition from assisted to unassisted pupil performance. Wood and Wood (1996) extended the definition of scaffolding and suggested the key responsibilities of teachers supporting the assessment process. Key elements include retaining the child’s engagements with the task, establishing and maintaining an orientation towards task-relevant goals, highlighting critical features of the task that the child might overlook, demonstrating how to achieve goals, and helping to control frustration. This definition is based on a key principle of dynamic assessment, that the amount of help provided by the teacher should be regulated according to the way in which pupils are coping (Wood and Wood, 1996: 7).
Bliss et al. (1996) explored teachers' scaffolding strategies and, in particular, the situations during which scaffolding is disabled. They suggested that scaffolding was precluded by the use of over-directive teaching strategies; when the pupils were left to tackle tasks without much help from the teacher; where the teachers engaged in interactions with the pupils but tended to ignore the pupils' existing knowledge; where the teacher and pupil had completely different lines of thinking; and where the teacher was unable to clarify their misconceptions because of time constraints.

**Performance Assessment**

During the 1990s, there was an increased interest particularly from academics in the United States (US), in defining educationally valued outcomes, in terms such as 'understanding' and 'higher order thinking skills' and in devising assessment tasks that have a clear relationship to these outcomes. This new emphasis became linked with a movement towards performance assessment in the US to which Popham (1994) and Linn (1994) made significant contributions. Its underlying motive is to define and assess what is really important in pupils' learning. The movement has shifted the emphasis of criterion referencing to accept that important educational outcomes may not be definable in behavioural ways.

Performance assessment covers tests and assessment tasks of very many different types. The tasks represent a broadening of the types of behaviours sampled by tests to include extended writing, oral and practical responses. Gipps (1994) noted that, in the US, the term performance assessment is often used to mean any form of assessment that does not use multiple-choice questions. Resnick and Resnick (1992) define performance assessment as that which makes judgments in a direct way on actual performance as opposed to conventional tests which consider indirect indicators of competence, as reflected in responses to artificially derived questions. A characteristic of both performance and authentic assessment is their reliance on human judgement and discretion in assessing pupils' work (Linn and Baker 1996). This could mean that teachers may devise their own criteria for assessing what constitutes a successful performance.
**Authentic Assessment**

The term authentic assessment was first recognised in the 1980s. The principal motive was the concern that traditional tests did not provide an accurate estimate of worthwhile knowledge, in particular they failed to assess the kinds of competences required in ‘real life’ or authentic situations beyond school (Archbald and Newmann, 1988). Authentic assessment involves a performance of some kind but it is not authentic unless it is extended into real life contexts (Zessoules and Gardner, 1991; Gardner, 1992; Johnson, 2002; Meyer, 1992). According to Meyer (1994:100), an additional important determining characteristic of authentic assessment is the locus of control. This refers to pupils taking the initiative by making decisions about the topic, the time allocated, and pacing.

Stimpson (1996) attempted to draw out some of the distinctive features of authenticity in assessment (see Table 2.1 below):

<table>
<thead>
<tr>
<th>Authentic Assessment</th>
<th>Non-authentic Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assessment of life related understanding in a ‘real world’ context</td>
<td>Assessment of largely discipline related knowledge or of issues in a subject discipline context</td>
</tr>
<tr>
<td>Strong ipsative element, pupil learning orientation</td>
<td>Largely normative and discipline/school/subject oriented</td>
</tr>
<tr>
<td>Over-riding concern with validity, particularly construct validity</td>
<td>Concern with validity (often mainly content validity in practice) but often dominated by issues of reliability</td>
</tr>
<tr>
<td>Wide range of classroom techniques used but an emphasis on those which are open ended</td>
<td>In general, a narrower range of techniques used with frequently an emphasis on more closed forms</td>
</tr>
<tr>
<td>An explicit desire to fully integrate learning and assessment</td>
<td>Learning and assessment as discrete activities</td>
</tr>
<tr>
<td>Large, detailed descriptive element in respect of pupil achievement</td>
<td>Less descriptive element but strong judgemental element</td>
</tr>
</tbody>
</table>

Having defined the alternative approaches to assessment for learning, the next section reviews the conceptualisation of formative assessment in relation to different theories of learning. This is followed by an exploration of the key components of formative assessment as teacher feedback, pupil self-assessment.
and assessment task design. The section concludes with an analysis of the contribution of alternatives approaches to formative assessment.

**Conceptualising Formative Assessment**

Since the production of the TGAT report, the research literature as reviewed by Crooks (1988) and particularly by Black and Wiliam (1998a) has added significantly to our understanding about the theory and practice of formative assessment. A variety of quantitative studies have established that formative assessment produces significant learning gains and in a few cases, the largest gains have been with pupils classified as the lowest achievers (Black and Wiliam, 1998a). However, surveys of formative assessment in classrooms reveal weak practice with scope for improvement. Assessments usually encourage rote and superficial learning even when the stated intention is to develop deep learning. The grading function of assessment is overemphasised and the learning function underemphasised, resulting in an emphasis on competition rather than personal improvement. Pupils are not encouraged to be responsible for their own learning which results in an unwillingness or inability to be involved in the assessment of their own learning.

Formative assessment is a ‘process’ rather than a ‘product’. This is comprehensively illustrated in key reviews by Wiliam and Black (1996), Black and Wiliam (1998a) and in contemporary research by Bell and Cowie (2001:12) which defines formative assessment as a cycle “dependent on teachers’ and students’ mutual engagement in a process which involves them eliciting, interpreting and acting on assessment information”. The authors suggest that the cycle applies to all assessments. However, what distinguishes the process of formative assessment is the application of feedback. The process must yield evidence of the ‘gap’ between the current and the desired level of pupil learning and suggest actions to close it (Wiliam and Black, 1996:543). Thus, assessment is formative when comparisons of actual and reference levels yields information which is used to alter the gap (McCallum, 2001) The key to this definition is that assessment can only be formative if a learning action follows the outcome; otherwise the practice is
summative (Black, 2003:3). The ‘formative’ aspect of practice therefore relates to diagnosing pupils’ needs, and informing teaching with a focus on achieving progress towards further learning (Assessment Reform Group, 1999).

Subsequent publications extended the definition and gave comprehensible accounts of how to alter the learning gap. Such definitions require specific roles of teachers and pupils to enable formative assessment to take place.

“An assessment activity can help learning if it provides information to be used as feedback, by teachers, and by their pupils in assessing themselves and each other, to modify the teaching and learning activities in which they are engaged. Such assessment becomes ‘formative assessment’ when the evidence is actually used to adapt the teaching work to meet learning needs” Black et. al., 2003, 2-3).

The role of the teacher is to help the learner understand the aims, elicit the evidence which can help the learner understand his/her present state of learning and enable the learner close the gap. Here, the teacher’s role is not as ‘assessor’ or ‘tester’ but as a ‘cognitive researcher’ diagnosing pupils’ learning processes and needs (Newman, 1998; quoted in Dann, 2002:36). There is strong evidence that the quality of teacher feedback to pupils has a profound impact on the quality of learning (Ammons, 1956; Elawar and Corno, 1985; Black and Wiliam, 1998a). The learning theories (behaviourist, social constructivist, situational) reviewed in the next section all support the view that feedback is essential for learning (Vygotsky, 1978, Bruner; 1986).

SECTION 2: FORMATIVE ASSESSMENT AND LEARNING THEORIES

The definitions of formative assessment by various authors show slight variations concerning the relative importance of the key components. For instance, the reviews by Black and Wiliam (1998a) place more emphasis on ‘teacher feedback’ than other components to regulate the learning processes. This was also outlined by Perrenoud (1998:85) who suggested that the mere presence of feedback is not comprehensive enough and that the cognitive and socio-affective mechanisms activated by pupils should also be considered in order to understand the effects of formative assessment. Similarly, Dann (2002) highlights the need to understand
socio-construction of feedback. Ecclestone (2002) conceptualises formative assessment by referring to the nature and varying degrees of pupil autonomy. Thus, the literature highlights different assessment languages in the proposed conceptualisations. This could be caused by the different ways in which various authors relate formative assessment to the learning process. Any attempt to conceptualise formative assessment, therefore, would be incomplete without exploring its links to the different learning theorises and the ways in which they manifest themselves in assessment terminology and practice.

The links between learning and formative assessment provided in the literature have involved discussions mainly on behaviourist, constructivist (Ecclestone 2002; Sadler, 1989; Torrance, 1993), and recently socio-cultural views of learning (Bell and Cowie, 2001). This section introduces and reviews the main characteristics of each learning theory and discusses their implications for the three components of formative assessment: assessment task design, teacher feedback and pupil self-assessment.

**Behavioural Learning Theory and Formative Assessment**

Dann (2002:13-4) suggests that the behavioural models of learning are constructed within a framework which consists of a series of highly specific learning objectives. She identifies the underlying assumptions of this model:

- What is worth knowing can be reduced to objective statements;
- Objective statements can be structured to present a sequential, logical, linear progression for learning;
- Learning can be accurately measured through rigorous assessment of achievement against the learning objectives.

If teaching is concerned with teaching objectives which are formed by discrete knowledge and skills, assessment is designed to assess the separate units of information and skills. Then the question arises ‘what does this suggest about the first component of formative assessment – the design of assessment tasks?’ Gipps (1995) and Resnick and Resnick (1992) argue that behavioural learning may result
in assessment tasks which mainly focus on lower order thinking skills (recall of recorded information). They point out the difficulty of integrating higher order thinking skills requiring pupils to interpret, analyse, apply, synthesise, and evaluate the discrete units to make sense of a new situation. Thus, in behavioural teaching, assessment tasks can encourage a trivial surface rather than a deep approach to learning.

In behavioural learning, teacher feedback, the second component of formative assessment, functions in specific ways. There is an assumption that teacher feedback has to make a summative judgement (Nieweg, 2002). It can be further assumed that the main function of feedback is to communicate whether a response is correct or not and whether a specific objective is met. Gipps (1996) suggests that as a result of the employment of traditional summative tasks within this learning model, assessment is more likely to encourage a norm-referenced type of judgement. Ecclestone (2002:41) suggests that the visible effects of the behaviourist tradition can be seen in ‘rewards’, ‘performance goals’ and ‘punishments’ which are deeply embodied in teachers’ assessment, particularly feedback strategies. Given the type of learning model, it could be argued that the formative function of teacher feedback, for the progression of pupils’ learning, is limited.

Scope for pupil self-assessment, the third component of formative assessment, is also limited. Pupils engaging in behavioural learning are viewed as ‘passive’ or ‘mechanical’ agents who react to the contexts and information given to them (Dann, 2002:13). So, the opportunities for pupils to interact with their teachers in the learning process, and therefore pupil self-assessment, are limited. When it is used, it may involve the use of pre-defined criteria since the learning model is driven by highly defined learning objectives. As Dann (2002) suggests there is little scope for negotiating and redefining the assessment criteria.

**Constructivist Learning and Formative Assessment**

As a learning theory, ‘constructivism’ is an umbrella term which encapsulates a multitude of perspectives (Fox, 2001:24). These include Piagetian (Piaget, 1967),
neo-Vygotskian (Wertsch, 1991), radical (Boyd, 1992; Hewson and Hewson, 1998), and social constructivist (Rogoff, 1990). At the centre of almost all of these perspectives is the suggestion that pupils learn through actively making meaning by building on their existing knowledge and schemata or structure (Gipps, 1996). Thus, the pupils' new knowledge or skill must be linked to the previous knowledge, schemata or structure that already exists.

Viewing learning through constructivism makes certain assumptions about the nature of knowledge. Fox (2001:24) suggests that, in this view, knowledge is constructed socially rather than accumulated as innate units. Furthermore, this view claims that all knowledge is personal and idiosyncratic (ibid). These two assumptions of constructivist theory affect the whole process of teaching, learning and assessment.

According to Taber (2002) and Hodson and Hodson (1998:34) if knowledge is constructed on the basis of idiosyncratic dispositions, constructivist teaching requires the teacher to:

- Analyse the logical structure of a topic to identify all the prerequisite knowledge required;
- Check that the learners have already learnt conventional versions of prerequisite knowledge;
- Create opportunities for pupils to develop their ideas;
- Provide stimuli for pupils to develop, modify, and where necessary change their ideas and views;
- Support their attempts to re-think and construct their idea and views.

What then are the implications of the requirements of constructivist teaching for formative assessment? One key prerequisite is that assessment should identify what pupils already know so that further learning can build upon their existing knowledge. This suggests that diagnostic assessment tasks are crucial, in particular during the introduction stage of a task. A further implication of this assumption is that unless pupils relate their prior knowledge or experience to a new situation, learning will not occur (Dann, 2002). This draws attention to the importance of the
context of the assessment tasks, in which the topic of unit is embodied, to enable pupils to engage in learning. This clearly creates a challenge to the different dispositions that pupils bring to a new learning situation. In order to be fair to a large body of pupils, assessment tasks should be built on a range of contexts so that the chances of pupil engagement are increased. A further point is that assessment tasks should be sufficiently challenging so that pupils’ prior ideas or concepts and assumptions are figured and configured (Dann, 2002).

In order to explore the scope for pupil self-assessment under constructivism, it is worthwhile revisiting the role of the learner. The constructivist view argues that the learner is an active knowledge maker (Calfee, 1995) and that he/she has control over this/her own learning. Gipps (1996) suggests that metacognition is a crucial component of constructivist learning since searching for connections and conflicts requires pupils to make ongoing judgements against what they already know and understand. Furthermore, the criteria for self-assessment can be discussed and negotiated with the teacher. It would seem that this learning model is better equipped to accommodate pupil self-assessment than the behavioural model.

The requirements for constructivist teaching outlined above suggest that teaching and assessment are inseparable which is one of the key assumptions of formative assessment. Assessment within this tradition takes more account of pupil cognition (i.e. how pupils engage with tasks and how they perceive and interpret feedback) in explaining learning. The learning and assessment are dynamic and socially constructed. Ecclestone (2002:44) suggests that constructivist assessment is less predictable, less amenable to regulation than behaviourist assessment and therefore more risky. Seen in this light, behaviourist approaches offer a comforting, low risk approach to learning and assessment (ibid).

**Formative Assessment as a Socio-Cultural Process**

One of the key criticisms of the constructivist view is that it theorises learning as a highly individualised activity by ignoring the socially and historically situated nature of learning. Drawing on Wertsch’s (1991) ideas, (Bell and Cowie, 2001) suggest that the main aim of a socio-cultural view of learning, is to create an
account of human mental processes that recognise the essential relationships between mental processes and their social, cultural and institutional settings. This learning model makes similar claims to the constructivist view about the nature of knowledge (Bell and Gilbert 1996, quoted in Bell and Cowie, 2001:115).

- Knowledge is constructed by people;
- The construction and deconstruction of knowledge is both personal and social;
- Socially constructed knowledge is both the context for and the outcome of human interaction;
- Social interaction with others is a part of personal and social construction and deconstruction of knowledge.

The above assumptions suggest that the cognitive activities which operate for learning are stimulated and created through social interactions in social settings. Furthermore, Bell and Cowie (2001) suggest that this theory of learning implies a model of formative assessment which is a ‘socio-cultural activity’ and which is ‘situated, distributed and mediated action’. Firstly, they argue that formative assessment is a highly contextualised and situated activity, so any attempt to theorise it should consider the context in which it occurs. So, learning in formative assessment is not solely a matter of acquisition and internalisation of skills and knowledge but a matter of increased access of learners’ roles to expert performances (Bell and Cowie, 2001). Secondly, cognitive activities are distributed across other people and the socio-cultural situation, and the social surroundings should not be seen only as sources of stimulation and guidance but as actual causes for construction of thought (ibid). Similarly, Cole and Engeström (1993:13) defines distribution of cognition as follows:

"...the ways in which mind is distributed depend crucially on the tools through which one interacts with the world, and these in turn depend on one’s goals. The combination of goal, tools and setting constitutes simultaneously ‘the context of behaviour and the ways in which cognition can be said distributed in context’.

The definitions above suggest distributed cognition is a socially supported process across the contexts in which it occurs. Finally, based on Vygotsky’s notion of social cognition, Bell and Cowie (2001) suggest that human mental functioning is
altered by various socially constructed factors (i.e. language), so the relationship between learning and assessment in this environment cannot be conceptualised without recognising the essential relationships between social, cultural and institutional setting (Wertsch, 1991). The authors conclude that formative assessment is a highly contextualised (setting dependant) and situated activity. Therefore, any attempt to understand it therefore should explore not the progression in pupils' learning, but consider the social dynamics of the context which give rise to such learning.

SECTION 3: KEY COMPONENTS OF FORMATIVE ASSESSMENT

Having explored the relationship between theories of learning and formative assessment, the next sections continue to describe the key components of formative assessment as teacher feedback, pupil self-assessment, and assessment task design.

Teacher Feedback

A major study by Tunstall and Gipps (1996) and Gipps et al. (2000) provided a typology for categorising assessment feedback in primary classrooms:

<table>
<thead>
<tr>
<th>Feedback Strategy</th>
<th>Evaluative Feedback</th>
<th>Descriptive Feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Type</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rewarding (positive)</td>
<td>Specifying Attainment (C)</td>
</tr>
<tr>
<td></td>
<td>Punishing (negative)</td>
<td>Specifying improvement (C)</td>
</tr>
<tr>
<td></td>
<td>Approving (positive)</td>
<td>Mutual construction of achievement (D)</td>
</tr>
<tr>
<td></td>
<td>Disapproving (negative)</td>
<td>Mutual construction of improvement (D)</td>
</tr>
<tr>
<td></td>
<td>Giving rewards</td>
<td>Telling children they are right/wrong</td>
</tr>
<tr>
<td></td>
<td>Giving punishments</td>
<td>Describing why an answer is correct</td>
</tr>
<tr>
<td></td>
<td>Expressing approval</td>
<td>Telling children what they have/have not achieved</td>
</tr>
<tr>
<td></td>
<td>Expressing disapproval</td>
<td>Specifying or implying a better way of doing something</td>
</tr>
</tbody>
</table>

Table 2.2: A Typology of Assessment Feedback
Based on their analysis of the purpose, style and the nature of the varying feedback strategies, Tunstall and Gipps (1996) concluded that evaluative feedback is more associated with judgements on the basis of explicit or implicit norms with reference to actual performance. This feedback type is not linked to formative assessment (ibid). Conversely, descriptive feedback is regarded as having a closer association with formative assessment with slight variations:

"Type C (specifying attainment and specifying improvement) shows a mastery oriented approach to formative assessment. It illustrates: teachers' acknowledgement of specific attainment; the use of models by teachers for work and behaviour; diagnosis using specific criteria; correcting and checking procedures. Type D (constructing achievement and constructing the way forward) approach to formative assessment. It illustrates teachers' use of both sharp and contextualised 'fuzzy' criteria; use of teacher and child work exemplars; teacher-child assessment of work; the use of strategies for self-regulation" (Tunstall and Gipps, 1996:189).

The above quotation demonstrates the important principle that each feedback strategy could be grounded on different learning theories. The nature of evaluative feedback appears to be 'judgmental', with the teacher being the sole provider. This is much consistent with the 'receptive-transmission' model of teaching with the teacher as the expert (Askew and Lodge, 2000) closely controlling the feedback process (Hargreaves et al., 2000). Here, the function of feedback is to judge and sum up the pupil achievement.

In contrast, the nature of descriptive feedback is based on a socio-constructivist theory of learning. In the mutual construction of achievement, the source of feedback is not only the teacher but it is also a product of the interaction between teachers and pupils. This interaction is characterised by:

"...an expended discourse in which 'expert' (teachers) enables others (pupils) to gain new understandings, make sense of experiences and make connections by the use of open questions and insights...feedback is a two way process" (Askew and Lodge, 2000:4).

What is implied by 'constructing the way forward' is one step beyond constructivism. Tunstall and Gipps (1996) recommend the flexible use of assessment criteria, the empowerment of pupils in finding ways to improve their learning through this strategy. Analysis by Askew and Lodge (2000:4) described this strategy as "feedback in the co-constructivist model" and suggested that this is
an “extended discourse involving a reciprocal process of talking about learning”. What distinguishes this strategy from the behavioural perspective is the melting of the boundaries of the power relationship between teachers and pupils. Here, pupils and teachers are interdependent.

The dialogue during which feedback is generated has been of interest and published studies have explored the dynamics of the process, with particular reference to provision of quality feedback. The establishment of trust between teacher and pupil is seen as a prerequisite (Chung, 2002) for enabling engagement in the dialogue, As Loughran (1997:60 quoted in Chung 2002:4) states: “…trust involves knowing and believing that individual ideas, thoughts and views can be offered and explored in challenging ways such that the challenge is professional not personal”. “Spontaneity”, “informality”, “reciprocity” are other qualities which have been attached to the descriptions of effective dialogue (Carnell, 2000).

**Pupil Self-Assessment**

The current theory of formative assessment states that one of the key prerequisites for pupils’ improvement is that pupils must develop the capacity to monitor the quality of their own work during its actual production (Sadler, 1989:119). In a constructivist view of learning, pupils should play an active part in assessment, both of themselves and of their peers, and this constitutes important evidence of their learning (Dunn, 2002; Ecclestone, 2002). This implies an active role for pupils by taking more control and responsibility for their own learning.

There is consensus among some authors about the definition of pupil self-assessment (Klenowski, 1995; Latham, 1997). This is typified by Klenowski (1995) who states that pupils assess their achievement by identifying their strengths and weaknesses with a view to gaining and improving their learning outcomes. Focusing on and reflecting back on the learning process are the key elements of this definition.

‘Self-reflection’ is also used to describe a similar process involving referring back after an achievement and considering the merits of the performance with a view to
evaluating (Moon, 1999). This is quite different from just recalling a performance.
Brown and Thomson (2000: 113) described ‘reflection’ with reference to co-
operative group work and suggested that it had three key elements. It involved
evaluation (how well pupils performed the task and how well they worked
together), analysis (what pupils did in order to identify what the factors were that
helped or hindered the smooth functioning of the group), and setting goals (that
will help the group function better in the future).

Pupil self-assessment makes several assumptions regarding the nature of learning.
The themes which are introduced here are that learning is actively and socially
constructed, and that reflection and metacognition are salient components of
learning. The role of the learner in his/her own learning is a key defining aspect of
pupil self-assessment. Studies suggest that pupil self-assessment provides the
opportunity for pupils to be actively involved in thinking about their learning with
a particular emphasis on how they are progressing in their own learning
(Klenowski, 1995). Instead of being passive receivers of assessment, pupils take
the initiative and responsibility to evaluate their performance. In such a context,
pupils are not the receiver of assessment feedback but are pro-active in setting,
monitoring and internalising standards (ibid).

A major concern in those studies was the form and use of assessment criteria for
pupil self-assessment. Some research studies linked ‘self-directed learning’ and
‘pupil autonomy’ to formative assessment, in particular its pupil self-assessment
component Ecclestone (2002) unpacked the theorisation of pupil self-assessment
and used a differentiated model of autonomy to develop her analysis. Following
the typology by Carr and Kemmis (1986), which distinguishes autonomy as
procedural, personal, critical and emancipatory, she argued that each type might
result in different pupil-self assessment and formative assessment practices as the
table indicates:
Table 2.3: Different Conceptualisations of Pupil Autonomy and Its Impact on Pupil Self-Assessment and Learning

<table>
<thead>
<tr>
<th>PROCEDURAL AUTONOMY</th>
<th>PERSONAL (PRACTICAL) AUTONOMY</th>
<th>CRITICAL and EMANCIPATORY AUTONOMY</th>
</tr>
</thead>
<tbody>
<tr>
<td>This refers to self-assessment activities involving the use of pre-determined assessment criteria, short-term goals and replication of information (ibid:36). The learner is proactive within well-specified rules.</td>
<td>Personal autonomy is based on one’s knowledge of his or her own strengths and weaknesses, learning habits and potential choices for actions and progress. Learning is more pupil-centred based on negotiation of intended outcomes how to achieve them. Here, self-assessment and mentoring is based on ipsative use of criterion, individualised feedback.</td>
<td>This type autonomy refers to critical engagement and the exercise of it. Here, pupils fully engage in decisions about assessment criteria.</td>
</tr>
</tbody>
</table>

When pupils construct their own evaluations, this does not happen in isolation, indeed it usually happens in classroom settings (Boud, 1995). Criteria are often formed through reflection that is based on interactions with peers and the teacher. Social dialogue also provides an opportunity and encouragement for pupils to reflect on their experience, to give and receive feedback, and agree on future actions (Latham, 1997). The interaction and dialogue between peers and between pupils and teacher are a crucial feature of self-evaluation and its effectiveness is determined by its quality (Boud, 1995). This is also linked to autonomy and self-direction in which learners are encouraged to intervene in their own learning process, obliged to make choices, and follow the consequences of their choices (Boud, 1995, p.34).

Having defined pupil self-assessment, this section continues to explore assessment task design. Here, emphasis is given in the ways with which alternative approach to assessment for learning (authentic and performance assessment) have provided distinct criteria for assessment task design.
Assessment Task Design

An assessment task is a crucial component of formative assessment since the way in which it is devised and implemented affects the quality of interactions and feedback that are generated during the assessment process (Black and Wiliam, 1998a). One of the criticisms in the literature reviewed was the limited amount of time teachers devoted to devising tasks for serving formative purposes as opposed to time spent for marking test papers and examinations (Jones et al., 1986). This occurred in spite of the recognised importance of assessment task design (Black and Wiliam, 1998a; Eisner, 1993). Indeed, this neglect is also reflected in the relatively small amount of published material on assessment task design in relation to formative assessment when compared to other components.

If formative assessment is concerned with the improvement of teaching and learning by finding out what pupils have learned and what they have not, assessment tasks that only allow the observation of failure and success is insufficient for formative purposes (Sadler 1989). So, the question arises “What assessment tasks will permit us [teachers] to obtain reliable and valid information about how much the learners really know, understand and can do, in a given subject area?” (Dumas-Carré and Larcher, 1987). McCallum (2000) and Dumas-Carre and Larcher (1987) suggested that some tasks have the capacity to encourage pupil self-reflection more than others. These types are as follows:

- A specific situation identical to the one studied;
- A ‘typical’ problem but not identical to the one studied;
- Quite a new problem requiring new reasoning and the construction of a new approach deploying established knowledge in a new way.

Guskey and Marzano (2001:4-6) suggest some key questions to guide the decisions about assessment task design:

- What is it that we [teachers] would truly like our pupils to be able to do?
- What do we do in classrooms to enhance learning, development and growth?
How can we align our assessment activities with our instructional goals?

The first question was regarded as the first step of task design (Jones et al., 1986; Guskey and Marzano, 2001). Jones et al. (1986) suggested that this involves decisions regarding the core objectives. In a criterion-referenced assessment framework, the objectives refer to specified key learning outcomes in the form of key facts, concepts, understandings or skills that are central for a unit of work. Indeed, this is not a new notion and the classification of educational objectives has received attention in the literature (Bloom, 1956; Ward and Murray-Ward, 1999).

The second and third key questions involve decisions regarding how to assess the task in a way that it contributes pupils' learning as well as it appears as invisible in teaching and learning. Guskey and Marzano (2001) suggested that assessment tasks should appear to be 'natural' so that assessment can be accommodated into the normal classroom activities. They also suggest that this stage also involves decisions about assessment task format.

The type of assessment format and its use can affect how pupils interpret, respond to and perform a given task. Assessment task formats vary from structured test types to open-ended enquiry. A multiple choice task format is the most traditional form of task format and it originates from the Intelligence Quotient (IQ) tests (Resnick and Resnick, 1994). However, such tests have been criticised for their limitations in eliciting the pupils' thought processes, and assessing, per se, simple testable skills rather than higher order thinking processes (Eley et al., 2001; Resnick and Resnick 1994; Haney et al., 1993; Wiley and Haertel; 1996; Badger, 1999). This is partly because its main focus is on the 'products' of learning rather than 'processes' (Rowntree, 1991).

**Task Design Criteria for Alternative Approaches to Assessment**

The varying assessment for learning movements have been in parallel with the evolution of thinking about formative assessment. Each alternative assessment approach has influenced how formative assessment is considered in a distinct way. Authentic and performance approaches to assessment had an impact on assessment
task design by providing a number of criteria to guide task design, which is reviewed in the next section.

**Task Design for Authentic Assessment**

Research by Zessoules and Gardner (1991: 51-63) and Gardner (1992) described the purposes of authentic assessment tasks. On the former, these are:

- nurturing complex understanding;
- developing reflection as a habit of mind;
- documenting pupils’ evolving understandings;
- assessment as a moment of learning.

Nurturing complex understanding could be made possible by devising assessment tasks which give opportunities for pupils to demonstrate what and how they understand rather than what they know (Zessoules and Gardner, 1991:51). One way of assisting the development of complex understanding is developing pupils’ ability to reflect on what they have learned (developing reflection as a habit of mind). How pupils have learned should also be monitored in order to explore how their learning changes and evolves through time (documenting pupils’ evolving understandings). Lastly, authentic assessment has the potential to be embodied within teaching and learning since this kind of assessment is not inserted at some point.

Developing the theme of nurturing complex understanding, Neumann (1994) describes authentic assessment as a learning experience which involves pupils in synthesising and creating new ideas during the assessment process. Newman and Archbald (1992: 72-75) suggest three criteria which authentic assessment must meet which relate to this learning experience:

- disciplined enquiry;
- integration of knowledge (requires interpretation, formulation and critique of the new information);
- value beyond evaluation (authentic achievements have aesthetic, utilitarian or personal value, apart from documenting solely the competence of the learner).
Disciplined enquiry refers to the creation of new knowledge which involves three steps for the learner (Archbald and Newman, 1988: 2-4). Firstly, it requires prior knowledge in understanding a problem. Secondly, it requires in-depth understanding of a problem. Finally, it requires the learner going beyond knowledge that has been produced by others.

Truly authentic assessment tasks depend on specific conditions being in place. Firstly, tasks need to be fully integrated into the learning process so that ‘situational learning’ is enabled. In other words, authentic assessment tasks are regarded as highly contextualised as opposed to standard testing which is considered to be highly de-contextualised (Gardner, 1992). The highly contextualised nature of learning in this approach raises questions about its transfer and application to schools. Newman and Archbald (1992) suggested that authentic tasks are not compatible with school contexts. They argued that such tasks require collaboration since achievements outside of schools are generally contingent on intense feedback interaction with others. Secondly, teachers must act as a ‘coach’, ‘advisor’ and ‘counsellor’ and create environments where pupils can honestly reflect back on their learning (Meyer, 1994). Thirdly, Archbald and Newman (1992) pointed to the need for flexible use of time for the implementation of such tasks and viewed this as a problem for many classrooms since there is only a certain period of time allocated for a particular piece of work.

Authentic and performance tasks appear to be very similar. Archibald and Newmann (1988) and Meyer (1994) classified authentic assessment tasks into 3 different categories:
Table 2.4: Authentic Assessment Task Types

<table>
<thead>
<tr>
<th>TYPE 1: TESTS OF DISCRETE COMPETENCIES</th>
<th>TYPE 2: EXHIBITIONS</th>
<th>TYPE 3: PORTFOLIOS AND PROFILES</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Definition:</strong> Application of previous learning to new-world situations or problems to measure proficiencies under standard conditions—the achievement is broken down in several discrete units. This type of tasks are more associated with the areas of language performance.</td>
<td><strong>Definition:</strong> Pupils present a learning product that they have already prepared or they could be asked to react to a new situation, applying what they have learned. This type of project work generally involves engaging in community based learning environments.</td>
<td><strong>Definition:</strong> A portfolio simply refers to the collection of pupils’ work. Portfolios can be authentic if they provide a learning experience where pupils reveal and monitor their understandings, abilities, progress and mental processes through reflection on an ongoing basis over a period of time. Profiles are not created by individual pupils and do not contain actual samples of work. These are forms that teachers, pupils, and sometimes parents fill out with ratings and summary judgements or description of achievements.</td>
</tr>
</tbody>
</table>

The three types of tasks described above differ in their format and nature. ‘Tests of discrete competencies’ refer to quantitative measures of proficiencies under standard conditions that also meet the criteria for authenticity (Archbald and Newmann 1988:20). Exhibitions refer to tasks which do not have a single clear solution and offer pupils an opportunity to demonstrate competence in practical skills, creativity, adventure, career exploration, logical inquiry, global awareness, voluntary service. Archbald and Newmann (1988) suggested that such tasks encourage reflective thinking and applications of understandings to a new situation. They suggested that exhibition tasks are the best representation of ‘authentic tasks’ since they offer more opportunities for the pupils to demonstrate in-depth understanding and integration of knowledge into real life situations (Meyer, 1994). Newman and Archbald (1992) argued that authentic tasks could be better expressed in the completion of long-term projects which involve performances of pupils to their peers and the public at large.

The preparation of portfolios is defined as an authentic assessment task which refers to the collection of pupils’ work in order to demonstrate competence (Archibald and Newmann 1988). The reason why portfolios suit authentic assessment relates to ‘real-life’ skills which are complex, holistic and cannot easily
be broken down into individual units (Losardo and Syverson, 2001). Portfolios are also regarded as ‘authentic’ since they have the potential to provide an authentic learning experience by “pupils setting goals for themselves (consistent with teachers’ goals), deciding which pieces best demonstrate competence in those goals, revising those pieces by using their peers’ and teacher’ input (Meyer, 1994:106). Gardner (1992) named this experience ‘process-folios’ and suggested that a portfolio should include moments of critical events; new understandings reached, self and peer-assessment reports, early drafts and initial ideas. Archbald and Newmann (1988) referred to the integration of reflection points at various stages of a task to encourage pupils to think critically about their work. Here, depending on their answers, pupils either proceed or go back to fill in a missing part of the process. Thus, reflection on the evolving learning process is regarded as the essence of portfolios for the purposes of authentic assessment (Reckase and Welch, 1999).

In the area of authentic assessment, Eisner (1993) developed a number of criteria for devising and appraising assessment tasks. In the following points, Sutton (2000) summarises the key criteria that were adopted from the article of Eisner:

- Assessment tasks need to allow pupils to illustrate what they know, understand and can do not only in limited school environments but in the world outside schools;

- Tasks should also reveal how pupils go about solving a problem, not only the solutions they formulate;

- Tasks could be devised in a way that it allows pupils work in groups. Here, there should be an assessment of the contributions of individuals to the group performance;

- Tasks should make possible more than one acceptable solution to a problem and more than one acceptable answer to a question;
- Tasks should be devised which enable teachers to observe the process of decision-making and to interrupt that process at strategic moments in order to seek an explanation for the choices they made and the strategies they employed;

- Tasks should permit pupils to select a form of representation he/she chooses to use to display or demonstrate what has been learned;

- Tasks should require pupils to display sensitivity to configurations or wholes, not simply to discrete elements.

For Eisner, creating a task means providing opportunities for pupils to engage in real life situations and gain holistic experiences through their engagement. To facilitate this experience, there is a need for teachers to analyse the complexity of out of school experiences and devise task for pupils not only to construct their own responses to the real life situations but also select a medium through which what they have learned can be made public (Eisner, 1993:232).

**Task Design for Performance Assessment**

Assessment tasks for performance assessment appear to be very similar to authentic assessment tasks with the main difference being that authentic tasks are linked to real life contexts and performance tasks involve a performance. Losardo and Syverson (2001:73) suggested that performance assessment tasks involve the application of skills and knowledge in practical situations. Ward and Murray-Ward (1999) suggested a typology to classify performance assessment tasks on the basis of the standardisation and control of the task. The first group consists of performance and production tasks which refer to pupils’ opportunity to demonstrate targeted behaviour under observation in controlled and standard conditions. Pupils’ performances in Physical Education Lessons, music and art classes could be examples of such tasks. Here, assessment is very much limited to pupils’ performance.
The second task group consists of project tasks which are very similar to authentic tasks. The only difference is that performance project work does not have to be connected to real life situations and can be conducted within school boundaries. Linn and Baker (1996) and Violato et al. (1998) suggested that performance tasks should be open-ended so as to give pupils some room to construct their performance. Performance tasks were also regarded as having the capacity to promote complex skills such as formulating problems, solving problems, reasoning, higher order thinking, and communication (Resnick and Resnick, 1992; Linn, 1993; Honig and Alexander, 1996).

Summary of Alternative Approaches to Assessment for Learning

Up to now, this chapter has reviewed the key definitions and characteristics of each type of alternative approaches to assessment for learning. The following table summarises how each approach contributes to the current understanding of formative assessment:
Table 2.5: Summary of Alternative Approaches to Assessment

<table>
<thead>
<tr>
<th></th>
<th>DYNAMIC ASSESSMENT</th>
<th>AUTHENTIC ASSESSMENT</th>
<th>PERFORMANCE ASSESSMENT</th>
<th>FORMATIVE ASSESSMENT</th>
</tr>
</thead>
</table>
| **DEFINITION AND KEY CHARACTERISTICS** | Broadly refers to a process of assessment of pupils' learning potential. Two distinct approaches are as follows:  
• Assessment of the ZPD;  
• Assessment in the ZPD. | Authentic assessment refers to a process of assessment which involves setting learning tasks in real life contexts. | Performance assessment refers to direct assessment of performance, in 'real time' rather than artificial retrospective analysis of previous performances. | Formative assessment refers to a process of closing a gap between current and desired levels of achievement of pupils' learning. |
| **UNDERLYING LEARNING THEORY** | Dynamic assessment is underpinned by:  
• Social constructivist learning based on Vygotskian notion of the ZPD. | Authentic assessment is underpinned by:  
• Situated learning;  
• Social constructivist learning;  
• Apprenticeship learning. | Performance assessment is underpinned by:  
• Situated learning. | Formative assessment learning is underpinned by:  
• Constructivist learning;  
• Socio-constructivist learning;  
• Behavioural learning. |
| **TEACHERS' AND PUPILS' ROLE** | Teachers' role heavily involves as 'mentor', 'helper' and 'regulator':  
• Diagnosis of the cognitive processes that leads pupil learning;  
• Promote learning through scaffolding. Through time scaffolding evolves from an assisted to a non-assisted form. | Teachers' role involves:  
• Setting tasks for real life experiences;  
• Collaboration with out-of-school agencies to enable 'authentic learning'. The last point implies that the power of teacher is diffused because of the existence and involvement of outside an agent. | Teachers' role involves:  
• Setting tasks for observing direct performances;  
• Observing the real performance;  
• Developing criteria for assessing the performance. | Teachers' role involves:  
• Elicitation of assessment evidence which indicates the existence of a learning gap and raising pupils awareness of it  
• Interpretation of the gap  
• Providing feedback in order to close the learning gap. |
| **ASSESSMENT TASKS** | Assessment tasks may involve:  
• Diagnostic tasks through which pupils' misconceptions and thinking can be understood;  
• Collaborative tasks which allow teacher / pupil and pupil / pupil interactions. | Setting the task in real life situations is crucial for authentic assessment. Tasks should provide opportunities for pupils to  
• Create knowledge;  
• Produce an original piece of work;  
• Enable deep learning rather than surface learning. | The nature of tasks is very similar to authentic assessment tasks apart from it not having a real life context. The rest of the principles remain the same.  
Performance type of tasks can vary from performance tests to project work types. | Assessment tasks should:  
• Include diagnostic tasks through which pupils' learning needs are identified.  
• Provide opportunities to reveal pupils' learning process. |
| ASSESSMENT FEEDBACK | Assessment feedback could be generated to meet the emergent learning needs of pupils during the interactive assessment process. | Similar to dynamic tasks, feedback could be situated in situations and is generated as a result of social interaction. | Assessment feedback should:  
• Provide suggestions on how to close the learning gap  
• Communicate the standards with a view to translating them as a means of how to improve pupil learning and setting new goals. |
|---------------------|-------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|
| PUPIL SELF-ASSESSMENT | Pupil self-assessment is an ultimate aim but not a major concern in the beginning. It is primarily teachers’ responsibility to assess pupils’ learning. | Theoretically, the locus of control in terms of defining the topic and pace of work rests with pupils. However, its practicality is dependent on the opportunities that are facilitated by the outside agents and teachers to integrate pupil self-assessment into the learning task. In this approach, portfolios could facilitate pupil self-assessment. | Pupil self-assessment is a crucial component of formative assessment. It enables pupils to evaluate their performance by identifying their strengths and weaknesses with a view to gaining and improving their learning outcomes:  
• To learn how to reflect on their learning;  
• To internalise standards which they are being judged. |
| CAPACITY FOR QUANTIFICATION | The two distinct approaches have different capacities for quantification.  
• Assessment of the ZPD leads quantified outcomes of learning potential;  
• Assessment in the ZPD leads qualitative descriptions of learning potential; However, quantification could be managed by considering the frequency and intensity of scaffolding that is provided to support the assessment process. | The quantification could be addressed through portfolios. Attempts to quantify achievement through portfolio could involve subjective decisions of what counts as assessment evidence. | Teachers are responsible for developing criteria for judgement. This can be varied to judge the performance as a whole (holistic) or break down each component so that they can be quantified separately. This is not a well researched area; however, the following strategies have been proposed in the literature:  
Criterion referenced quantification could be compatible with the underpinnings of formative assessment. |
| TO WHAT EXTENT IS IT FORMATIVE? | Assessment in the ZPD can serve formative functions because:  
• The concern is to diagnose and to develop a deep understanding of pupils’ learning process;  
• The action followed by diagnosis serves to improve pupils’ learning through scaffolding. | Authentic assessment can serve formative functions by  
• Portfolios providing opportunities for teachers and pupils to monitor and reflect on pupils’ evolving understanding. | Performance assessment could serve formative functions by encouraging teachers to offer feedback during the completion of task. The extent to which an assessment is formative is dependent on the learning theory that is based on and operated. |
As demonstrated in Table 2.5, there is considerable variation in the characteristics of alternative approaches to assessment for learning. Also, the respective learning theories show variation across different assessment types. However, what is common to all of them is that they confirm a shift away from traditional tests. Each has a particular emphasis. Both authentic and performance assessments are distinguished on the basis of their assessment task characteristics. Dynamic assessment is characterised by teacher/pupil interactions and by teacher scaffolding in relation to assessment tasks. In formative assessment, teacher feedback for closing the learning gap is key to its definition. What is common to all of the alternative approaches to assessment for learning is:

- The integration of assessment into the process of teaching and learning;
- Focusing assessment on the learning process as well as learning product;
- Increasing opportunities for pupils to engage in their own learning and assessment.

Each of the three alternative approaches to assessment can assist in conceptualising formative assessment and they extend the understanding of different components. The literature on authentic and performance assessment makes a contribution to the approach to devising tasks which enable deep learning. Dynamic assessment research examples provide the ways in which pupil learning can be supported during the learning process and reconceptualise the role of teachers in assessment. However, in each case, challenges are experienced in quantifying the qualities of the learning processes. The following review extends some of these points within the review of the key components of formative assessment. This is then followed by an account of how formative assessment appears in the classrooms.

SECTION 4: FORMATIVE ASSESSMENT IN PRACTICE

The previous sections defined the key components of formative assessment and explored assessment tasks in the light of their underlying assumptions mainly from a theoretical perspective. The section below reports on research evidence from
classrooms to enhance the understanding of teacher feedback and pupil self-assessment by limiting the focus to only selected areas.

**Review of the Characteristics of Effective Teacher Feedback**

Published research on feedback has focused on its specific characteristics and functions. The research evidence reveals that achieving effective feedback is a complex issue, and feedback is not always beneficial, depending on its nature (Coe, 1998). This is evident in a major review by Kluger and DeNisi (1996) which noted wide variations in the effect of feedback. They concluded that "feedback interventions do not always increase performance and under certain conditions are detrimental to performance" (Kluger and DeNisi, 1996:275). In the next section, the characteristics of effective feedback are reviewed based on three key areas; its specificity, timing, and focus.

**Specificity of Feedback: General or Focussed**

In research on formative assessment and teacher feedback, much attention has been focussed on the specificity of the feedback, whether feedback should specifically correct individual errors or include general comments regarding pupils’ errors. Bangert-Drown* et al.* (1991) used a meta-analysis of 58 experiments to investigate the corrective function of feedback on the improvement of learning and concluded that specific feedback is more effective than minimal feedback in indicating whether a response is right or wrong.

Effective feedback should also be transferable, for example from one task to another one. Research evidence suggests that pupils are not very good at bringing their previous experience to bear on solving related problems (Kahney, 1986). However, how a problem is presented has a powerful effect on a child’s ability to understand and relate the problem to his/her previous experience (Fisher, 1990).
Timing of Feedback

Research has offered some insights into the importance of timing on the effectiveness of feedback by comparing the relative advantages of delayed and immediate teacher feedback. In a review of published studies, Bangert-Drowns et al. (1991) demonstrated that delayed feedback resulted in increased time for pupils to reflect back on their own work and therefore provide opportunities for pupil self-assessment. The authors suggested that immediate teacher feedback can act to restrict or interrupt the learning experience and can become cognitively or motivationally inhibiting.

A contrasting result was gained from a review of quantitative studies by Kulik and Kulik (1988) who investigated the effect of immediate versus delayed feedback in the applied classroom studies using classroom quizzes, programmed materials, and experiments on the acquisition of test content. In this review, delayed feedback was effective in situations where subjects were encouraged to acquire test content or learn lists with feedback that would repeat the stimulus word (Bangert-Drowns et al., 1991:216). Immediate feedback was evidenced usually being superior to delayed feedback when the tasks presented greater cognitive demands, such as more conceptual learning in applied situations or list learning where the feedback gave only the correct response without repeating the stimulus (Bangert-Drowns et al., 1991:216). A contradictory result was gained by Mory (1992), who reported that when the assessment task corresponded to a higher level of cognitive demand, a delayed feedback was more beneficial, as it gave more time for students to think. Such studies suggest that feedback delay is beneficial only under specific conditions, and that it is also dependent upon the type of assessment of task conducted.

Focus of Feedback: Task-Referenced or Self-Referenced

Another important factor in the use of feedback is the type of reference used when giving performance comparisons - in other words, whether pupil performance is referenced against those of other pupils or against the requirements of the task. This is well illustrated in a study by McColskey and Leary (1985) in which the
effects of self- and task-referenced feedback were investigated with pupils at different levels of achievement. It was concluded that self-referenced feedback which conveyed the message of failure, led to lower self-esteem, expectations and motivation. Whereas if the feedback was task-referenced, it produced an increase in effort and higher expectations for future performance.

The effects of task-referenced and self-referenced feedback on interest and performance was also investigated by Butler (1988). In this study, 48 eleven year old pupils with differing abilities received either self-referenced numerical grades or task-referenced individual comments or both. The results indicated a 30% increase in the performance of pupils who received comments. Pupils given only grades showed no improvement in their scores. Pupils who received both comments and grades performed similarly to those given grades alone and did not improve in their scores.

Research Studies on Pupil Self-Assessment

The scope of research studies reviewed here on pupil self-assessment can be categorised into two main areas. Firstly, there are quantitative studies judging the merits of pupil self-assessments on the basis of selected criteria (Boud and Tyre, 1995; Daniels and Welford, 1990) and their effects on pupils’ academic achievements (Fontana and Fernandes, 1994). Secondly, there are qualitative studies which focus upon defining and conceptualising pupil self-assessment practices in classroom contexts (Klenowski, 1995; Tanner and Jones 1994; Towler and Broadfoot, 1992). A specific emphasis is also given the tools and methods that are applied in the implementation of pupil self-assessment (Tamir 1999; Latham, 1997).

Quantitative Studies of Pupil Self-Assessment

Published studies have focused upon investigations of whether pupil self-assessment could be trusted. Such a focus links to concerns about whether pupils can be ‘trusted’ to be objective when making judgements about their own performance (Tuck, 1995). To elucidate this, a research study by Boud and Tyree
(1995) examined university first year law school students’ ability to make self and peer-assessments by being able to establish appropriate criteria for this assessment and their ratings of themselves and their peers against those criteria. Independently, the teacher used the same scheme to produce equivalent sets of marks for each student. Results showed high level of agreement between the marks given by peers and those given by teachers and in a small number of cases pupils assessed themselves less favourably. The finding by Boud and Tyree (1995) that peer and self-assessment and teacher assessment renders similar results is also supported by other researchers in this area (Heuston, 1990; Daniels and Welford, 1990; King, 2000).

The relationship between pupils’ ability to mark themselves and their perceived ability was investigated by (Raynor, 1995) who studied the possible discrepancies between groups of different ability levels in their marking of themselves by comparing marks given by the students and the staff with year 11 students. The results obtained from 167 (n=52 for the more able group, n=115 remainder of the group) indicated that the more able students tend to underestimate their achievement (46% underestimated the effort grades and 53 % underestimate their attainment grades) whilst the average group had a strong tendency to overestimate (60 % overestimated their effort grades and 65 % their attainment grades).

Pupils’ ability to mark themselves was also addressed on the basis of gender differences. Daniels and Welford (1990) investigated 15 and 16-year-old pupils’ estimates of their performance and teacher predictions for completed science tasks. The students were divided into five ability groups. Results indicated that regardless of their ability the students were conservative in their estimates and described their performance as ‘average’. Boys however of all performance bands made over-estimates and might be described as possessing inflated expectations of performance. Higher achievers were more able to accurately assess their own performance than the lower achievers. This result is consistent with other research studies (Bloom, 1976; Wylie, 1979). Overall, the conclusions reached reinforce the view that pupils are capable of self-assessment and find it a positive experience.
Fontana and Fernandes (1994) were interested in the relationships between the regular use of pupil self-assessment techniques and learning enhancement and autonomy at primary school level. In this experimental study, the experimental group were instructed to achieve self-assessment through the following components:

- judging their answers against right-wrong criteria;
- assessing their own learning outcomes against the defined objectives and assessment criteria of the activity;
- discussing the clarification of these assessment criteria and even providing their own personal input.

Comparisons of the pre and post-test maths results indicated that 307 pupils in the experimental group (88.2 %) performed better in the post-test than in the pre-test and 41 pupils (11.8 %) performed worse. In the control group case, 222 pupils performed better (72.3 %) and 85 performed worse (27.7%). On the basis of their results, the authors suggest that pupils encouraged to use self-assessment showed significant gains in academic achievement compared with the control group (Fontana and Fernandes, 1994). In a subsequent publication, Fernandes and Fontana (1996) linked earlier findings to the concept of locus of control and regular self-assessment activities. ‘Locus of control’ refers to a person’s generalised expectancy as to whether or not he or she has power over events (Rotter, 1966, quoted in Fernandes and Fontana, 1996:302). A person with an internal locus of control, views himself or herself as a causal actor in determining events; whereas a person with external locus of control as due to external factors such as luck, fate and the intervention of the powerful others (Fernandes and Fontana, 1996:312). The authors reported the relationship between regular use of self-assessment and less dependency upon external sources as explanations for academic success when compared to control group pupils who did not use self-assessment.
Qualitative Studies of Pupil Self-Assessment

Qualitative studies on pupil self-assessment have focused on in-depth explanation and analysis of the process. In one study, Klenowski (1995) focused on the processes involved in pupil self-evaluation and applied case study approaches in an Australian Secondary High School and an English Further Education (FE) College. The study described two broad types of self-evaluation processes. The informal processes were more prevalent in everyday teaching and learning and were generally performed in oral forms. The use of group and class discussions and other practices related to target setting constituted the informal student self-assessment. Based on her observations she suggested that opportunities for informal self-assessment assisted pupils in becoming ‘self critical’. Students were shown to develop better understandings of their strengths and weaknesses. She described the formal and structured self-assessment processes as being more paper based and this involved teacher-student negotiating with students, and implementation of agreed procedures after reviewing written evidence of attainment (portfolios).

Notwithstanding degrees of formality Klenowski (1995:155) suggested that the key dimensions of pupil self-assessment processes are:

- The use of criteria by students to assess their own learning;
- The interactive dialogue which occurs between pupil and teacher, during the analysis of pupils’ self-assessment;
- The ascription of a grade by the pupils for their own work.

Klenowski (1995) urged the importance of interactive dialogue between pupils and teachers so that pupils develop understandings of what constitutes quality in given tasks and use those for self and peer evaluation. Ongoing feedback interactions and negotiations were identified as the main properties of this dialogue (Klenowski (1995:157). Finally, Klenowski (1995) pointed out the value of these activities as an active learning environment through which pupils have chances to develop their metacognitive skills.
Published research has illustrated the processes, challenges and potential of pupil self-assessment in an English primary school context (Dann, 1996). In a case study, Year 5 and 6 pupils were encouraged to assess their own work. Firstly, the teacher and the pupils discussed and reached agreement on what constitutes success for the task at hand. Then the pupils assessed their own and a peer’s work in relation to the criteria identified and set targets for improvements. On the basis of her analysis of the negotiation of the assessment criteria stage, Dann (1996), reported differences in pupils’ perceptions of negotiations on the assessment criteria. For example, one group of pupils did not regard the criteria as negotiated or shared but as ‘contrivance for conformity’ (Dann, 1996:57). Furthermore, some personal factors were involved in pupils’ assessment of their peers. Subsequently, she urges that we need to distinguish carefully between pupils’ ability and their willingness to use such criteria. This study illustrates the challenges associated with the implementation of pupil self and peer-assessment.

The ways in which pupils developed metacognition was addressed by the research studies. In a computer based text study Puntambekar and Boulay (1999) suggested two key actions to develop metacognition. Firstly, the authors emphasised the importance of differentiated feedback to communicate next steps on how to complete the task. Secondly, collaboration with peers and reflections on what has been done constituted an important stage for learners developing assessment criteria, an outcome which was also reported in early studies by Brown and Palincsar (1978) as a means of developing metacognitive skills.

**Research Studies on the Use of Specific Tools of Pupil Self-Assessment**

**Recording and Profiling**

Another salient area in reviewing pupil self-assessment relates to the investigation of the specific tools and methods used to carry it out. Adams and Burgess (1992) suggest that pupil self-assessment could be promoted through an appropriate recording process in which records of what pupils have done and can do, provides the potential for pupils to manage their own learning. This is consistent with Raven’s (1992) suggestions on pupils keeping records of selected events such as
where things went particularly well or alternatively lead to frustration and discomfort.

Latham (1997) considers profiling as an appropriate instrument for self-assessment and focuses on its contribution to learning. Here, profiling is described as a process during which a record of a range of assessment evidence is collected from a variety of contexts. From this perspective, profiling is a process where individuals develop charts and check their achievements against learning objectives. Latham (1997) further suggests that profiling provides a framework which has the potential to encourage self-assessment.

Latham (1997) described a practical implementation of profiling during which pupils were encouraged to evaluate their strengths and weaknesses. The conclusions suggested that the programmes' effectiveness depended upon:

- Pupils realising the inherent value within the system and being prepared to be fully involved;
- Committed teachers who can guide the tutorial discussion towards fostering pupils' personal development, including an understanding of the learning process and assisting in the planning of future targets. To achieve this level of expertise some tutors may need training which must be provided;
- The availability of a 'private space' and uninterrupted time for the tutorial discussion;
- The setting of relevant and realistic action plan which can be achieved within the time frame, structure and opportunities readily available.

As the above statements indicate, a key element of profiling and self-assessment is the review discussions. Pupils developing understandings of their strengths and weaknesses during reviewing processes is also recognised by Klenowski (1995).
Portfolios and Process-folios

The use and potential value of portfolios was discussed earlier on within task design for alternative approaches to assessment. In this section, their use for pupil self-assessment is discussed.

Zessoules and Gardner's (1991) described examples of pupil self-assessments through 'process-folios' that are collections of examples from pupils' work in a visual arts class. Here, the pupils were encouraged to build a portfolio to capture and represent their varying and evolving achievements. The authors described this experience as 'authentic', stemming from the genuine experience which pupils gain by exploring and monitoring their learning, and it is suggested that pupils developed their reflection skills by documenting their evolving understanding. Other published studies, in the context of pre-service teacher education (Klenowski, 2000; Lyons, 1998; Vavrus and Collins; 1991), primary and secondary education (Hewit 1995; Porter 1995; Engel; 1994; Borthwick, 1995) also reported similar uses of portfolios for pupil self-assessment purposes.

A further benefit of using portfolios was reported by Klenowski (2000) as teachers (lecturers) becoming more flexible in their teaching and developed better understanding of the standards by which students' progress can be judged. The following table (Table 2.6) illustrates 2 grades and how they were defined:

<table>
<thead>
<tr>
<th>Grade</th>
<th>Description of the grade</th>
</tr>
</thead>
</table>
| A (distinction) | - Evidence selected is outstanding and surpasses module objective.  
- Reflective statements show evidence of deep and original thinking and interpretation supported by well references.  
- Work is well structured and concise. |
| D (fail) | - Evidence selected does not meet module objectives.  
- Reflective statements are simple recall of facts without research evidence and references.  
- Work is poorly structured and organised. |

Despite the formalised grading system exemplified above, Klenowski (2000) reported difficulties and disagreements amongst the lecturers in the grading of students' portfolios. The students were found to have concerns regarding the
clarity or targets while the lecturers stated concerns over the ambiguity of the grading system and they requested more specific criteria and examples of graded work.

**Self-Reporting/ Diaries**

One study by Tamir (1999) has investigated the use of self-report as a way of promoting student self-assessment in Hawaiian and Israel high and lower secondary schools as a part of an innovative science course. This study encouraged students to reflect upon a list of concepts and skills that the students were expected to master before the actual teaching started. The researchers suggested that the activity was helpful for empowering students to build a framework of concepts and skills so that they could reflect back at a later stage.

Palmer (1992) explored teacher trainees’ use of questionnaires and diaries for self-assessment and the utility of this process for course evaluation purposes. This was in the context of a three-week teacher training course comprising language teaching in Norway. The participants were asked to fill in questionnaires which aimed to identified their approach, strategies and attitudes towards learning. Diary entries were made on a daily basis and loose guidelines were provided as to what aspects should be considered in this evaluation process. Results suggested that self-assessment was achieved when participants are given guidance and time to reflect (Palmer, 1992:232). In the case of course evaluation, the information in the diaries provided opportunities for formative course evaluation. By reviewing a written record of student experiences, the teachers were able to see how they responded to different types of classroom events on different occasions and thus developed a better perspective of learning and teaching processes (Palmer, 1992:234). Furthermore, the author suggested that student self-assessment is motivative since it empowers students in shaping the direction of the course programme.

**Reassembling the Components of Formative Assessment**

This section (Section 4) reviewed the research evidence on the two key components (teacher feedback and pupil self-assessment) of formative assessment.
The current part connects how those individual components relate to in the process of formative assessment in classrooms.

The research studies reviewed describe formative assessment as a process involving a series of ‘regular planned assessment incidents’ which involve eliciting assessment evidence, interpreting it and acting upon it for improving learning. (Bell and Cowie 2001; Gipps et al., 2000; Burke; 2000, Wiliam and Black 1996). In their research in classrooms, Torrance and Pryor (1998), identified three important stages ‘teacher introduction, pupil activities and plenary session’. During each stage, teachers used a variety of assessment and feedback strategies to facilitate formative assessment. They summarised these ‘processes of formative assessment’ (Torrance and Pryor, 1998:160-161) in the below table:

<table>
<thead>
<tr>
<th>A</th>
<th>Teacher observes pupils at work</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>Teacher examines work done</td>
</tr>
<tr>
<td>C</td>
<td>Teacher asks questions (seeks to elicit evidence of what pupil know, understand and can do)</td>
</tr>
<tr>
<td>D</td>
<td>Teacher asks for clarification about what has been done, is being done or will be done; pupil replies.</td>
</tr>
<tr>
<td>E</td>
<td>Teacher questions pupil about how and why specific action has been taken (meta-process and meta-cognitive questioning); pupil responds</td>
</tr>
<tr>
<td>F</td>
<td>Teacher communicates task criteria (what has to be done in order to complete the task) or negotiates them with pupil</td>
</tr>
<tr>
<td>G</td>
<td>Teacher communicates quality criteria or negotiates them with pupils</td>
</tr>
<tr>
<td>H</td>
<td>Teacher critiques a particular aspect of the work or invites pupils to do so</td>
</tr>
<tr>
<td>J</td>
<td>T supplies information, corrects or makes a counter suggestion</td>
</tr>
<tr>
<td>K</td>
<td>T gives and/or discusses evaluative feedback on work done with respect to; task, and or effort and or aptitude, ability (possibly with reference to future or past achievement)</td>
</tr>
<tr>
<td>L</td>
<td>Teacher negotiates with pupil what to do next time</td>
</tr>
<tr>
<td>M</td>
<td>Teacher suggests or negotiates with pupils what to do next time</td>
</tr>
<tr>
<td>N</td>
<td>T assigns marks, grade or summary judgement on the quality of this piece of work or negotiates an agreed one with pupil</td>
</tr>
<tr>
<td>P</td>
<td>T rewards or punishes pupil, or demonstrates approval or disapproval</td>
</tr>
</tbody>
</table>

The first stage of the cycle, eliciting the assessment evidence, is important to gain an understanding of pupils’ prior knowledge, skills and understanding, as well as their real life experience with this context (Eggen, 2002). This is necessary in order to identify pupils’ learning weaknesses and misconceptions. This stage is represented by steps A, B, C, D, E. These steps are facilitated primarily by observation of pupils’ (listening and questioning). Bell and Cowie (2001) suggest that formative assessment is enabled by not only considering how pupils respond in pupil-teacher interactions, but also how they behave in pupil-pupil interactions.
In the research examples given here, teachers also sought to engage in meta-
cognitive questioning to encourage introspective thought on how and why pupils
approached tasks in certain ways. Bell and Cowie (2001) suggest that the quality of
this assessment information was dependent on how it was elicited (the co-operation
between pupil and teacher) and the classroom atmosphere (the level of interest in
the topic and learning activities). They suggested that the elicitation of assessment
evidence is therefore a planned and setting dependant activity since what teachers
elicit was a result of what they ask (content) and how they asked it.

A second step in the implementation of formative assessment is the interpretation
of assessment evidence which will then be formulated as feedback. This
interpretation process is required where teachers seek to identify a gap between
current and desired levels of learning (Wiliam and Black, 1996). There are several
ways in which teachers interpret assessment evidence. Teachers can make an
'ipsative' interpretation when a pupil is assessed against her or his previous
performance (Knight, 2001). The interpretation can be 'criterion referenced'
which involves a pupil's performance being compared against a set of pre-defined
criteria of what constitutes success for a task (ibid). Bell and Cowie (2001)
indicated that these two methods of interpretation were the most appropriate for the
facilitation of formative assessment. Conversely, norm-referenced approaches,
involving relative comparison of pupils, are not as well suited, since they are
insensitive to changes in individuals’ learning.

Various authors viewed taking action on interpreted evidence of assessment as the
third stage of the formative assessment cycle. Black (2003) suggests that in order
for assessment to be formative, the actions taken should help to close the gap
between the current and desired level of achievement. This action is regarded as
‘feedback’ and ‘feedforward’. The appearance of this feedback in a classroom is
illustrated by steps H, J, K, L, M, T. Torrance and Pryor (1998) suggested that
possible teacher intentions of these actions were related to the enhancement of
quality of future work and creating influence on pupils’ attributions and
motivations through their increased involvement in assessing their work and
negotiating the next steps. The study by Bell and Cowie (2001) defined the action
as ‘retro-active, interactive and delayed’. It was retro-active because it involved
the revisiting of concepts and other requirements of the tasks which were not fully understood. The action was ‘interactive’ in the sense that the assessment information, which fed into the action, was gathered during an ongoing and informal interaction between the teacher and pupils (ibid). On many occasions, the action was formed immediately after the elicitation of assessment evidence (ibid). The authors suggested that the action was ‘delayed’ in the sense that the teacher did not give further information until there was a reasonable level of interest and knowledge in the class. Bell and Cowie (2001) suggested that the formative nature of this activity was evident in the increased waiting time for the increased pupil input. Further formative evidence occurred in the ways in which the teacher regulated her teaching (delaying for new knowledge introduction). This was exemplified by the inclusion of pupils’ questions as a further task requirement. The findings of Bell and Cowie (2001) are consistent with the study of ARF (1999), who reported the increased waiting time as an improvement in teachers’ formative assessment practices, as a result of their collaboration with the researchers.

The studies reviewed here indicate a strong positive effect of pupil involvement in decisions on the development of their learning. They suggest that formative assessment in the classroom appeared as both planned as well as intuitive (tacit) interactive process (Torrance and Pryor, 1998; Bell and Cowie, 2001). When teachers were not conscious of doing formative assessment, they talked about their gut feelings which lead them to change the flow of the lesson and the regulation of activities, knowing when to stand back and observe the lesson and when to initiate (Bell and Cowie, 2001).

The above part reviewed the research evidence on the appearance of formative assessment in classrooms. The following section continues to review the literature on pupils’ views of issues that are related to formative assessment.

SECTION 5: PUPILS’ VIEWS ON VARIOUS ASSESSMENT ISSUES

Pupils’ active involvement in learning and assessment is crucial for assessment for learning. The whole concept of formative assessment is based on a model in which
pupils actively engage in making sense of their learning and assessment. The way in which pupils engage with formative assessment affects its success (Black and Wiliam, 1998a). The following review covers those studies on pupils’ views that are relevant to their formative assessment experience.

Published studies investigated pupils’ engagement in tasks. Wallace (1996a) indicated that the fact that pupils carry out a task does not necessarily mean that they really engage with a task. In Wallace’s study, pupil interviews suggested that the nature of the task (activities involved) and the difficulty level were important and affected pupils’ involvement and sustaining their interest. Pupils also valued their degree of empowerment over the task (Wallace, 1996a). Additionally, if pupils had no control over their learning, the experience constituted little meaning to them thereby disabling engagement, a point also made in a similar study by Cooper and McIntrye (1993).

Subsequent studies explored pupils’ perceptions of what make good conditions for learning and teaching (Rudduck 1996a; Rudduck et al., 1996; Wallace, 1996b; Cooper and McIntrye 1993; Brown and McIntrye 1993). Wallace (1996b) interviewed pupils aged 8-10 about their interpretations of their relationship with their teacher. This was repeated when they were aged 13-15 (Rudduck, 1996a). One of the discernible aspects of this relationship was teachers’ willingness to listen to what pupils really say and respond to their learning needs (Wallace, 1996b).

Published studies explored the classroom conditions which in pupils’ view promotes learning. At primary level, McCallum et al. (2001:281), Year 2 and 6 pupils valued teaching strategies such as warm up activities which help them ‘get ready to learn’. This can be linked to Wallace’s (1996a) concept of ‘engagement to tasks’. At Key Stage 3 level and General Certificate of Secondary Education (GCSE) level, Year 9 and 10 pupils mentioned teachers establishing ‘order, security and trust’ as promoting factors in class learning. Pupils’ descriptions of order included mutually understood patterns of behaviour between teachers and pupils. Pupils also valued their teachers’ willingness to listen to and respond their needs.
The Distinctiveness of Geography

Geography offers a distinctive way of thinking about and analysing the world. Geography encourages students to identify where physical and human features of the earth are located, and analyse why they are located where they are (Rogers, 1992). The geographer identifies spatial patterns (where things are today, or were in the past, or might be in the future), and the spatial processes (why things are where they are, or were, or might be in the future) that help explain the patterns (ibid). Geography helps to better understand physical and human environments and the forces that affect them.

If students are to become better geographers, it means they learn meaningfully, think flexibly and make reasoned judgements rather than memorising facts about the world around them (Waugh, 1998). The following sections attempts to draw the potential of formative assessment to promote geographical learning. In order to achieve this, it is essential to have an appropriate means of defining the key concepts and skills that are distinctive in the learning of geography in Key Stage 3.

Geographical Concepts

Concepts in geography consist of the abstracted criterial attributes that are common to a given category of objects, events or phenomena (Ausebel et al., 1968). There have been some attempts by geographers to identify and classify the concepts that are important in the learning of geography. Catling (1976) reduced geographical concepts to three different but fundamentally interrelated items. These are spatial location, spatial distribution and spatial relations. Naish (1982) classified geographical concepts as concrete and abstract. Concrete concepts are the ones that are observable directly through senses in daily life such as traffic flows and friction of distance (Gagne 1966; Naish, 1982). However, abstract concepts are the ones which help to organise key geographical ideas such as ‘spatial interaction’ which is not directly observable (Naish, 1982: 38-9). In the following, a comprehensive...
list of key overarching geographical concepts and the relevant geographical questions that they raise are provided (Sutton, 2002).

<table>
<thead>
<tr>
<th>Key Concepts</th>
<th>Relevant Geographical Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location</td>
<td>Where are places in absolute and relative terms?</td>
</tr>
<tr>
<td>Distribution</td>
<td>What are the patterns of physical and human features on the earth’s surface?</td>
</tr>
<tr>
<td>Processes</td>
<td>What are the physical and human processes which underlie spatial patterns?</td>
</tr>
<tr>
<td>Cause and Effect</td>
<td>What are the background and trigger causes?</td>
</tr>
<tr>
<td></td>
<td>Are the effects short or long-term, permanent or temporary?</td>
</tr>
<tr>
<td>Relationships</td>
<td>How do people interact with each other and their environment, including an appreciation of opportunities, challenges and constraints?</td>
</tr>
<tr>
<td>Change</td>
<td>How and why are places changing?</td>
</tr>
<tr>
<td>A sense of place</td>
<td>What is it like to live in a particular place?</td>
</tr>
<tr>
<td>Distinctiveness</td>
<td>How and why are places different from other places?</td>
</tr>
<tr>
<td>Interdependence</td>
<td>How and why are places and people linked to and affected by processes in other parts of the world?</td>
</tr>
<tr>
<td>Sustainable development</td>
<td>How can we conserve places for future generations?</td>
</tr>
</tbody>
</table>

One of the crucial questions in the learning of geographical concepts for geography teachers could be, “What sort of skills do pupils need in order to answer the relevant key questions?” (Naish, 1982). In order to answer the question it is appropriate to analyse the nature of geographical skills that are required in the current Key Stage 3 GNC.

**Geographical Skills**

Geographical learning invites pupils to cope with a huge and diverse subject matter to think about, analyse and explain geographical patterns and issues. This requires pupils being able to analyse, synthesis, associate and classify geographical concepts (Wood, 1987). Developing key skills for the learning of geography is important in developing a specific type of reasoning that is crucial in learning about geography (Leat, 2000). A comprehensive list of “key skills” that may be applied in geography KS 3 classrooms are categorised as follows (Sutton 2000):
Table 2.9: The Key Skills in the Learning of Geography

<table>
<thead>
<tr>
<th>Skill Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Investigative</td>
<td>Ability to identify geographical questions and issues.</td>
</tr>
<tr>
<td></td>
<td>Ability to identify and collect evidence to make use of a wide variety of</td>
</tr>
<tr>
<td></td>
<td>sources of information.</td>
</tr>
<tr>
<td></td>
<td>Ability to observe, select and record information accurately.</td>
</tr>
<tr>
<td>Conceptualising</td>
<td>The ability to identify concepts and to understand the inter-relationship</td>
</tr>
<tr>
<td></td>
<td>between concepts.</td>
</tr>
<tr>
<td>Hypothesising</td>
<td>The ability to use hypotheses to predict, assess trends and evidence in order</td>
</tr>
<tr>
<td></td>
<td>to make judgements.</td>
</tr>
<tr>
<td>Planning and</td>
<td>The ability to establish an appropriate sequence for an investigation.</td>
</tr>
<tr>
<td>Monitoring</td>
<td></td>
</tr>
<tr>
<td>Interpretative/</td>
<td>The ability to give meaning to data presented in a variety of forms.</td>
</tr>
<tr>
<td>Analytical</td>
<td></td>
</tr>
<tr>
<td>Communicative</td>
<td>The ability to present information in a clear and appropriate way and</td>
</tr>
<tr>
<td></td>
<td>describe it clearly using correct geographical terminology and techniques.</td>
</tr>
<tr>
<td>Social</td>
<td>The ability to work co-operatively with others.</td>
</tr>
<tr>
<td>Evaluative</td>
<td>The ability to evaluate the validity and limitations of evidence.</td>
</tr>
<tr>
<td></td>
<td>The ability to evaluate methods of collecting and presenting evidence</td>
</tr>
<tr>
<td></td>
<td>The ability to analyse evidence and draw valid conclusions leading to</td>
</tr>
<tr>
<td></td>
<td>reasoned judgements.</td>
</tr>
</tbody>
</table>

Published articles which offer overlapping taxonomies of skills are available in the learning of geography. These taxonomies usually include some reference to sequencing and sorting, classifying, comparing, making predictions, relating cause and effect, drawing conclusions, generating new ideas, problem solving, testing solutions and making decisions (Leat, 1999). Leat suggests that this list of skills are very much associated with 'thinking skills' and suggest that their application is important for pupils to develop a clear conceptual structure of the subject (Leat, 2000). Leat (2000:141) suggest that a conceptual structure is different from a factual knowledge since it provides a mental model that could change in different contexts. Implicit in this definition is the transferability of the use of skills to understand new issues.

The communicative skills outlined in the above table require decisions on the use of appropriate tools to communicate geographical information and understanding.

The common tools available in presenting geographical information are graphicacy (Maps, diagrams, photographs, models) literacy and numeracy (Balchin, 1973). Additionally, oral activities could be a medium for communicating geographical knowledge and understanding.

The GNC (1995) that was available when this research was conducted emphasise geographical enquiry skills for all Key Stages with the emphasis on the geographical questions, enquiry skills and related specific skills. Indeed, the
geography level descriptions which constitute a major focus of this doctoral thesis appear to be constructed on research skills. The following table illustrates the required skills to achieve each level and how they progress (QCA1998, p5):

<table>
<thead>
<tr>
<th>Levels</th>
<th>Enquiry Questions</th>
<th>The use of enquiry skills</th>
<th>The use of specific skills</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 1</td>
<td>respond to questions</td>
<td>make observations use the resources provided</td>
<td>express views</td>
</tr>
<tr>
<td>Level 2</td>
<td>ask and respond to questions</td>
<td>select information from sources provided</td>
<td>express views and begin to use appropriate vocabulary</td>
</tr>
<tr>
<td>Level 3</td>
<td>respond to a range of questions</td>
<td>use skills and sources of evidence</td>
<td>offer reasons for observations and judgements</td>
</tr>
<tr>
<td>Level 4</td>
<td>suggest suitable geographical questions for study</td>
<td>use a range of skills and evidence</td>
<td>communicate findings using appropriate vocabulary</td>
</tr>
<tr>
<td>Level 5</td>
<td>identify relevant geographical questions</td>
<td>select and use appropriate skills and evidence</td>
<td>reach plausible conclusions and present findings both graphically and in writing</td>
</tr>
<tr>
<td>Level 6</td>
<td>identify relevant questions and suggest appropriate sequences of investigation</td>
<td>select and use a wide range of skills and evidence</td>
<td>present conclusions that are consistent with evidence</td>
</tr>
<tr>
<td>Level 7</td>
<td>with growing independence, identify geographical questions, establish a sequence investigation</td>
<td>select and use accurately a wide range of skills and evidence</td>
<td>begin to reach substantiated conclusions</td>
</tr>
<tr>
<td>Level 8</td>
<td>show independence in identifying appropriate geographical questions and implement an effective sequence of investigation</td>
<td>select and use effectively and accurately a wide range of skills</td>
<td>reach substantiated conclusions</td>
</tr>
<tr>
<td>Exceptional performance</td>
<td>undertake geographical enquiries independently</td>
<td>use accurately a wide range of skills and evidence</td>
<td>reach substantiated conclusions, presented effectively and accurately, evaluate work by suggesting improvements in approach and further lines of enquiry</td>
</tr>
</tbody>
</table>

As indicated earlier, the use of skills enable pupils to grasp a subject specific reasoning to develop their understanding. Having briefly outlined the range of geographical skills, the next stage involves devising assessment tasks which incorporate the use of these skills. The following section provides the links between geographical learning and formative assessment and explores the potential and pitfalls that formative assessment could face in the learning of geography Key Stage 3.
Implications for Formative Assessment in Key Stage 3 Geography

The literature reviewed on formative assessment has established that learning and assessment are symbiotic. Having accepted that the role and value concept and skill development in geography education is inevitable, then teachers must find ways to develop it. However, developing pupils’ geographical concepts and skills could depend on the degree of teachers’ understanding as well as their ability to promote their development. Assessment has the potential to promote geographical skill development by making pupils’ geographical thinking more visible and explicit (Leat, 2000). This suggests that teachers need to develop means of discovering and understanding pupils’ structures of thinking in geography. Therefore, it is also vital for teachers to understand the nature of geographical concepts, skills and how they progress. The final part of the section attempts to find explanations for how two components of formative assessment (assessment tasks and teacher feedback) could be used to support geographical learning.

Assessment Task Design

The previous section defined geography as a subject which has a distinctive way of analysing issues, as having a problem-solving methodology and involving research skills in understanding concepts and their relationships (Day, 1995:90). The literature reviewed highlighted a number of different approaches to assessment for learning which could have implications for the promotion of pupils’ geographical learning. An obvious implication of alternative approaches for assessment for geographical learning could be the potential of implementing ‘authentic assessment’ (Stimpson, 1996). Authentic assessment is defined as an attempt to evaluate learners’ ability to use previously acquired knowledge in the solving of problems or in the completion of specific tasks that are set in real life contexts (Stiggins and Bridgeford, 1982). As opposed to some other subjects, geography lessons could be extended into real-life contexts through fieldwork where pupils engage in active investigation of their environment. The value of such tasks relates to engaging students in the challenge of producing, rather than reproducing knowledge (Newmann and Archbald, 1992).
A strong justification for geography’s connection to real world could also be related to the very nature of the subject, which is concerned with the critical environmental and ecological issues. The increasing importance of conflicts of land use and the increasingly complex web of human and environmental interactions are (Day, 1995:90) all in the scope of KS 3 Geography Devising tasks to offer pupils opportunities to understand these geographical issues could not be possible by only gathering and using facts concerned with places and ideas (Stimpson, 1996). Rather it requires the design of assessment tasks which offer opportunities for pupils to analyse what is going on, explaining why, and proposing alternative strategies in dealing with conflicting issues (Day, 1995:90). Here, the focus is not only the quality of solutions that pupils generate but also the learning process that give rise to these solutions. The literature refers to certain tasks such as concept mapping as having the potential of exploring the extent to which pupils engage in thinking abstractly and logically about their physical, economic, social and political environments (Resnick and Resnick, 1992; Novak and Govin, 1984). Therefore, tasks for the purposes of formative assessment which are concerned with promoting and improving pupils’ learning should enable teachers to observe how pupils approach the task (Eisner, 1993). In other words, formative assessment tasks for the learning of geography should aim to develop pupils’ ability to learn how to learn about the world around them (Stimpson, 1996:393).

Geography as a school subject no longer deals only with the facts related to “facts” and “theories” for explaining the physical world but encapsulates more socially sensitive issues such as immigration, citizenship and racism (Morgan and Lambert, 2001). This requires geography teachers to devise tasks which enable the exploration of different values and points of views and acknowledge the possibility of accepting more than one explanation or solution to a geographical issue or problem. Assessment tasks of such nature also require teachers to develop assessment criteria for valuing and assessing pupils’ thoughts when dealing with such delicate issues.
Teacher Feedback

The literature on formative assessment has not specialised to explore subject specific aspect of each component of formative assessment. However, it is possible to make suggestions on the nature of formative assessment feedback to support geographical learning. The geographical learning described in this section emphasise the importance of understanding of geographical concepts and the use of skills in exploring geographical issues and addressing geographical questions. Feedback supporting this process therefore should address specifically pupils’ conceptual understanding and promote their skill development. This thesis argues that the whole model formative assessment is based on a model of progression which is concerned with promoting advancements in pupils’ learning. It is also recommended that the three components of formative assessment including teacher feedback should be tailored around a well-defined map of geographical learning within the current frameworks available.

The context for the implementation of formative assessment in the classroom is the Geography National Curriculum (GNC) in Key Stage 3. The following section explores the issues of continuity, progression and differentiation within the GNC and discusses their implications for formative assessment.

Continuity & Progression in the GNC and Formative Assessment

A requirement of any curriculum planning model is the need to plan for continuity and progression (Hughes, 1996). Continuity relates to the maintenance and development of a range of aspects of geographical education within the geography curriculum over a period of time (Butt, 2002:69). The most widely recognised aspects of continuity are geographical content, teaching and learning activities and assessment procedures (ibid). The broad use of the term ‘progression’ refers to the measurable advances in knowledge, understanding and skills made by pupils (Butt, 2002:76).

The GNC generates a model of progression which was informed to some extent by the Draft Proposals for Geography in the National Curriculum, (SCAA, 1994).
This document stated that progress in pupil attainment throughout the Key Stage (quoted in Bennetts, 1996:85) could be measured by the extent to which pupils:

- Broaden and deepen their knowledge and understanding of places and themes;
- Make use of a wide range and precise vocabulary;
- Analyse, rather than describe, geographical patterns, processes and change;
- Appreciate the interactions within and between physical and human processes that operate in any environment;
- Appreciate the interdependence of places;
- Become proficient at conducting and comparing studies at a widening range of scales and in contrasting places and environments;
- Apply their geographical knowledge and understanding to unfamiliar contexts;
- Select and make effective use of skills and techniques to support their geographical investigations;
- Appreciate the limitations of geographical evidence and the tentative and incomplete nature of some explanations.

Planning for progression in pupils’ learning within the GNC is facilitated by the use of level descriptions (Rawling and Westaway, 1996). Within the current GNC, the level descriptions highlight the range of expected performance that pupils should demonstrate when they progress from one level to another. This requires teachers to make sense of level descriptions in order to plan for progression both between different levels and within one level. Exploring how teachers make sense of level descriptions and how they use them for planning for progression within Key Stage 3 Geography is one of the objectives of this study.

**Planning for Continuity and Progression in Pupils’ Learning**

Planning for progression in pupils’ learning in geography is regarded as one of the challenging aspects of geography teaching and learning (Butt, 2002). Although the existence of guidance materials by the Geographical Association (Butt et al., 1995), the literature lacks research studies in the area of progression in geographical learning. Bennetts (1996:82) outlined some principles to guide planning for progression.
Whenever possible, teaching should build upon pupils’ existing knowledge and previous experience;

- Learning tasks should be matched carefully to pupils’ capabilities;
- The overall scheme for a subject should take account of ways in which pupils mature (intellectually, socially and physically) over their period of secondary education;
- Special attention should be given to progression in those aspect of a subject which are likely to be important to pupils’ future learning, for example, ideas and skills which have wide application and which underpin more advanced learning in that subject.

If planning for progression is concerned with promoting advancement in pupils’ learning, then assessment has a central role to play since its very basic definition relates to measuring such advancements (Butt, 2002). Indeed, the whole model of formative assessment is built upon promoting progression in pupils’ learning. Central to the achievement of this, teachers need to know what constitutes learning for individual pupils and how it is progressed throughout the key stages.

**Differentiation and Formative Assessment**

The focus of this section is an exploration of the curriculum planning principle of ‘differentiation’ and a review of the strategies available for differentiating classroom practices. A review of the literature shows that differentiation is closely linked to formative assessment.

One of the key underpinning principles of the GNC is the need for it to be accessible to all pupils. This is evident in the opening statement of the document “The programme of study for each key stage should be taught to all or the great majority of pupils in the key stage, in ways appropriate to their abilities” (DES, 1995:1). This requirement implies a differentiated classroom with provision matching differing learning needs (Waters, 1995). The fundamental principle of differentiation relates to maximising and optimising learning conditions in
consideration of variations amongst pupils and their learning needs (Battersby 2000; 2002).

In describing differentiated classroom practice, there are a number of possible strategies. In geography teaching, Waters (1995) outlined a number of principles for differentiation under the headings of ‘planning’, ‘teaching’, ‘resources’ and ‘pupil needs’. Planning involves clarification of the aims and objectives of the tasks and making decisions about task design. Differentiated teaching involves the use of a wide range of activities, styles, regulation of the pace of the lessons, the balancing of the questions and the creation of a positive learning environment. Differentiation by resource involves making available a range of materials taking account of the need for clarity, format and language to increase access for pupils with differing ability groups, background and gender - a point which was reiterated by Norton (1999). Finally, all of the three differentiation strategies require diagnosis of pupils’ differing needs. Battersby (1995) extended some of the principles and in discussing some practicalities of its application, placed differentiated assessment practices and the provision of assessment feedback as key factors in achieving effective differentiation. Differentiation strategies were also reviewed in an influential article by Davies (1990).

The following table summarises the different strategies and their capacity to support the process of formative assessment on the basis of the work of Davies (1990) and Battersby (1995).
<table>
<thead>
<tr>
<th>Different Differentiation Practices</th>
<th>Definition</th>
<th>To what extent is it formative?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Differentiation by outcome</td>
<td>Differentiation is represented by the different responses that pupils make to common stimuli and tasks.</td>
<td>This practice of differentiation is not formative since there is no intervention during the process of learning to change the learning outcome.</td>
</tr>
<tr>
<td>Differentiation by resources</td>
<td>All pupils could be working on the same task but with a range of different resources.</td>
<td>Providing a variety of resources by taking into account of differences in pupils' reading, understanding and interpretation certainly could increase their engagement in the task. However, just providing a range of resources does not necessarily result in improvements in learning. Increasing access for pupil engagement through varied resources could be just an initial stage of formative practice.</td>
</tr>
<tr>
<td>Differentiation by graded tasks</td>
<td>Pupils may use the same stimulus, resources but they will follow a series of tasks or questions, which become increasingly demanding and difficult.</td>
<td>Just altering questions will not guarantee improvement in pupils' learning. Teachers could gain some insightful information about the common mistakes that pupils make and they can act upon this evidence. What matters for formative assessment is that the need of developing the ways in which pupils go about answering such questions, supporting them during their answering with a view to understanding pupils’ learning process and altering their learning during the learning process.</td>
</tr>
<tr>
<td>Differentiation by task</td>
<td>Pupils may use the same stimulus, resources but they will have some options available to produce their own product (i.e choosing from a choice of writing a letter, producing a pamphlet or a poster).</td>
<td>Amongst the other strategies available this can serve formative purposes by increasing opportunities for pupils to demonstrate what they are capable of doing. Devising tasks for differentiation is about thinking about optimising and maximising learning opportunities and it could inform medium and long term strategic planning of formative assessment.</td>
</tr>
<tr>
<td>Differentiation by stimulus and by task</td>
<td>Here, not only the tasks but also the teacher initial stimulus are varied.</td>
<td>Planning differentiation at a stimulus level can be achieved only if teachers have developed a detailed and accurate knowledge of pupils differences in their learning (Lambert and Balderstone, 2000). Again, stimulus is the first stage of pupils' engagement in the task. For formative purposes, pupils need to be followed up, have their learning needs met and the whole learning process supported.</td>
</tr>
</tbody>
</table>

What becomes apparent from the above analysis is that not every differentiation practice can support formative assessment. Differentiation for formative purposes should be a ‘process’ involving a detailed analysis of pupils’ differences and learning needs and the development of ways of supporting pupils’ learning. Just devising tasks involving items of varying difficulty levels or a variety of modes of presentations also will not guarantee formative assessment. In order for differentiation to serve formative assessment, teachers need to develop and provide differentiated feedback for tasks. They also need to develop an understanding of the ways in which the outcome of such tasks could differ and how they could inform individual target setting for improving pupil learning. To sum up, the
extent to which differentiation enables formative practice is dependent upon whether it contributes to teachers’ understanding of the pupils’ learning process and whether it can provide an appropriate route that teachers follow to monitor and support the learning process. Differentiation for the support of formative assessment also requires teachers to be aware of the occurrence of a variety of assessment evidence which they need to interpret based on their own discretion. This clearly puts more demands on them.

Thus, formative assessment is about how teachers handle and manage differentiation effectively. Lambert and Balderstone (2000) suggest that the success of differentiation is dependent upon teachers’ ability to constantly monitor the classroom and to know when to step back and when to intervene. They suggest that this is possible when teaching becomes ‘a reflective practice’ and when teachers also become active assessors of their own lessons.

CONCLUSION

The literature review chapter provided a discussion on theory and current developments in research into assessment for learning. As a first step, alternative approaches to assessment for learning were explored with a view to elucidating how each of them contributed to the current understanding of formative assessment. This chapter further illustrated the ways in which different learning theories lead to particular models of formative assessment. It then used the model which was illustrated in the introduction chapter and explored the specific components to deepen the current understanding of formative assessment.

The literature review demonstrated that the current understanding of formative assessment is focused on three key components, namely, assessment tasks, teacher feedback and pupil self-assessment. The research studies as well as the major reviews in the area, mainly focused on the last two components by investigating their specific characteristics. The aspect which is given less emphasis in the literature is the nature of assessment tasks and the ways in which they interact with the other components in shaping the quality of formative assessment. This thesis
explores how teachers approach task design and investigates their motives and priorities in devising tasks for enabling formative assessment.

One aspect which became visible in the literature reviewed was how formative assessment is shaped by the national framework within which it operates. As indicated in the introduction chapter of this thesis, within the National Framework in England, for the core subjects, formal end of Key Stage 3 tests accompany teacher assessment. In non-core subjects including geography, the assessment for the end of the key stage is entirely in the hands of teachers. Deriving a summative judgement about pupils’ progress has presented a challenge and it has not been addressed comprehensively in the literature. The key works in this area (Black and Wiliam, 1998a) recognised the issue but the ways in which this manifests itself in current formative assessment practices has not been researched. This thesis tackles this issue by disentangling the role that assessment tasks play over the entirety of Key Stage 3 in answering the feasibility and compatibility of formative assessment with summative assessment within the current national framework.

The theoretical models of formative assessment feedback reviewed in the chapter are very much informed by the oral forms of it, which stemmed from classroom observations. The potential role of written feedback has not been addressed extensively. With the introduction of formative assessment discourse, there has been a movement from ‘feedback for recording achievement’ to ‘feedback to improve learning’. Addressing the changed perceived role of written feedback and the ways in which it is used to alter the learning gap therefore could contribute to filling the existing gaps in the assessment literature.

The conceptualisation of formative assessment as well as alternative approaches reviewed in this chapter stated that the role of the teacher in assessment is crucial. The theoretical positions implied certain responsibilities for teachers to perform assessments in distinct ways as ‘mentors’, ‘apprentices’ and so on. Parallel to this, the National Curriculum requires that teachers fulfil certain requirements. A number of valuable hypotheses on how teachers should support the assessment process have been proposed. However, the literature has paid relatively little attention to how teachers process and interpret such requirements that take the
form of their formative assessment practices in their classrooms. The literature reviewed lacks the middle link in the chain, which is teachers’ own identity as assessors and teachers, that is positioned between policy statements and guidance materials and their formative assessment practices in classrooms. The ways in which teachers themselves conceptualise and reflect back their formative assessment practices is therefore a major objective of this research project.

Finally, published research on formative assessment tends to focus on how it appears in classrooms in a limited periods of time by focusing on key events and strategies that enable formative assessment. Not enough emphasis has been devoted to ‘strategic planning of formative assessment over an extended period of time. Thus, one of the priority objectives of this thesis is to fill in this gap in the context of geography teaching by exploring the processes involved in the medium and long term planning for formative assessment.

**Summary Outline of the Literature Review Chapter**

This chapter reviewed the theoretical and research literature on formative assessment. Particular emphasis was given to alternative movements to assessment for learning and the ways in which they aided the understanding of formative assessment. The chapter explored formative assessment by breaking it up into three key components: assessment tasks, teacher feedback and feedforward, and pupil self and peer-assessment. This chapter also examined the GNC as a context for formative assessment and identified the concepts of differentiation and progression as crucial in implementing formative assessment in Key Stage 3 geography. Before presenting the findings on how teachers and pupils viewed the selected aspects of formative assessment, the following chapter (Chapter 3) gives an overview of the research approach and methods used for this study and describes the research process in detail.
CHAPTER 3: METHODOLOGY

Introduction

The focus of this section is a personal expression of the philosophical underpinnings of this research project, in particular the epistemological and ontological assumptions that underlie the methods applied. It has been argued that neglected and unjustified epistemological assumptions result in perceiving research as a ‘technology’, more precisely the mechanistic application of research methods (Usher, 1996:9). The aim of this chapter is not to completely divorce research assumptions and methods but to create links between epistemology and method with reference to the practicalities of research.

The above-mentioned position requires the researcher to answer questions such as ‘What does constitute research evidence?’, What is reality, if there is any?’ , ‘What does count as research data?’, ‘What is the researcher’s role in generating research evidence?, ‘What is the relationship between the researcher’s assumptions and the collection and analysis of data?’ Through reflecting on the questions, the chapter then provides the rationale for evaluating this research with a focus on validity and reliability and shares the challenges of finding answers to these significant questions and issues.

Description of Research Approach and Methods

By looking at the character of the overall purposes and sub-purposes, a researcher could make some preliminary judgements regarding an appropriate research approach. Research can be conducted with a highly structured fashion or a totally grounded one and each of the approaches makes certain assumptions about what counts as research data. The choice of research approach is, therefore, a critical decision. Before explaining the rationale for a qualitative approach for this research, the following figure summarises the research process and methods employed:
The above figure describes the major steps of this research project. The research did not follow a complete linear process between the steps and it was intertwined. For instance, although the research questions for teacher interviews were mainly informed by the literature review, the initial interviews helped the researcher to refine the questions throughout the research process. New questions were added and some others were abandoned. Data analysis was ongoing and it started from the first interview. Overall, the research data was generated constructively, around a semi-structured framework, by each step informing the next. Data collection and analysis were, therefore, iterative processes.
Having completed the teacher interviews, the researcher had to make a choice about the nature of classroom observations. She could have either observed each teacher for a limited period of time or agreed with one teacher and observed for an extended period. Each would provide different understandings of the research problem. Observing one teacher through an assessment task seemed to offer better opportunities to extend some emerging issues from the teacher interviews. Teacher 3 was selected and he agreed to participate. Classroom observations were followed by pupil interviews. The pupils for this study were selected by the teacher and the researcher and they were interviewed after the completion of classroom observations.

This research data was generated through the use of a number of research methods. To reiterate, the research methods used were as follows:

- Teacher interviews were conducted to explore teachers’ views of their formative assessment practices;
- Classroom observations were conducted to describe one selected teacher’s assessment practices as well as to contribute to the understanding of selected assessment issues that emerged from teacher interviews;
- Pupil interviews focused on pupils’ views of assessment issues and explored their reflections upon their assessment experiences.

SECTION 1: EPISTOMOLOGICAL AND ONTOLOGICAL ASSUMPTIONS OF THIS RESEARCH PROJECT

The starting point of this section is to reflect upon the extent to which assumptions affect the methods chosen - the ways in which the underpinnings of research methods manifest themselves within the entire research processes.

Research paradigms are sets of ideas that constitute a system which gives us some judgement about the nature of reality, or a reason why we must be content with knowing something less than the nature of reality (Reese, 1980). Paradigms can be characterised by the way their proponents respond to the following three concepts:
- Ontology: What is the nature of the knowledge? or What is reality?
- Epistemology: What is the relationship between the knower and the known? How it is possible, if it is, for us to gain knowledge of the world? (Hughes and Sharrock: 1990:5)

With their underlying assumptions and philosophies, paradigms comprise sets of ideas which underpin a model or general approach to be taken in breaking down the complexity of the real world (Patton, 1990). These ideas or beliefs set the guidelines for investigative methods to be used and for the means of interpreting and analysing the research which is conducted (Husen, 1994).

There are two main paradigms employed in social science research (Husen, 1994). On one side, logical positivism uses quantitative approaches to test hypothetical-deductive generalisations, and on the other hand phenomenological inquiry uses qualitative and naturalistic approaches to inductively and holistically understand human experience in context-specific settings (Patton, 1990). Each approach has implications not only on what counts as research evidence but also on how a researcher finds out about it.

The first way in which research assumptions and methods are related can be justified on the basis of how human beings are viewed. The question here is whether humans are active meaning makers or whether their meanings are determined by the social structures (Holdaway, 2000). Each way of viewing human beings implies a particular orientation and structure to the research process. One view, suggested by Durkheim (1952) on the basis of his study of suicide rates, suggested that individual action is a result of changes in social structures. Effectively, he argued that social “facts” such as rates of suicide could be treated as objects with changing proportions which are measurable (Holdaway, 2000:156). Here, individuals are seen as passive, and they are more determined by social structures than determining their own actions (Holdaway 2000). The second view of humans is as active meaning makers. Meanings are created through humans perceiving and responding to situations which involve negotiation with others.
So, the ramification of this is that any relationship which can be conceptualised in terms of social processes can have diverse meanings for those involved in it (ibid).

So, what does all of these mean for a research situation? The research method which would follow from the first position (viewing humans as passive) would have a reductionist approach to explain the social processes. In order to measure the social processes, the predefined variables will be placed in the equation. What constitutes research data is then what fits strictly in this predefined equation and structure. For the second position, the researcher has a task of explaining and describing the individuals and situations in consideration of the context in which they construct their meaning.

The Role of Researcher

Producing social explanations through qualitative research requires explanations about the role of researchers in the research process. Lincoln and Guba (2000) refer to unpacking the researchers’ control over the research process as a critical issue in eliciting the role of researchers. They identify a number of key questions for researchers to ask:

“Who initiates? Who determines salient questions? Who determines what constitutes findings? Who determines how data will be collected? Who determines in what forms the findings will be made public, if at all? Who determines what representations will be made of participants in this research?” (Lincoln and Guba, 2000:175).

From a positivist paradigm the researcher’s task is to find out about the social phenomena by keeping a distance to eliminate their own assumptions and values. They are not considered to be a part of the research and their presence is not expected to affect the quality or type of data gathered (Scott, 1996:155).

Implicit in the above questions is that researchers’ control over the research process has implications as to what counts as research data and how researchers generate and communicate it. In order to disentangle different voices in research, Lincoln and Guba (2000) suggest reflexivity. They define reflexivity as ‘the
process of reflecting critically on the self as a researcher, the “human as instrument” (Lincoln and Guba, 2000:183). This process involves an exploration of the researcher’s multiple identities that the researcher brings into the research and the ways in he/she shapes the research process (ibid).

The research process in this study has been affected by the researcher’s decisions at every stage. Denzin (2002:349) defines these steps as an interpretive process in the following way:

- Framing the research question;
- Deconstructing and analysing critically prior conceptions of the phenomenon;
- Capturing the phenomenon, including locating and situating it in the natural world and obtaining multiple instance of it;
- Reducing the phenomenon, or reducing it to its essential elements;
- Constructing the phenomenon back in terms of its essential parts, pieces and structures;
- Contextualising the phenomenon, or relocating the phenomenon back.

In this research, the researcher was an external being to research participants. Coming from a higher education institution as well as being educated in another country contributed to this situation. However, the research was neither value free nor totally grounded. The researcher had brought her own intellectual ‘baggage’ consisting of her conceptualisation of what constituted formative assessment, which was mainly informed by the literature review. Furthermore, her agenda was partly defined and structured in the same way that Denzin defines the process of formulating research questions. This gave the researcher partial control over the research process, not only in terms of the kind of data generated but also what counted as research findings. However, this was not a fully controlled process. During the data collection, the researcher tried to create room for participants to prioritise their answers. During the analysis, their points were fitted into the structure and were compared and contrasted with other teachers’ responses, regardless of whether they fit in the structure. Reflecting upon the research experience, participants were given procedural autonomy to contribute to the
research process. The meaning making was mutually constructed and participants’ meanings were interpreted by the researcher.

**Conceptions of Social Reality and Knowledge**

Social reality is treated differently by different research paradigms. The status of social reality is linked to ontological assumptions, which question its nature and form (Hughes, 1990). Positivist social scientists suggest that social reality is external to the individual and imposes itself on human consciousness from outside (Cohen and Manion, 1985). This refers to a realist view of objective reality, which is related to the assumption that the world exists and is knowable as it really is (ibid). Here, the world has a real existence outside the human experience (Lincoln and Guba, 2000:176).

An epistemological formulation that connects positivism suggests that knowledge is hard and tangible (Cohen and Manion, 1985). Here, the social world consists of a system of variables, which are distinct and analytically separable parts of an interacting system (Popkewitz, 1984:37). In positivist ways of knowing “variables are measured independently of other elements” of the social system (ibid) and human phenomena can be explained by counting frequencies with which the phenomena occur and establishing causal relationships (Glesne, 1990).

According to the interpretivist paradigm, the nature of social reality is internal to the individual and reality is a personal construct (Cohen and Manion, 1985). Here, individuals and social settings are interdependent (ibid). The interpretative ontological assumption suggests that “realities are wholes and cannot be understood in isolation from their contexts, nor can they be fragmented for separate study of the parts” (Lincoln and Guba, 1985:39). This assumption is informed by a specific way of viewing human nature, that humans construct their interpretations of the world they experience (Cohen and Manion, 1985). Therefore, social reality is context dependent and its interpretation is pluralistic (Greene, 1990).

The underlying assumptions about the form and nature of interpretivist knowledge points out values that connect individuals to contexts around them (Grumet, 1990).
Interpretative knowledge is value impregnated (Scott, 1997). Data is grounded in the context (Strauss and Corbin, 1998). Also, interpretative knowledge represents emic knowledge that is the understanding of the perspectives and meanings of those in the setting being studied (Stake, 1983, quoted in Greene, 1990:235). Knowledge is dependent on people’s values, meanings and interpretations. Therefore, knowledge exists in relation to contexts and discourses and it is partial and perspectival (Scott, 1997:31).

A Rationale for Qualitative Methodology

In the previous part, the ontological and epistemological underpinnings of this research are defined. This part examines the rationale for a qualitative methodology on the basis of its various capacities and their relevance to the current project.

Methodology refers to the logic of a research process in terms of how a researcher develops a strategy to find out the research issue (Blaikie, 1993). Methodological decisions are informed by ontological and epistemological assumptions as well as the nature of the research issue. Denzin and Lincoln (2003:4-5) define this logic for qualitative research as follows:

“Qualitative research is a situated activity that locates the observer in the world. It consists of a set of interpretative, material practices that make the world visible. These practices transform the world. They turn the world into a series of representations, including field notes, interviews, conversations, photographs and memos to the self. At this level, qualitative research involves an interpretative, naturalistic approach to the world”.

The “situated nature” of qualitative research implies that decisions about design and strategy are grounded in the practice, process and context of the research itself (Mason, 2002:24). Grounding research in a context requires study of situations in their natural settings (Denzin and Lincoln, 2003). Hammersley (1992) refers to this as a preference for naturally occurring data. On the basis of the same assumption, Manson (2002:24) suggests that qualitative researchers apply specific thinking about research design and strategy:
“Thinking qualitatively means rejecting the idea of a research design as a single
document which is entire advance blueprint for a piece of research. It also means
rejecting the idea of a priori strategic and design decisions, or that such decisions can and
should be made at the beginning of the research process. This is because qualitative
research is characteristically exploratory, fluid and flexible, data-driven and context-
sensitive”.

This research study is interested in teachers’ and pupils’ interpretations of
assessment issues. Chapter 1 defined formative assessment as one of the
Government’s key strategies in raising the standards of education. It is also the
interest of this research to explore how geography teachers and pupils construct
and deconstruct the policy, how they adapt in their daily classrooms and what
meaning they make out of it. In bridging policy interventions and classroom life,
Hargreaves (1978) advocates qualitative research for its “immunological capacity”
and suggests that “qualitative research helps to provide us with the necessary
immunological understandings; for only when we understand the precise nature of
the host body can we design our innovatory grafts with confidence that they will
prove to be acceptable” (Hargreaves, 1978, quoted in, Hammersley 2000:397).
Then, qualitative research could be an appropriate way of understanding and
documenting teachers’ and pupils’ views and experience of formative assessment
in terms of how it fits their daily classroom life.

In order to enable the exploration of teachers’ and pupils’ perspectives, the design
of the current research allowed space to construct their meanings. This was
achieved by a semi-structured framework for teacher and pupil interviews. The
same situation applied to the observational research. Here, qualitative research is
also better positioned in understanding formative assessment within the context in
which it occurs, and the interplay between the characteristics of the context and
how various agents of the context respond to it. As Hammersley (2000:395)
suggests “the aim is not to judge, but to understand”.

Having described the rationale for qualitative methodology, the following section
explores the reasons for the use of interviews and observation as data collection
methods in this study.
SECTION 2: A RATIONALE FOR THE CHOICE OF RESEARCH METHODS

A Rationale for the Use of Interviews as a Research Method

Interviewing is most commonly defined as a research method to provide data on understandings, opinions, attitudes, and feelings (Kvale, 1996). Interviews are varied in terms of their structure and the level of power that is exercised by the interviewer (Denscombe, 1998). On one hand, interviews could have no structure with the respondents having flexibility over how they construct their meanings (Arksey and Knight, 1999). In this case, the flow of the interview is determined by the respondent and the spontaneous generation of questions during the interview (Patton, 1990). Conversely, interviews could be oral processing of questionnaires, with a strict structure (structured interview). In between, there are semi-structured interviews in which the respondent constructs her/his answer within a loosely framed structure.

The degree of structure of interviews affects the interview process in terms of how the interactions are shaped, the extent to which interview is controlled and the nature of research data generated as a result. Arksey and Knight (1999:7) suggest that in highly structured formats, interview questions are determined in advance and interviews must strictly stick to a script. Here, the interview process is highly controlled by the content (Briggs 1986). The aim of structured interviews is therefore to generate research data which tries to minimise the effect of the interviewer and the research setting (Silverman, 2001:88). Therefore, the validity of structured interviews depends on the successful application of the same procedures (Kirk and Miller, 1986). Here, the assumption is that the reality is “out there” and the interviewer should elicit it without contaminating what is being said. Holstein and Gubrium (2002:115) defines the interviewees’ roles as “vessels of answers” who are also epistemologically passive, and not engaged in the production of knowledge.

The latest trends in interviewing have come some distance from structured questions; as qualitative researchers are realising that interviews are not neutral...
tools for data gathering but are active interactions between two or more people leading to negotiated, contextually based results (Denzin and Lincoln, 2000:646). Holstein and Gubrium (1997; 2002; 2003) view interviews as active and collaborative meaning making occasions. Silverman (2001) defines such interviews as constructionism and suggests that meaning is created actively by the participants of the interview process. In this process, social reality is constructed as a result of the interview process (Holstein and Gubrium, 1997). What distinguishes this approach from a structured one is the dynamics of the interview process, as suggested by Mishler (1986:53-4):

"Rather than serving as a stimulus having a predetermined and presumably shared meaning and intended to elicit a response, a question may more usefully be thought of as part of a circular process through which its meaning and that of its answer are created in the discourse between interviewer and respondent as they try to make continuing sense of what they are saying to each other....respondents learn from how interviewers respond to their answers – restating or rephrasing original question, accepting the answer and going on to the next question, probing further information – what particular meanings are intended by questions and wanted in their answers in a particular interview context".

The way in which the data is gathered through interviews has further implications for data analysis. From a structured point of interview, a common approach refers to the process of categorising, coding and then summarising descriptions on what has been said. If the interview process is an essential part of the meaning making, then reflections upon the interview process is necessary to unpack not only what has been said but also how it has been said (Holstein and Gubrium, 2002).

The researcher’s previous arguments of what constitutes research evidence, and her way of viewing human beings as ‘meaning makers’ strongly suggested that interview would be an appropriate research tool. However, matters were not only related to the researcher’s identity and assumptions (‘who I am as a researcher’), but it is also the actual research questions (‘what I want to find out’). Interviewing could be a valuable option when the research is concerned with the understanding of people’s experience and the meaning that they make out of their experience (Seidman, 1998:3). As outlined in Chapter 1, this research is interested in finding out individual teachers’ and pupils’ stories about assessment. The key research questions were mainly concerned with how individual teachers viewed their assessment practices, how they judged themselves and what meanings they
attached to their assessment. So, the research is concerned with experiences, motivations and self-reflection on these. Therefore, any research approach involving testing predefined theories or filling strictly defined categories of what constitutes this experience was rejected. For this reason a survey approach was not adopted in this study. When compared to other means of research methods, interview seemed to be better suited to absorbing individuals’ insights and reflections because its flexibility renders the capacity to allow respondents to express their thoughts and feelings. However, the extent to which this potential was fulfilled was determined by the dynamics of the interviews, as reviewed in detail in the next section.

The appropriateness of semi-structured interviews to this study is also justified by the nature of the present incomplete understanding of formative assessment in Key Stage 3 Geography. As Chapter 2 highlights, the conceptualisation of formative assessment is ongoing and there is not an agreed definition that encapsulates its meaning for all of subjects. Having outlined the lack of research describing the nature of formative assessment for Key Stage 3 Geography, grounding this research on a highly structured framework neither was possible nor desired. This research rather aimed at exploring how it looks like in Key Stage 3 Geography. However, the researcher did take into account what is known about formative assessment and built the exploration on a loose framework which was informed by the initial literature reviewed.

Another reason for the use of semi-structured interviews related to the researcher’s past research experience. Given the researcher’s lack of interviewing experience, the use of semi-structured interviews was valuable in providing prompts for fluently perpetuating the interview process. Thus, being a novice researcher, as well as being an outsider, meant that the use of semi-structure interviews was considered a relatively safe option.

A Rationale for the Use of Observations as a Research Method

Observation is commonly defined as a research method which involves the immersing of researchers in a ‘setting’ to generate data through experiencing and
observing that setting at first hand (Manson, 2002:84). Here, the immersion in the social setting is crucial to capture naturally occurring phenomena within the setting in which it occurs (Denzin, 1970; Erikson, 1990), rather than understanding it in deliberately created environments such as within experimental studies. Selecting observations for data generation coincides with a view that social situations can be understood through making sense of them in consideration of the setting specific factors. Here, the epistemological assumption suggests that the social world can be understood by observing, by participating and by experiencing “real life” settings, and interactions (Manson, 2002:85). The ontological perspective recognises the centrality of interactions, actions, and behaviours to make sense of social reality (ibid).

Observations can vary in terms of their level of structuring. Structured observation, which has associations with the positivist paradigm, involves assigning certain types of events or behaviours to previously defined categories (Croll, 1986). By using a pre-determined classification, the aim is to eliminate possible variations in the perceptions of individuals (Denscombe, 1998). Observations can also be conducted without a structure, sometimes the researcher being participant, with a view to understanding the meanings of social relations and social processes in the settings in which they occur. Again, the choice is dependent on how the social world is understood in relation to the research questions.

Generating research data through observations raises a number of key issues mainly in relation to the researcher’s role and the notion of ‘naturally occurring’ social settings (Coffey, 1999; Atkinson and Coffey, 2003). Atkinson and Hammersley (1994:249) recognise the importance of defining the identity of researchers and raising a number of key questions for researchers to reflect upon:

- Whether the researcher is known to be a researcher by all of those being studied;
- How much, and what, is known about the research and by whom;
- What is the orientation of the researcher, and how completely he or she consciously adopts the orientation of insider or outsider.
Although the above questions are commonly raised in relation to participant observation, Atkinson and Coffey (2003) recognise the importance of reflecting upon the researcher's self and the its relation to in creating relationships within the social setting not for all qualitative research. The authors suggest that through this reflexivity:

"we should recognise that we are part of the social events and processes we observe and help to narrate... To deny our being “there” misunderstands the inherent qualities of both methods[participant observation and interviews] -in terms of making sense of social worlds of which we are a part (either through participant observation or as facilitators of shared accounts and narrative strategies" (Atkinson and Coffey, 2003:426).

Any research method choice requires the considerations of methodological assumptions (hows) as well as the nature of research questions (whats). The literature reviewed highlighted the lack of research in classrooms in exploring formative assessment and suggested that if formative assessment is going to survive and prosper, it is necessary to conceptualise its image from practice (Torrance, 1992). Furthermore, Bell and Cowie (2001) defined formative assessment is a setting dependent concept and suggested the importance of identifying the setting-specific aspects to make sense of it. These considerations suggest that classroom observations would be helpful in understanding formative assessment with a particular view to identify the ways in which it manifests itself within classroom situations. Particularly, an investigation of how teachers move pupils forward in their learning to a desired level through identifying teachers' strategies and what conditions need to be met for this purpose has the potential to enhance the current understanding of formative assessment.

The use of classroom observations for this research is also derived from a motivation to understand selected key issues which emerged from the initial analysis of the interview data. It aimed to compliment the interview data by allowing the observation of the appearance of formative assessment in a selected classroom setting. However, the claim here is not to triangulate “what people say” and “what they do” but to identify observable formative assessment events to illuminate the understanding of formative assessment in classroom situations. This is because the researcher assumes that neither method is more valid than another.
To elaborate on that, it is appropriate to reflect upon what is meant by making sense of “social events” through observations. In this regard, Atkinson and Coffey (2003) suggest the following:

“an ‘event’ in the social world is not something that just happens: it is made to happen. It has a beginning, middle, and an end. It is differentiated from the surrounding stream of activity. Its structure and the observer’s capacity to recognise it are essentially narrative in form.”

For both methods, success is dependent on the researchers’ capability to recognise social events and situations and “extract” them (Manson, 2002) either from an interview or an observation transcription. So, the relationship between the two research methods was not only based on a concern for validating each other, but also to enrich the conceptualisation of formative assessment by filling each other’s respective gaps.

SECTION 3: DESCRIPTION OF THE RESEARCH PROCESS

Selection of Research Participants for Teacher Interviews

Teacher interviews were conducted with 12 Key Stage 3 Geography teachers who worked in Leicestershire schools. Apart from one teacher (Teacher 4), all of the teachers were heads of geography or humanities departments. The majority of teachers were experienced, as the following table summarises:
Table 3.1: The Interview Participants

<table>
<thead>
<tr>
<th>Interview Participant</th>
<th>Head of Department</th>
<th>Subject Specific</th>
<th>Experience (years)</th>
<th>Gender (F: Female, M: Male)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher 1</td>
<td>✓</td>
<td>✓</td>
<td>8</td>
<td>F</td>
</tr>
<tr>
<td>Teacher 2</td>
<td>✓</td>
<td>✓</td>
<td>10</td>
<td>F</td>
</tr>
<tr>
<td>Teacher 3</td>
<td>✓</td>
<td>✓</td>
<td>35</td>
<td>M</td>
</tr>
<tr>
<td>Teacher 4</td>
<td>✓</td>
<td>✓</td>
<td>4</td>
<td>M</td>
</tr>
<tr>
<td>Teacher 5</td>
<td>✓</td>
<td>✓</td>
<td>9</td>
<td>F</td>
</tr>
<tr>
<td>Teacher 6</td>
<td>✓</td>
<td>✓</td>
<td>13</td>
<td>M</td>
</tr>
<tr>
<td>Teacher 7</td>
<td>✓</td>
<td>✓</td>
<td>18</td>
<td>F</td>
</tr>
<tr>
<td>Teacher 8</td>
<td>✓</td>
<td>✓</td>
<td>15</td>
<td>M</td>
</tr>
<tr>
<td>Teacher 9</td>
<td>✓</td>
<td>✓</td>
<td>20</td>
<td>F</td>
</tr>
<tr>
<td>Teacher 10</td>
<td>✓</td>
<td>✓</td>
<td>19</td>
<td>F</td>
</tr>
<tr>
<td>Teacher 11</td>
<td>✓</td>
<td>✓</td>
<td>5</td>
<td>M</td>
</tr>
<tr>
<td>Teacher 12</td>
<td>✓</td>
<td>✓</td>
<td>15</td>
<td>M</td>
</tr>
</tbody>
</table>

The selection of teachers for this research was not on a random basis but involved discussions with the supervisor in identifying a number of considerations. Firstly, it was important to select heads of geography departments since they were assumed to be the key people in their departments in implementing assessment arrangements. Secondly, all of the teachers interviewed had links with the Leicester University in terms of supporting Post Graduate Certificate in Education (PGCE) Geography students in gaining their teaching experience. So, the teachers had connections with one of the researcher's supervisors, a factor which affected the selection process. Overall, the criteria that were used to select the teachers were as follows:

- willingness to participate;
- interest in assessment issues;
- being head of geography departments.

The anticipated interest in assessment issues was an important criterion in selecting the research participants. For instance, Teachers 5 and 1 had completed Masters degrees related to teacher assessment. Teacher 2 was a co-author of an article about target setting processes that was applied in her school. Teachers 1, 2, and 7 showed their interests in assessment by attending a number of conferences. By
including such participants, the researcher expected to gain rich data on formative assessment.

**Ethical Agenda**

The issue of ethics in educational research raises special concerns at every stage of research and has implications for the whole research process. The British Educational Research Association adopted a set of guidelines to propose ethical standards for educational research. This guideline implies certain responsibilities for educational researchers to the research profession, research participants and wider public.

A researcher’s ethical responsibility to the research profession implies that researchers are under obligation to describe accurately and fairly what their research participants say (Morrison and Scott, 1992:97). This implies that researchers should aim to avoid fabrication, falsification and misinterpretation of evidence, data, findings and conclusions (BERA, 1992). This requirement creates an extra challenge for qualitative researchers since they become research instruments (Miles and Huberman, 1994). It is therefore salient for all researchers to report research conceptions, and analysis procedures accurately and in sufficient detail to allow other researchers to understand and interpret them (BERA, 1992). In this thesis, the researcher attempts to communicate openly about the conceptual underpinnings of this project as well as justifying how the literature reviewed shaped her theoretical understanding of the subject and how the analysis of research data was processed. Furthermore, she reflects upon her role as a researcher and attempts to unpack her intellectual luggage that she brought into this study. Great care was also taken to choose representative examples when actual participants’ accounts are used.

The researchers’ ethical agenda also implies certain responsibilities towards research participants. Participants in a research study have the right to be informed about the aims, purposes and likely publication of findings involved in the research and of potential consequences for participants, and to give their informed consent before participating in research (BERA, 1992). In the current project, all the
participants were sent letters which communicated the aims, recording procedures and potential form of research outcomes (Appendices A, B). Furthermore, the researcher assured the participants about the anonymity of their identities and institutions. This was achieved by researcher using a coding system in using research participants in this thesis.

Researchers also have ethical responsibilities to the wider research and other interested parties such as educational practitioners to report their findings (BERA, 1992). This also allows the researchers to test the relevance, accuracy and comprehensiveness of their findings as they emerge within the process of enquiry (ibid). The researcher in this study aims to disseminate the findings by publishing a number of articles within the journals that are accessible for the practitioners and research community.

Selection of Research Participants for Classroom Observations and Pupil Interviews

Having completed eight interviews the researcher analysed the primary findings of teacher interviews to assist the development of the research questions for the observational study. The initial findings of teacher interviews suggested the importance of assessment tasks as a facilitator for formative assessment. This finding informed the selection of research participants for observational study in terms of working with a small number of teachers and conducting observations within the entirety of assessment tasks. Teachers 1 and 3 were selected as potential candidates for observational study because of their ability to elaborate on issues that were experienced during teacher interviews. Teacher 3 responded positively to the suggestion and accepted the researcher as a non-participant observer for 15 lessons within two Key Stage 3 Geography classrooms. Therefore, the sampling decision for classroom observation was shaped by both the initial findings of teacher interviews as well as the interview experience.

The teacher selected for the observations is male and has been teaching geography for thirty six years. He is a subject specialist teacher and head of the geography department. One classroom was a "mixed ability" class of Year 9 level while the
other was a “higher ability” class in Year 8. The same teacher taught both classes. He has been with the same classes for two years.

The school is a local authority maintained secondary school in which pupils are aged between 11 to 14. It is located in a small village in Leicestershire. Levels of academic achievement are high at this particular school. This is confirmed by the Year 2000 Standard Assessment Tasks (SATs) results conducted at the end of Key Stage 3. The results in the core subjects of English, Science and Mathematics are among the highest in the county in 2001 when this study was conducted.

As mentioned earlier, pupil interviews ensued from the classroom observation phase of the research. Twenty Key Stage 3 pupils were selected from the two classrooms that the researcher observed. The selection of pupils was decided jointly by the teacher and researcher and involved a representation of the following:

- Differing ability levels;
- Gender (boys and girls);
- Year 8 and 9 classrooms.

The following table provides information on pupils interviewed, regarding their gender, year group and working levels which were provided by their teacher:
Table 3.2: Information on Participants for Pupil Interviews

<table>
<thead>
<tr>
<th>Interview participant</th>
<th>Year Group</th>
<th>Working level</th>
<th>Gender</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pupil 1</td>
<td>9</td>
<td>5</td>
<td>M</td>
</tr>
<tr>
<td>Pupil 2</td>
<td>9</td>
<td>5</td>
<td>M</td>
</tr>
<tr>
<td>Pupil 3</td>
<td>9</td>
<td>6</td>
<td>M</td>
</tr>
<tr>
<td>Pupil 4</td>
<td>9</td>
<td>5 to 6</td>
<td>F</td>
</tr>
<tr>
<td>Pupil 5</td>
<td>9</td>
<td>5 to 6</td>
<td>F</td>
</tr>
<tr>
<td>Pupil 6</td>
<td>9</td>
<td>6</td>
<td>F</td>
</tr>
<tr>
<td>Pupil 7</td>
<td>9</td>
<td>6</td>
<td>F</td>
</tr>
<tr>
<td>Pupil 8</td>
<td>9</td>
<td>6 to 7</td>
<td>M</td>
</tr>
<tr>
<td>Pupil 9</td>
<td>9</td>
<td>6 to 7</td>
<td>M</td>
</tr>
<tr>
<td>Pupil 10</td>
<td>9</td>
<td>5 to 6</td>
<td>F</td>
</tr>
<tr>
<td>Pupil 11</td>
<td>8</td>
<td>5 to 6</td>
<td>F</td>
</tr>
<tr>
<td>Pupil 12</td>
<td>8</td>
<td>5 to 6</td>
<td>F</td>
</tr>
<tr>
<td>Pupil 13</td>
<td>8</td>
<td>5 to 6</td>
<td>F</td>
</tr>
<tr>
<td>Pupil 14</td>
<td>8</td>
<td>5 to 6</td>
<td>M</td>
</tr>
<tr>
<td>Pupil 15</td>
<td>8</td>
<td>6</td>
<td>M</td>
</tr>
<tr>
<td>Pupil 16</td>
<td>8</td>
<td>6</td>
<td>M</td>
</tr>
<tr>
<td>Pupil 17</td>
<td>8</td>
<td>6</td>
<td>M</td>
</tr>
<tr>
<td>Pupil 18</td>
<td>8</td>
<td>6</td>
<td>F</td>
</tr>
<tr>
<td>Pupil 19</td>
<td>8</td>
<td>6 to 7</td>
<td>F</td>
</tr>
<tr>
<td>Pupil 20</td>
<td>8</td>
<td>6 to 7</td>
<td>M</td>
</tr>
</tbody>
</table>

A Summary of the Research Process

The introduction section outlined the nature of this research study. Here, the following table summarise the links between the use of research methods and the aims and objectives of the study.
Table 3.3: The Description of Research Process

<table>
<thead>
<tr>
<th>Research Methods</th>
<th>Nature</th>
<th>Research Aims and Objectives</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher interviews</td>
<td>• semi-structured</td>
<td>• To understand how teachers view their formative assessment practices</td>
<td>• With 12 Geography Key Stage 3 Teachers</td>
</tr>
<tr>
<td></td>
<td>• audio recorded</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Classroom observations and</td>
<td>• semi-structured</td>
<td>• To describe formative assessment in class situations</td>
<td>• 15 lessons (8 lessons with Year 9 class, 7 lessons with Year 8 class)</td>
</tr>
<tr>
<td>the follow up interviews</td>
<td>• video recorded</td>
<td>• To enrich the findings of teacher interview data</td>
<td>• The teacher was interviewed after each lesson on emerging issues.</td>
</tr>
<tr>
<td>with the teacher</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pupil interviews</td>
<td>• semi-structured</td>
<td>• To find out how pupils view a number of selected assessment issues</td>
<td>• 14 pupils (7 from Year 8, 7 from Year 9 class)</td>
</tr>
<tr>
<td></td>
<td>• audio recorded</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Documentary Analysis</td>
<td>• Personal</td>
<td>• To support other research methods by providing background information</td>
<td>• Voluntary basis</td>
</tr>
</tbody>
</table>

SECTION 4: IN THE FIELD

The Application of Teacher Interviews

Having decided on the interview questions, the researcher wrote a letter to teachers asking for their participation and conveyed the aims of the research (Appendix A). Having accepted involvement, two pilot interviews were conducted with Teacher 1 and 2 in June 2000. Following initial analysis and improvements, teacher interviews were conducted between November and May 2001, with each interview taking approximately 60 to 120 minutes. The engagement with the data was iterative and each interview informed the next. All dialogue was recorded on audiocassette. The interviews took place primarily in the geography rooms of the respective schools.

Before starting the interviews, the teachers were assured of the confidentiality of the interviews. All of the teachers were given the interview questions prior to the respective meetings. The rationale behind this was to make them more comfortable and familiar with the issues as well as to allow them to develop ideas and speak more widely on the issues that were raised by the researcher. However, they were also asked whether the researcher had missed any important information that might...
be salient for the research. During the interview phase, teachers generally organised their answers according to the list of questions. However, they were free to change the sequencing of the main issues and questions. Some prompts and some additional questions were also used in the interview which emerged during the natural flow of the conversation.

The semi-structured nature of questions allowed a developmental research approach to be adopted through the entirety of teacher interviews. The prompts were not used in a mechanistic fashion. Some of them were abandoned depending on the dynamics of the interviews, but some others explored further. Some emerged during the course of the interviews which were not initially considered. New emerging ones were used in further interviews. So, there was a constructive process of meaning making from the first to the last interview and each one contributed to those ensuing ones by creating new horizons to the enquiry.

The interview data generated in the fieldwork was a result of the questions raised as well as interaction between the researcher and the teachers. This realisation arose after reviewing the transcribed data. At various occasions, the researcher realised that the conversation was shaped on the basis of preferred issues when the interviewee provided several options for its direction. It was also clear that the researcher had the control to decide the length of an answer and when to move to the next point. This was sometimes the result of a lack of researcher skill in managing note taking efficiently, in particular in the first series of interviews. So, during the initial stages, more “management of meaning” occurred as opposed to meaning negotiation. This realisation highlights the importance of reflecting upon the interview process.

**The Application of Classroom Observations**

The observation phase was a 2-month study, between January to March 2001, involving 1 teacher in 1 school in two Key Stage 3 classrooms, one of which was Year 8 and the other was Year 9. The researcher spent one lesson each week in two classrooms and observed 15 lessons as a non-participant qualitative observer. All of the lessons observed were video-recorded.
The observational study was piloted through recorded observation of the two lessons in the autumn 2001. This helped the researcher in the refinement of research objectives as well as making decisions about how and where to position on the camera. On the first day of actual interviews, the teacher introduced the researcher to the two classrooms and the researcher described the aim of the project and requested their approval for the use of the camera. On a number of occasions, the researcher illustrated to the pupils how she was going to use video recording by illustrating extracts of the recorded lessons. Throughout ongoing communication, a culture of acceptance was achieved for the existence of not only the video camera but also the researcher.

As outlined in Chapter 1, the observational study aimed to explore the appearance of formative assessment in two selected classrooms as well as extend the findings of initial teacher interviews. After each observed lesson, the teacher was interviewed to clarify the emerging issues. Overall, the teacher was interviewed seven times. On two occasions, the teacher requested the recorded data and offered his reflections on this own practice. In this sense, the use of recorded data assisted the process of constructive meaning making.

The use of video recording requires the consideration of a number of issues, some of which related to technical (Silverman, 2001), in particular where it is located (Sacks, 1992, quoted in Silverman, 2001:24). In this study, the video recorder was positioned in the far rear area of the class to minimise distraction. Therefore, it did not capture the facial expressions of pupils. The video-recorder was placed on a tripod during the recordings. This gave the researcher opportunities to observe the classroom and take written notes to record her initial reflections.

**The Application of Pupil Interviews**

Approximately one month after completion of the classroom observations, the researcher commenced the analysis of observational data to generate questions for pupil interviews. The literature reviewed also informed the constructing of questions. Before proceeding with the pupil interviews, the researcher wrote a
letter to pupils to invite them to participate, which also stated the aims and the nature of interviews (Appendix B). The pupil interviews took place in May 2001 and all of them were audio recorded. Each interview approximately took 15 to 25 minutes and pupils were interviewed individually. The interviews were recorded with a small tape-recorder. A semi-structured schedule was applied to generate research data. The pupils were asked to reflect upon their previously completed assessment task, as well as to answer open-ended questions.

A number of challenging issues emerged during the pupil interviews. As indicated above, there was a month time gap between the observational study and pupil interviews. This created the problem that some pupils were not able to recall their experience and therefore they could not articulate issues in generic terms. The researcher used a high number of prompts and examples to recall their experience. Another challenge related to dealing with situations during which pupils asked what other pupils had said about certain questions as if searching for a ‘correct answer’.

Collection of the Documentary Evidence

During the interviews and classroom observations the researcher collected documents to support the other forms of research data. This was achieved on a voluntary base, which was entirely dependent upon teachers’ willingness to share their documents. Teachers appeared to be generous and provided assessment task examples, targets sheets as well as assessed work to exemplify issues as well as to support their views.

SECTION 5: DATA ANALYSIS AND ISSUES OF VALIDITY AND RELIABILITY

Data Analysis

Qualitative data collection and analysis tend to be continuous rather than discrete, and it is generally in the form of messy and unconnected lengthy written texts, photographs, conversations and transcriptions (Ely et al., 1991; Manson, 2002).
This requires the use of strategies to bring order to the data (Morrison, 2001). One strategy is called grounded theory and involves the generation of categories for use in analysing the data (Glaser and Strauss, 1967; Strauss and Corbin, 1990; 1998). Here, theoretical categories are developed from the analysis of collected data and must fit them. These categories need to explain the data they subsume (Charmaz, 2003:251). Hypotheses also need to be developed and tested in the field (Bartlett and Payne, 1999). Therefore, grounded analytical procedures do not use predefined codes that are informed by other sources, and analysis should be totally grounded in data (ibid).

Miles and Huberman (1994) suggest three concurrent stages of data analysis - data reduction, data display, and conclusion drawing. They suggest that this is an interactive model and there is not a linear sequence. The following illustrates the analytical processes involved (adopted from Miles and Huberman, 1994:10).

**Figure 3.2: Components of Data Analysis: Interactive and Iterative Model**

The main data reduction tool for this research was coding which is addressed in the following part.

**Organisation of Data through Coding**

Following the fieldwork phase, the raw data was prepared for analysis. The first stage of data preparation was the transcription of the audio and video recordings
into written text. Subsequently, the data needed to be collated and organised. This was accomplished by developing codes to categorise data. Codes are labels for assigning units of meaning to descriptions of inferential information compiled during a study (Miles and Huberman, 1994). In the development of codes, the literature reviewed helped in providing ideas for the development of a list of initial codes (Appendix E). This was not a detailed list, but was merely informed by the literature reviewed. The construction of codes was also partly grounded in the sense that the codes also emerged from the raw interview data. The researcher also used her reflective remarks which accompanied codes. This allowed the researcher to seek new explanations.

Whether codes were predefined or developed through the texts, it was useful to have clear definitions in order to provide consistency in the meanings of the codes (Miles and Huberman, 1994). However, this was not easy to achieve. The working definitions of codes evolved throughout the analysis. Arksey and Knight (1999:163) raise a number of questions for the emergent categories:

- Do they [codes] cover all of the data that are relevant to the research aims?
- Are new categories needed?
- Do existing categories need to be split into sub-categories?
- Are there too many categories, so that you are regularly giving the same segments more than one code and finding the distinctions implied by the different codes do not really describe significant differences in the data?

Coding was not a straightforward activity; it rather involved analytical decisions about the meaning and capacity of codes in explaining the research aims. At this stage, the researcher also used memos to remind herself of her ideas, assumptions and potential explanations and use of codes.

**Data Retrieval and Conclusion Drawing**

Having organised the data around the codes, the next stage involved searching for information about a topic or theme that had been indexed under the same code (Arksey and Knight, 1999). This stage mainly involved the use of a number of
tactics for interpreting data. The main strategies involved (Miles and Huberman, 1994; Dey, 1993; Ritchie and Spencer, 2002:309):

- Noting of patterns and themes;
- Making contrasts and comparisons of different pieces of data;
- Examining regularities, variations and singularities;
- Finding associations between experiences and attitudes, between circumstances and motivations etc;
- Counting;
- Writing analytical descriptions.

Miles and Huberman (1994:68) describe patterns as explanatory or inferential codes to identify an emergent theme, configuration, or explanation. Patterns constituted the analytical units, which prompted the researcher to look for common occurrences. For instance, the use of self-reports as a means of facilitating pupil self-assessment was noted as a pattern. However, this did not mean just the naming of the situation. The researcher wrote analytical descriptions to describe the specific characteristics of its use; in other words, she looked for explanations within the data. Then this pattern was looked for in other transcriptions and compared and contrasted with the other reported practices. At this stage, the researcher also counted the number of occurrences, such as how many teachers or how frequent they used the self-reports to make sense of the overall practice of using self-reports to draw conclusions.

Miles and Huberman (1994) caution researchers against naming a pattern too early and the tendency to draw immature conclusions on the basis of an early definition. They warn about checking initial impressions through other tactics to drawing conclusions (plausibility). This task was achieved by the researcher by checking the data throughout the process rather than taking meanings as guaranteed. In this sense, the analysis process was iterative (Dey, 1993), and did not follow a linear sequence, and involved moving backwards and forwards for meaning clarifications.
Validity and Reliability

Questions about reliability refer to the degree of replicability that might be achieved from one’s study to another; in other words, the extent to which other researchers studying the same or similar settings generate the same findings (Seale, 1999:140). As this traditional definition suggests, reliability is concerned with the standard application and consistent use of research instruments, which are associated more with quantitative methods (Mason, 2002). Although it is argued that this definition is incompatible with an interpretivist epistemology, qualitative researchers have suggested criteria to address the issue. LeCompte and Goetz (1982) argue for researcher reflexivity which suggests that researchers need to be open about their fixed attributes, their adopted identities, as well as being aware of the social situations within which the data is generated (Seale, 1999:141).

A traditional definition of validity refers to the accuracy of researchers’ claims to accuracy of the extent to which they measure what they think they are measuring (Campell and Stanley, 1963). The literature distinguishes between at least two types of validity. Internal validity is concerned with whether the findings are genuine for the group they claim to represent; and external validity is concerned with the applications of findings to other groups and situations. External validity has been a concern for qualitative researchers who have argued that traditional definitions of validity and reliability are developed mainly from a positivist standpoint, and therefore, they are not considered applicable to qualitative research (Denzin, 1983; Watling, 1995; Simco and Warin, 1997). As Denzin (1983:133) suggests:

“The interpretivist rejects generalization as a goal and never aims to draw randomly selected samples of human experience. For the interpretivist every instance of social interaction, if thickly described (Geertz, 1973), represents a slice from the life world that is the proper subject matter for interpretive inquiry...Every topic...must be seen as carrying its own logic, sense of order, structure, and meaning”.

As illustrated above, Denzin (1983) does not view generalisation as a priority for qualitative research. This is because qualitative research is more ideographic in its nature with a concern for understanding the social world by considering the
circumstances of a specific setting (Williams, 2002). In this view, each setting is context and time bounded and, therefore, generalising to others settings is not a concern (Lincoln and Guba, 2000b). From this perspective, rejecting external validity has also implications for sampling since qualitative researchers work with smaller samples to describe social phenomena more fully (Seale, 1999; Manson, 1996).

Validity has been reconceptualised by qualitative researchers. This has led to the development of a number of criteria for judging the value and quality of qualitative research. In relation to the generalisability of case studies, for example, Stake (2000:23) coined the term “naturalistic generalisation” which suggests making generalisations to similar cases rather than to a population of cases. In the work of Schofield (2002), Goetz and LeCompte (1984) introduce the concepts of “comparability” and “translatability” which advocate the potential applicability of qualitative studies to other similar situations. Guba and Lincoln (1982) and Lincoln and Guba (1985:316) advocate ‘transferability’, which suggest that generalisability should be achieved by looking for similarities with other situations.

Underpinning the above terminology is the need for the researcher to provide thick and rich descriptions about the settings so that there will be a room to argue for transferability of findings (Seale, 1999; Janesick, 2003). However, qualitative research is not only about descriptions; it is also about explanations and interpretations. Simco and Warin (1997:662-3) provide a set of criteria which raise a number of issues in relation to the validity of qualitative research. These are as follows:

- **Completeness** is the extent to which a narrative or transcript is a ‘full’ record of the event which it attempts to represent.
- **Adequacy of interpretation** signifies a process of description, questioning and redescription, the seeking out of the missing bits of the picture.
- **Self-reflection** is the researchers’ capacity to situate themselves within their own process.
In the current research, the researcher's concern was not centred upon a positivist external validity. This was justified on the basis of epistemological and ontological assumptions, which acknowledged the possibility of multiple realities since the very notion of positivist validity would have prioritised the researcher's claim of knowledge at the expense of others (Watling, 2000:4). However, this does not suggest that the research conducted is invalid. It provides insightful information about teachers' assessment practices which were at the core of the study's intentions. There is generally an accepted trade off between breadth and depth within limited resources, and rich data might not have been gathered if the scope of study had considered greater numbers. However, it is argued that some findings could be followed up at a later date, testing those with larger samples.

Neither is this research primarily concerned with replicability to judge its reliability. Here, the researcher was interested in providing representations of events and not reproductions (Watling, 2000). The situations and people studied in this study are not time or context free, and replicability itself, therefore, was not considered appropriate to address. However, the researcher provided information on her assumptions about social reality, about the research process, as well as the research questions and analytical framework, which explains overall how the conclusions were drawn. Through self-reflection of researchers, there could be a better scope for qualitative research to achieve reliability.

**Summary Outline of the Methodology Chapter**

This chapter provided an overview of the research methodology used in this research and described the research process. The first section explored the philosophical underpinnings of this study and linked the choice of data collection methods to the epistemological and ontological assumptions of the research. The second section described the rationale for the use of interviews and classroom observations as the main source of research methods. The third section described the research process and provided information on the participants. The subsequent section then illustrated the process of data collection and focused upon the researcher's field experience. The chapter ended with the information on the data
analysis procedures before addressing the issues related to validity and reliability of this research.

Having addressed the overall research process in Chapter 3, Chapter 4 presents findings and analysis of this research in three sub-sections, relating to teacher interviews, classroom observations and pupil interviews respectively.
CHAPTER 4: FINDINGS AND ANALYSIS

PART 1: FINDINGS AND ANALYSIS OF TEACHER INTERVIEWS

Semi-structured interviews were conducted with twelve Leicestershire geography teachers in order to examine the practice of formative assessment in the classroom Geography in Key Stage 3. This part presents the findings by considering each key component of formative assessment: assessment task design, teacher feedback and target setting, student self- and peer-assessment. This part ends with a consideration of how teachers make summative assessments of pupils’ attainment at the end of Key Stage 3.

SECTION 1: THE DESIGN OF ASSESSMENT TASKS

The interviews provided information on the assessment tasks used in pupil assessment in Key Stage 3 geography. Also they gave insights into teachers’ thinking when they designed assessment tasks. Teachers mainly considered progression, differentiation and enjoyment as crucial task design criteria. Their task design was also informed by the optional tasks available. Each of these criteria is discussed in the following subsections.

Planning for Progression in Pupils’ Learning

The research literature on the implementation of formative assessment in classrooms has tended to focus on key events in limited periods of time. Lambert (2000:138) highlights the importance of the strategic departmental planning of formative assessment over an extended period of time in order to enhance the day-to-day practice. This leads on to the need to consider the curriculum planning principle of progression. Planning for progression in pupils’ learning requires both a conceptual understanding of the nature of progression in Geography in Key Stage 3 and an ability to relate that understanding to an interpretation of the level descriptions.
Conceptual frameworks to help teachers understand and articulate the nature of progression in pupils' learning have been produced by notably Bennetts (1995; 2002) but these are not necessarily known to teachers and used by them to help them plan the curriculum over the medium or long term. As it was mentioned in the literature review, the level descriptions are an attempt to define progression in pupils' learning based on several progression strands. Progression in pupils' learning in the enquiry and skills strand is the most clearly articulated and is based on pupils' growing independence in planning and selecting and using skills for enquiry.

The teachers appeared to have less confidence in elaborating on the nature of progression in pupils' learning to help them plan a sequence of assessment tasks. This finding supported Robert's investigation (1995) in which planning for progression was largely intuitive and were based on their experience. Teachers appeared to rely exclusively on the level descriptions to unpack progression and were unaware of the guidance which became available from the Geographical Association (Butt et al., 1995).

In line with Bennetts' (1995; 2000) interpretation of key elements of progression, teachers considered progression with reference to three key elements - "breadth of geographical knowledge", "depth of geographical understanding" and "skill development". Of the three elements, progression in the breadth of pupils' knowledge was the easiest for teachers to articulate; for instance, eight teachers were able to refer to, "an increase in the scale of places" and "variety of geographical contexts". They were comfortable with the planning for progression in enquiry and the related skills. Teacher 3 was able to describe his interpretation of a progressive development of research skills from Year 7 to 9. A typical example involved describing advancement in research skills from Year 7 to 9. In this example, at Year 7, enquiry work involved the use of secondary data with simple methods of investigations. At Year 8, pupils used multiple resources to substantiate their research. At Year 9, pupils were expected to carry out independent work by establishing the sequence of steps of an investigation and applying different research methods in their investigations.
Teachers were also able to articulate progression in map skills as was evidenced by specific examples given by 7 teachers. Teacher 3 exemplified a typical example of progression in map skills from Year 7 to 9, which demonstrated that he used a planned approach to the progressive development of his pupils’ map skills.

"...in Year 7 we expect them to become familiar with the fundamentals of an ordinary survey map; for example what individual symbols mean and the information on scale. In Year 8, they then go on to interpreting such concepts in contexts like settlement and use this knowledge of scale, distances and directions to look at settlement types. In Year 8, they also have to identify more advanced features of an ordinary survey map. In Year 9, they use skills they have learned earlier on and apply them to a new topic”.

In contrast, teachers found it difficult to articulate what might be meant by increasing depth in pupils’ understanding. Identifying what constituted progression in geographical skills was another challenging issue and all the teacher responses were brief although they did recognise a broader category of cognitive skills such as ‘describing, analysis, explaining’ which they used in association with the enquiry skills.

All teachers were aware of the need to design assessment tasks which related directly to the criterion referenced statements in the level descriptions and this appeared to be the number one priority. This process involved teachers interpreting the level descriptors in order to understand what pupils were required to achieve. They found the interpretation of the level descriptions difficult. For example, Teachers 4 and 5 referred to this requirement as “frustrating” and commented that level descriptions were “not clear” and “vague”. In total, 10 out of 12 teachers expressed dissatisfaction with the level descriptions.

They found it particularly difficult to grasp what distinguished one level from another, such as a strong Level 4 versus a weak Level 5, as Teacher 2 commented:

“...the difference between one level and another is something very hard to understand. You can see that there is progression but how we should translate them [the levels] for an individual task is a big problem”.

Teachers experienced difficulty in understanding the extent of progression in enquiry skills from one level to another. Teachers 7, 8 and 12 specifically
commented on the difficulty in identifying the difference between 'reaching plausible conclusions', 'begin to reach substantiated conclusions' and 'reach substantiated conclusions'. Teacher 12 further suggested that:

"...without guidance materials which specifically exemplify how we can offer these, we are given a very hard job in our hands".

The third difficulty experienced by the teachers was the problem of identifying criteria to enable pupils to achieve the higher levels. With one exception, - " being able to conduct independent work" - teachers seemed to struggle in finding criteria.

The interview data identified planning for progression as involving a series of steps. All of the teachers suggested that the hierarchy of statements in the level descriptions provided the guidance for progression from one year to another. As Teacher 2 pointed out:

"...in Year 7, tasks cater for levels Level 3 to 6. For Year 8, it is Level 3 to 6, and for Year 9 groups, tasks will cover between Levels 4 to 8".

Another important issue for all teachers was transforming level requirements to a form which could be communicated to pupils. This was achieved by the use of checklists and assessment record sheets, all of which served similar purposes. The below example (Figure 4.1) typifies an assessment record sheet which consists of level descriptions and corresponding expectations, a checklist and a list of targets for pupils to choose from. Teacher 4 provided the following assessment record sheet:
**Figure 4.1: An Example of an Assessment Record Sheet**

<table>
<thead>
<tr>
<th>ASSESSMENT RECORD SHEET</th>
<th>Teacher Checklist</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Student Checklist</strong></td>
<td><strong>Teacher Checklist</strong></td>
</tr>
<tr>
<td>Have you included any named examples of volcanoes/earthquakes?</td>
<td>Have you included any named examples of volcanoes/earthquakes?</td>
</tr>
<tr>
<td>Have you thought about the best order to put your project in?</td>
<td>Have you thought about the best order to put your project in?</td>
</tr>
<tr>
<td>Is your work presented clearly?</td>
<td>Is your work presented clearly?</td>
</tr>
<tr>
<td>Have you used different methods of presentation e.g. maps, diagrams, graphs, pictures?</td>
<td>Have you used different methods of presentation e.g. maps, diagrams, graphs, pictures?</td>
</tr>
<tr>
<td>Was it handed on time?</td>
<td>Was it handed on time?</td>
</tr>
</tbody>
</table>

**COMMENTS:**

3. You have described volcanoes/earthquakes and suggested explanations for their location. You have used evidence to answer key questions, using some appropriate vocabulary.

4. You have described how volcanoes/earthquakes work, the effects they have on places and the lives and activities of people who live here. You have suggested suitable key questions to research and used evidence and appropriate vocabulary in your write-up.

5. You have selected good key questions. As well as describing how volcanoes/earthquakes work and the effects they have, you have explained some of the processes at work in volcanoes/earthquakes and included information about how people try to manage the danger of volcanoes/earthquakes. You have selected evidence from your research, reached a conclusion and used a variety of methods of presentation in your write-up.

6. You have selected good key questions and put them in a good order. As well as describing how volcanoes/earthquakes work and the effects they have, you have explained a range of the process at work in volcanoes/earthquakes and shown that these produce lots of different types of volcanoes/earthquakes. You have described and compared different approaches to managing areas where volcanoes/earthquakes occur. You have selected a range of evidence and used it effectively to come to a good conclusion.

**TARGETS:**

- To include more detail in your description.
- To explain your ideas/to explain your ideas more clearly
- To use a variety of methods of presentation, including maps, graphs, diagrams, pictures etc.
- To try and present your work more clearly

Such assessment record sheets were crucial for teachers’ assessment practices. Firstly, such materials guided teachers by giving them a vision of the relatively important concepts and skills, which provided a framework to observe and monitor pupils’ advancement in their learning over a period of time. For instance, on the basis of his checklist, Teacher 4 suggested that:

"... in Question 4.1 asked them to describe the effects of volcanoes or earthquakes. Over a period of time, I look at the frequency and relevance of pupils’ use of, 'because', 'as a result' and 'caused by' wording in their descriptions. If there is an increasing pattern within a year, I would think that they are progressing in their descriptions".
This teacher was using connectives which is an aspect of the National Literacy Strategy to make a judgement on improvement in the quality of his pupils’ written descriptions.

Secondly, such checklists enabled teachers to communicate to pupils what constituted progression within an assessment task. For instance in the above checklist, in order for pupils to achieve a Level 5, they were asked to describe “some of the processes at work in volcanoes / earthquakes”. In order to achieve a Level 6, they needed to describe “...a range of processes at work in volcanoes / earthquakes.” Regardless of what teachers thought about the practicality of level descriptions, they recognised the importance of communicating their expectations of pupils’ achievement within a task. As Teacher 5 described “…you can’t expect pupils to progress unless you show where they need to progress towards”.

Guidance on planning for progression in the Literature Review chapter identified the issue of pupil maturity (Bennetts, 1996). Three teachers recognised the issue of pupils’ cognitive maturity as an important consideration when they devised tasks. Half of the teachers agreed that pupils in Year 7 were limited in their ability so that they did not introduce “complex geographical skills”, and they only offered basic map skills. In Year 8, enquiry and other project work as well as more complex skills were offered.

As the literature indicated, formative assessment and planning for progression overlap in the sense that both are concerned with promoting pupils’ learning. Although teachers never used the term ‘formative assessment’ while they were talking about progression, their strategies for planning for it revealed how both concepts were related. In teachers’ eyes, planning for progression related to having a map of learning against which pupil learning could be monitored and supported over periods of time. However, they also recognised that progression was salient but was just one aspect of supporting pupil learning – differentiation and promoting pupil engagement were both regarded as important as the following subsections explain.
Planning for Differentiation

Teacher interview data revealed teachers’ planned use of a variety of differentiation strategies, by task, outcomes, stimulus and resources. These strategies are similar to those proposed by Davies (1990) as discussed in the Literature Review. The section summarises the implementation of different differentiation practices in the design of assignment and suggests the ways in which they supported teachers’ formative assessment in Geography Key Stage 3.

Differentiation by Graded Task and Differentiation by Task

The most common means of differentiation appeared to be through tasks which were structured to include questions with an incline of difficulty. These were designed to be accessible to pupils of different levels of ability. Eight teachers reported regular use of structured tasks. Teacher 7 provided an example of a task which began with a cloze procedure, was followed by the interpretation of diagrams and ended with explanations of complex issues. Another example was drawn from an end of a Year 7 geography written examination, in which questions were increased in cognitive demand, from a simple sequencing to a ‘why’ type explanation. Part of this document is shown below (Teacher 8, end of Year 7 geography exam paper, p4):

![Figure: 4.2 An Example of Differentiation by Graded Task](image)

33. Put the forms of transport in order from fastest to slowest: Train, Ship, Lorry, Plane.

34. What advantage would a ship have over a plane?

35. Give two reasons why lorries and not trains are often used to transport goods.

The priority for differentiation for pupils with different abilities was generally regarded as maximising learning opportunities for lower achiever pupils; nine teachers justified differentiation on this basis. Three teachers reported that differentiation was also helpful for preserving the self-esteem of low achieving
pupils. Teacher 8 suggested that by this means “... the least able also have some success from it [the task] and do not feel frightened and intimated”.

Differentiated tasks also involved additional activities, mainly for the purposes of extending opportunities for high achiever pupils. Three teachers adopted this practice, and Teacher 7 described the aim as “to extend their [high achievers] thought”. Such tasks were characterised as involving opportunities for pupils to take the initiative in conducting research, to identify or extend research questions. Such elements seemed to be driven or inspired by the high GNC level descriptions – particularly the comments in Level 7 and 8 descriptions regarding the identification of geographical questions and sequencing of investigations.

The notion of pupil independence in carrying out geographical enquiry was perceived to be problematic with four teachers agreeing that pupils did not perform well when they decreased the level of instruction and increased their freedom. Indeed, one teacher discontinued such differentiation because of this reason. He (Teacher 4) mentioned that when there was less help and more independence, very few pupils progressed and achieved higher levels which was his original motive.

Another way of differentiation by task related to offering pupils a number of choices to demonstrate the product of a task. Teachers viewed this as increasing opportunity for pupils to demonstrate what they had learnt. For instance, Teacher 5 gave the example of pupils making choices between preparing a poster or writing an assignment. This practice was regarded as “stimulating” (Teacher 11), “fair” (Teacher 2), and “inclusive” by increasing the chances of pupil participation (Teacher 8). In Teacher 8’s class, in a unit of work on ‘Tourism’, pupils had the choice of either writing a tourist pamphlet or preparing a poster. Nine teachers mentioned that differentiated task design requires extra time and effort and that it was difficult to find this time within their work schedules.

*Differentiation by Stimulus*

The second most common strategy for differentiation was the provision of the instructions for completing tasks. This involved providing additional sheets to
guide pupils in their completion of different stages of the task. Such sheets were provided as an option in two teachers’ classrooms. The other ten teachers interviewed also seemed to be in favour of such “guidance sheets” but did not use them regularly. Another form of task differentiated instruction was the use of writing frames. Its degree of scaffolding and frequency varied among the assessment practices of different teachers, largely on the basis of their perceived ability of pupils. Four teachers interviewed reported that they used structured writing frames only for pupils regarded by them as low ability. These teachers mentioned that such frameworks were useful in terms of organising pupils’ thoughts and as a useful means of scaffolding for weaker pupils. One teacher in this group, Teacher 9, mentioned that this device should not be used with the high achievers since, “they need to let their words fly when they write” (Teacher, 3).

**Differentiation by Resources**

Relatively less emphasis was given to differentiation by resources. Only two teachers mentioned it, and in these cases it was a differentiation practice mainly for pupils with lower ability. This included the choice of less text and simplified language. Teacher 7 described this principle:

> “less able pupils’ work sheets contain less text, simpler wording, a more structured format like simple fill in the gaps and true and false or fill the missing words”.

Pupils’ language skills were an issue in one of the schools which had pupil populations in which the majority came from a non-British ethnic background. Some of these pupils were bilingual, with English being their second language. For Teacher 5, this required a further level of differentiation on the basis of language. This was done by modifying the level of English to accommodate pupils with different English literacy skills. Teacher 5 reflected back on her experience of language differentiation;

> “We are not assessing only geography... We have a boy who is Portuguese and the first three pieces of work he did for me was in Portuguese. In my first lessons, I wanted him to tell me what he knows about geography. I gave him lots of magazines and told him to cut things up about our unit of work. At the end of the unit, he gave me his work. Remember, the National Curriculum says I have to assess them and does not address or guide us as to how we deal with such incidents”
Differentiation by Gender

In general, teachers recognised gender related performance and made an apparent link between gender and preferred task type, as the following table summarises (Table 4.1):

<table>
<thead>
<tr>
<th>Views on gender related ability</th>
<th>“Boys tend to be better in explaining processes and patterns than girls who are much better in describing and presenting”. (Teacher, 4).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Views related to task specific characteristics</td>
<td>“Girls like writing and like presenting their work with diagrams and maps. However, boys more like presentations in diagrammatic formats” (Teacher, 11).</td>
</tr>
<tr>
<td></td>
<td>“Girls do better at the end exam” (Teacher, 3).</td>
</tr>
<tr>
<td></td>
<td>“Boys are much better when there is ICT in the task, in particular, in presenting their work through ICT” (Teacher, 9).</td>
</tr>
<tr>
<td></td>
<td>“Girls are better in project work” (Teacher, 2).</td>
</tr>
<tr>
<td></td>
<td>“Boys do better in oral presentations and tasks including ICT (Teacher, 1)”.</td>
</tr>
</tbody>
</table>

The common characteristics of the above responses suggest that tasks including Information and Communications Technology (ICT) work were something that boys succeeded in as opposed to extended writing pieces in which girls were perceived to do better. This relationship is consistent with the claims by Butt (2001) and Wood (2002), who suggested that boys perform less well when there is extended writing. Seven teachers expressed this view. Three teachers suggested that there was no gender specific achievement for the high ability groups, and addressed the issue with the low achiever students. It was also noticeable that female teachers tended not to mention girls in explaining gender specific achievement and addressed the questions generally with reference to boys; the converse was not the case with the male teachers.

Although teachers recognised differences in achievement linked to gender, they reported that gender was not taken into account when planning differentiation for
assessment tasks; thus, boys and girls were given the same tasks. However, differentiation in relation to gender occurred to a greater extent within the task implementation stage. The style of teaching was often tailored with gender in mind and was concerned with sustaining boys’ engagement in written tasks. For instance, in Teacher 7’s case, this was helping boys through individual tutorials to assist in their essay writing work. Another teacher recognised that boys were not as responsive as girls to the instructions and so more systematic instructions were administered on how to complete the task. In all situations, differentiation rendered more assistance and advantages to boys than girls.

The overwhelming recommendation from teachers on dealing with the gender related achievement issue was that of offering a variety of tasks throughout the Key Stage 3 was likely to be very beneficial. It was generally accepted that this would provide both genders with opportunities to access tasks that they were ‘naturally’ capable with and through which they could gain achievement. Three teachers specifically commented that differentiation within one task appeared to address this problem (i.e. offering pupils some freedom to chose the nature of the finished product). Teacher 3 mentioned “creativity” to deal with boys’ underachievement in written tasks. He explained his relevant experience:

“In one task, rather than asking them to write a formal essay, I asked the pupils to communicate their findings through a child’s diary. It was quite surprising to recognise what actually boys could do even in emotional matters. I think the potential is there, if you give them a chance to develop it”.

Although teachers acknowledged differences between boys and girls in achieving different tasks, differentiation for different genders was not prioritised. Similar to the concerns expressed regarding variety, the requirement for consistency across the department created the biggest obstacle. For nine teachers, preparing instructions and tasks for different sexes was not acceptable since it was perceived to be complicated, and indeed chaotic. Another reason was that differentiation might prevent pupils from developing what they were weak at, a claim which was made by two teachers. These teachers considered this practice to be unfair and considered that everybody should be treated equally regardless of their sex.
Differentiation and Formative Assessment

This section explores how differentiation links to the current conceptualisation of formative assessment. Teachers considered differentiation in two distinct ways, for the short term as well as for the entirety of the Key Stage 3. For the short term, differentiation was provided for within a task by structuring instructions and questions. To a lesser extent, it took the form of producing different modes of outcomes for one task. For the long term, variations in task type facilitated differentiation. This was also intended to be fair to all pupils. Regardless of the time frame, differentiation appeared to be a planned practice.

How could differentiation contribute to the success of formative assessment? In the literature, differentiation is about maximising learning opportunities for all pupils. With reference to the model proposed in the Introduction chapter, differentiation seems to fit in the first stage of the formative assessment process by increasing the chances of revealing pupils' learning. This is explained in the following figure (Figure 4.3):

Figure 4.3: The Links Between Differentiation and Formative Assessment

As Figure 4.3 indicates, differentiation could provide constant support throughout the formative assessment practices. Differentiation promises good practice both
for teachers and pupils with the potential to sensitise teachers to different forms of learning evidence. It may help assist them to recognise different avenues through which learning can be monitored and supported. For pupils, it means that they have increased opportunities to demonstrate what they are capable of doing.

How much of the potential benefits noted above was confirmed by the current differentiation practices that the 12 geography teachers reported? For all of the teachers, differentiated practice was about increasing learning opportunities. However, this was a concern more associated with the first stages of formative assessment, that is, task design which aimed to portray differentiated learning outcomes (i.e. choices of an essay or an oral discussion at the end of a task). This was due to the difficulty teachers experienced dealing with implementation, recording and marking of differentiated learning evidence. For instance, a role-play task reported as requiring articulating assessment criteria and feedback preparation within oral discussions; whereas an enquiry project required assessment integrated within written work. Thus, there was a trade-off between the benefits of variety to pupils and the need to achieve consistency and reliability in marking within a department. Therefore, what was differentiated was the learning outcome. The research evidence also suggests that differentiation was primarily for the benefit of low achiever pupils. For the high achievers, differentiation was more problematic since teachers seemed to have difficulty in finding ways to stretch higher achievers’ abilities.

Other Task Design Criteria

*Enjoyment*

Pupil enjoyment was an important criterion for teachers to consider when devising tasks. Teachers commented that enjoyment increased pupil engagement with tasks. Teacher 9 suggested that unless pupils were interested in the tasks and enjoyed them, there was little chance that their motivation would be sustained. Similarly, Teacher 8 commented that when pupils were interested in tasks, they tended to spend more time and were more enthusiastic to complete the follow up work. Examples varied due to teacher subjectivity about what constituted ‘interest’
and ‘enjoyment’ in terms of topics and task types. However, a consensus still existed in favour of enquiry, fieldwork, and tasks involving group work such as role-plays.

Tasks involving ICT work were regarded as a means of promoting stimulation and interest. This was not on the research agenda, but it was specifically mentioned by three teachers within the discourse of task design. Each of these three teachers agreed that managing and assessing tasks involving ICT were problematic when the task involved the use of the Internet. Teacher 8 explained his concerns with reference to assessment of ICT:

"The children just type the words, print the information from the web and then hand them in to me as being their homework. It is the same, if they just copy articles from journals, cut and copy it and submit to me".

Teacher 8 continued:

"...when you go the ICT room, there is always two children in front of one computer together, sharing a piece of work. It is almost impossible to tell that who has done what. Although it is a terrific resource, assessment in this area is beset with difficulties and problems”.

The second problem above related to the management of group work within tasks involving ICT. More precisely, the problem was in the identification of individual achievement within a group task. This was also a strong disincentive for the use of group-based learning in five other teachers’ practices.

**Creativity**

In comparison to other task design criteria, creativity was almost the least mentioned. This is not surprising given that it is a relatively new task design issue within the literature on Key Stage 3 geography. From 12 teachers interviewed, only Teacher 3 mentioned it as a salient element. In his case, this was achieved by giving pupils freedom so that they had space to construct and communicate their own ideas. This is very similar to the elements of expression and imagination,
articulated by Rawling and Westaway (2003). In a follow up interview, he suggested that tasks such as role-play where pupils had to write their own stories rendered the most creative responses.

**Use of School Curriculum and Assessment Authority (SCAA) Materials**

One of the most significant features of assessment in Key Stage 3 geography is that it is entirely determined by teachers’ ability to design assessment tasks to assess pupils’ learning since there is no externally devised end-of-key-stage exams. As stated in the introduction, only a few published materials (Exemplification of Standards for Geography Key Stage 3) are available for guiding teachers’ assessment practices. The literature describes two distinct functions of the SCAA (1996a; 1996b; 1996c) materials. Firstly, they can guide teachers on how to design their tasks (Rawling and Westaway, 1996); secondly, they can help teachers to make their end-of-Key-Stage 3 judgements (Tidmarsh and Weeden, 1997). It is important to consider how these documents have been used and how they affect teachers’ assessment task design.

This research did not address how individual SCAA materials were used. Rather, it aims to present their overall within teachers’ practices. Teachers reported two different ways of using the SCAA (1996a; 1996b; 1996c; 1996d) materials, as the Table 4.2 presents:

<table>
<thead>
<tr>
<th>DIFFERENT TYPES</th>
<th>USE OF SCAA</th>
<th>NUMBER OF TEACHER</th>
</tr>
</thead>
<tbody>
<tr>
<td>TYPE 1</td>
<td>• The SCAA is used as a reference material for moderation purposes.</td>
<td>6</td>
</tr>
<tr>
<td>TYPE 2</td>
<td>• The SCAA is used as a guidance material for assessment task design.</td>
<td>4</td>
</tr>
</tbody>
</table>

In the first type, SCAA materials were just used for moderation purposes. This included teachers checking their levels and comparing whether they matched the levels that appeared in the SCAA materials. This is reflected by Teacher 3’s
comment that, “They are used as reference materials to justify our levels”. For two teachers in this group, this was important since they felt that there were no other ways of checking whether they interpreted the levels in the appropriate way.

In the second type, the SCAA materials were used to guide teachers in their task design. Four teachers in this type suggested that they paid particular attention to how differentiation was achieved and how the assessment methods were used. These teachers did not use the materials in their original form. They were “adapted” (Teacher 7) “changed” (Teacher 5) and “adjusted” (Teacher 9). The SCAA (1996a) task (Brazil) was regarded as the most useful and the one which was used with the least modification. Indeed, all of the four teachers referred to the SCAA(1996a) task during the interviews.

Teachers in Type 2 identified the ways in which they found the SCAA materials useful. Teachers referred to the SCAA materials as “free resources” (Teacher 9) and “rich resources” (Teacher 6), “good stuff” (Teacher 9), “useful” (Teacher 7) and “good resources” (Teacher 5). However, their use was dependent on whether the topic matched with the intended year group. For instance, if Brazil was a Year 9 topic, then there was a great chance that it would be used with that year group since it was originally devised for Year 9 groups. When it did not correspond to the course content, some of the resources (maps, diagrams) were used with other year groups.

Two teachers did not use the SCAA materials in any way. One representative response was by Teacher 1 who suggested that, “We looked at them but we wanted to do something else which is basically what we did”. For Teacher 11, task design was a reflective process that he would prefer to engage in on his own initiative. He suggested that:

“I enjoy making these tasks up and I think there is no better way of understanding what to expect from students unless you work on the task yourself”.

For him, this engagement in task design was important in understanding whether the material would work for his pupils.
SECTION 2: TEACHERS’ VIEWS ON TARGET SETTING

The current understanding of formative assessment suggests that the success of this process is not only dependent on identifying pupils’ strengths and weaknesses but also increasing their awareness of what they need to do in order to improve their learning (Black and Wiliam, 1998a). There has been a growing body of research evidence which suggest that qualitative target setting in the classroom could enable this and could contribute to raising standards of pupil attainment. This potential was exemplified in geography Key Stage 3 by Hamson and Sutton (2000) and by Thompson (2000) and at primary level by Clarke (1998). The following subsection also explores this potential by illuminating the practice of teacher and pupils target setting and nature of their targets before analysing how target setting links to formative assessment.

Teacher Target Setting

The teacher interview data suggested that target setting practice was guided by written sheets which offered a tick-box of pre-set targets consisting of several levels of possible pupil attainment. Nine out of 12 teachers reported the use of assessment sheets or record sheets to manage pupil self-assessment and target setting for a task. An example of such a sheet (Figure 4.1) is given in the Planning for Progression subsection of this thesis and analysed the extent to which it helped teachers in facilitating progression. Here, another example is provided to illustrate how it contributed to teachers’ target setting practices. The following assessment checklist was provided by Teacher 1:
The example above demonstrates a similar practice of target setting and use of such assessment sheets, to that reported by Hamson and Sutton (2000; Indeed, the same analytical framework is used). As noted in the above figure (Figure 4.4), pupils ticked the relevant boxes when they completed the task. In a way, the checklist with the assessment criteria provided them with an understanding of what they were expected to achieve and how they were going to be judged. Teachers 3, 8, and 9 specifically suggested that in this way pupils were encouraged to become...
more responsible for their own learning, a reported outcome stated in the study by Hamson and Sutton (2000).

For seven teachers, assessment checklists or other documents with similar functions which were given to pupils at the beginning of tasks, were the main tools for their target setting practices. These sheets were kept by pupils and reviewed at various stages during the implementation of tasks. Teacher 12 described the main function of her assessment checklists:

"This assessment sheet helps pupils to assess themselves by giving a score from the list, each of which indicates descriptions of achievement. Having considered what they have achieved and not achieved, they could set their targets. A possible list of targets is available for pupils at the end of each geography unit. These checklists are also attached to their student self-assessment report".

In the above practice, in a way, target setting and pupil self-assessment were used simultaneously. Here, pupils were encouraged to make a selection from these targets on the basis of the outcomes of their self-assessment. In the case of Teacher 3, this sheet provided space for pupils to set their own targets. However, this appeared to be an isolated case amongst the other teachers interviewed.

In the practices of Teachers 1, 6 and 9, pupils were given a record sheet, which identified learning outcomes in the form of levels or marks, future targets and assessment criteria. Teacher 1 suggested that indicating levels in this sheet helped her to set specific subject targets which are also quantifiable by monitoring whether the level was improved at a future point. She advocated her system since it was manageable for her.

In Teachers 6 and 8’s practice, there was no written assessment checklist or sheet. The target setting seemed to be an informal short-term practice and was driven by the pupils’ emergent needs. Teacher 6 explained this:

"...we feedback and feedforward to the pupils depending on their immediate needs so there is no formal procedure in this process"

and commented that:
"I do not want to be driven by outsiders about targets I should set and when and how I should do this. I need to consider what is happening in the classroom and use this evidence to feed my target setting practice”.

The above quote reflects dissatisfaction with the situation where the level descriptions drive the target setting process. A similar attitude was reported by Teacher 8 who claimed that the level descriptions dominated his overall assessment practices. He described them as a “straight jacket” and declined to use them in his target setting practice. Stemming from this issue, the following section provides an analysis of the nature of targets, as indicated by teacher interviews and school policy documents.

**Nature of Targets**

The teacher interview data showed differences in the nature of targets depending on the respective time intervals. These differences are illustrated in the following figure:

*Figure 4.5: Differences in the Nature of Targets According to the Respective Time Intervals*

<table>
<thead>
<tr>
<th>SHORT-TERM</th>
<th>MEDIUM TERM</th>
<th>LONG-TERM</th>
</tr>
</thead>
<tbody>
<tr>
<td>From lesson to lesson and week to week</td>
<td>Over a topic</td>
<td>Over a year or the key stage</td>
</tr>
<tr>
<td>- Targets were produced on the basis of pupils' emergent needs.</td>
<td>- Targets were informed by main assessment objectives of a unit of work</td>
<td>- Targets in this category were numerical in their nature and expressed by the National Curriculum levels (i.e. level 6, 7..)</td>
</tr>
<tr>
<td>- Targets in this category were not level driven</td>
<td>- Targets for medium term were level driven</td>
<td>- These targets were set by the Geography Departments in accordance with schools expected targets</td>
</tr>
<tr>
<td>- These targets were informed by teachers in class observations</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The differences in the three categories of targets (Figure 4.5) are in accordance with Hopkin’s (2000) typology, which identified different planning for assessment
for different time intervals. In the current research, teachers described their medium targets as primarily associated with the content of the level descriptions. This was apparent in the previously provided example of the target checklist (Figure 4.4), which included statements such as “I described and analysed the problem”. Teachers agreed that such statements were taken directly from the GNC. Teacher 10 explained the reasons:

“It is school policy and we are very much associated with the level descriptions. As a part of this process, we need to identify what level they [pupils] are currently working at and which they should be aiming for ... Really what we are saying is that we set targets on the basis of what they are capable of doing and by looking at the next level”.

Teacher 10 was one of the other nine teachers interviewed, who saw target setting as identifying the next steps for the “improvement of learning”, as expressed in the level descriptions.

Six teachers mentioned that they also set targets for pastoral care in addition to their subject specific targets. This was seen in the documentary evidence collected, which indicated additional written pastoral targets in four teachers’ assessment sheets. For Teacher 3, this was very important since he saw learning as, “not limited to subject specific knowledge and skills”. Having the “right attitude” was equally valued by him. Similarly, for Teacher 6 promoting personal and social development was valued and he suggested that “the current system does not give us much opportunity to promote these skills”.

**Pupil Target Setting**

One of the issues addressed was how pupils viewed target setting and whether they encouraged pupils to take the initiative and set their own targets. The teacher interview data indicated that there was not much scope for pupils to set their own targets; only 3 of 12 teachers permitted this (Teachers 1, 3 and 11 encouraged pupil target setting). It was made possible by providing space for pupils to write down their own targets on teachers’ assessment checklists. Teacher 3 strongly believed in the importance of pupil target setting:
"The important thing about pupil target setting is that children themselves have the chance to be aware of their learning needs and be convinced that they need to set targets to improve that...this will hopefully make them work harder and result in improvement".

Apart from these three, the teachers agreed that pupil target setting should be guided by teachers. This was seen in the assessment checklists collected, in which teachers defined pupils’ targets. In this target setting practice, pupils’ freedom in their target setting was limited exclusively to the possibilities set by teachers.

Teachers’ domination over the pupil target setting could be related to their assumptions about pupils’ ability to set effective targets. Teacher 7 reported:

"Pupils tend to set very generic targets such as ‘I need to work harder’ or ‘I should not talk too much’. To me it is quite doubtful that pupil target setting will raise their level of achievement. Yes, this is rather doubtful".

Teachers’ ideas on pupils’ ability to set high quality targets was related to their perception of pupils’ ability. Teacher 4 referred to lower achiever pupils and suggested:

"Most students will set targets which are not necessarily achievable or realistic. When they set, they won’t set it for a particular time span".

Teacher 2 described lower achieving pupils’ targets:

"...focusing on the subsidiary aspects of the work more than the ones that promote their learning, such as the presentation".

Here, the differences between teachers in linking ‘perceived pupils’ academic achievement’ with their ‘target setting ability’ become apparent. Teachers 1 and 3 who encouraged pupil target setting, made the point that the pupils’ ability to set high quality targets was determined by the quality of the teachers’ whole assessment practice, and in particular, their ability to effectively communicate assessment criteria. On the other hand, teachers who did not encourage pupil autonomy in target setting linked this to pupils’ ability in identifying meaningful and realistic targets. Lastly, teachers did not make any distinction between boys and girls of their differing ability to set targets.
Target Setting and Formative Assessment

The literature reviewed suggested that communicating an achievement gap and developing strategies to close this gap are the key aims of formative assessment (Black and Wiliam, 1998a). The research evidence demonstrated that target setting was a salient strategy for teachers to achieve this aim in communicating the targeted attainment level. However, there were several weaknesses in practice:

- The majority of teachers appeared to have difficulties in identifying subject specific targets to direct pupil learning;
- Targets in the assessment sheets appeared to be too numerous;
- The wording used in the target sheets appears to be very similar to in the geography level descriptions and teachers perceived that the pupils would find them vague;
- Targets were not reviewed systematically by teachers;
- Targets did not lead to a sequence of actions which could assess the development of pupils’ learning;
- It appeared to be a top-down approach which was teacher dominated and did not involve pupils’ input into decisions about their targets.

A key weakness in teacher target setting appeared to be teachers’ insecurities about what targets should be set. This then leads to further considerations on ‘whether what is set constitutes improvement in pupils’ learning’ or in other words ‘whether meeting these targets will genuinely represent an improvement in pupils’ learning’ (ATL, 1998).

In identifying targets for improvement, teachers relied completely on geography level descriptions, in some cases even on those whose potential to reflect pupils’ learning was viewed as doubtful. In their eyes, this resulted in a mechanistic approach over which they felt that they had little control. Targets were set in a prescribed manner rather than on the basis of meeting emerging individual learning needs. Indeed, this could mean that how raisin standards in a nationwide consistent language interferes in classrooms and affects teachers’ in class assessment practices.
SECTION 3: TEACHERS’ VIEWS ON PUPIL SELF and PEER - ASSESSMENT

This section illustrates the findings and analysis of teacher interview data on the practices of pupil self and peer-assessment. It starts by describing pupil self-assessment and is followed by peer-assessment with a focus on the specific tools employed. This is followed by teachers’ views about the validity and reliability of such practices and their perceived contributions to pupils’ learning. The section ends with a debrief of emerging issues some of which constituted a barrier in the regular use of pupil self and peer-assessment. How pupil self and peer-assessment appear in classroom situations is addressed in the section on classroom observations.

Practice of Pupil Self-Assessment

Teacher interview data revealed three major tools for facilitating pupil self-assessment as “assessment sheets”, “self-assessment reports” and “oral discussions”, as indicated below table:

<table>
<thead>
<tr>
<th>Specific tools used</th>
<th>Type 1</th>
<th>Type 2</th>
<th>Type 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>As a part of assessment sheets and checklists</td>
<td>Self-assessment reports</td>
<td>Oral discussions between teachers and pupils</td>
<td></td>
</tr>
<tr>
<td>Timing</td>
<td>At the end of a unit</td>
<td>At the end of a unit</td>
<td>Any time during lessons</td>
</tr>
<tr>
<td>Type of assessment task to which it is applied</td>
<td>Any task</td>
<td>Any task</td>
<td>More equipped to activities with oral discussions</td>
</tr>
<tr>
<td>Frequency</td>
<td>Once within a unit of work</td>
<td>Once within a unit of work</td>
<td>Day to day basis</td>
</tr>
<tr>
<td>Format</td>
<td>Written</td>
<td>Written</td>
<td>Oral</td>
</tr>
<tr>
<td>Assessment Criteria</td>
<td>Against predetermined criteria</td>
<td>More ipsative and self-focused</td>
<td>Against the objectives of activity or unit</td>
</tr>
</tbody>
</table>

Teachers reported the use of pupil checklists as the most common means of facilitating pupil self-assessment. Nine of the 12 teachers interviewed applied Type1. Here, pupils were asked to assess themselves on the basis of their expected learning outcomes, during the completion of tasks, by judging their attainment
against the predefined statements. Once pupils finished it, they were required to tick the items which they thought they had successfully completed. Teachers reported that pupils received a sheet for each assessment task and kept it with them during the completion of the task. In addition, five teachers required pupils to give themselves a grade on their attainment in non-subject-specific aspects such as “effort, co-operation with others (if it was a group task), personal organisation, attitude, concentration, speaking and presentation”.

The teacher interview data also revealed the use of pupils’ written self-assessment reports, which they were required to write at the end of certain tasks. Three teachers (out of 9 who used assessment sheets) reported the occasional use of pupil self-assessment reports. Here, teachers encouraged pupils to reflect upon their experiences by asking them to identify their strengths and weaknesses as well as the aspects of tasks that they found enjoyable and challenging. The following extracts exemplify two Year 7 pupils’ self-assessment reports, derived from the documentary evidence provided by Teacher 3. They were written by pupils about a debate, which was addressed in a role-play task.

“I think my strengths were my facts and props. I supplied lots of props not only me as a fisherman but also for other people as well. You can see that in my speech. I think I was a bit brilliant. I had a net with plastic fish in, sunglasses, a hat and a picture of the dam... My only weakness was that I stopped and I started a couple of times when I got to the end of a paragraph”. (Pupil A, year 7, Teacher 3’s class).

“I think that my strengths were: I worked on a longer piece of writing to show they I understood different situations of people relative to the Aswan Dam. I used a selection of props and co-ordinated head-dress (co-ordinated with the rest of the group of fisherman). I think my weaknesses were: I tried to talk in Arabic accent but I couldn’t. I could have had an outfit that looked like an Arab” (Pupil B, year 7, Teacher 3’s class).

As revealed earlier, teachers were concerned about pupils’ ability to find the appropriate assessment criteria to set their own targets. Consistent with this concern, in the above quotes, pupils tended to focus on their performances and costumes rather than critically evaluating their learning process for that particular unit of work (i.e what were my strengths and weaknesses in solving conflicting demands about a dam). This was not the intention of Teacher 3, as he suggested:
...this [pupil self-assessment] is not about asking them to summarise what they did, I am actually inviting them to recall their learning experience”.

This mismatch necessitates the exploration of how assessment criteria are communicated and reinforced. This is addressed in detail in the findings and analysis of classroom observations in which Teacher 3’s practices are considered.

A third tool for pupil self-assessment was reported to be the use of oral discussions during lessons. This activity was described as “very informal” (Teacher 8) and it was not a planned activity. In such moments, Teacher 8 invited pupils to talk about their experience of the task and particular difficulties that they experienced. Teacher 8 suggested that this activity generally happened during whole class discussions.

The employment of assessment criteria against which pupils evaluate their attainment showed variations among different tools used. In Type 1, which reported the use of checklists for pupil self-assessment, assessment criteria appeared to be in the form of qualitative statements. These were very much associated with the geography level descriptions (See Figure 4.4). In pupils’ assessment reports, pupils were asked to assess their performances on the basis of their weaknesses and strengths without much guidance on assessment criteria.

**Practice of Pupil Peer-Assessment**

The teacher interview data suggested that pupil peer-assessment was associated with assessment tasks that included group work. Teachers’ examples of such tasks involved role-plays, group presentations, group discussions and group preparations of a written piece, such as a newspaper article or a pamphlet. Teacher 5 suggested that, “it was the case where each group had prepared a pamphlet and marked other group’s pamphlets”. In Teacher 7’s practice it was:

“...they [pupils] prepared a news bulletin about the Taiwan earthquake. They [pupils] presented what they had prepared, everybody looked at the other groups’ and looked at the quality of the bulletin and their presentation performances”.
In another example, Teacher 3 mentioned a role-play task after which pupils had to assess each other on the basis of their strengths and weaknesses and select the best and the weakest five amongst their peers.

Overall, peer assessment was not a formalised part of teachers’ assessment practices and its frequency appeared to be low. All of the teachers mentioned the use of peer assessment only once or twice in a year. The low appearance could be explained on the basis of three factors. Firstly, teachers had other priorities for pupil self-assessment, as Teacher 5 explained:

“Pupil peer assessment occurs occasionally, but this is not a formal part of the tasks... It doesn’t happen a lot because we literally do not have time to finish all the set of specifications for Key Stage 3. So, things like pupil peer assessment fit in where they can. If we are convinced that it is necessary, it tends to happen on Friday afternoons as the last thing”.

For Teacher 5, there were other important priorities which were related to the GNC requirements. Teacher 5 was not alone in her thinking. The majority of the teachers interviewed perceived pupil-peer assessment as a ‘luxury’, which was a common justification for its omission in teachers’ practices. Secondly, the low frequency could relate to difficulties associated with its practical implementation and management. Teacher 9 suggested that:

“It is a time demanding thing ... we spent a good half of the lesson organising the class for the group work to make sure that they are doing well co-operatively. It is very time consuming”.

The time demanding task of arranging the appropriate groups (the arrangement of the group members / the organisation of the class) appeared to be challenging. In addition to the practical difficulties outlined above, teachers’ perceptions of its value, validity and reliability also shaped teachers’ peer-assessment practices as the table below demonstrates.
Table 4.4: Teachers’ Perceptions of Problematic Aspects of Pupil Self and Peer-Assessment

<table>
<thead>
<tr>
<th>Problems in Validity and Reliability</th>
<th>Teachers own accounts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Objectivity</td>
<td>“…they may not be quite objective in judging their friends” (Teacher 6).</td>
</tr>
<tr>
<td></td>
<td>“The reason why pupil peer assessment is challenging is because pupils make judgements on the basis of personality and other factors” (Teacher 10).</td>
</tr>
<tr>
<td></td>
<td>“When they assess themselves, they can be over critical, sometimes which ends with them becoming demoralised about their work” (Teacher 5).</td>
</tr>
<tr>
<td>Use of Assessment Criteria</td>
<td>“High achiever pupils are able to assess themselves and they generally use the right criteria rather than focusing on the trivial aspects of the task” (Teacher 4).</td>
</tr>
<tr>
<td></td>
<td>“They are not necessarily looking at the right sort of things so that pupil self and peer-assessment become another weapon that I would not use that much” (Teacher 8).</td>
</tr>
<tr>
<td></td>
<td>“Many of the them, in particular low achievers, will end up with comments which will be very trivial such as the neatness of the work” (Teacher 7).</td>
</tr>
<tr>
<td></td>
<td>“Them working with the National Curriculum levels in their own assessment is the hardest bit” (Teacher 8).</td>
</tr>
<tr>
<td>Maturity of Pupils</td>
<td>“We tend to not use such assessments in Year 7 because pupils are not mature enough in that year group” (Teacher 1)</td>
</tr>
<tr>
<td></td>
<td>“You need to be mature about it but some cannot be. They can easily say ‘Oh! I don’t like that person or I like that person so I should give her higher marks’. (Teacher 5).</td>
</tr>
</tbody>
</table>

Teacher responses indicated objectivity as the greatest problem. Nine teachers interviewed specifically used the word “problem”. Here, more concerns were reported about pupil peer-assessment than pupil self-assessment because of the involvement of personal factors when pupils evaluated their peers. Interestingly, Teacher 5 linked objectivity to pupil maturity and suggested that younger pupils were more likely to consider personal factors when assessing their peers. Teachers reported that that a certain level of maturity was required in order to achieve successful pupil self and peer assessment. Specifically, two teachers did not use pupil peer assessment with Year 7 pupils since they regarded them as not ready for such engagements. This finding is consistent with the previously reported teacher perceptions of Year 7 pupils, who were described as not mature enough in their skill development.
In the case of pupil self-assessment, five teachers reported students being over critical in assessing themselves, which was illustrated in the above table. The teacher interview data did not indicate that pupils’ objectivity in assessing themselves differed according to their own gender.

The use of assessment criteria by pupils has been investigated by Klenowski (1995) and was identified as a key contributor to the success of pupil self and peer-assessment. In the current study, teachers recognised the importance of assessment criteria but reported difficulties in understanding and using them by pupils. For teachers, converting level descriptions to a pupil friendly language to be used in self and peer-assessment seemed to a demanding task, as Teacher 8 suggested;

“we teachers find it quite hard to understand what level means, I don’t think it is realistic to expect that pupils will understand and use them to judge themselves”.

A second perceived difficulty reported by teachers related to pupils’ ability to use the appropriate assessment criteria. Seven teachers pointed out their concerns about pupils focusing on the appropriate criteria in assessing themselves and their peers. Here, three teachers reported a distinction between high and low achiever pupils and suggested that high achievers tended to use the appropriate assessment criteria, whereas low achievers focused on trivial aspects of their own and their peers’ work. Regarding the use of assessment criteria, teachers did not report pupils’ willingness to use assessment criteria, as was illustrated by Dann (1996) as a factor in explaining the involvement of personal factors.

Perceived Contributions to Learning

The interview data identified teachers’ perceptions about the contributions of pupil self and peer-assessment to pupils’ learning. Some distinctive quotations are tabled below.
Table 4.5: Teachers’ Perceptions about the Contributions of Pupil Self and Peer-Assessment to Pupils’ Learning

<table>
<thead>
<tr>
<th>The Learning Outcome</th>
<th>Teachers’ Accounts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reflection</td>
<td>“Student self evaluation encourages reflection. Being reflective also signals the areas of development, your new targets (Teacher 12).”</td>
</tr>
<tr>
<td>Empowerment</td>
<td>“They are put into the teachers’ role and something novel for them to think about. They look at things in a slightly different way, they respond to that quite well (Teacher 3)”</td>
</tr>
<tr>
<td>Interaction</td>
<td>“Very often they do not get the opportunity to see what the other people have done, it is good in this sense because they look at differences between their work and the others (Teacher, 1)”</td>
</tr>
<tr>
<td></td>
<td>“What enables self and peer assessment is discussions during a group work. What I call this is more like a learning process (Teacher, 2)”</td>
</tr>
<tr>
<td></td>
<td>“Definitely group work. Through sharing knowledge and skills, the children become aware of the standards and pick up from other children what is missing or what isn’t very useful” (Teacher 3).</td>
</tr>
<tr>
<td>Genuine Assessment Experience</td>
<td>“It is important to make them mark each other. It is more important than just telling them what constitutes quality of a particular piece of work (Teacher 5)” (as suggested, it is not good telling them what constitutes quality, they need to be interactive),</td>
</tr>
</tbody>
</table>

The teacher interview data revealed a consensus among teachers about the contribution of pupils’ self and peer-assessment to their pupils’ learning. Reflection was perceived as a key contributor. Six teachers out of 12 valued pupil self-assessment practices for its encouragement of reflection. Examples were given by teachers of how reflection developed pupils’ ability to understand their weaknesses, so that they better appreciated their learning. This recognition is consistent with the findings of Klenowski (1995), who regarded reflection as a key property of pupil self-assessment and a prerequisite for formative assessment.

Teachers valued the interactive dialogue which occurred between pupils during their peer-assessment activities. Three teachers viewed this as a powerful tool for making pupils aware of the assessment criteria. The use of group work and oral discussions were perceived as appropriate for this purpose. Furthermore, teachers linked metacognition to pupil peer-assessment, which was achieved by the internalisation of
assessment criteria by pupils and the subsequent use in assessing their friends and reflecting upon their own learning. The least recognised benefits were the empowerment of pupils in their learning.

Pupil Self / Peer-Assessment and Formative Assessment

The following points summarise the identified problems and explain the reasons for its omission:

- Teachers had difficulties converting level descriptions to a form that pupils could make sense of and use for their own and their peers' assessment.

  This was an apparent challenge that teachers faced and appears to be an area which they require external assistance.

- The amount of group work appeared to be low due to the difficulties related to its management.

  As reported previously, pupil peer-assessment appeared within tasks which involved group work. For teachers, group work was regarded as being time consuming and challenging and was not a priority that they considered in their task design. The major challenges were related to arranging classrooms and groups and sustaining efficient group work through constant monitoring.

- Teachers felt that they were relinquishing their power as assessors - a situation which they were reluctant to concede control.

  Teachers appeared to struggle in handing the power of being an assessor from themselves to pupils. In order for teachers to accept self and peer-assessment, a fundamental change in their perceived identity as assessors is required.
Teachers felt that they needed to preserve pupils’ self-esteem by limiting the exposure to the general classroom, of assessment information on individual pupils’ work.

Teachers reported that when pupils assess each other the privacy of the assessment outcomes were lost and pupils exposed what they had done to the others. In an extreme case, this was perceived as ‘naming and shaming’. A few teachers tended to use pupil peer-assessment with the higher ability classes or “top sets” and deliberately avoided using with pupils with lower confidence and lower ability.

Pupil self-assessment appeared to be a highly structured and teacher-controlled activity within which pupils had procedural autonomy.

As stated previously, pupil input in deciding on assessment criteria for their self and peer-assessment was not a priority for teachers. The reported practice of self and peer-assessment appeared to be a ‘top down’ approach with teachers controlling the overall practice. Thus, this process was not a situation of teacher – pupil negotiation.

SECTION 4: TEACHERS’ VIEWS ON FEEDBACK

Assessment feedback is a key component of formative assessment (Black and Wiliam, 1998a), and the discussions in the literature describes this potential for improving pupils’ learning. In the following, the first subsection describes teachers’ reported feedback strategies. Then, the subsequent subsections address the factors affecting the construction of feedback – these factors are time, teachers’ dilemmas of communicating pupils’ weaknesses, departmental policies and teachers’ control over the feedback process. The section ends with a debrief of emerging problematic issues which constituted a barrier in the accomplishment of formative assessment feedback. In this section, most of the interview relates to teachers’ feedback on written tasks.
Teachers’ Feedback Strategies

In the literature review, recent typologies of teacher feedback were provided which helped to conceptualise formative functions of feedback (Gipps et al, 2000; Tunstall and Gipps, 1996). The current study suggests that teachers responded to their pupils’ work according to different feedback strategies depending on the assessment task format. For written tasks, the most common form of assessment feedback was reported to be written comments and marks at the end of each task. Here, the majority of teachers distinguished between quantified and written qualitative feedback. In teachers’ examples which stem from both interviews and documentary evidence, their written comments somewhat resembled “the specification of attainment”, which are described as ‘feedback communicating why pupils’ work were correct and what they have and have not achieved of’ by Gipps et al. (2000) and Tunstall and Gipps (1996). In explaining such feedback strategy, eight teachers also mentioned communicating ‘pupils’ strengths and weaknesses’. These teachers made the point that feedback was more effective in situations in which students were specifically made to see their mistakes rather than feedback which simply states a grade or mark or whether their response was right or wrong. A similar conclusion which was reached by Butler (1988).

The use of written comments, specifically for the purpose of suggesting the way forward in learning, appeared to be less common then teachers’ oral in-class feedback. In these cases, the content of such feedback involved suggestions to pupils of other ways of doing things. This appeared to be similar to the feedback strategies advocated by Gipps et al. (2000) and Tunstall and Gipps (1996) which referred to “specifying improvement by suggesting or implying a better way of doing something”. In the current research, this type was exemplified by teachers asking, ‘Have you thought about ...?’ (Teacher 7) or “How about doing …next time” (Teacher 11).

When describing feedback derived from oral interactions, teachers often mentioned the ‘feedforward’ function. Eight teachers suggested that oral activities were better
equipped to accommodate such functions since they could encourage pupils to talk, to allow teachers to make a quick assessment and to suggest how they could improve in a short time and with greater frequency. Teachers often did not attach the same properties to written feedback.

Teachers also provided other feedback strategies within their practices. Such examples constituted approvals and the provision of rewards when achievement was satisfactory and disapprovals when achievement was not satisfactory. These types of feedback, classified as ‘evaluative’ in the Literature Review, were regarded as not having direct links to formative assessment (Tunstall and Gipps, 1996). In this research, therefore, the scope in this section is limited only to feedback that had a descriptive function.

**Time as a Concern**

Time limits appeared to be a key issue for teachers in achieving regular feedback. Nine teachers reported the perceived advantages of high frequency feedback in raising pupils’ awareness of how they were doing in their learning, as exemplified following comment by Teacher 8:

> “Regular feedback is important. The more you tell people how they are doing and where they are at, the more responses you are going to get”.

In discussing frequency, teachers made a distinction between their oral and written feedback. All of the teachers interviewed agreed that the time available for regular written feedback was limited and pointed to the impossibility of checking every single piece of work that pupils wrote. For instance, Teacher 3 commented that:

> “When you teach 330 children a week there are limits. How much can you do? It becomes impossible to write comments on everything they produced”.

Limitation of time was not only a concern for written feedback but it was also reported to cause a limitation on “assessment practice” as a whole. Ten teachers specifically
mentioned that teaching was a very demanding profession. In expressing this, Teacher 5 drew analogy to a marathon runner, saying that:

“They need to finish one activity, move to the next one, collect the work, provide comments and marks then make the arrangements for forthcoming task”.

During the interview, 6 teachers did not indicate that they made a connection between assessment to teaching, as Teacher 11 commented:

“...assessment should not be bolted on to the end of units but if you are not careful you end up doing everything for assessment and where is the learning, and the fun, teaching and interacting. There has to be body of knowledge and understanding before you can say how do I find out what you know”.

In addition to this difficult schedule, assessment had to fit in somewhere in the unit. Many times teachers implied that they had very little time for debriefing sessions in which they could reflect back on the learning process and on pupils’ experiences. In these teachers’ practices, they needed to move on to the next topic or task, without having the time for critical reflections.

Feedback Content

Interviews sought to understand how teachers described the content of their feedback and what shaped it. Here, teachers interviewed made a distinction between their feedback on written tasks and feedback generated during the oral interactions. Essentially, when feedback was written, it was regarded as standard driven derived from the assessment checklists. Conversely, feedback generated during the oral interactions was regarded as more individualised and not necessarily related to the standards or objectives, “but related to pupils’ emergent needs” (Teacher, 6).

Feedback – The Dilemma of Communicating Pupils’ Weaknesses

The extent to which feedback should convey messages of pupils’ failure has been reported as a dilemma for teachers in the literature reviewed. For instance, Torrance
and Pryor (1996) reported teachers’ concerns to protect pupils which caused them to provide feedback which only partly communicated pupils’ errors. Similar occurrences were reflected in the teacher interviews in the current research. Six teachers commented that they should communicate pupils’ strengths more than their weaknesses. Teachers justified this position by relating positive feedback to pupil encouragement and motivation. Teacher 4 explained this as follows:

“Feedback should have a positive element since it should encourage and improve pupils’ positive attitudes towards learning and get them to feel that they have done successfully”.

For Teacher 7 “pupils’ readiness” for accepting their weaknesses was a factor affecting feedback practice while departmental policy also had a bearing. This is expressed in the following quote:

“Many pupils are more prepared to hear their strengths than their weaknesses. However, every piece of work that they get back must - that is the faculty policy – have a written comment on it and it must have something positive. So, even if it is the worst piece of work, you must find something good about it.”

Pupil readiness was also related to pupil maturity, as Teacher 3 suggested:

“...they are at the age of 11 to 14. They are adolescent and many of them have quite low self-esteem. We need to help them to build up their confidence...yes I need to find something positive”.

Feedback and Departmental Policies

Teacher interview data suggested links between the ways in which teachers conveyed feedback and departmental policies. Firstly, consistency across a department was a key driver for determining the feedback format. Ten teachers explicitly reported that feedback format was mainly determined by departmental policies, which was justified in terms of the need for consistency and precision. All of the teachers mentioned that they were currently increasing their use of levels for summing up pupil achievement at the end of tasks. Four teachers suggested that using levels exclusively was more helpful than taking results in different forms and later harmonising them and
converting them into a final form. Seven teachers supported a consistent marking scheme on the basis that it was helpful for pupils to develop an understanding of their performance.

Indeed, 'consistency' was one of those words which was very evident and frequently used in schools' policy documents. The following example is an extract from the 'Assessment, Recording and Reporting Policy, which Teacher 6 used:

"Marking and assessment of all types should be consistent (assessment policy, use of grades/marks, marking exercise books). Consistency is vital within an individual's classroom and classes. It would also be recommended that assessment and marking procedures within a department should show a high level of standardisation. Mark schemes should always be shared and similar assessment tasks used this helps to keep up the shared understanding. This is particularly important where children may have more than one teacher in their school career".

The use of the same marking scheme as a departmental policy was not universally welcomed by all of the teachers interviewed. Two teachers took a negative standpoint about the use of this marking scheme. At an extreme end, this was seen as a threat to teacher autonomy. Teacher 8 viewed this as, "another example of how levels dictate to us". The concept of 'threat to professional autonomy' was also visible in the following discourse:

"I am not free in the way in which I can mark my pupils and keep my assessment record. I have to operate in a way that does not come from me. Indeed, it is the imposition of certain things from outside" (Teacher 4).

There have been also positive comments about levelling every major assessment tasks. Three teachers commented that it helped them to moderate the validity and reliability of their judgement (final level decisions) with other staff in their schools. Overall, feedback format was not solely teachers' choice, but was also a matter of departmental decisions.

Departmental assessment policies also shaped feedback content, although to a lesser extent than format. This was mentioned by both Teachers 3 and 4. For instance,
Teacher 3’s written feedback specifically addressed corrections of grammar and corrected spelling mistakes. The motive behind that was the request from the English Department that aimed to improve pupils’ use of grammar in humanities lessons. A similar intention was reported by Teacher 4 who taught pupils for whom English was their second language.

**Teachers as Owners of Assessment Feedback**

Teacher interview data suggests that teachers had individual ideas regarding means of feeding forward to pupils to benefit their learning. Tomlinson (1999) called this teachers’ tacit knowledge and described it as their personal experience of how to react and respond in situations. Overwhelmingly, teachers’ language in describing their assessment practices indicated that they perceived the feedback process as something that they should control, for the pupils’ best interests:

"I should set their targets for their improvement…” (Teacher 4)

"I check their work when I can, they also keep the assessment sheets that I devised …. (Teacher 2)

“…they tend to focus on wrong criteria if I don’t give..”(Teacher 11)

The feedback practices of all teachers were congruent with a domination discourse with them being the dominant individuals. To reiterate on the dominant practice; pupils were given predefined assessment criteria, which were devised by teachers. Then teachers checked whether pupils were on track, and then set targets at the end. When pupils set their own targets, teachers used assessment sheets or checklists which were provided pupils a list of possible targets.

This practice implies a ‘receptive-transmission’ model of teaching during which experts (teachers) drive the whole process and give advice on how to improve (Askew and Lodge, 2000). This research portrayed this model, in which the way forward was achieved by teachers ‘specifying attainment’ rather than through ‘mutual construction of improvement’ where pupils and teachers engage in discussions and where power
relations are invisible. Even when they engaged in discussions regarding the results of their assessment tasks, a situation which reported by only three teachers, the nature of interactions still appeared to be one-way, as suggested by Teacher 11:

“...when I give marks, I go to the pupil and have a brief conversation and tell them whether I think they did better or worse than their last performance…”

Having described the factors influencing the construction of teacher feedback, the following summarises the key conclusions of this section.

Emerging Issues

The analysis of interview and documentary data provided insights into teachers’ understanding of formative assessment feedback. Teachers felt that feedback could have a profound effect on the quality of learning if it provided pupils with specific information on how to improve their learning. Their perceptions of feedback for learning were grounded in their beliefs (such as preserving pupils’ self-esteem) and experiences rather than from guideline materials of what constitutes good assessment feedback. Teachers appeared to like following their “gut feelings”. However, they also felt that they needed to fit these in the current national framework. Therefore, what was portrayed as ‘formative assessment feedback’ was a controlled process, but still kept the key intentions as were outlined in the literature.

- **Written assessment feedback was standardised and individualized on the basis of level descriptions.**

This was mainly in response to the need to satisfy the requirements for departmental consistency and concerns for the manageability of final level decisions at the end of the Key Stage 3. The result of that was only limited assessment evidence (i.e. just the items appear in the assessment checklists and sheets) were elicited. There was a possible loss of other learning evidence, which, therefore, was not acted upon.
Teachers had a struggle in giving up some of their power and they were the only agents in giving feedback.

Regardless of feedback format, teachers had the control over the feedback process. There were very limited opportunities for after task debriefing or tutoring to increase pupil input. Pupils had no involvement in the construction of assessment criteria. Therefore, the dialogue for feedback construction was appeared to be ‘one-way’ communication.

SECTION 5: TEACHERS’ VIEWS ON LEVEL DESCRIPTIONS

Another focus of teacher interview is an exploration of how geography teachers used their assessment evidence which they had elicited throughout Key Stage 3 to derive a summative level. In order to illuminate this aim, the first subsection discusses how teachers interpret level descriptions. This is followed by teachers’ experiences of communicating levels to pupils. The third subsection explores how teachers give levels to assessment tasks. The fourth subsection highlights the quantification procedures used and continues by describing the recording of assessment evidence. Subsequently, the ways in which final level decisions are made is explored.

Interpretation of Level Descriptions

Teacher interviews revealed that working with level descriptions is beset with problems and concerns. The most common problem was the difficulty of understanding what was required to achieve each level, as Teacher 2 suggested:

“...they are abstract statements...I think levels are very difficult to understand, you really need to pick out what they mean. We spent quite a bit of time trying to understand what they really mean” (Teacher 2).

“Picking out” what level descriptions ‘really’ meant was hard because:
"... the National Curriculum mentions the relevant geographical issues and questions, where is the line between them? By the way, can someone tell me what is a geographical question and what makes it geographical anyway?" (Teacher 5).

Understanding the individual statements of the level descriptions was a common and distinct problem from which almost every teacher suffered. For some, one-way of resolving this was to breakdown the statements so that some meaning could be gained; this approach is exemplified in the ‘quantification of assessment evidence’ subsection. A less common way which was reported, involved considering the overall meaning or impression of the level rather than the individual components.

As commented previously, half of the teachers tended to divorce teaching from assessment. Teacher 7 commented that levels had a role to play in this, as follows:

"The National Curriculum shows what you need to teach and assess. It does not tell how you do that. But often these two do not come together. When you think too much about levels, you tend to assess more. I think there is a tension between teaching and assessment which is actually caused by the existence of levels" (Teacher 7).

One of the challenging aspects of working with levels appeared to be their communication to pupils, as the following subsection reports.

**Communicating Level Descriptions to Pupils**

The communication of levels between teachers and pupils was referred to in previous sections. As pointed out, teachers devised their own tools to ease this task such as checklists and assessment record sheets. These sheets included specific information on what pupils needed to do in order to achieve each level as well as on what basis they were judged using level terminology. Additionally, only Teacher 1 mentioned the use of portfolios to make pupils aware of the levels by actually demonstrating levelled examples of previous tasks.
For every teacher, communicating levels to pupils involved translation into a pupil friendly language. This was one of the biggest challenges that the teachers encountered. Several teachers mentioned their concerns as follows;

“Even we (teachers) can get bogged down with what exactly each one [level] means and exactly where the boundary line between them is. How are we going to explain them to the pupils?” (Teacher 8).

“It is up to us as teachers to try and make them [level descriptions] pupil friendly so they can use them and see where they are going...this is a very difficult area and that we haven’t been any guideline or documents as to how we are going to actually achieve this” (Teacher 5).

“Now all children receive a copy of these statements which will be pinned down in the back of their note books. This says that increasingly the children are going to be levelled... what we generally end up doing is just copying and pasting the statements from the National Curriculum. There is no point saying that this is Level 6 when the kids have no idea what Level 6 is in this particular work” (Teacher 1).

“We need to communicate what we mean by each level to pupils as well as how they can improve and reach to the next level. This issue has not been faced yet by the government... In terms of explaining theory regarding progression from one level to another, explaining them in language that they really make sense is left to us, indeed that is something quite difficult to achieve” (Teacher 11).

“Levels cannot be understood by the pupils, They are too big for them to monitor their progress anyway” (Teacher 7).

The data above clearly highlights teachers’ difficulties in translating levels into a language which pupils understand. Having defined the challenges associated with the communication of level descriptions, the following subsection describes the ways in which teachers approached the levelling of assessment tasks.

**Working with Level Descriptions: Levelling Assessment Tasks**

As described previously, teachers reported that more and more assessment tasks are being levelled. However, there was an apparent dissatisfaction which related to their interpretation in judging learning outcomes, in particular identification of boundaries between consecutive levels. In order to compensate for this, five teachers subdivided
levels to increase the precision of their judgement. Teacher 2 mentioned the use of a marking scheme based on levelling each assessment task. As a head of department, with the new requirements from the Local Educational Authority (LEA) she found herself using a more complicated marking and recording scheme. This scheme which was advocated by the LEA involves each level being divided into three categories, being As, Bs and Cs. For instance, 4A corresponds to a weak level 5, 4C means a weak level 4 while B grades are in the middle of the respective classifications. Before this arrangement, the teachers in the school were free in the way in which they marked and recorded the assessment data.

Without being aware of the LEA’s requirements, 4 other head-of-departments also used this ‘divided level system’. Teachers expressed mixed views regarding this division. For these four teachers, it was necessary for precision. In particular, it was helpful in monitoring progression within a given level. For other teachers who did not use a divided level system, it was viewed as just another detail that the central government and the LEA made in order to keep them busy. The findings suggests that variations mentioned above were a result of teachers’ differing ability in interpreting and being able to adapt and use them.

The teachers also expressed views on what was the perceived formative capacity of dividing levels into sub levels. For teacher 5, this subdivision was found to be “a truly formative record of assessment” since it created a precise record of where a pupil was located within one level. For teacher 7 and 11 this subdivision had a key role which was more than a ‘marking scheme’. In these two cases, this was viewed as being ‘very motivating’ since the pupils had a chance to view their progress within one level. Teacher 2 viewed pupils’ experience in the following way:

“They come here in Year 7 and they are Level 4 ... they finish a whole year they are still Level 4. Even though they may be Level 4 in the first bit of Year 8. Levels are too big for them to understand whether they have made any progress at all”.

146
In the above paragraphs, teachers' experiences of levelling assessment tasks are explored. The following subsection describes how teachers recorded assessment evidence towards making a final level judgement.

**Recording of Assessment Evidence**

Marked or levelled work was subsequently aggregated or accumulated before final level decisions were made. Overwhelmingly, teachers used their mark book to store the assessment data that they had recorded for the three-year duration of the key stage.

Despite the slight differences which exist, the mark books contained similar features. They were generally presented with the pupils' name on the left hand side and the accompanying marks or levels for each unit of work adjacent to each name in a separate cell. Such tables were applicable to one assessment task which generally took three to four weeks. Levels, marks and grades were included. In some cases, pupils' efforts were also marked.

When asked about the potential advantage of such summary records of pupils' achievement, teachers' answers indicated three broad categories. These were:

- Being able to monitor one pupil's achievement throughout the year;
- Being able to monitor a class of pupils;
- Being able to compare achievement.

The first category above was given a higher emphasis in the majority of teachers' responses. Monitoring pupils' performances 'in the long term' and 'medium term' were also the words that were used to explain the first item. Teacher 6 also had a system in which she monitored pupils' performance within one assessment task, again in quantitative terms. 'Computerised tools' were also reported by Teachers 5 and 12. For Teacher 6, this was a computerised tool which was used when parents requested the assessment information of their children.
What was striking about such mark books was their highly numerical nature, generally in the form of levels and marks. There was almost no space for teachers’ reflective comments on specific aspects of pupils’ learning. During the interviews, the quantitative nature of the recordings was described as ‘a more manageable way of recording’. However, all teachers highlighted that the ‘real assessment information’ was in their head and the mark books were just a summarised record of pupil attainment at particular times. So, in a way teachers made a distinction between formative assessment evidence which is recorded in their heads, and summative assessment evidence which was recorded in marks books. So, then the real issue here is how the teachers use these two different forms of recorded data, and which one they rely on more and how they actually make a final level. The following section explores the marriage of these two different recorded information and identifies the ways in which teachers manage the situation.

**Final Level Decisions**

From the interview data, it is possible to gain a glimpse of the individual variations in teachers’ practices in deriving final level decisions. Building on what has been said with regard to their marking and recording and how they derive a final level, the practices of the twelve teachers’ interviewed were categorised as ‘flexible’ and ‘cautious’.

**Flexible Levelling**

The flexible type of levelling is where professional judgement has the greatest influence over decisions. Teacher 2, one of the seven in this group, summarised the core of this type of practice in the following way:
“There is neither a mathematical formula nor a method of calculation. There is no clear cut. You look at the National Curriculum levels, you look at your record of particular pupil’s levels, marks or grades. More importantly, you have been with this child for the last three years. You know her or him more than the levels. The pupils also have ups and downs. It is actually your professional judgement which tells you the point where you need to stop and which level actually your pupils worked within”

Explicit in the above statement are echoes of the views expressed by other teachers. Here, the teacher quoted above placed a high priority on the ‘flexibility’ and ‘freedom’ of using her professional judgement. More experienced teachers’ reflections indicated comments such as ‘it is gut reaction at the end of the day’. In Teacher 12’s view:

“My professional judgement is the main drive to decide a final level. It is the main drive because I know the child the best. It does not matter what they have done, in particular in the summative exams. The result of the tasks just reinforce what is already in my head about the child”.

Professional judgement needed to be the main criteria for the teacher above and it should be used with care. He continued:

“Professional judgement is the tool that keeps your balance in this chaotic system. In the end, what we found with the levels is that professional judgement is really the only way, there is no other way of determining the levels”.

Professional judgement was not the exclusive means of decision-making for all of the teachers who were flexible in their final level judgements. There was evidence to suggest that teachers also looked at the level descriptions and actually decided whether pupils meet the requirements of a particular level. For three teachers in this group, meeting almost half of the statements was enough to allocate a pupil to a certain level. For three teachers, the yardstick was two thirds. Teachers operating in this flexible mode seemed to be cautious but still flexible enough to place their professional judgement as the major contributor to their decisions.

Cautious Levelling

“There are quite a number of variables that affect my final decision about pupils. I go back to my mark book, look at their attendance, their marks, grades and their effort. Then I look at their end-of-year summative exams. In a way it is the backing up. We need to consider how well they have done during the Key Stage 3 as a whole” (Teacher 10).
The above comment is characteristic of teachers engaging in ‘cautious’ levelling. Five out of twelve teachers were considered to be in this category. Interestingly, ‘cautious levelling’ behaviour occurred in the schools in which summative end-of-year exams were used in addition to other assessment tasks that were implemented during the year. For teachers in this category, end of key stage summative exams had a key role in confirming or justifying their decisions. Then the issue here was whether the summative result matched the average level that a pupil has achieved during a given year. As Teacher 3, one of the teachers in this group, commented:

“Surprisingly enough, the levels coming from the summative assessment matches with the levels coming from the pupils’ other projects. This surprises me because I expect the summative to be harder. In the majority of the cases, it matches with the other assessment evidence”.

The ‘convenient match’ was the case for another teacher in this category. However, summative exam results were more guarded when this match did not happen, as Teacher 6 suggested:

“When we have different types of assessment evidence - that we have called formative and summative - to me the formal side of assessment is more valid. However, it does not mean that I give the weight to the summative result. I try to find the overall balance of really”.

Beliefs regarding the validity of different assessment was also a salient intervening parameter in teachers’ decision making. For cautious teachers, having a summative exam was a vital part of this decision making. In this view, the absence of such a measure would result in unreliable final level decisions.

Summary Outline of the Findings and Analysis of Teacher Interviews

In the first part of the Chapter 4, findings and analysis of teacher interviews focused on teachers’ thinking and use of formative assessment. Here, the findings were presented in five sub-sections. The first section explored teachers’ task design criteria as well as how they planned for progression and differentiation in their practices. The second section presented findings on how teachers used target setting and explored the
nature of targets that they set for their pupils before identifying the weaknesses in this area. The third section presented teachers’ views on pupil self and peer-assessment and outlined how teachers used these activities. The fourth section presented teachers’ reported feedback strategies. The last section described teachers’ experience of working with Key Stage 3 geography level descriptions and described how they derived a final level decision.

Having presented the findings and analysis of teacher interviews, the next part (Part 2) of this chapter illustrates the findings of classroom observations.
PART 2: FINDINGS AND ANALYSIS OF CLASSROOM OBSERVATIONS

One of the difficulties in identifying what constitutes formative assessment is that it mainly takes account of the theoretical understanding of the issues involved (Torrance, 1993). Therefore, one of the problems relates to knowing formative assessment well theoretically, but not identifying what it actually looks like in day to day practice in classrooms (ibid). There have been a few projects which have characterised formative assessment through classroom interactions in primary schools (Torrance and Pryor, 1998). Later studies explored formative assessment in classrooms in secondary schools in science and mathematics (Wiliam and Lee, 2001). However, there remains a lack of research on geography Key Stage 3 classrooms.

This part of this thesis presents the research findings and analysis of classroom observations with an overall aim of exploring and exemplifying its appearance in two selected Key Stage 3 classrooms. The classroom observation data is also supported by the findings of follow-up teacher interviews with the teacher observed. This part consists of four sections. The first section outlines the context of classroom observations and describes the nature of assessment tasks that were observed. The second sections presents research findings and provides insights into the implementation of the assessment tasks. This section also explores the elicitation of assessment evidence and pays attention to the factors influencing elicitation of assessment evidence. The third section presents the findings of teacher feedback by describing feedback strategies which the teacher used before describing and contrasting the situations in which formative assessment feedback was and was not achieved. The fourth section describes the observational findings on the communication of geography level descriptions. The final section describes the use of pupil self and peer-assessment by analysing the nature and use of assessment criteria before highlighting a number of key issues related to the use of such activities.
SECTION 1: THE CONTEXT OF CLASSROOM OBSERVATIONS

As stated in the Methodology Chapter, the observation phase was a 2-month study involving 1 teacher in 1 school in two Key Stage 3 classrooms one of which was Year 8 and the other was Year 9. The researcher spent one lesson each week in two classrooms and observed 15 lessons as a non-participant qualitative observer. One group was a “mixed ability” class of Year 9 level while the other was a “higher ability” class in Year 8. The same teacher taught both classes. The teacher is male and has been teaching geography for thirty six years. He is a subject specialist teacher and head of the geography department. He has been taught the same classes for two years.

Description of Assessment Tasks

Year 9 Group Assessment Task

For the Year 9 group, the unit of work was completed within 15 lessons, over a 7 weeks period. It was part of a wider unit relating to ‘Ecosystems and Sustainable Development’. The following table (Table 4.6) summarises the medium term plan:
### Table 4.6: The Medium Term Task Plan for the Year 9 Group

#### YEAR 9 ECOSYSTEMS AND SUSTAINABLE DEVELOPMENT

<table>
<thead>
<tr>
<th>LESSON</th>
<th>LEARNING OBJECTIVES</th>
<th>TEACHING/LEARNING ACTIVITIES</th>
<th>ASSESSMENT OPPORTUNITIES</th>
</tr>
</thead>
</table>
| 1       | ♦ Describe, and explain what affects climate  
         ♦ Identify the names of climate types in general.  
         ♦ Identify and explain the 4 main factors – latitude, distance from the sea, prevailing winds, relief | ♦ The teacher set the task  
         ♦ The teacher explained the task in depth  
         ♦ The teacher surveyed pupils’ prior knowledge and understanding to assess a baseline  
         ♦ Pupils answered questions from their textbook and prepared a diagram for answering a number of questions | ♦ Questioning  
         ♦ Teacher questioning  
         ♦ Teachers observing and intervening when pupils answered the written questions  
         ♦ Teacher provided feedback for Individuals, and the whole class during the written activity |
| 2       | ♦ Discuss further features that give rise to different types of climate  
         ♦ Relate these factors in explaining the variations in vegetation.  
         ♦ Introduce people of tropical rain forest (TRF). | ♦ Pupils copied a diagram in showing vegetation in TRF  
         ♦ Pupils and the teacher interpreted the diagram  
         ♦ Pupils used their atlases to find out further geographical features of the selected TRF area | ♦ Open and close ended questioning during the lesson  
         ♦ Teacher observed and provided feedback to pupils while they were drawing and interpreting it |
| 3       | ♦ Explore further issues in relation to TRF people’s life and environment  
         ♦ Explain how TRF people’s life are shaped by their environment | ♦ Pupils watched a video about the TRF  
         ♦ Pupils wrote a short essay under the four selected headline on the life of TRF people | ♦ Teacher’s and students’ immediate reflections on the video |
| 4       | ♦ Discuss the nature of case study  
         ♦ Discuss with pupils about possible ways of structuring and presentation of the case study | ♦ Pupils started writing up their case studies | ♦ Establish the assessment criteria |
| 5       | ♦ Reviewing of case studied which were submitted  
         ♦ To recap the conclusions and findings of the TRF unit of work | ♦ Whole-class discussions about the case studies | ♦ Teacher reinforced the assessment criteria by exemplifying what constituted success in this case study |
| 6       | ♦ Teacher introduced the concept of ‘finite resources’ and provided a context (the water problem in Phoenix)  
         ♦ Describe the nature of the problem | ♦ Students watched a video tape on the water problem in Phoenix | ♦ Whole-class questioning  
         ♦ Reflections on the video |
| 7       | ♦ Teacher continued to investigate the problem  
         ♦ Teacher encouraged pupils to set hypothesis which to be investigated through an enquiry project. | ♦ Pupils brainstormed and wrote their own questions for the enquiry  
         ♦ The teacher and pupils discussed about the questions and possible avenues of investigations | ♦ The teacher commented on the enquiry questions  
         ♦ Oral feedback interactions between the teacher and pupils about next steps o |

154
The lessons observed were part of a wider unit on ‘Natural Environment and Sustainable Development’, which was addressed by studying climates and life in Tropical Rain Forests (TRF). The unit of work for the TRF was carried out over a classroom period of five weeks (ten lessons) with this extending into an enquiry project about the conflict between humans and environment in the last two weeks (Phoenix water problem). Towards the end of the third week, pupils were asked to prepare a written report (case study) on the geographical characteristics of Brazilian rainforests and the lifestyles and the culture of its inhabitants.

For the Phoenix water problem, the teacher spent two lessons introducing the water problem within a developing area. Then he asked pupils to develop a hypothesis in order to address the problem of finite resource (water) in a developing area. This led to an enquiry project during which pupils were asked to identify their own questions to investigate the issue. They were also asked to develop their own solutions to solve the identified problems and discuss the feasibility of their solutions in their report as well as in the classroom. This task involved class discussions and answers to the written questions from the chosen textbooks. Each pupil was asked to submit an enquiry report to communicate his or her research and findings.

*Year 8 Group Assessment Task*

For the Year 8 group, the unit of work was completed within 14 lessons. It was part of a wider unit relating to ‘Population and Migration’. The following table (Table 4.7) summarises the medium term lesson plan:
Table 4.7: The Medium Term Lesson Plan for the Year 8 Group

<table>
<thead>
<tr>
<th>WEEKS</th>
<th>LEARNING OBJECTIVES</th>
<th>TEACHING/LEARNING ACTIVITIES</th>
<th>ASSESSMENT OPPORTUNITIES</th>
</tr>
</thead>
</table>
| LESSON 1 | ♦ Describe, and explain global population distribution  
♦ Identify features that affect the population density of an area                                                                                                         | ♦ Students answer the questions from the textbook to further explore global population distribution  
♦ Teacher explains differences in population density in local area                                                                                       | ♦ Questioning  
♦ Individual, group and whole class feedback on the written activity                                                                                       |
| LESSON 2 | ♦ Identify features that give rise to high and low population density  
♦ Explain the reasons for changes in population over time                                                                                                 | ♦ Students answer the questions from the textbook  
♦ Students and the teacher interpret graphs and data about changes in population (from the textbook)                                                 | ♦ Open and close ended questioning during the lesson  
♦ Homework assignment: 'Why Leicestershire has more people than Northern Scotland?'                                                                        |
| LESSON 3 | ♦ Explain why people migrate  
♦ Describe different types of migration  
♦ Identify push and pull factors                                                                                                                                             | ♦ Students fill in a work sheet on push and pull factors                                                                                                             | ♦ Feedback on the work sheet                                                                                                                                     |
| LESSON 4 | ♦ Introduce the role-play  
♦ Explain how rapid town growth affects people  
♦ Understand the living conditions of a favela                                                                                                                               | ♦ Students watch a video on a family who lives in a favela in Brazil                                                                                               | ♦ Identify evidence on the video about the services available and the standards of living.                                                                  |
| LESSON 5 | ♦ To support students to give oral presentations  
♦ To help the students to assess the other groups                                                                                                                                  | ♦ Organise the group(s) for the role-play  
♦ Establish the assessment criteria for the role-play  
♦ Prepare the observation sheets (use of facts, speaking, rehearsal and comments)  
♦ Students write the scripts in groups                                                                                                                        | ♦ Teacher monitors groups while they are working                                                                                                                |
| LESSON 6 | ♦ To support the students to carrying out the role-play  
♦ Debriefing                                                                                                                                                                      | ♦ Students in groups give oral presentations adopting a role                                                                                                       | ♦ Teacher’s and students’ immediate reflections on the role-plays  
♦ Students fill in the observation sheets  
♦ Homework task: Prepare the self and peer assessment reports                                                                                                  |
| LESSON 7 | ♦ To give students an opportunity to demonstrate in an alternative form what they have learned                                                                                                                               | ♦ Prepare a poster for summarising what the students have learned about the living conditions in favelas                                                            | ♦ Teacher comments on self and peer assessment reports  
♦ Feedback on the posters                                                                                                                                           |
For the Year 8, the lessons observed focused on part of a wider unit Population and Migration. The teacher spent the first three weeks of the unit addressing causes and effects of changes in populations of regions and countries. This was followed by the teacher providing a case example which involved a story of a family, who migrated from a village to a big city in Brazil. The pupils watched a video illustrating the new life of the family. They were asked to prepare a role-play in groups in order to reveal their understanding of the causes and implications of migration for the selected family. Pupils were asked to write the script, act it, and then evaluate themselves and their peers after the role-play. Pupils submitted their evaluations in the form of written reports to the teacher.

SECTION 2: FINDINGS AND ANALYSIS IN RELATION TO ASSESSMENT TASKS

This section provides the findings and analysis of classroom observations on assessment tasks. The teacher interview data revealed teachers’ task design criteria. This section extends these findings by providing insights into the problems and possibilities in relation to task implementation in two selected classrooms. In order to achieve this, the first subsection reports findings regarding pupils’ responsiveness to assessment tasks. The second subsection provides an analysis of how assessment evidence was elicited through teacher questioning, teacher observations, and review of homework. The third subsection continues to identify other factors that were influential in the process of elicitation of assessment evidence - a trusted relationship, class culture and differentiation. This section ends with a brief of the key findings in relation to elicitation of assessment evidence in the classrooms observed.

Pupils’ Responsiveness to Assessment Tasks

Pupil engagement in tasks has been identified as a critical aspect of evaluating assessment tasks in the Literature Review chapter. Researchers explained this by exploring the relationship between task context and pupils’ social experiences
In the current research, the scope of the classroom observations is limited to pupils' perceived responsiveness to the tasks.

The observed responsiveness of the Year 9 pupils to the task changed throughout the task. Pupils' responsiveness was high in a number of instances when pupils were invited to contribute and interact with the teacher. This was very visible during the introduction to the task during which the teacher invited pupils to brainstorm and also after the video when they reflected upon the documentary. The pupils were eager to contribute and take part in what was going on. However, the level of enthusiasm was not sustained throughout the assessment. This was observed by the high frequency of teacher feedback concerned with the management of the class or keeping the class on task during the later stages of the task. In particular, pupils seemed to be less enthusiastic about answering written questions and drawing diagrams from their exercise books during which they were asked to keep quiet and work on their own.

A different picture of pupil engagement was observed in the Year 8 classroom. The general response of Year 8 pupils to the assessment task was positive throughout. Year 8 pupils seemed to be more engaged and motivated throughout the task, even if it required written answers to questions during which they had to work on their own for longer periods of time. Overall, the Year 8 assessment task offered more opportunities for group work and involved a role-play activity which was welcomed by pupils.

Understanding pupil responsiveness requires a fuller understanding of the nature of dynamics, pupils' backgrounds, perceptions and all of the other social dynamics within classrooms as well as the perceived task difficulty level (Wallace, 1996a), and the relationship between teachers and pupils (McCallum, 2001). Pupils' perceptions of some of these aspects are given in the Findings and Analysis of Pupil Interviews section. However, the observational findings suggest a link between task format and pupil engagement. Pupils appeared to be more engaged when there were opportunities for them to contribute within a social context group such as role-play and group work more than tasks involving solely individual work, such as the case study. This is
consistent with the findings of Wallace (1996a) and Cooper and McIntyre (1996) who suggest that pupils value their control over their learning process, which also makes learning more meaningful for them.

Elicitation of Assessment Evidence

This part of the thesis illustrates the findings and analysis of classroom observations, as well as the follow-up teacher interviews, regarding the elicitation of assessment evidence. The following table indicates the ways in which assessment information was gathered.

<table>
<thead>
<tr>
<th>Definition</th>
<th>Teacher Strategy</th>
<th>Its relation to Assessment Tasks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher questioning</td>
<td>Teacher asks pre-defined and grounded questions to elicit pupils' learning.</td>
<td>This happened generally when the teacher introduced a new concept, or started a new task.</td>
</tr>
<tr>
<td>Teacher observing pupils in the learning process</td>
<td>The teacher observed pupils on how they completed the tasks.</td>
<td>Opportunities for pupil observation were created during when they completed the written activities and group-work.</td>
</tr>
<tr>
<td>Teacher reviewing pupils' homework</td>
<td>The teacher provided individual written feedback for specific pieces of work.</td>
<td>These type of written homework constituted the form of short written essays, and the preparation of summary posters.</td>
</tr>
</tbody>
</table>

Teacher Questioning

The most common form of eliciting assessment information was observed as oral teacher questioning. In particular, teacher questioning was evident in the introduction to the task for the purposes of exploring pupils' prior knowledge, experience and understanding. This type of questioning was characterised by the teacher inviting pupils to do a brainstorm for a topic. The following extract illustrates this type of teacher questioning:
Table 4.9: Classroom Extract (Obs. 3, Week 3, Y8)

In the beginning of the lesson.
T: Who knows something about migration? Have a go?
Pupil 19: It is something where people go to another place in order to work.
T: People go from one country to another in order to work. That is an example of migration.
Pupil 18: It is when people move to another country to live there.
T: Somebody moves to another country to live there. Well done both of you, there are very good examples, you also mentioned the work which is very important that we will talk about later on.
Pupil 8: When somebody moves from one area to another one.
T: When somebody moves from one area to another area, he hasn’t mentioned the country, notice. So what we mean by area?
Pupil 8: It could be a city where you live in
T: It could be a city,
Pupil 10: A region.
T: It could be a region but it doesn’t necessarily mean another country.

The above example represented the most common ways of teacher questioning, which occurred during the introduction of the task when new concepts or a new topic were introduced. Here, the teacher aimed to elicit how pupils defined migration which he considered as an indication of their prior knowledge about the concept (Teacher 3, interview 2). The above dialogue continued with the teacher asking questions about push and pull factors which cause migration.

Another example of teacher questioning was observed during the whole-class discussions, after pupils watched videos and read passages from their work. Here, the teacher encouraged pupils to reflect back upon what they had seen. The following extract exemplifies a follow-up videocassette on a native tribe in Brazil Rain Forest (Obs. 2, Week 2 Y9 group):
Table 4.10: Classroom Extract (Obs. 3, Week 3, Y8)

<table>
<thead>
<tr>
<th>T: Now, we are going to discuss the possible subheadings of your case study. I want you to remember the video that we watched the other day... What points can you remember from the film - what aspects as we saw it can be put into subheadings which characterise their way of life?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pupil 2: Canoes</td>
</tr>
<tr>
<td>T: Canoes (he wrote the word on the white board) How they build the canoes and use them.</td>
</tr>
<tr>
<td>Another one?</td>
</tr>
<tr>
<td>Pupil 4: Fishing</td>
</tr>
<tr>
<td>T: ...what are the methods for fishing? That makes it an interesting paragraph.</td>
</tr>
<tr>
<td>Pupil 6: Ceremonies</td>
</tr>
<tr>
<td>Pupil 2: Hunting</td>
</tr>
<tr>
<td>T: Hunting. Good.</td>
</tr>
<tr>
<td>T: Right, let’s take the hunting. You extend this a little bit, what can you tell?</td>
</tr>
<tr>
<td>Pupil 4: What are their shot guns?</td>
</tr>
<tr>
<td>T: Ok, what have they got and what do they do with them? Ok describe how they hunt...What are our famous questions - what, when, how and why etc. We have got description which is telling us what it is like, what happens. Then we have got explanation why. So, describe something and explain why they do it in this way. Let’s think about farming. What questions can we ask about farming to describe and explain it in your case studies.</td>
</tr>
</tbody>
</table>

The above examples illustrated the ways in which the teacher attempted to elicit pupils’ knowledge and understanding through questioning in the teacher-whole class interactions. The teacher suggested that pupils’ answers to these questions provided him with, “general understanding of pupils’ knowledge”, which then gave him a “feel of where the class was” (Teacher 3, interview 1). Teacher 3 continued:

“...after a quick review of what they know, I adjust my teaching, the pace in a way, for the rest of the lesson. The quality of their answers guides me...” (Teacher 3, interview 1).

The teacher did not only use questioning as a strategy for elicitation of pupil learning but also for the regulation of his teaching.

**Teacher Observing Pupils**

The findings of classroom observations suggested that teacher observation of pupils appeared to be another tool used by the teacher to elicit assessment information. It was not possible to identify specific events to explain this situation since it happened in every stage of the lessons. However, there were two obvious situations during which the teacher planned to observe pupils. The first was when pupils completed
written work (i.e. answering questions in their exercise books, preparing a poster or a diagram). In such occasions, the teacher walked around the class, checked each pupil’s work briefly and provided feedback on their work.

Teacher observation of individual pupils was important for elicitation of assessment evidence since the teacher had opportunities to see how an individual pupil was getting on with a given task. Many instances were observed where the teacher asked for clarification about what the pupils had done. This was also the moment where the teacher questioned individual pupils about why and how specific actions had been taken (metacognitive questioning).

The teacher was also sensitive to the questions that pupils raised during his systematic observation, seeing it as an important source in eliciting their understanding. The teacher commented that

“When I walk around them and look at their work, it is an opportunity for them [pupils] to ask their own questions. There are children who lack the confidence to ask or speak in front of the class. This individual attention allows them to communicate with me...When they come up with their own questions it is important evidence to demonstrate what they have understood...” (Teacher 3, interview 3).

The teacher considered that listening to pupils’ own questions and explanations was an important indication of their learning. Therefore, the teacher deliberately created space and planned for observing pupils (Teacher 3, Interview 3).

**Teacher Reviewing Pupils’ Homework**

Another means for eliciting assessment information was reported to be the “reviewing of pupils’ homework” (Teacher 3, interview 1). Pupils were given a number of homework tasks ranging from writing a short essay to preparing a poster. For the teacher, a quick review of pupils’ homework contributed to the picture of individual pupil’s understanding. Their homework was also an indication for the teacher to see
whether the pupils had a common misunderstanding or misconception. Therefore, he attached a diagnostic function to pupil homework in eliciting information about their learning.

**Factors Influencing Elicitation of Assessment Evidence**

The classroom observations and follow-up teacher interview data suggested that what was elicited, as assessment evidence, was a product of the interactions and the social situations at a given time. The teacher summarised a number of key factors that shaped the dynamics of such moments, which he saw as having a direct impact on what was elicited as assessment evidence. The following describes these factors as ‘a trusted relationship’, ‘class culture’, ‘classroom management’ and ‘differentiation’.

**A trusted relationship**

Teacher 3 considered eliciting what pupils knew about a topic as a complex matter and suggested that a number of social factors influenced the quality of assessment information in classrooms. He recognised the importance of “a trusted relationship” between himself and the pupils, in order for them to be able to “stand up and talk in front of the class” (Teacher 3, Interview 2). He suggested that building such a trust took time and it was important to make sure that when pupils did not provide the correct answers they were not intimated.

**Class Culture**

The teacher suggested that each class has its own “cools” and “in and out trends” (Teacher 3, interview 4). He made a direct link between pupils’ willingness to answer a question and current class trends. He suggested that if pupils did not want to answer something, it could mean that they simply did not want to contribute rather than not knowing the answer, as reflected in his following comments:
"...Mary, she is a very able girl but she could not answer my question. This does not mean that she doesn't know the answer. Perhaps she was talking to the boys in the front row. Answering teacher's questions may not be very cool in this class [Year 9] perhaps...This creates a tension in terms of understanding whether a pupil knows something..." (Teacher 3, interview 4).

**Classroom Management**

Teacher questioning for the teacher served more purposes than the elicitation of assessment evidence. This was also observed on a number of occasions when the teacher selected pupils on the basis of their engagement with the task. For instance, if a pupil was talking with other pupils, she or he was more likely to be picked out for attention by the teacher. He explained this situation as follows:

"...if somebody is talking in the class, yes I do it [ask questions to them] subconsciously. I do it to remind them that they need to go back to the task... and the child has not been criticised and got back on the task and probably feels successful and motivated..." (Teacher 3, interview 3).

The nature of questions for managerial purposes were observed as more fact-recall type of questions which asked pupils to recall some pieces of information.

**Differentiation**

The teacher interview data in the previous section described how teachers differentiate in order to maximise learning opportunities for pupils with differing abilities. Teacher 3 appeared to differentiate his questions for the same purposes. His differentiation appeared to be informed by two key factors. Firstly, he suggested that his initial questions for exploring pupils' prior knowledge shaped the rest of the questions that he asked subsequently in the lessons (Teacher 3, Interview 5). Thus, this early assessment was used to structure differentiation. Secondly, he directed questions with differing demands in accordance with pupils' perceived abilities, as he explained as follows:
"...I am very conscious of doing that [asking questions in accordance with their perceived ability]...there must be a balance in the classroom for a mixed ability group. I consider their abilities and previous performances before I ask them a question...So, more open-ended questions might go to cleverer children...” (Teacher 3, Interview 5).

For Teacher 3, perceived different ability was not the only criterion for tailoring his questions for different pupils. He also considered gender in his differentiation practices:

"...Pupil 4, she is a very quite girl. I asked her the question about the wedding tradition of the village people since I thought that that would increase her interest in the lesson. I think girls are more willing to respond to things like traditions, customs in this topic...” (Teacher 3, interview 5).

It appears that Teacher 3 did not only consider gender but also personality. Although he did not mention ‘personality’ explicitly during this interview, his examples involved descriptions attached to pupils’ personality in explaining his in-class differentiation practices.

The Key Findings on Elicitation of Assessment Evidence in Classrooms

The findings in the previous section suggested that Teacher 3 employed a variety of techniques to elicit the assessment information. The key outcomes of this study are summarised below:

■ Elicitation of assessment evidence appeared to be a partly planned activity.

It seems that the elicitation of assessment evidence was a partly planned activity during which the teacher was sensitised to gathering evidence of pupils’ learning. Among the various practices described, teacher questioning appeared to be the least planned. However, the teacher had a number of key predetermined concepts on which the questions were grounded. Furthermore, he planned to allocate time to conduct his systematic observation at various stages of the task to gain snapshots of pupils’ learning.
The quality and nature of elicited evidence of pupils’ learning were determined by the nature of the assessment task and how it was implemented.

Regarding the elicitation of learning evidence, a link was seen between what was planned through the assessment task and what was achieved. For instance, through poster preparation, the teacher was able to gather learning information on whether pupils were capable of categorising information and present its key elements in condensed form (this is the predetermined function of a poster presentation). As a further example, written assignments enabled the teacher to assess pupils’ ability to comprehend the entirety of a situation and to present an argument in a written response.

There were setting-related factors which affected the quality of elicited assessment information.

The examples in this research illustrated the setting-bound nature of formative assessment, which is consistent with the claims of Wiliam and Black (1996). The authors view the elicitation of assessment evidence as a complex process and points out the risks of misinterpretation (of the reasons) for not eliciting evidence. This finding is also consistent with a definition of formative assessment as ‘setting grounded’ (Dann, 2002; Bell and Cowie, 2001) and it is shaped by the dynamics of the settings within which it is constructed.

SECTION 3: THE FINDINGS AND ANALYSIS OF TEACHER FEEDBACK

This section continues with a description of the findings and analysis of classroom observations and follow-up teacher interviews on teacher feedback. The section starts with a description of feedback strategies which the teacher used. The section continues to describe the situations which enabled and inhibited formative assessment feedback before highlighting the emerging points.
Description of Teacher Feedback

In the two classrooms observed, the teacher constructed assessment feedback in a number of varying contexts and with different functions. Such contexts included the correction of errors, the communication of standards and the feeding forward for the direction of pupils' learning. Eliciting the process of teacher feedback was a challenging task with it being situated within a social context and it was a part of what the teacher did as a whole. However, the typology proposed by Tunstall and Gipps (1996) and Gipps et al. (2000) provided a framework to describe the nature of teacher feedback observed. This study noted a variety of feedback strategies ranging from evaluative to judgemental functions.

Evaluative Feedback

The core practice of evaluative feedback appeared to comprise the teacher correcting pupils' answers to written or oral questions. Although this study did not measure the frequencies of each feedback type, feedback with corrective function was the most common type identified. This type was observed to be typified by the teacher approving or disapproving a given answer in short interactions. Such occurrences were noted during:

- End of lessons when the teacher checked pupils' work systematically;
- Introduction of a new topic when the teacher surveyed pupils' previous knowledge and understanding;
- Any stage of lessons when the teacher observed pupils' completion of a part of a task.

This type of feedback was most noted in situations when the teacher checked the written parts of the task, during the last 10 minutes of the lesson. The teacher would have spent 30 to 40 seconds with each pupil to review quickly what they had done during a lesson. An example of such a moment is shown in the vignette below.
(Obs3/Year 9). Here, the teacher was checking pupils’ written answers to the questions in their exercise books.

Table 4.11: Year 9 Group Classroom Observation Extract (Obs. 3, Week 3, Year 9)

<table>
<thead>
<tr>
<th>Teacher is checking Pupil 9’s work.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher is in the back row, checked James’ work and now he is with Pupil 23.</td>
</tr>
<tr>
<td>T: Right, a bit scruffy, it is pretty good, all those labels should go. Did you see the board (T refers the instructions that he wrote) and how I exemplified it? Have you followed the instructions?</td>
</tr>
<tr>
<td>Pupil 23: No</td>
</tr>
<tr>
<td>T: Why? So, follow my instructions next time.</td>
</tr>
<tr>
<td>T is talking to the whole class now.</td>
</tr>
<tr>
<td>T: Pupil 23 worked very quickly, so he could not read my instructions here (T shows the board where his instructions were written) full page, it says full page but he only did a quarter of the page, it is a waste of time anyway.</td>
</tr>
<tr>
<td>T is now looking at individual pupil’s work that they did today. Pupils are showing their work one by one and the teacher is spending 20 – 30 seconds with each pupil.</td>
</tr>
<tr>
<td>T: Good, correct, ok</td>
</tr>
<tr>
<td>T: Pupil 5, there is too much information for this question, try to be selective and just put the key answer, the most important aspect.</td>
</tr>
<tr>
<td>T: Correct, correct, correct, correct. Yes, yes, yes, yes, very neat as well, well done.</td>
</tr>
<tr>
<td>T is quickly scanning pupils’ answers and questions.</td>
</tr>
<tr>
<td>He didn’t give any written feedback, he just put a little tick to three pupils that he checked so far.</td>
</tr>
<tr>
<td>T: Next, correct…. Here, you need to put full stop here. Be careful about punctuation and the full stops. Next</td>
</tr>
<tr>
<td>T: Yes, a bit slow, but it is accurate.</td>
</tr>
<tr>
<td>T put ticks for all the questions. Wrote ‘Good’ for the last two work.</td>
</tr>
<tr>
<td>T: See what happened yesterday, you had a hard day today, I put a four point for your improvement today.</td>
</tr>
<tr>
<td>T: Yes, very good.</td>
</tr>
<tr>
<td>T: Full stops please and here as well. You are getting mixed up with rain fall …look at the correct graph, rain fall is on this side and temperature is the other, ok. Good.</td>
</tr>
<tr>
<td>T is now talking to the whole class.</td>
</tr>
<tr>
<td>T: A lot of people read the graph wrong….. the temperature is the red line, the centigrade side of your graph, many of you got 30 degrees and 9 degrees, it should have been 23 to 29 degrees.</td>
</tr>
<tr>
<td>T: Good.</td>
</tr>
<tr>
<td>T almost checked more than half of the class. T now turns to me and says the following.</td>
</tr>
<tr>
<td>T: Through my end-of-lesson checking I found that they have all done ok, except the temperature which they read the wrong side of the graph from the book but I drew their attention to it.</td>
</tr>
<tr>
<td>The bell rings and the lesson finishes.</td>
</tr>
</tbody>
</table>

During this brief scanning, the teacher and pupil interactions were just one-way, were short and were managed by the teacher. In addition to correcting pupils’ work, the teacher also checked whether pupils had followed the instructions and completed the task during the given time. Such checking resulted in either the teacher praising pupils or identifying quickly what went wrong and providing brief information about what needs to be done differently in the future. As stated previously, the literature does not consider solely evaluative feedback as having formative potential since it communicates judgements about pupils’ performances rather than specifying or
mutually constructing expected pupil achievement (Tunstall and Gipps, 1996; Gipps et al., 2000). However, for the teacher observed in this study, this quick scanning was not only about expressing his judgements, as illustrated below:

"...especially checking the written work, it is time demanding. It would probably take me an hour to sit down with the individual pupil, and write and mark on their books. ...in the last ten minutes, I can scan, I can comment on their handwriting, structuring of their work, presentation, paragraphing and I can see what they missed. ...Pupil 5, she had too much information, I said you put too much information really you should read and write more selectively, I told her to delete some of the information. She was standing there, I was giving her individual attention, it might be 30 seconds but I think it means a lot to them, ...I think that personal touch has a lot of effect on them" (Teacher 3, week 3, interview 3).

As the teacher argued above, through this short checking he created a platform where he was able to individualise his feedback, which, he suggested, had a positive impact on pupils.

Feedback with evaluative functions was also noted in the beginning of lessons mainly in response to the closed teacher questioning. Here, the teacher acknowledged the correct answer and moved to the next point. When the answer was incorrect, the teacher usually provided the correct one, gave some clues to guide pupils' thinking, or at least, he asked another pupil. There have been key repeated events which occurred during such feedback. Firstly, the teacher tended to choose pupils whom he considered to be off-task. Therefore, the choice of pupils was affected by classroom management purposes. Secondly, clueing involved the rephrasing of a question second time with different wording. The teacher suggested that this method seemed to give clues to pupils regarding the correct answers but also allowed them to gain the satisfaction of completing the question themselves (Teacher 3, Interview 2, Week 2). Generally, the clueing observed was heavy, sometimes to the extent of mentioning the first letter of the expected answer. This type of assistance generally appeared to be associated with the closed fact recall questioning.

Overall, the evaluative feedback which was observed had the function of approving or disapproving a given pupil response. Such interactions were observed as short and
curtailed in the sense that the teacher was so keen to provide the correct answer. However, amongst the various strategies described, when the teacher rephrased questions, pupils had more time to construct their answers and this provided longer interactions, some of which lead to descriptive feedback. The following section describes descriptive feedback in detail.

**Descriptive Feedback**

In the typology of Tunstall and Gipps (1996), descriptive feedback is defined as having formative functions since it specifies pupils’ current level of attainment and specifies a means of improvement. This definition attaches a different duty to feedback, which goes a step beyond the basic correction of what has been done. In the Literature Review chapter, this feedback strategy was explored by linking it to the learning theories. The following section describes its appearance in the classrooms observed, identifies when such feedback was used and highlights a number of key issues which affected its use.

For both year groups observed, descriptive feedback mainly emerged when the teacher reviewed a completed piece of task. This feedback was observed when the teacher:

- returned completed tasks to individual pupils and explained his judgements;
- debriefed the whole task at the end of a unit and invited pupils to reflect upon their experiences of the task;
- showed a completed task to the whole class which exemplified what constituted good quality work.

It is important to illustrate what happened when descriptive feedback was used. Here, the key action was the teacher explaining to pupils whether their work was successful or not and justifying the reasons behind that (specifying attainment). Sometimes this was followed by suggestions of how to improve the work in the future and what needs to be done differently in order to achieve this (specifying improvement).
following vignette illustrates the occurrence of this feedback (Obs. 4, Week 7, Year 9) when the teacher was reflecting upon submitted work, to the pupil who prepared it:

Table 4.12: Year 9 Group Classroom Observation Extract (Obs. 3, Week 3, Year 9)

<table>
<thead>
<tr>
<th>T</th>
<th>now picks up another essay and asks Pupil 7 to come.</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>Leicester versus Scotland. This essay is about why Leicestershire has a much bigger and denser population than the North of Scotland. So, you successfully paragraphed it. I recommended you to actually write in short, sharp paragraphs. And you got the introduction, setting the scene of what Leicestershire and Scotland look like. Some judgmental comments on the beauty scenes of Marble Arc. But Leicestershire can be noisy and ugly. You talk about the good points of living in a village where the shop is nearby, you don’t have to move far, and the problems such as living in a village might mean travelling to school. So, you draw the pros and the cons of living in a rural area - public transport is a problem in living in a rural area. You mention your village – it is about five, six miles away. So you are using definite and concrete points there and connect the points to your life experience ‘Above you summarised the bad sides of living in a village in North of Scotland but there are some good points. Ok, this is really good but can you please show me any points towards where it mentions about the jobs Pupil 7?</td>
</tr>
<tr>
<td>Pupil 7</td>
<td>I don’t think it does.</td>
</tr>
<tr>
<td>T</td>
<td>It mentions the entertainment facilities, being able to shop, it mentions the advantages of living in a city. … Good, the key thing you missed out there is being able to earn living- a point that we made during the lesson. That is the only missing point I can think of. So, next time, consider the whole points we summarise during the lesson and check whether it is important for your project. If it is, include it in your work. Pupil 7: yeah</td>
</tr>
<tr>
<td>T</td>
<td>And it should have been your main point. First of all the farming, secondly manufacturing, you have mentioned correct factors for example public transport and being near to things. So, this one is very well written but she missed out a couple of key points so I will probably give 15 to 16 for this because of the way it is written. You hit several very important points although you have missed a couple of others. If you had included the couple of others, it would have been achieved a full mark. Thank you.</td>
</tr>
</tbody>
</table>

The above vignette typified the use of descriptive feedback during debriefing sessions with a pupil. Here, the teacher communicated pupils’ strengths and weaknesses in their learning as well as modelled what constituted a more successful essay.

The nature and use of assessment criteria for descriptive feedback is an important issue. In the above example, the teacher explained and judged pupils’ work against certain criteria, the majority of which were derived from previously set lesson objectives. On a number of occasions, the level descriptions were used to judge pupils’ work. However, not all of the time were the assessment criteria objective-related and focused on the learning processes. There were other notable occasions during which quantity and presentation of work were praised. The implications of this for pupils’ perceptions of what constituted good quality in a given task is also touched

171
upon in the Findings and Analysis of Pupil Interviews part of this chapter. Although, the teacher did not want to focus on subsidiary aspects of the task, he suggested that the very nature of geography tasks, involving maps and diagrams, resulted in some focus on presentational aspects (Teacher 3, interview 5, week 5).

The literature reviewed suggests that the formative capacity of feedback also depends on whether it specifies the way forward or how to improve. The teacher observed used various strategies to enforce this aim. As the vignettes illustrated, he directly specified the next steps for improvement to individual pupils (i.e. “you should consider the key points that we talked about during the lesson”). On two occasions, he displayed successfully completed tasks on the classroom walls. Furthermore, with the Year 8 group, he encouraged pupils to make suggestions on how to improve their peers and their own work through self and peer-assessment. The common theme in all of these three examples suggests that the existence of such feedback for specifying improvement was dependent on whether time was allocated to reflect upon what had been done.

**Enabling and Disabling Formative Assessment Feedback**

The literature reviewed attached a prominent role to feedback in achieving formative assessment. What this has raised clearly for this research is the key question of what formative assessment feedback is and what it looks like in two selected geography classrooms. The following table summarises the findings of classroom observations and describes the situations in which formative assessment feedback was, and was not achieved.
Table 4.13: Inhibiting Situations for Formative Assessment.

<table>
<thead>
<tr>
<th>INHIBITING FACTOR</th>
<th>IMPLICATIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>A tendency to focus on the quantity of work.</td>
<td>This promoted quantity of work at the expense of quality of learning in the activity.</td>
</tr>
<tr>
<td>A tendency to pay attention to the presentation of the work.</td>
<td>Pupils spent a lot of time with the drawings, colouring and shading rather than discussing the points which the diagrams illustrated.</td>
</tr>
<tr>
<td>Assessment feedback tended to have a focus on the product rather than the process through which the product was generated (finishing the work within the given time was emphasised greatly).</td>
<td>Rather than encouraging pupils to consider their learning processes, a product oriented classroom atmosphere was created.</td>
</tr>
<tr>
<td>Spending too much time on didactic teaching in the form of transmitting information rather than engaging pupils with the concepts.</td>
<td>This resulted in limiting the number of occasions where the teacher entered the ZPD of the pupils and engaged in their thinking and understanding in depth.</td>
</tr>
<tr>
<td>On various occasions, the teacher tended not to communicate pupils' weaknesses in order to preserve their self-esteem and motivation</td>
<td>Opportunities to explain where pupils went wrong were missed.</td>
</tr>
</tbody>
</table>

When looking at the various levels of inhibiting factors within classroom practices, the focus of feedback appeared to have the most direct effect on the generation of formative assessment feedback. As suggested previously, there were occasions noted during which quantity and presentation of work received acknowledgements over the quality of tasks.

An interesting outcome was the teacher’s tendency to communicate pupils’ strengths in order to protect their self-esteem. Strengths were observed to be communicated more explicitly, whereas weaknesses were not articulated fully and were explained less. This was motivated by the teachers’ beliefs on what constitutes good teaching, as explained in Part 1 of this chapter (Teacher 3, Interview 1, Week 1).

In contrast to the inhibiting factors, there were moments which were constructive in building formative assessment feedback. These moments are explained on the basis of ‘enabling factors’, a term which is coined by Singh (2001). The following table illustrates these factors and their observed implications:
Table 4.14: Situations which Enabled Formative Assessment Feedback

<table>
<thead>
<tr>
<th>ENABLING FACTOR</th>
<th>IMPLICATIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explicitly sharing the assessment criteria with the pupils.</td>
<td>This gave clear ideas and focus to pupils about the expected level of attainment.</td>
</tr>
<tr>
<td>Feedback conveying information not only on pupils’ strengths and weaknesses but also directing them to the next steps that they need to take in order to improve their learning. Such feedback was achieved during the debriefing sessions and moments of reflection.</td>
<td>Understanding next steps would have moved pupils on their learning and would have given them the confidence that they actually could improve their competence and success. Thus, such feedback could have been motivating.</td>
</tr>
<tr>
<td>Involving pupils in the assessment process by giving them opportunities to use assessment criteria in assessing themselves and their peers.</td>
<td>This enabled more space for the pupils to contribute to the whole process of assessment. This also gave opportunities to the teacher to understand whether the pupils internalised the assessment criteria.</td>
</tr>
<tr>
<td>The understanding of individual pupils with a view of expected reactions to teacher feedback.</td>
<td>The teachers’ good understanding of individual pupils made him aware of their possible expected reactions. This resulted in the teacher customising the delivery of the feedback to different pupils.</td>
</tr>
</tbody>
</table>

Overwhelmingly, feeding forward appeared to be a salient factor in increasing the formative potential of assessment feedback. The first stage of this process involved the teacher identifying the problems and weak points in pupils’ learning. Feeding forward was built on this specific information. The observational findings highlight the importance of modelling success not only by telling pupils what needs to be done but also demonstrating examples of high quality work.

Summary of Key Findings Regarding Teacher Feedback

- **Teacher feedback emphasised the evaluative over the descriptive function.**

  The teacher commonly used evaluative feedback that approved or disapproved what pupils had done. Such feedback appeared to have an end point, and at many of the occasions it did not lead to any action on how to improve the work.

- **Teacher feedback was one-way communication during which pupils had little involvement in its construction.**
The observations suggested that the process of giving and receiving feedback was affected by the distribution of power between the teacher and pupils. The teacher was not ready to give up his power and this resulted in not increasing pupils’ responsibility for their learning. Furthermore, the teacher was observed as being very keen to provide the right answer straightaway which decreased the time for pupils to construct their answers.

- A link appeared between the nature of the task and focus of teacher feedback.

The nature of feedback was linked to the nature of the assessment task. When the instructions on how to complete the task were highly structured (the poster example for the Year 9 group), teacher feedback appeared to confirm whether what had been produced was right or wrong, a finding which is consistent with the suggestions of Gipps (1996). On the other hand, when pupils were given more freedom (i.e. when the Year 8 pupils prepared the scripts for the role-play), the teacher feedback appeared to involve a strategy of specifying attainment which also involved discussions with pupils.

SECTION 4: COMMUNICATING LEVEL DESCRIPTIONS

This section continues with a description of observational findings on the communication of Geography level descriptions. Such understanding is important since the current formative assessment practice is operated under the National Curriculum. The teacher interview data provided information on how teachers perceived and used level descriptions. One common finding of teacher interview data indicated the difficulties of communicating levels to pupils through a specific assessment task. The following paragraphs explore how such communication was facilitated by
- Identifying the stages of the tasks at which the levels were mentioned;
- Examining the ways in which the levels specified the standards of achievement for the tasks observed.

The classroom observations revealed two specific occasions when the levels were communicated explicitly. Firstly, the levels were communicated at the beginning of the task when the teacher introduced a unit of work. A second type was also observed at the end of a unit when the teacher judged the finished product against level descriptions.

The teacher’s use of levels at the beginning of new units showed variations. For the Year 8 group, the teacher did not mention the levels at all. For the Year 9 group, when the levels were introduced at the beginning, the teacher either related a level to a key objective of the unit or explained it comprehensively. The following extract exemplifies a situation where the levels were introduced by the teacher in the beginning of the task for the Year 9 Group. The below information was written on the white board and was left on the board for the entire lesson.

**Table 4.15: The Introduction of Level Descriptions**

<table>
<thead>
<tr>
<th>Unit 9.2.1: Equatorial Rain Forest-The Aims of the Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>- How do the equatorial climate and vegetation develop?</td>
</tr>
<tr>
<td>- How is the equatorial rain forest ecosystem influenced by nature and people?</td>
</tr>
<tr>
<td>- How can people create changes in the TRF? <strong>Level 6</strong></td>
</tr>
<tr>
<td>- Recognising the conflicting demands on the environment –<strong>Level 6</strong></td>
</tr>
</tbody>
</table>

The above information was written on the white board and all of the pupils wrote the points in their notebook. Having done this, the teacher continued:

“...so if you understand how people create changes in the TRF, how they actually cause different demands...and if you understand the implications of that then you are doing well and you will get a level 6” (Teacher 3, interview 5).
The above communication lasted only a couple of minutes. During the ensuing five weeks, the levels were not mentioned at all. During the beginning of fifth week, the teacher mentioned levels again but this time explained them in more detail in relation to the planning of a new enquiry project as the following vignette (Obs6, week 5) illustrates:

<table>
<thead>
<tr>
<th>Table 4.16: Year 9 Classroom Vignette Illustrating Justification of Levels</th>
</tr>
</thead>
</table>
| **T:** You are going to ask 6 to 10 questions about these three things, what is the problem, how this occurs and how to solve the problem... it is not just as simple as to say that there is a problem. There are several problems. Therefore, you need to be selective about the problems that are created by the water. Level 6, it says that you appreciate many links and the relationships that make places dependent on each other. Forinstance, Phoenix depends on the river what? Pupils: Colorado **T:** Colorado, River Colorado depends on Rockies - the water supply. You recognise how the contradicting demands on the environment may arise. How can we explain this one? Pupil 8? Pupil 8: It is like, if you do something it affects the environment. **T:** Yes, the environment, but demand on the environment was the Colorado River Valley an environment?… **T:** Gland Canyon was 250 km long, that Canyon like Grand Canyon was full of wild life, you would have concerns for conservation and you would say ‘don’t drawn on that Canyon and don’t make it a reservoir because all animals live near the river might be in danger... but people wanted the water won the argument, so dams have been being built. **T:** So all those things happened and destroyed the environment because people thought it was worthwhile. So do you understand that there is a conflict, an argument between people. **T** is again reading the National Curriculum Document. **T:** Can you describe and compare different approaches of managing the environment? What does it mean to manage environment? **T:** Do you appreciate different values and attitudes that result in different approaches and different effects on different places and the people. So do you understand and appreciate why they have done what they have done to the River Colorado and the desert? Now, this is the key one of the day, can you draw on your knowledge and understanding and suggest some questions and appropriate sequence of investigation. Can you invent 6 to 10 questions and can you put them a special what to make sense? Pupils: Order **T:** Well done Pupil 8 again. A sequence of things and the order of things… what can you suggest about the order? Tom: How to solve it? **T:** Yes, it comes after how the problem occurs, well done… That was the Level 6. Now for the Level 7, you appreciate sustainable development affects the management of environmental resources… At Level 7, you need to understand these things and realise that if the Americans want Phoenix to grow if they want more people live there, more factories, more houses…what do they need to do in order to achieve that? Pupils: Water **T:** And how do they manage water. Can they get hold of any more water? So we have got to understand, so I will read it again. So you appreciate the consideration of sustainable development and how this can affect the management of the environment and the resources. So, they have to manage their environment and they have to manage their water if they want to keep growing. Next one, With growing independence that is by your self—you draw on your knowledge and understanding, you will devise your own questions and establish a sequence of independent enquiry. Level 8 is a bit more complicated than this. Let’s listen: Drawing on your knowledge and understanding show that you have independence in identifying questions and issues and you can use them in your questions. So, it is virtually from Level 6 is being helped, Level 7 is being able to ask questions but Level 8 is virtually by your selves and you should be able to answer these questions, well I think you are doing well so far. So, we looked at two things. Water supply—it is limited and which is going down. We have sustainable development which means Phoenix is continuing to grow and the people live there are continue to live there. There are some problems, can anybody start rolling—suggesting some of the problems, ask the questions, it is not just the water supply problem. Pupil 19? Pupil 19: How hot it is. **T:** How hot it is—ok, what is the contribution of this question to understanding the problems of water and sustainable development.
In the above vignette, the success in communicating levels was dependent on the teacher’s ability to relate them to the water problem of Phoenix. The teacher seemed to relate the individual statements of levels to the requirements of levels. For instance, he was able to explain why there were conflicting demands and how this could have affected the sustainable development in the area. Furthermore, he invited pupils to contribute to the interpretation of levels. So, this was not a top down approach and under the teacher’s guidance, the levels were made more meaningful to pupils.

Another means of communicating levels consisted of showing pupils previously levelled pieces of work to illustrate how their work had been judged at the end of a unit. The following vignette (Obs.7, Week 7) illustrates a moment when the teacher was sharing his reflections on a levelled task with the whole class.

Table 4.17: Year 9 Classroom Vignette Illustrating Justification of Levels (Obs.7, Week 7, Year 9)

| In the beginning of the lesson, T is exemplifying high quality work to the whole class. |
| T: Pupil 6, you have a certificate for this work, I am very impressed, a personal best |
| Right, what I am going to say will be a little bit embarrassing to you, but something like this has come from your thinking and in depth understanding of problems in a number of areas which is a high level of skill ...you really stretched your thinking and I am giving you a Level 7, it is the best work of you I have seen best improvements from you this year. |
| I am very very pleased. |
| T is now checking another submitted work. |
| "...this work demonstrated to me that she was very organised and she planned it very carefully. She has clear sections and sub-sections that provided a coherent structure for the enquiry report. She improved the diagram work, and she put extra effort all the way through. ....With this work, she demonstrated to me her commitment and interest in what she has been doing. She tried really hard...and she received a Level 7 for this. |

The above vignettes illustrate the ways in which the levels were interpreted and used to form a judgement on a piece of submitted work. In the above examples, the levels were not the only assessment criteria; the teacher also praised effort, organisation and planning. Furthermore, the teacher judged the pupil’s performance with reference to her previous performances and in comparison to their previous levels.
A problem which was identified in the teacher interviews, was the perceived difficulties of communicating levels to pupils. One main explanation related to perception of levels as being abstract definitions of pupils' ability. In particular, providing opportunities for pupils to achieve higher levels was problematic since teachers themselves expressed difficulties in understanding the specific requirements. Pupils becoming independent were one of the most problematic aspects. This was confirmed through observational findings since the teacher was seen to only mention it briefly while he was reading the National Curriculum. Furthermore, he did not mention the ways in which pupils could gain independence through completing the task.

SECTION 5: PUPIL SELF AND PEER-ASSESSMENT

The only obvious pupil self and peer-assessment was used with the Year 8 group within the Population and Ecosystems task during the Week 5 and 6. As stated previously, this task ended with a role-play during which pupils used their scripts which they had prepared previously within groups. Before the actual enactment of the role-play, the teacher asked pupils to evaluate themselves and their peers on the basis of a number of determined criteria; these were the use of facts and understanding, speaking skills and props and rehearsal. There was also a space for pupils’ personal comments. Pupils watched each group acting and they took notes. After each role-play, the teacher debriefed each group briefly. The teacher gave two extra days for pupils to submit their assessment reports.

In the following part, the self and peer-assessment are evaluated on the basis of the choice and use of the assessment criteria and its relation to assessment task type.

The Nature and Use of Assessment Criteria

The literature reviewed addressed the form and use of assessment criteria (Klenowski, 1995), and its implications for pupil autonomy (Ecclestone, 2002). Sadler (1989)
emphasises the importance of the internalisation of assessment criteria by the pupil, which then could be used for the closure of the learning gap (Black and Wiliam, 1998a).

In the current study, the pupil self and peer-assessment observed involved the use of predetermined criteria comprising the use of facts and understanding, props and rehearsal. These criteria gave pupils a focus. For instance, to assess the use of facts, pupils were advised to take notice of when a group mentioned factual information such as a weekly wage or the proportion of rent in the whole wage. Understanding constituted pupils’ ability to relate push and pull factors to make their point. The last two criteria related to the clarity of speech and acting skills.

In order for pupil self peer-assessment to promote formative assessment, it is critical to evaluate the extent to which the selected criteria for pupil self and peer-assessment reflect the overall learning objectives. The teacher summarised the key objectives of the unit as follows:

“I want them to understand the migration, push and pull factors as well as developing empathy for the immigrants. When they see the video on favelas I want them to think and feel what it is like to be in that part of the world. Before laughing, I want them to think. I want also them to do it through a group work …” (Teacher 3, interview 1, week 1).

The use of facts as a criterion encouraged pupils to select and use statistics to support their argument during the play. The understanding criterion encouraged pupils to provide examples and explanations of the push and pull factors in the context of a favela. Pupils’ ability to reflect on the living conditions in the favela, their understanding of what those people experience, promoted the teacher’s objective of increasing pupils’ empathy skills. The evidence for the achievement in these respects was very clear. The majority of the groups clearly stated some factual information and the teacher praised all of the groups for their success in explaining the difficulty of living in favela.
However, the use of facts and understanding constituted one out of four criteria. The other three were concerned with acting such as the use of props, and speaking both of which were related to the dramatic aspects of the role-play. In the follow up short debriefing, the teacher praised these dramatic aspects and sometimes the choice of costumes. Not surprisingly, when pupils selected the best five in their reports, they valued acting as being almost the most important criterion for their decisions. Therefore, the criteria related to the dramatic aspect of role-play outweighed the other criteria and created a mismatch situation in relation to the teacher’s overall objectives.

Having explained the nature of assessment criteria, the following subsections, introduce a number of issues that could be important in implementing pupil self and peer-assessment.

Assessing Pupil Self and Peer-Assessment: Further Thoughts

Group or Individual?

One of the teacher’s broad aims was to achieve effective group work with the Year 8 pupils. To promote this, pupils worked, prepared the script, rehearsed and acted in groups. However, the assessment criteria for pupil self and peer-assessment was only concerned with individual achievement, and they failed to create team recognition. Slavin (1995) suggests that such an assessment approach may promote competitive behaviour since everyone would try to improve their own performances regardless of what the rest of the group does. Although this is beyond the scope of this study, what it could indicate is that the teacher was not ready to give up the individual assessment and to assess pupils in-groups, a point which was revealed by the teacher interviews reported previously.
Quantification of Assessment Criteria

One of the issues which deserves a brief note relates to the quantification of assessment criteria. In the activity observed, pupils gave a score (one to five) for each pupil for their peer-assessment. There was an absence of guidance on what levels of achievement corresponded to each score and therefore consistency between different students was not achieved with respect to the scoring system. However, the time framework within which the pupils worked was not long enough for pupils to form a qualitative judgement since they had so little time after each role-play. However, they were given two days to complete their assessment reports to justify their scores. So, there were opportunities for pupils to discuss in the class and to justify their scores.

Focus: Product or Process

The previous part suggested that the assessment criteria emphasised the role-play activity itself. It could be suggested that this placed focus on the outcomes of the lesson, which is a product of a six-week process. Pupils could have been encouraged to talk about their whole learning experience including their group work experience.

Building Reflection Moments

There were also occasions during the observations when the teacher quickly created an opportunity to reflect back on what happened. This was not a formalised part of the tasks (Teacher 3, interview 6). The following vignette (Obs. 1, Week1, Year 8) typifies one such moment during which the class was encouraged to talk about the skills that they gained during the construction of their last task:
Table 4.18: Classroom Vignette which Illustrates an Example of Reflection Moments (Obs.1, Week 1, Year 8)

| T: ...before starting our new unit I want you think about what skills you have learned in your last project. Now who has found out the new skills, the new skills that you had to learn to do the Blackney Project? |
| Pupil 12: I used the computer |
| T: Ok, so you used the computer the first time to type out this sort of report, well done. What skills did you use? What was daunting for you or a bit of a struggle for you? |
| Pupil 14: When I read something and when I needed to write the actual report, I had to change it into my own words. T: So reading another document which is complicated...big words, so you need to write it in your own words, it is a very good point, excellent. Who else did something new that they haven’t done before? What about the account bit? |
| Pupil 17: Finance bit. T: Finance, tell me a bit more about it. Pupil 17: It is like working about how much you should pay and how much you should invest as well as explaining the numbers that you write down. T: How much would it cost you to run the nature reserve, and how are you going to raise the money to run it as well as explaining and justifying your financial decisions. Pupil 19: Thinking about what to write, I mean deciding about what to choose. T: Thinking about what to write, good. When I marked your work, I looked at our lesson objective that we set in the beginning. Ok, about two-three weeks time you will be doing a role-play about people who live in a shanty town, Favela, which is a poor slum in a city in Brazil which is a part of your study of this new topic, people and population of the world. When I mark your work at the end, I will consider the points that you made today, how selective you are and whether you include the key objectives and answer the key questions. |

Brief Summary of Findings and Analysis of Classroom Observations

This part presented the research findings and analysis of classroom observations with an overall aim of exploring and exemplifying its appearance in two selected Key Stage 3 classrooms. The classroom observation data was also supported by the findings of follow-up teacher interviews with the teacher observed. This part consists of five sections. The first section outlined the context of classroom observations. The second section described the issues related to the elicitation of assessment evidence and paid specific attention to the factors influencing this process. The third section presented the findings of teacher feedback by describing feedback strategies which the teacher used before describing and contrasting the situations in which formative assessment feedback was and was not achieved. The fourth section described the observational findings on the communication of geography level descriptions. The final section described the use of pupil self and peer-assessment by analysing the nature and use of
assessment criteria before highlighting a number of key issues related to the use of such activities.

Having presented the findings and analysis of classroom observations, the next part (Part 3) focuses on pupil interview data.
PART 3: FINDINGS AND ANALYSIS OF PUPIL INTERVIEWS

Introduction

Chapter 2 suggests that a key element of the success of formative assessment lies in pupils’ active involvement in their assessment and learning. This chapter also describes existing research evidence on the ways in which the assessment process is affected by pupils’ understanding and interpretation of situations. Up to now, the current chapter has discussed how teachers viewed assessment. In the second section, classroom observations have been reported by the researcher, building on the interpretation of her reflections as well as the teacher’s explanations. What is missing is how pupils themselves experience the formative assessment process and what meaning they make out from their experience of it. Pupils’ views are important in this research as this assists in complimenting the conceptualisation of formative assessment.

In this study, pupil interviews were conducted to gain insights into their perceptions and understanding of the key components of formative assessment, which were also addressed in the teacher interviews. Having completed the classroom observations, 20 (10 from the Year 8 and 10 from the Year 9 group) pupils were interviewed individually. The interviews took place after they completed their assessment tasks. This part of the thesis presents the findings and analysis of pupil interviews and focuses on pupils’ reported experiences of formative assessment. This part consists of five sub-sections. The first section presents pupils’ views of what they valued in an assessment task. The second section explores their perceptions of their help-seeking behaviour and their interpretations of teacher’s initiations and help. Section three and four address pupils’ views on self and peer-assessment. The fifth section describes pupils’ perceptions of geography level descriptions.
SECTION 1: PUPILS' VIEWS ON ASSESSMENT TASKS

This section provides pupils’ views on assessment tasks. To explore this issue, pupils were asked what mattered to them regarding this issue. Pupils valued enjoyment and expressed preferences towards certain task formats and the value attached to personal input, as the following subsections present.

Enjoyment and Interesting Nature

For the pupils interviewed, the most desirable characteristic of an assessment task was that it should be interesting. Fifteen out of 20 pupils mentioned this criterion as the first thing that mattered for them when commencing a task. Consistent with teacher interview findings, pupils related the interesting nature of tasks primarily to the topic of a unit and then secondly to task format. Both year groups found it interesting to learn about new places and cultures in different parts of the world and consequently found the topics quite interesting and stimulating. Pupil responses linked being interesting to enjoyment and suggested that if they find a task interesting, they have more opportunities to enjoy it. As indicated previously, having fun and raising pupils’ interest were important criteria for teachers in their task design.

Task Format

Pupils expressed their preferences for certain task formats. Both Year 8 and 9 pupils reported specific preferences towards group-work. When compared to other task types, role-play was clearly regarded more positively, as a Year 8 pupil, Pupil 18 explained:

"In the essays you find the information from the different resources and try to write it down. But when we use what we write in a physical activity like we did in role-play it is so good and so fun. So, this makes us love and learn more about geography”.

In contrast to the enthusiasm shown for role-play activities, all of the pupils found essay writing to be uninteresting. Pupils used adjectives such as “boring”,

186
“challenging” and “hard” to describe essay writing. As explained in the above quote, one problem related to pupils’ difficulty in collating information from different sources into written essays. Pupils particularly objected to the high frequency of written work and they appreciated the inclusion of other more oral and interactive activities. In conclusion, they appeared to be more in favour of group work and oral tasks.

Pupil responses on preferred task types varied between boys and girls. Boys indicated that they prefer tasks that included ICT work and those involving diagrammatical work. The majority of the girls reported that they were particularly pleased with role-play and group work.

**Pupil Involvement**

Pupils' involvement in their learning and assessment is a key aspect of formative assessment (Black and Wiliam, 1998a; Klenowski, 1995). The literature reviewed reported studies to illustrate how this is achieved through pupil self and peer-assessment. The teacher interviews reported how teachers viewed pupil involvement and what kind of opportunities were created to achieve it. In this section, the ways in which pupils related their involvement to tasks is explored within the tasks that were described in the observational study.

Pupils' responses highlighted the perceived value of their involvement as active learners to tasks. Their assumptions of involvement were expressed as “participating in what is going on” (Pupil 10, Year 9) and “contributing to discussions” (Pupil 6, Year 9). Three pupils also mentioned that they felt involved in lessons when they could use their experiences and examples from their own lives.

Pupils were asked to describe the kind of tasks which they felt offered them opportunities to be involved. The consensus of pupil views was also that group work
was beneficial for this purpose, but on conditions that they are grouped with other pupils whom they liked. Pupils further indicated that situations where the teacher dictated activities from the front of the class was not positive for learning, since it did not allow the pupils to actively “use their brains” (Pupil 2, Year 9). Six pupils also mentioned that when the teacher talked for “too long”, they lost interest. Pupils’ responses suggested that their own active involvement was important for their learning. The views of pupils described here shows a considerable overlap with published research evidence, which suggest that children value their active involvement and empowerment over tasks (Wallace 1996a) so that learning becomes meaningful to them (Cooper and McIntrye, 1996).

In summary, pupils interviewed were more in favour of tasks that were associated with contemporary learning theories such as social constructivist learning and experiential learning while they opposed traditional directive teaching, where the teacher transmits knowledge.

Relevance of Tasks to Everyday Life

Pupils questioned the use of tasks in real life when they reflected upon the tasks that they had. Here, Year 8 pupils valued the unit on population since they considered it as important and up-to-date issue, as pupils commented below:

“Immigration is important. Everyday in the news we hear something about it. I am glad that we learnt why people turn into immigrants and what they experience” (Pupil 12, Year 8).

“Playing Jose [a character in the role-play] made me to think differently when I look at immigrants” (Pupil 14, Year 8).

There was a considerable concern in Year 9 pupils’ responses about the usefulness of the topic in their real lives. Six out of ten pupils reported the value of understanding new cultures. However, for four pupils, the unit on Tropical Rain Forests was not useful. To support this view, one pupil suggested:
"I don't know I will ever the need the information about the Kayapo Village for the rest of my life. It seems a bit useless, doesn’t it?" (Pupil 8, Year 9).

SECTION 2: PUPILS' HELP-SEEKING BEHAVIOUR AND THEIR IMPLEMENTATION

The classroom observation section illustrated the ways in which the teacher deliberately used strategies to elicit evidence of pupils' learning. This section reports on whether pupils themselves initiate teacher help when they experience difficulty in the completion of a task and how they interpret teacher help. The knowledge in this area could complement to the understanding of elicitation of pupils' learning by providing insights into the pupils' contribution to this process.

The pupil interview data suggest that pupils did not hesitate to ask for help when they had difficulty. They related difficult situations as not being able to understand an issue, concept, or instructions in terms of how to complete a task. The most common way of seeking help was putting their hands up. However, this was not easy for 5 pupils since they considered themselves as “shy”. The pupils explained their difficulty as follows:

"Sometimes he [the teacher] just picks someone’s point and explain to the whole classroom. It could be embarrassing if it is an easy thing” (Pupil 4, Year 9).

"There are some very bright people in this class. My question could be terribly easy for them. Then, I pretend that I know what I am doing. I wouldn’t stand in front of the whole class” (Pupil 8, Year 9).

The reserved approach taken by pupils was dependent on the perceived difficulty level of the question. If it was a very easy one for other pupils in the classroom, then they were hesitant to ask for help.

Pupils valued their teacher's help in tackling their difficulties in their learning – they defined him as a ‘helper’ or ‘saver’. The teacher’s reaction was perceived in two ways, either explaining, or, the most commonly, telling what needs to be done to remedy the situation. The observational study identified the teacher’s strategies as rephrasing or
clueing to give pupils’ time to find their own answers. Regardless of the slight variations in their perceptions, teacher involvement was generally regarded positively, as the following pupils’ responses suggested:

“He explains it so that you can understand it” (Pupil 8, Year 9).

“He will assist you and tell you what you need to do. If it is a question where most of the class get it wrong, he will talk to the whole class” (Pupil 1, Year 9).

“If we ask for help then he tells you what you need to do but he never gives the right answer” (Pupil 16, Year 8).

The literature illustrated pupils’ varied perceptions of teachers’ help. For example, Pollard *et al.* (2000) report that, across a range of curricula within UK schools, pupils in Year 6 (Key Stage 2) reported negative views of teachers’ involvement (the majority of pupils have a negative or mixed view of teachers checking their work). Pupil responses in this thesis suggested that this was the case for some pupils who were apprehensive about their work being under public scrutiny. However, pupils indicated other opportunities during which they felt more comfortable about seeking the teacher’s help. In the light of pupil interview and classroom observation data, the following section summarises the classroom conditions that promoted a positive attitude towards help-seeking behaviour as well as affected their behaviour.

**Conditions which Promoted Positive Help-Seeking Behaviour**

- **Pupils felt comfortable about revealing their questions and needs when they had opportunities to interact with the teacher individually.**

As reported in the observational findings, the teacher reviewed pupils’ work individually at the end of lessons and he was in favour of such moments and described it as a “personal touch”. For pupils, this was an opportunity to raise questions if they did not ask them in front of the whole class. Furthermore, pupils appeared to value this personal touch.
• Pupils trusted their teacher perceiving that he was going to do the right thing for them.

Pupils trusted their teacher since they saw him as "knowledgeable" and "very experienced" and "understanding". The pupils commented that

“He is teaching over thirty years. He knows what he is talking about" (Pupil 8, Year 9).

“He helps us. He reassures us and tells us what we are doing is right. He says we could do it better if we do it another way” (Pupil 14, Year 8).

• Group tasks appeared to give pupils opportunities to clarify their questions with other pupils thereby decreasing the consultation with the teacher every time.

Pupils reported to be in favour of group tasks. One major reason was the greater opportunities to interact with their peers. Pupils appeared to be in favour relying on their peers for their questions as long as the group composition was “right” for them.

SECTION 3: PUPILS’ VIEWS ON PEER-ASSESSMENT

Pupil interviews aimed to understand what pupils thought about the practice of assessing their peers and how they interpreted their peer-assessment experience. This issue was addressed only with the Year 8 pupils since they used peer-assessment as a part of their task and the researcher observed the activity. In the beginning of this chapter, the teacher interview section highlighted that pupil peer-assessment was not widely used because of teachers’ perceptions of the difficulties that pupils have in forming a fair judgement when they assess their friends. However, overall, teachers felt that peer-assessment was useful. In order to enrich this finding, the following section reports on whether pupils consider themselves objective in assessing their peers and how useful they found the activity.
Objective or Subjective?

When pupils were asked whether they were fair when they assessed their peers, six out of ten pupils regarded themselves as objective. Here, pupils suggested:

“...I try to forget that they are my friends and I just look at them and how well they have done. When I assess my friends, it should not be important whether they are my friends...” (Pupil 11, Year 8).

“...you just focus on who puts the real effort and who does not...” (Pupil 15, Year 8).

“...I tend not to think of their personality and focus on their performance and who did the best. I don’t think people look at this [peer-assessment] as something that we use to put each other down...” (Pupil 12, Year 8).

In the above quotes, pupils suggested that objectivity was achieved by focusing on assessment criteria (effort, performance) rather than who the individuals were. Here, boys used a more concrete language in expressing their objectivity. However, girls appeared to use a more tentative language such as “I tend to...” and “I try...”. For both groups, objectivity constituted using the assessment criteria and eliminating personal views and relationships.

Four out of ten pupils highlighted the difficulty of achieving objectivity when assessing their peers. For these pupils, friendship was an obstacle in achieving a fair assessment, as they commented:

“I think my friends would be offended if I give them bad marks. I wouldn’t give them really bad marks. I wouldn’t give them bad marks even if they were really terrible. So, I wouldn’t be that objective I think” (Pupil 13, Year 8).

“It is hard to assess other people, because they might be offended. For this reason, I am sometimes reluctant to put things down” (Pupil 14, Year 8).

“Even though I know that they could have done it better and could have tried hard, you cannot give them bad marks, realistically. If I give them bad marks they would be offended. How could I say that they are terrible when I see them and share my time with them in the breaks?” (Pupil 20, Year 8).

The pupils’ comments above reflected their concerns of being interpreted as offensive if they are critical to their friends. In particular, assessing close friends was
problematic. Pupils' out-of-class relationships had an impact on their decisions about their peers' success during peer-assessment activities.

The findings presented here draw a varied picture of pupil perceptions in achieving objectivity. This finding is somehow different to the findings of teacher interviews that highlight teachers' strong concerns about pupils' objectivity. Such concerns were used to explain why they did not use it in class time. Considering the fact that the pupils interviewed in this study had relatively more regular peer-assessment activities, it could be interesting to examine whether there is a link between achieving objectivity and the regular use of peer-assessment.

**Other Factors Influencing Pupil Peer-Assessment**

**Assessment Task Type**

Pupils reported their preferred task type in relation to pupil peer-assessment. Four pupils specifically suggested that group work suited peer-assessment for two reasons. Firstly, pupils reported that they felt more comfortable about assessing their peers within group tasks. One pupil commented that:

"When it is a group activity, it is easier to assess our friends because you generally put your mark or comment as a group and then you don’t have to notice as much as individuals. You look through the whole group, as a group and assess them as a whole" (Pupil 17, Year 8).

For this pupil, group tasks also helped him to achieve objectivity in peer-assessment. He continued:

"..By this way [group task], I can be really fair to individuals..." (Pupil 17, Year 8).

Another reason why pupils appeared to be in favour of group tasks for peer-assessment related to the increased opportunities to understand other pupils' attainment through group work, as is explained:
"When we work in groups, I came to know whether a person is trying hard, helping the group to achieve. So, when I mark him or her, I really had an idea about what they have done" (Pupil 19, Year 8).

Value of Peer-Assessment

Pupil interviews aimed to explore pupils’ perceived value of peer evaluations. The majority of pupils’ responses involved adjectives such “important” and “useful”. Pupils regarded the activity as a tool through which they develop their self-assessment. One of the pupils recalled his experience:

“I think it is very important that we are given the chance to evaluate our friends...This then makes us to think about how well and bad we did. We compare ourselves with them” (Pupil 15, Year 8).

For this pupil, assessing their friends increased their self-reflection.

SECTION 4: PUPIL’S VIEWS ON SELF-ASSESMENT

The ability to assess one’s self was reported as a hard task by all of the pupils interviewed. Even pupils who regarded themselves as objective in assessing their peers appeared to experience difficulty about assessing themselves objectively, as reflected below:

“I found it very tough. It is totally different from assessing other people - it is harder. You never know whether you are modest enough to yourself. You may end up with putting yourself somewhere really high that you do not deserve at all”. (Pupil 13, Year 16).

Other pupils outlined similar concerns by describing self-assessment as “tough” (Pupil 19, Year 8), “challenging” (Pupil 15, Year 8) and “really hard” (Pupil 14, Year 8). Pupils reported that for a successful self-assessment, being objective about oneself was the main struggle. One pupil commented that: “there is no other chance – everybody will be biased about themselves” (Pupil 20, Year 8). Overall, pupils appeared to find self-assessment challenging.
Pupils were asked the reasons why achieving objectivity was hard in their own assessment. Pupils’ responses suggested that they could identify and report their strengths easier than their weaknesses. One pupil commented that:

“I don’t like it to sound like my work was wasted and I am really terrible” (Pupil 12, Year 8).

Pupils’ tendency to report their strengths at the expense of weaknesses could be related to the teacher’s feedback strategies. As reported in the observational findings section, the teacher tended to communicate pupils’ strengths more often to preserve their self-esteem, which is also reported in the Literature Review chapter as a factor preventing teachers from communicating pupils’ weaknesses (Torrance and Pryor, 1996).

Pupils had their own solutions about trusting their judgements of themselves. Five pupils in Year 8 mentioned that they needed to refer to their teacher’s evaluation to confront the objectivity of their assessment. Two pupils preferred checking their judgement with their peers in addition to their teacher’s judgements.

Despite the acknowledged problematic aspects, pupils commented positively about their self-assessment experience and described it as “useful” (Pupil 16, Year 8) and “an activity which made them understand better” (Pupil 14, Year 8).

**SECTION 5: PUPIL’S VIEWS OF THE GEOGRAPHY LEVEL DESCRIPTIONS**

This thesis tries to understand formative assessment within the GNC. Much has been written about the underlying theory and key principles of formative assessment, but its links to the National Framework has come low on the list of priorities, and has not been addressed fully. To understand this link, the teacher interview data examined the effects of levels on teachers’ task design. Furthermore, it illustrated teachers’ difficulties in communicating levels to pupils. The observational data described the
stages in which the levels were communicated to pupils and examined the ways in which levels specified the standards of achievement. A major focus of this section is to explore pupils’ awareness about the levels they were working at and their perception of what needed to be done in order to achieve a higher level. The section also addresses the meaning of levels for pupils.

**Pupils’ Awareness of their Attainment in Relation to Level Descriptions**

Pupils’ awareness of levels was investigated by asking them whether they were able to state at which level they were working. Pupils were able to mention their levels apart from one Year 9 pupil. Four pupils mentioned that “they were not sure exactly” but still “accurately” guessed their levels. Nine pupils were able to specify what level they were working at actually by using words such as “towards” and “to”. Pupils were also described their levels by specifying whether they were working at a strong or weak level. Overall, pupils appeared to be sure about their levels numerically.

Translating their working levels to what it meant in terms of their learning appeared to be a hard task for pupils, as expressed in the following:

“Sometimes he [the teacher] comes in and praises me and says I am given a high level. It would be better if he says what I do well in order to get it” (Pupil 8, Year 9).

“I just remember I am working at level 6. I can’t tell why” (Pupil 2, Year 9).

“When I receive higher levels, I know that it means I understand the unit in depth. But I am not one hundred percent sure” (Pupil 16, Year 8).

Pupils’ understanding of levels was further prompted by seeking their explanations of what constituted higher and lower levels. Regardless of their year groups, pupils appeared to justify levels in more generic terms without referring to their individual attainment or exemplifying it within their finished task, with an exception of three Year 8 pupils explaining their levels in relation to their last task. For all of the pupils “effort” was the key factor in achieving higher levels. This was followed by “quantity”
and then "presentation of the work", as key criteria for determining performance in relation to levels. In the following, pupils commented on what makes a higher level:

"A level 7 could be a very thick work, a very detailed one. If you do everything that we asked for, even more...lots of presentation. If we try really hard, that is level 7. So, it really depends on how much effort you put in" (Pupil 2, Year 9).

"I think it is the presentation again" (Pupil 3, Year 9).

"More information, more pictures and more graphs to show things" (Pupil 20, Year 8).

When compared to generic criteria, very few examples were given to justify levels in cognitive terms. One pupil mentioned the importance of being selective, as follows:

"What makes a higher level is our ability to choose the most important bits of information and putting them in our work" (Pupil 9, Year 9).

Another criterion was "understanding". For two pupils, understanding a topic was necessary to reach a high level. One pupil commented that:

"If we show understanding of the topic and if we relate to the other things that we have learnt". (Pupil 10, Year 9).

Appropriate in-class behaviour was also an element that was perceived to be important by two pupils to achieve high levels. This was characterised by comments:

"...doing what our teachers tell us to do in class" (Pupil 2, Year 9).
"do not get trouble in the class" (Pupil 19, Year 8).
"do whatever the teacher says in the class" (Pupil 8, Year 9).

Overall, pupils linked effort to achieving higher levels. Submitting "thick" and well-presented work was interpreted as being important to receiving a high level. Conversely, pupils commented that not "trying enough" and submitting "scruffy work" could cause low levels.
Perceived Meaning and Value of Level Descriptions

Pupils appeared to attach a number of meanings to levels. Levels were reported as a tool to monitor their progress, a motivational benchmark to aim for, as well as a potential demoraliser. The following paragraphs present these findings in detail.

Levels for Monitoring Pupils’ Progress

Pupils regarded levels as helpful to monitor their progress. Here, pupils described progression as the increase in their levels in quantitative term. Six pupils interpreted levels as useful for this purpose. One pupil suggested that:

“They are useful in the sense that when we get a better level, it shows that we have improved our work” (Pupil 1, Year 9).

Pupils monitored their levels annually. For instance, it appeared that pupils did not monitor their levels from one task to the next or one month to the next. However, they referred to previous levels in which they were working at Level 7, when discussing their progress.

Levels as a Motivational tools or Demoralisers

Pupils appeared to value levels since it provided them with a reference point to aim for. Pupils commented:

“...you push your self towards them. You feel much better about yourself when you achieve a new level. It feels really successful after a long time” (Pupil 9, Year 9).

“I know that it is so hard to get a high level but I am working towards Level 6. I am more far away to a Level 5 and closer to Level 6. I hope that I will achieve it this year” (Pupil 6, Year 9).

Not every pupil perceived levels as a motivator. For pupils who achieved a lower level than they expected, the levels not a topic with which they wanted to engage. Two
pupils specifically reported their negative feelings towards levels. One pupil reflected his experience as follows:

"I did not want to join the breaks if I got a low level. All people will talk about at break will be their levels" (Pupil 1, Y9).

The above comment suggests that levels could go beyond the individual pupils and became a social issue. At an extreme case, as reflected above, levelling left a pupil with a feeling of exclusion since he did not meet the teacher's expectations.

Three pupils also reported that working at the same level over a period of time left with them which a feeling of "non achievement" and "hopelessness". The following reflection typifies such feelings:

"I don't think I will achieve a high level. I have been in level 5 over a year" (Pupil 14, Y8).

The above reflection confirms teacher concerns about levels as being "too big for pupils to monitor their progress".

**Brief Summary of Findings and Analysis of Pupil Interviews**

This part presented the findings and analysis of pupil interviews and focused on pupils' reported experiences of formative assessment. This part consists of five sub-sections. The first section presented pupils' views of what they valued in an assessment task. The second section explored their perceptions of their help-seeking behaviour and their interpretations of teacher's initiations and help. Section three and fourth addressed pupils' views on self and peer-assessment. The fifth section described pupils' perceptions of geography level descriptions.

Having presented the findings and analysis of pupil interview data, the next chapter (Chapter 5) draws the conclusions of this research and discusses the implications for policy and practice and also provides suggestions for further research.
CHAPTER 5: CONCLUSIONS AND IMPLICATIONS

The central aim of this thesis is to explore the planning of formative assessment over the whole of Key Stage 3 geography with the selection and design of assessment as a central focus. Important understandings regarding teachers’ views and practices in formative assessment were gained through this research. The chapter begins with a summary of the key findings and analysis that inform the main conclusions and suggestions for ways forward. The first section of the chapter begins with a consideration of each component of the formative assessment process, the selection and design of assessment tasks, feedback and target setting and self and peer-assessment. This is followed by a brief examination of how teachers use assessment tasks to make summative judgements of pupil attainment. The chapter ends with a consideration of the implications of the findings for policy, practice and further research.

The Design and Selection of Assessment Tasks

The thesis has focused on the design and selection of assessment tasks as the basis for teacher feedback, target setting and self and peer-assessment. The extent to which assessment is rendered formative is governed by the type and selection of assessment tasks implemented by teachers. In this section, conclusions are discussed in relation to several assessment task design criteria which affect the formative nature of assessment. The key design criteria identified were:

- Responding to the level descriptions,
- Planning for progression;
- Planning for differentiation;
- Fostering engagement and motivation;
- Providing variety in assessment task formats.
Responding to the Level Descriptions

Responding to the GNC level descriptions was the main concern of the teachers in the design of assessment tasks. This was related to the perceived external pressure to set annual quantitative targets in order to raise standards of pupil attainment. Some teachers also valued the guidance that level descriptions provided for their assessment practices. Level descriptions were used to provide the current and desired level of pupil attainment. Almost all teachers found assessment more manageable if they incorporated a relevant range of levels into the design of their assessment tasks. One clear benefit afforded by this approach was that the decision about final level judgements was easier to infer from the information provided by these tasks.

Although all teachers saw the incorporation of level descriptions into assessment tasks as being important, this requirement was challenging. The major problem was interpreting and using levels in assessment task design. This is because the level descriptions are abstract, theoretical and content free statements which are not easily transferable to actual assessment tasks for a specific topic. Whilst teachers were able to translate the level descriptions into research-based inquiry tasks, they found it more difficult to use them in other tasks, for instance, those which required pupils to engage in oral activities such as role-play. Some teachers perceived that level descriptions were a threat to their professional autonomy. In particular, they felt that responding to level descriptions made their assessment task design and implementation more mechanistic.

The literature reviewed suggested that planning for progression in pupils’ learning requires an understanding of the nature of progression in Geography Key Stage 3 and the ability to interpret level descriptions in the light of this. Teachers knew that they had to provide an opportunity for their pupils to reach the highest levels of which they were capable by the end of Key Stage 3. Although they recognised the importance of addressing progression in their task design, planning to achieve this for the entirety of
Key Stage 3 was challenging. Importantly, challenges arose because teachers did not have a coherent map of progression in pupils’ geographical learning and skill development for Key Stage 3; they only relied on their experience and an intuitive feel for progression as well as the use of the level descriptions. Moreover, some teachers were not convinced about the capacity of the level descriptions to encapsulate a comprehensive portrayal of geographical learning in Key Stage 3.

Teachers had developed their own tools to monitor and record progression in pupils’ learning throughout the Key Stage 3. Some used assessment checklists and record sheets in association with a series of assessment tasks over the key stage. These teachers valued these tools for their ability to help them to monitor progression. At the end of tasks, assessment information was mainly recorded as marks and grades in their mark books. Teachers made a distinction between formative assessment evidence which was mainly retained in their head, and a summary of assessment evidence which was recorded in their mark books. This suggests that teachers lacked a systematic way of recording in a written form detailed descriptions of specific aspects of pupils’ progress which they could act upon in a planned way.

**Differentiation**

Differentiation was also seen as an important consideration in the selection and design of assessment tasks. Differentiation was achieved differently in the short and long term. In the long term, it involved offering a variety of assessment task formats such as essays, letters and replies, newspaper reports, debates, oral presentations and posters. In the short term, it was implemented through structuring questions within an incline of difficulty and by providing additional guidance to help pupils to tackle tasks. In both the short and long term, teachers aimed to maximise pupils’ access to enable them to complete the tasks successfully. This was an important practice in terms of implementing formative assessment, since it increased the opportunity for all pupils to show their strengths and weaknesses so that their potential could be realised. In the long term, a downside of the attempt to provide a variety of assessment task
formats was that it presented an additional burden for teachers; this required them to develop variations in their marking and recording strategies, particularly in harmonising attainment levels across different types of assessment tasks.

The research concluded that differentiation by assessment task was more widespread for the “lower-achieving” pupils than the “more able”. Differentiation allowing high achievers to excel, existed to a limited extent since teachers seemed to have difficulty in finding ways to stretch the ability of “higher achiever” pupils. This was mainly grounded in the problem of articulating what constituted independent work for an assessment task. At the same time, it was apparent that the whole concept of differentiation could create a paradox in that it offered tasks involving a high degree of independence only to “higher achiever” pupils, while low-achieving pupils were denied the opportunity to achieve high-level tasks since they were restricted to simpler, structured tasks. However, teachers argued that this distinction was necessary since low-achievers simply would not obtain educational benefit from attempting difficult tasks which required a high level of independence.

While the majority of teachers acknowledged the differences between girls and boys in terms of their preferred assessment task formats, they did not administer different tasks by gender to compensate. Instead, they coped by applying a variety of assessment task formats over the key stage so that both gender groups had an opportunity to succeed.

**Engagement and Motivation**

Trying to ensure that the assessment tasks were enjoyable and stimulating to their pupils was regarded by all teachers as very important. One way of ensuring this, was to use a variety of assessment task formats. Teachers believed that unless their pupils sustained their interest in the assessment task, they were unlikely to demonstrate their true potential. Ultimately, to achieve effective formative assessment, pupils need to
engage positively with the process of completing the task in order for their potential to be developed and assessed.

Overall, teachers' assessment task design criteria were informed by their own experience in terms of what worked for their pupils, which is consistent with the work of Cooper and McIntryre (1996). They appeared to be unfamiliar with assessment task design criteria which are available in the research literature on authentic and performance assessment (Newman and Archbald; Meyer, 1994; Eisner, 1993). On a pragmatic level, some teachers used the SCAA (1996a; 1996b; 1996c; 1996d) Optional Assessment Tasks to help them with their moderation of pupils' work rather than to help them to improve their design of assessment tasks. This confirms the observations of the ARG (1999) who suggested that guidance for teachers on assessment was limited to helping them to make confident judgements about their pupils' level of attainment.

Teacher Feedback

Teacher feedback is a critical component of formative assessment, since, to a large extent, in conjunction with target setting, its function is to give direction for future learning and to enable development towards the desired level of pupil attainment. These provided insights into teachers' understanding of formative assessment feedback. Teachers felt that feedback could have a profound effect on the quality of learning if it provided pupils with specific information on how to improve their learning. Teachers' perceptions of feedback for learning were grounded in their pedagogical beliefs such as the importance of preserving pupils' self-esteem and in their experiences rather than on the use of guidance materials or research literature of what constitutes good assessment feedback (Cooper and McIntryre, 1996). Teachers appeared to prefer to follow their "gut feelings". Teachers' medium term feedback appeared to be related to the level descriptions and this was planned and facilitated through their own devised assessment record sheets and checklists. In class feedback
appeared to be more ad hoc and based on teachers’ perceptions of pupils’ emergent needs.

Teachers reported the use of various feedback strategies, which ranged from descriptive to evaluative as identified by Tunstall and Gipps (1996). Teachers identified different feedback strategies for their oral and written feedback. They attached descriptive functions to oral feedback more than their written feedback. Oral feedback was delivered more frequently and, for teachers, it provided more opportunities to suggest ways of improvement. The observational findings also confirmed the potential of oral interactions to give pupils guidance for future learning development. This demonstrates that oral feedback was a preferred strategy. However, they emphasised that there were time and resource constraints on individual feedback.

Teachers’ individual feedback in a written form, tended to be more standardised as it was based on predetermined, level-driven criteria. This was mainly designed to achieve departmental consistency and manageability. Teachers reported that more assessment tasks were in the process of being levelled. This trend reflected the need for them to be able to aggregate their judgements about pupils’ attainment to make a final ‘best-fit’ judgement at the end of Key Stage 3. A consequence of this standardised and routinised approach to assessment using criteria directly related to the level descriptions, was that only limited evidence of pupils’ learning was elicited. (i.e. the items appearing in the assessment checklists and sheets). There could have been a loss of other evidence of pupils’ learning, which, therefore could have been acted upon. This suggest that teachers need to prioritise the key learning objectives for an assessment task and make sure that they include them in their assessment and record sheets.

Teachers’ written feedback took the form of brief comments, and marks and grades. These were attached to small items of written work within an assessment task as well as at the end of tasks. Teachers suggested that providing detailed written comments
was not manageable within their high workload, since they dealt with large a number of pupils with limited resources. Furthermore, the evidence from this research suggested that the use of levels, marks, and, grades, rather than detailed comments, better facilitated the process of aggregation towards a final level decision. This is consistent with the major review of formative assessment Black and Wiliam (1998b:5) who suggested that marking generally appears to be unfocused and lacking the specific information on how pupils could be improved. This situation conflicts with Butler’s evidence that comments rather than grades are necessary to improve pupils’ learning (Butler, 1988).

The nature and form of teacher feedback was also shaped by departmental policies. The emphasis on achieving consistency within a given department caused individual teachers to change their practice to align themselves with other teachers. This was not welcomed by all of the teachers, and some perceived this as a threat to professional autonomy.

In their written and oral feedback on assessment tasks, teachers appeared to overemphasise the trivial and subsidiary aspects of pupil performance, although they claimed they did not do so. This conclusion is consistent with the findings of Black and Wiliam (1998b:6) who suggested that teachers’ feedback overemphasises the quantity and the presentation of work and neglect its quality. As a result, opportunities were lost to help pupils to improve their learning since the focus of assessment was on subsidiary aspects of work such as presentation. It was clear from the pupil interviews that their perception of “what mattered” in their written work was quantity and presentation rather than the quality of their thinking. This was related to the nature of the feedback that they had received on their written work. This was an inhibiting factor for formative assessment because it prevented teachers from focusing on those aspects which could make a difference to pupils’ learning.

Effective feedback depends on the elicitation of evidence of pupils’ learning as they complete an assessment task. Classroom observations demonstrated that pressure was
often applied by the teacher to his pupils to finish a task within a certain time frame. In this study, it was observed that more emphasis appeared to be placed on the assessment of the final product (in this case a pupil role-play) rather than how pupils approached the task (process). This focus on 'product' rather than 'process' resulted in the identification of a limited number of assessment criteria for the pupil role-play and consequently opportunities were lost to assess the learning process that preceded it.

There was also evidence from both the teacher interviews and the observations that teachers’ fundamental beliefs about pupils’ self-esteem could inhibit the effectiveness of feedback, a point which is consistent with the findings of Torrance and Pryor (1996). This research suggests that teachers tended to over-emphasise the strengths of “low-achieving” pupils in order promote their self-esteem. This concern to promote positive reinforcement could detract from the quality of formative assessment feedback due to omitting to communicate pupils’ weaknesses in their learning to pupils.

There was also evidence that the concern of teachers to maintain a high level of pupil engagement can also militate against effective feedback. The observational findings revealed that concerns for classroom management can work against the objectives of formative assessment. Specifically, in some situations, oral feedback was disproportionately directed towards misbehaving pupils, in order to keep the class on task and maintain discipline.

The observational data suggested that the teacher’s reluctance to relinquish his power and authority in the classroom meant that he controlled the feedback process. The teacher interview data also suggested that teachers had ownership of all aspects of the assessment process.
Target Setting

Target setting was a particular form of feed forward which operated in the medium and long term. It appeared to be powerful, and was the commonest strategy used by teachers for communicating to their pupils the next steps in their learning. Teachers facilitated target setting through assessment checklist and record sheets which they had devised and these were made available to pupils at the beginning of tasks. These assessment checklists included an ‘aide memoire’- a checklist to help pupils to complete the task, assessment criteria for the task expressed mainly in the form of statements relating to the level descriptions and a pre-specified list of subject specific and generic targets to help pupils to improve. These teachers used the same format which they had devised for each assessment task throughout the year. They perceived that such tools made the whole process more manageable and sustainable. As was the case with assessment feedback, target setting practice was essentially a teacher controlled and managed activity. Teachers perceived that their pupils lacked the ability to set high quality targets which meant that pupils had very little freedom to set their own targets.

There was an over-riding concern by teachers to identify relevant subject-specific targets which could directly assist pupils’ learning. As was the case for planning for progression, teachers relied on level descriptions to guide their thinking for their medium term target setting. However, for this purpose, the interpretation and use of level descriptions was beset with problems. The majority of teachers reported that they used the same statements transposed from the level descriptions in their assessment checklists and record sheets. Whilst, some teachers successfully managed this task, others expressed their concerns that level descriptions did not leave much space for them to exercise their professional autonomy in their target setting practices.
Pupil Self and Peer-Assessment

It was well established in the literature review chapter that one of the principles of effective formative assessment is that pupils come to understand the standards against which they are being judged (Sadler, 1989). Pupil self and peer-assessment are acknowledged as being important for developing this understanding with the research evidence suggesting that these approaches encourage pupils to think about their learning and develop their metacognitive skills. Sharing these assessment criteria with pupils is also recognised as an essential characteristic of effective practice in formative assessment and is featured in the guidance provided in the National Strategy Key Stage 3 materials.

The teacher interviews revealed that the main variations which occurred in teacher practice related to the nature and frequency of pupil self and peer-assessment. The majority of teachers used assessment checklists to guide pupils' self-assessment. The remainder either required pupils to prepare self-assessment reports, or followed a less formalised pattern and merely asked for pupils' reflections during lessons. Compared with self-assessment, pupil peer-assessment was much less prominent but was occasionally used with group work.

Pupil self and peer-assessment afforded a number of benefits that contributed to pupils' learning. A major benefit was that pupils developed their skills of self-reflection. This is in line with the literature reviewed which identified the value of self-reflection in equipping pupils with assessment criteria to evaluate their own learning (Klenowski, 1995). Furthermore, teachers valued the interactive dialogue which enabled their pupils to develop a better understanding of the assessment criteria for a particular task. From the pupils' perspective, they described their self-assessment experience positively and suggested that it helped them to understand the assessment criteria as well as their targets. When they assessed their peers, they gained insight into other pupils' approaches, which they would not have done otherwise.
Despite recognising the benefits of pupil self and peer-assessment, teachers in general did not make extensive use of self-assessment and their use of peer-assessment was even rarer. This was essentially related to three factors; teachers’ perception of the ability of their pupils to engage in the assessment process; the difficulty that some experienced in translating assessment criteria into a language that their pupils could understand; and the difficulty of finding time for self and especially peer-assessment when there was a perceived need to meet the statutory requirements of the Programme of Study.

When compared with the other components of formative assessment, effective pupil self and peer-assessment requires teachers to transfer some of their power as assessors to pupils. In this research, teachers were very reluctant to give up their power and authority to their pupils; they were given freedom within well-defined parameters. Teachers were sceptical about the ability of their pupils to assess their own and their peers’ progress accurately. In particular, they felt that their pupils in Year 7 were not sufficiently mature to conduct self and peer-assessment.

Teachers found it difficult to translate the assessment criteria into language their pupils could understand so that they could use them to assess themselves and their peers. The majority of teachers used assessment criteria which were lifted directly from the National Curriculum. Some translated the level descriptions into pupil friendly language to make their more accessible. Some had simplified the language of the level descriptions whilst attempting to retain their meaning; they allocated the statements to several progression strands to produce a generic document which they used for all assessment tasks.

Although teachers recognised the merits of pupil self and peer-assessment, they were concerned about its feasibility in the reality of the classroom situation. Overall, the research revealed a pattern of pupil self-assessment which was well structured and within which pupils had to operate predefined assessment criteria to assess themselves and their peers. This is consistent with “procedural autonomy”, one of the categories
identified by Ecclestone (2002) to describe autonomy in pupils’ learning, in which she stated that “the learner is proactive within well-specified rules”. In this research, the assessment criteria for pupil self and peer-assessment appeared to be very standardised and limited in focus, which left little scope for pupils to reflect back on their overall learning experience. As a consequence, pupil engagement in thinking about their own learning was very much teacher directed and lacked the critical dialogue between teachers and pupils as a vehicle for further pupil development. Overall, pupil self and peer-assessment were not generally a priority for teachers and were not proactively designed into assessment tasks.

Using Level Descriptions to Decide on a Final Level at the end of Key Stage 3

This thesis has highlighted the importance of context for implementing formative assessment and the Literature Review chapter explored the scope of the GNC as a context for formative assessment. The Findings and Analysis chapter drew attention to the impact of level descriptions on teachers’ formative assessment practices. What became clear was that level descriptions affected teachers’ overall practice significantly in a number of ways. One issue was how they used the results of assessment tasks over the key stage to make summative judgements.

This thesis briefly explored the ways in which (formative) assessment data informs the final level decision (summative) and looked at the tensions caused by the different functions of assessment. However, this was a minor objective of this research. Teachers’ practices in deriving final level judgements were found to fall into two broad categories, which could be described as a “cautious approach” and a “flexible approach” respectively. The cautious approach was characterised by teachers utilising less professional judgement when making a final level decision and instead aggregating the results of individual assessment tasks to derive a final outcome. In this practice, the weight of the final level decision evidence was derived from recorded assessment data for the three years of the key stage. Furthermore, these teachers used a summative end of year assessment task to back up their final level judgements.
However, they did not view themselves as being highly mechanistic; they were rather cautious about their final level decisions. For those teachers who used the flexible approach, professional judgement exerted the greatest influence on their judgement. For these teachers, assessment evidence was not only determined by the results of their tasks but also by their judgements of pupils’ learning throughout the Key Stage 3.

The two different approaches were based on different ways of thinking about the role of assessment in teaching and learning. In the cautious approach, teachers tended to separate assessment from teaching and learning. For them, assessment had to have a summative function and when they assessed their pupils, there was a clear distinction between teaching and learning. In the teacher flexible approach, assessment was thought of somewhat differently. Here, the assessment language that the teachers used appeared to be more formative and summative assessment appeared to be a less of a priority. These teachers also tended not to value the summative tests as sole indicators of pupil achievement.

Irrespective of the different emphasises in making final level judgements, all of the teachers expressed reservations about working with level descriptions. For some teachers, levels were not the “right” criteria for judging pupils’ progress; for some others, they did not match teachers’ ‘gut feelings’ about pupils’ learning. To reiterate, the underlying reasons for this situation were as follows:

- Some teachers had doubts about the capacity of level descriptions to reflect pupils’ progress in their learning.

- Interpreting and devising level descriptions for specific units of work was identified by teachers as a weak aspect of their assessment practices.

- Teachers felt that since they lacked the ability to interpret the level descriptions, the communication of levels to pupils was a hard task. In some cases, teachers just ended up copying and pasting statements in pupils’ work sheets.
In cases where some teachers felt directed and restricted by the level descriptions, they tended to stand by their professional judgements as a way of expressing autonomy about their assessment practices.

**Limitations of the Thesis**

This section addresses the limitations of this thesis on two levels. The first section highlights the methodological limitations. This is followed by the limitations in relation to the conceptualisation of formative assessment.

The findings in this research are mainly developed on the basis of teacher interviews. This was a methodological decision prior to the study and the researcher set out to explore teachers' interpretations of what had happened in their classrooms and how they had used assessment. Throughout the interviews, the need for a documentary analysis became evident since teachers had to reflect on their assessment tasks and marked pieces of work many of which were not available during the interviews. A documentary analysis would have been beneficial for this study not only to enrich the understanding of assessment tasks and teacher feedback by exemplifying them but also by complementing teacher interview data for methodological triangulation.

The use of multiple research methods in a single study require some articulation on the ways in which triangulation is used. In this study, the triangulation was not achieved by each method checking the other's validity by addressing the same questions; but by enriching the understanding of a number of selected assessment issues. For instance, observational study involved one teacher in order to deepen the understanding of formative assessment by illuminating its implementation within geography classrooms. The findings drawn from the observational study is therefore limited to one Key Stage 3 geography teacher and two classrooms. Similarly, the pupil interviews involved the selection of pupils from the two classrooms which were also used in the observational study. Based on the use of research methods, it could be
suggested that the classroom observations and pupil interviews are complementary since the findings were generated from the same contexts. However, the findings from teacher interviews were derived from the practice of twelve teachers and therefore their overall compatibility with the findings from the other two methods is limited. Being aware of this limitation, the research design was a deliberate choice and the researcher preferred to observe the continuation of a task over a period of time rather than observing different teachers within shorter periods of time.

The last limitation relates to the understanding of formative assessment. This research has sought to build on the emerging conceptualisation of formative assessment by examining the key components and the relationship between each other in the entirety of this thesis. The three-component model was used to interpret and synthesise published research in Chapter 2. The same model was used to make sense of teachers' interpretations of their assessment practices in Chapter 3. By including certain criteria and excluding some others, the overall conceptualisation of formative assessment would be incomplete. However, working with selected criteria (the components) was again a strategic decision for in-depth conceptualisation of formative assessment. This research also acknowledges other factors that could contribute to the understanding of formative assessment such as classroom culture, teacher and pupil relationship and beliefs regarding how pupils learn. Thus, the conceptualisation of formative assessment would only provide a partial view.

Implications for Policy, Practice, and Research

The final section addresses the implications of this research for:
- Policy, specifically that which relates to the Key Stage 3 GNC and guidance given to practitioners;
- The professional practice of geography teachers as well as those who lead and manage their work in schools;
Future research, specifically as it relates to Key Stage 3 GNC, but as importantly to a wider and deeper understanding about formative and summative assessment as key aspects of learning.

**Implications for Policy**

The notion that assessment should be an integral part of teaching and learning was promoted by TGAT as the basis for the assessment arrangements for the National Curriculum. The report referred to the need to plan for the incorporation of assessment opportunities and introduced the idea of teacher ‘feedback and feed-forward.’ Integrating assessment into the planning of teaching and learning proved to be difficult to implement and weaknesses in assessment practice were a persistent feature of OFSTED reports throughout the 1990s. Concerns about teachers’ lack of understanding of assessment were articulated by Black and Wiliam (1998a). The changes in the arrangements for assessment in the National Curriculum, notably the replacement of the content driven SOAs by level descriptions, provided a new context and challenge for teachers implementing formative assessment.

It was the publication of a research survey by Black and Wiliam (1998a) which by emphasising the potential of formative assessment to improve learning and raise standards of attainment, placed formative assessment firmly on the policy agenda. The key message was that assessment could improve learning provided that weaknesses in classroom practice were addressed. The survey investigated the separate components of formative assessment, paying particular attention to teacher feedback but did not explore the issue of the design and selection of assessment tasks as a vehicle for planning and implementing formative assessment in the short and long term which was the focus for this study.

There was an urgent need for guidance on the theory and practice of formative assessment. Some of this was generic and was designed to help teachers appreciate the standards to which their students were working (SCAA, 1996a). The same
organisation was responsible for the production of optional assessment tasks. The Assessment Reform Group (1999) referred to a number of initiatives aimed at helping teachers with classroom assessment such as Consistency in Teacher Assessment (QCA, ACAC). These provide materials to help teachers to making judgements of pupils' work.

Since the publication of the Black and Wiliam survey the Government has shown increasing interest in providing practical support. Unfortunately, this material was not available when the data was collected for this research. The most concerted support for teachers so far has come from the National Key Stage 3 Strategy which has made formative assessment (referred to as assessment for learning) a key priority. The strategy has resulted in the production of guidance materials which are designed to help teachers to become aware of the characteristics of good practice, to appreciate the links between these characteristics, their impact on students' learning and the teaching strategies which are necessary to achieve these learning outcomes. This guidance represents a significant improvement in support of the implementation of day-to-day formative assessment in classrooms. So far the practical guidance from the National Strategy has been generic. In 2004, further generic work is planned on assessment for learning but there are also plans to launch subject specific material.

In Geography, teachers were supported by an active subject association which produced guidance to help teachers interpret and use the level descriptions (Butt et al., 1995). Later articles explored formative assessment in practice by identifying the purposes of assessment and its use for improving practice by Lambert and Balderstone (2000), Butt (2002), differentiation by Battersby (1995) and Waters (1995), target setting by Hamson and Sutton (2000) and Howes (2003), diagnostic and formative assessment of students' thinking by Leat and McGrane (2000), continuity and progression by Bennetts (1996), planning for progression by Hopkin and Telfer (2000). The majority of articles are published by the journal, Teaching Geography, which provides valuable advice on practice and theory of assessment. In addition to this, the Geographical Association (GA) published a number of materials which
specifically address assessment for learning and provides strategies on how to use level descriptions (Butt et al., 1995) and how to plan assessment for medium, short and long term (Hopkin, 2000; Howes, 2000a; Howes, 2000b). However, this research demonstrated that teachers were not aware of the materials the guidance materials produced by the Teaching Geography; nor had they consulted the literature on the use of level descriptions published by the GA. Evidence suggests that Geography teachers need practical guidance to implement formative assessment. The production of guidance materials to assist assessment task design and selection should be the one of the key priorities. Assessment task design was the component of formative assessment that challenged teachers the most. In planning assessment tasks over the Key Stage 3, teachers felt that they would have benefited from the provision of guidance on progression in skill development. A need remains to provide guidance to help teachers plan for formative assessment as a long-term strategy over the key stage.

This study highlights a number of issues which have implications for policy makers:

**Assessment Task Design**

This research identified assessment task design as a challenging aspect of Key Stage 3 geography teachers’ assessment practices. The main problem appeared to be designing tasks to promote geographical learning strategically for the entirety of the Key Stage 3. Teachers appeared to view task design as a stand-alone issue and they tended to plan for tasks as a unit rather than over a period of time which could constitute a context to promote geographical learning. This finding suggests that teachers need practical guidance on

- How a series of assessment tasks, which are interrelated in order to promote continuity and progression, could be devised over the KS 3 to promote pupils’ geographical learning?
Progression in Geography Skill Development

One of the key reasons why the above question challenged the geography teachers was that lack of a comprehensive map of progression in geographical learning for Key Stage 3. The above question could be addressed by policy makers by identifying the key issues that matter in the learning of geography. This research suggests that geography teachers require external help and guidance in finding out

- What skills are important in the teaching and learning of geography?
- How specific skills progress throughout the Key Stage 3 and how geographical skill development could be supported throughout the Key Stage 3?

This research concluded that geography teachers appeared to be skilful in defining what constituted progression for enquiry and map skills but not for the development of social skills and empathy. They also viewed geography as more comprehensive than enquiry and map skills. This need perhaps was met with the introduction of a new thinking skills category (creative skills and evaluation skills) in the Secondary Curriculum; however, the need still remains to help teachers to interpret and incorporate these skills in their planning. Exemplifying how and when a specific skill is promoted through a series of assessment tasks could provide valuable guidance for geography teachers. Such materials could also be helpful for teachers to interpret the policy requirements and alleviate some of the tensions that this work could potentially bring.

This research has identified the issue of planning for skill development as a crucial area. This is difficult for all teachers because of the different skill frameworks of a generic type which have been introduced to which teachers in Key Stage 3 will be obliged to respond. Skills for enquiry featured prominently in the 1995 version of the National Curriculum and their importance has been further emphasised in the latest version. However, in addition the 2000 National Curriculum includes key skills and thinking skills. It is the teachers who are obliged to decide for themselves how these
skills will be incorporated in the short, medium and long term lesson plans. This is an area where guidance would be very useful.

**Differentiation for High Achiever Pupils**

Differentiation was one widely used practice to support pupils’ learning. What became evident was teachers’ difficulty in maximising learning opportunities for higher achiever pupils. Teachers identified one way of achieving this as promoting pupils’ independence in their learning. Understanding and enabling pupil independence constituted a challenging area and geography teachers could benefit from a number of practical strategies on how to achieve this. Key questions that policy makers could address are as follows:

- **What does ‘independent learning’ mean in the learning of geography?**
- **What should be the nature of teachers’ support throughout the years of the Key Stage 3 to achieve pupil independence in a number of aspects of the learning of geography?**

This thesis provides insights on possible solutions to the above questions from one angle by illustrating different ways of using assessment criteria and how independence could be achieved through pupil self and peer-assessment. As outlined previously, this involves pupils’ in the construction and use of assessment criteria and in the negotiations with their teacher during this process. Depending on the dynamics of this process and the sharing of power, one could conceive of different degrees of pupil independence in ‘evaluative skills’. Similar guidance could be provided to teachers how to develop pupils’ independence in other aspects of the learning of geography in Key Stage 3.
Supporting Pupils’ Learning through Target Setting

This thesis identified target setting as a powerful tool to enable formative assessment. However, teachers had concerns about their target setting and sought guidance to make this practice more effective. The following areas remain as grey areas in teachers’ practices:

- **Identification of subject specific targets, which could assist pupils’ geographical learning.**

  In addressing the above issue, the teachers used geography level descriptions which they found difficult to interpret. In addition to the existing exemplary materials, geography teachers would benefit from further guidance in this area.

  Teachers viewed target setting not only as a tool to give pupils subject specific direction but also to help them to develop more generic skills when they were learning geography. Teachers valued organisation skills and wanted to develop pupils’ social skills within their Key Stage 3 experience. However, teachers did not place enough emphasis in the development of these skills and their place could be secured if they are included in guidance documents.

Accessibility of Guidance Materials for Practitioners

This research highlighted that teachers were not aware of the guidance materials produced by the Geographical Association or the SCAA to help them to plan for progression and differentiation in pupils’ learning. Consideration needs to be given to the ways in which these materials are made available to teachers. One way of increasing teachers’ awareness would be to strengthen the links between the research community and practitioners. This could be achieved by involving practitioners in research in collaboration with the research community.
Implications for Practice

The importance of this research is that it brings together the theory and practice of formative assessment and provides insights into the implementation of formative assessment in Key Stage 3 geography classrooms. On the basis of teachers’ experiences, a number of suggestions are made to overcome the current challenges for the effective implementation of formative assessment. The following sections discuss the implications of this research for practice for the key components of formative assessment.

Assessment Task Design and Selection

This research highlighted the importance of assessment tasks in shaping the nature and quality of formative assessment and suggested that the existence of the other components is dependent on whether they are planned when assessment tasks are designed. This requires that teachers consider feedback, target setting, pupil self and peer-assessment together, rather than separate components and plan the ways in which each component could support the rest. This thesis suggests that each component of formative assessment is crucial and the success of formative assessment is dependent on whether the components are planned together in the task design stage.

This research suggested that teachers tended to consider each component separately and did not address them altogether during their task design. For instance, in the majority of the cases, pupil peer-assessment was used only if teachers had some extra time at the end of lessons. Overwhelmingly, target setting was not linked to feedback and pupil self and peer-assessment since teachers used different assessment criteria with each component. This could be explained by teachers lacking a well-defined map of progression in pupils’ geographical learning and tailoring their overall planning for assessment on that basis. In order to plan formative assessment strategically teachers need to consider the aspects of progression in pupils’ learning and how they can support them throughout their tasks.
Differentiation also appeared to be crucial to achieving formative assessment. In this research, teachers tended to consider differentiation prior to their task implementation; an example was the use of different modes of presentation for an assessment task for pupils. Relatively less consideration was given to how differentiated learning evidence could be scaffolded and could inform target setting. Thus, teachers appeared to be skilful in planning for differentiation but not in using the differentiated learning outcomes for improving pupils’ learning. This suggests that when teachers plan for differentiation, they should consider how differentiated learning outcomes could provide varying scope for feedback, target setting, and pupil self and peer-assessment.

**Suggestions for Target Setting**

This research revealed the challenges in geography Key Stage 3 teachers’ target setting practices in identifying subject specific targets to improve pupils’ learning. Some teachers mainly used the statements in the level descriptions to guide their practice, which created the problem of interpretation and communication to pupils of the standards which they were aiming. For instance, in extreme cases teachers used level statements as they appear in the GNC without any interpretation or additional explanation. This suggests that teachers should develop ways of illustrating what constitutes success for a piece of work, for instance either by modelling or presenting a piece of successful work. This could not only assist pupils’ understanding of ‘what’ is expected them to achieve but also give them ideas about ‘how’ to achieve. Furthermore, teachers could involve their pupils in the decision making process during which they set targets and construct the expected achievement. This could help teachers to communicate targets to their pupils more effectively as well as empower their pupils in their assessment and learning.

This research demonstrated that teachers used different approaches for target setting for different periods of time, a finding which is consistent with Hopkin (2000) who differentiated assessment practices for different time intervals. Target setting should
support pupils both emergent short-term and long-term learning needs. This requires some strategic planning decisions. For the strategic planning of formative assessment, teachers need to consider targets, which potentially could lead to improvements in closing pupils’ learning gap, which are also grounded into a map of progression in pupils’ learning throughout the Key Stage 3.

**Suggestions for Pupil Self and Peer-Assessment**

Pupil self and peer assessment is an essential component of formative assessment and promises a number of benefits, as were suggested in the Literature Review chapter. Although the recognised potential benefits of pupil self and peer-assessment to enhance learning, teachers found the implementation challenging. The main explanation was the difficulty which teachers experienced in identifying the “appropriate” assessment criteria for pupils to use in their assessment. As in the previous suggestion, a defined map of progression in pupils’ geographical learning’ could inform teachers’ choice of assessment criteria over a series of assessment tasks.

A prominent reason why pupil peer-assessment was rarely used was related to teachers’ doubts about their pupils’ ability to be objective when they judged their peers’ work. This problem could be overcome by developing pupils’ peer assessment skills by increasing opportunities for its use. The Literature Review defines pupil self and peer-assessment as “skills”, a competence which could be developed through time and experience. Thus, the improvement in this area could be achieved if such activities are planned in advance so that their frequency is secured.

**The Way Forward: New Research Areas**

This research suggested that the understanding of formative assessment has been evolving. In order to contribute to the debate, this research has provided insights into the current practice of formative assessment in Key Stage 3 classrooms by using a three-component model. However, it is also acknowledged that the application of this
model might have provided a partial view of formative assessment. In the light of the research data, it is clear that there are a number of important and relevant issues in the understanding of formative assessment which have not been addressed in the literature. Further research could address these issues by addressing the following questions:

- **How do teachers make their pupils aware of the standards for which they are aiming?**

As this research suggested, pupils need to understand the standards on which they are being assessed. This is a key principle of formative assessment and understanding how teachers manage this process would be beneficial.

- **How do teachers and pupils record pupils’ evolving understanding and skill development?**

In the assessment cycle, the literature identified a three-stage model, elicitation of evidence, interpretation of evidence and acting upon evidence (Wiliam and Black 1996). This research suggests that there is a key stage missing, the recording of pupils’ evolving understanding and skill development. Teachers used a hybrid of methods to record pupils’ progress, a mixture of marks and grades for effort and attainment on an ongoing basis, punctuated by varying number of assessment tasks for a particular unit of work. There is a need for further research to identify the recording process to help teachers to improve their practice.

- **How can formative assessment be used to develop the thinking skills identified in the National Strategy?**

The 1997 Government White Paper, Excellence in Schools, emphasised the need for deep learning and the teaching of thinking skills, and both of these feature prominently in the National Strategy for Key Stage 3. This new emphasis has important
implications for the whole process of formative assessment. How teachers’ thinking and practice of formative assessment change as a result of the strategy could be addressed by further research studies.

- **How do teachers aggregate the achievement of pupils in a structured way to make a summative judgement?**

The current research paid some attention to the relationship between formative and summative assessment in Key Stage 3. The TGAT Report (1987) suggested that it was possible to aggregate the results from assessment tasks throughout the key stage to make an accurate judgement. Wiliam and Black (1996) draw attention to the possible tension between the formative and summative use of assessment evidence. The research conducted briefly explored how teachers used assessment tasks throughout the key stage to make summative judgements, it did not explore the issue in detail. Such research would be clearly important in resolving how teachers could reconcile these two functions of assessment.

- **How does the current debate about the nature of geographical learning in the Secondary Curriculum impact on formative assessment?**

This research suggests that the key action of formative assessment is the closing of the gap between current and desired levels of performance (Chapter 2). To enable this in key stage 3 geography, teachers need clear ideas about what constitutes learning in geography as well as how it progresses.

- **How do teachers change the learning cultures of their classroom to facilitate formative assessment?**

Black (2003) introduces the notion of the need to ‘engineer the learning environment’ to facilitate formative assessment and this could be investigated further in Key Stage 3 geography.
How do teachers and use different assessment task formats to provide opportunities for feedback, feedforward and pupil self and peer-assessment?

In recent years, the range of assessment task formats has widened in Key Stage 3 geography. How the nature of a task affects the other components of formative assessment has not been adequately explored and this is an area for further research. Such information could be valuable in deepening the understanding of the relationship between the key components of formative assessment.

How do pupils perceive specific components of formative assessment?

The literature reviewed highlights the absence of studies on pupils’ views on assessment issues. This research provided the evidence that their interpretations were crucial in the implementation of formative assessment. Although there have been a number of publications which addressed pupils’ views related to aspects of assessment issues (Pollard et al., 2000), the majority are mainly from teachers’ perspectives. In particular, there is potential to explore pupils’ views on different assessment task formats. Further research on assessment should make pupils’ views more visible.

The Way Forward: Research Approaches in Understanding Formative Assessment

In understanding formative assessment, particular attention ought to be given context-related issues. The literature review provided research evidence and suggested that formative assessment is a highly contextualised process and shaped by the dynamics of the context within which it occurs (Bell and Cowie, 2001). The findings of this research also identified the relationship between pupils and teachers, classroom culture and a number of context related factors that affected the implementation of formative assessment. In the light of these findings, case studies, which offer opportunities to understand situations in depth through the application of a number of research methods, could aid in understanding formative assessment in classrooms. There has
been also increasing research evidence suggesting that improving the practices of formative assessment through research could be achieved by engaging practitioners in the research process. The ARG (1999) reported the improvements in practice when practitioners engage in action research where teachers reflect on their own practice.
REFERENCES


Assessment Reform Group (1999), Assessment for Learning: Beyond the Black Box, Cambridge, University of Cambridge School of Education.


BERA (1992), Ethical Guidelines, Edinburgh, published by SCRE for BERA.


Davies, P. (1990), *Differentiation in the Classroom and in the Examination Room: Achieving the Impossible?*, Cardiff: Welsh Joint Education Committee.


DfEE (1995), Geography in the National Curriculum, London, HMSO.


Department for Education and Employment and Qualifications and Curriculum Authority (2000), A Scheme of Work for Key Stage 3: Geography, London, QCA.


237


SCAA (1994), *Geography in the National Curriculum*, Draft Proposals, HMSO.

SCAA (1996a), *Key Stage 3 Optional Tests and Tasks: Unit 4, Brazil. Geography Key Stage 3*, London, SCAA.

SCAA (1996b), *Key Stage 3 Optional Tests and Tasks: Unit 2, Earthquakes. Geography Key Stage 3*, London, SCAA.

SCAA (1996c), *Key Stage 3 Optional Tests and Tasks: Unit 1, Coastal Processes and Environmental Management. Geography Key Stage 3*, London, SCAA.

SCAA (1996d), *Key Stage 3 Optional Tests and Tasks: Unit 3, Settlement. Geography Key Stage 3*, London, SCAA.


Wylie, R. C. (1979), The Self Concept, vol.,2, Theory and Research on Selected Topics, Lincoln, Nebraska, University of Nebraska Press.

APPENDICES
APPENDIX A

Introductory Letter on Research Project-Addressed to Teachers
I am a Doctorate student currently studying at the University of Leicester, Department of Education, under the supervision of Alan Sutton and Marlene Morrison. I am presently conducting research to investigate certain aspects of assessment in Key Stage 3 Geography. The title of the research is “An investigation into the Theory and Practice of Formative Assessment’.

The main aim of the research project is to gain understanding of how teachers devise a series of assessment tasks and conduct their short, medium and long-term assessment practices. This research also touches on the issues of target setting. Although some relevant theoretical discussion exists in the literature, I am seeking to complement this with some anecdotes stemming from practical experience. I hope also that this research can help to raise awareness issues relating to assessment arrangements, in the eyes of educational policy makers.

The main part of my research project comprises of one hour, semi-structured, recorded, interviews with teachers. The second part of the project is to conduct classroom observations in order describe formative assessment practices in day-to day classrooms

I am very much hoping that you will be willing to be involved. Actually, I especially would like to work with you for the observational stage because of your knowledge on the assessment issues. Describing your assessment practices is an extremely important one and I feel that there is much to be learned from you.

We would be very happy to hear from you and to be to arrange a meeting. My contact details are included below.

Thanking you in advance for your assistance.

Best Wishes.

Yours sincerely,

Yonca Tiknaz

University of Leicester,
School of Education,
21 University Road
LE2 7RF Leicester
APPENDIX B

Introductory Letter on Research Project-
Addressed to Pupils
Dear Student,

08.05.2001

I am a Doctorate student currently studying at the University of Leicester, Department of Education. I am presently conducting research to investigate certain aspects of teacher assessment in Key Stage 3 Geography level. With this letter, I am inviting you to take part. I am looking for a sample of about 15-20 Key Stage 3 students.

The main aim of the research project is to gain understanding of what do you think about the assessment of the last project you have just completed (Population for year 8 students, ecosystems & sustainable development for the year 9 students). I am really interested in how you evaluate your own piece of work and how do you interpret your teacher’ comments. It is also important for me to know what you think about the level descriptions in the Geography National Curriculum.

The main part of my research project comprises of one 15-30 minutes, recorded, interviews with you, which I hope to conduct between 28 May and 11 June during your lunch breaks. It would be also very useful if you take the project with you for the interviews. The main reason for the tape-recording of the interviews is that I would benefit from review of the dialogue, particularly as English is my second language and I would like to listen to the tapes to later on. The interviews will be conducted by myself.

I am very much hoping that you will want to participate. The topic of the research is important and I feel that there is much to be learned from your experience as a student.

I would be very happy to hear from you in order to arrange a meeting. Could you please submit this letter to your teacher by identifying a possible interview date and time for the interviews.

Thanking you in advance for your help.

Best Wishes.
Yours sincerely,

Yonca Tiknaz

University of Leicester,
School of Education,
21 University Road
LE2 7RF Leicester

Possible interview day:
Time:
APPENDIX C

TEACHER INTERVIEW QUESTIONS

Questions prepared for use in semi-structured interviews with teachers to investigate key areas of assessment practice.
Teacher Interview Questions on Assessment Practices

Could you please give me a basic outline of assessment practices which you use during the course of Key Stage 3?

PART 1: QUESTIONS ABOUT ASSESSMENT TASKS

1.1 When you devise assessment tasks, what are the most important criteria for you to consider for a Key Stage 3 Geography classroom?

1.2 How important is differentiation in the setting of assessment tasks?

   Do you differentiate girls and boys when setting particular assessment tasks?

1.3 Do you use the SCAA exemplar assessment tasks?

   Has the SCAA actually helped you to devise your assessment tasks?

   Do you adopt them?

1.4 To what extent are assessment tasks devised to facilitate progression in geographical learning?

1.5 How frequently do you set assessment tasks over the key stage?

PART 2: QUESTIONS RELATED TO TEACHER FEEDBACK

2.1 How do you use the results of the assessment tasks to inform students of their progress?

2.2 Do you make your students aware of the standards by which they are being judged, i.e. level descriptions?

2.3 When you provide feedback, what form does it take? (i.e. marks, grades, general comments and statements). Are they related to the level descriptions?
PART 3: QUESTIONS RELATED TO TARGET SETTING

3.1 Do you set targets for pupils as goals for their improvement?

3.2 What types of targets are set for pupils?

3.3 How do you communicate these to pupils?

3.4 Do students set targets for their own improvement? What types of targets do they set for themselves?

PART 4: QUESTIONS RELATED TO PUPIL SELF AND PEER - ASSESSMENT

4.1 Do your pupils assess their own progress?

4.2 Do you use peer-assessment?

4.3 How do you use peer-assessment?

4.3 How often is peer-assessment is used?

PART 5: QUESTIONS RELATED TO FINAL LEVEL JUDGEMENTS

5.1 How do you use performance data from the results of the assessment tasks to decide on a final level for each student at the end of Key Stage 3?

5.2 How do you aggregate information from the assessment tasks to decide on a final level?

5.3 How do you derive final levels in consideration of the statement of attainments in the level descriptions?
APPENDIX D

PUPIL INTERVIEW QUESTIONS

Questions prepared for use in semi-structured interviews with pupils to explore a number of key areas of assessment practice.
INTERVIEW QUESTIONS

1. PUPILS’ HELP SEEKING BEHAVIOUR
   - What do you do if you are stuck in your work?
   - What does your teacher do?

2. PUPILS’ UNDERSTANDING OF LEVEL DESCRIPTIONS
   - Do you know which level you are working at?
   - What level has been awarded to you for this work? Why do you think it was awarded?
   - What do you need to do in order to achieve a higher level?

3. PUPILS’ IDEAS ON HOW TO MOVE FORWARD IN THEIR LEARNING - WHETHER THEY HAVE CLEAR TARGETS
   - Do you need to know what you need to do in order to make your work better?
   - If you do it next time, what would you do?

4. PUPILS’ IDEAS ON ASSESSMENT TASKS
   - What have you found interesting, useful and challenging about the assessment task you have just completed?
   - If you had to organise a geography lesson on this topic what would you do?

5. PUPILS’ VIEWS ON SELF AND PEER-ASSESSMENT
   - Could you be fair when you assess your self? (your friends?)
   - In what situations do you feel comfortable about assessing your friends?
APPENDIX E

PRELIMINARY LIST OF CODES FOR
DATA ANALYSIS
# Teacher Assessment in Geography Key Stage 3

## Code List

<table>
<thead>
<tr>
<th>Name of the Code</th>
<th>Illustration</th>
<th>Defining Questions-Issues</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Internal Context</strong>&lt;br&gt;IC: Characteristics&lt;br&gt;IC: Characteristics of School</td>
<td><strong>IC</strong>&lt;br&gt;IC-CHAR-SCH</td>
<td>Background information about each school&lt;br&gt;*Differentiate the code on the basis of</td>
</tr>
<tr>
<td><strong>IC: Organisational Procedures</strong></td>
<td><strong>IC-ORG</strong></td>
<td>The existence of a humanities or geography department&lt;br&gt;The number of humanities teachers&lt;br&gt;The number of geography subject specific teachers&lt;br&gt;The population of school&lt;br&gt;Other characteristics in relation to pupil population</td>
</tr>
</tbody>
</table>

## Characteristics of Geography Department: DEPT

<table>
<thead>
<tr>
<th>Characteristics of Geography Department: DEPT</th>
<th>DEPT-COM</th>
<th>Moderation issues&lt;br&gt;Focus on the emphasis of communication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communication between Staff</td>
<td>DEPT-COM</td>
<td>Focus on the nature of assessment information that is communicated to parents</td>
</tr>
<tr>
<td>Communication with Parents</td>
<td>DEPT-INT</td>
<td>How important is formative assessment in departmental development plan?</td>
</tr>
<tr>
<td>Departmental Development Plan</td>
<td>DEPT-DDP</td>
<td></td>
</tr>
<tr>
<td>Other Characteristics of Geography Department</td>
<td>DEPT-CHR</td>
<td></td>
</tr>
</tbody>
</table>

## External Context

<table>
<thead>
<tr>
<th>External Context</th>
<th>EC</th>
<th>Enquiry the existence of any endorsement related to the LEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>EC: LEA related Issues&lt;br&gt;EC-END&lt;br&gt;EC-END-COUNTTY</td>
<td><strong>EC-END&lt;br&gt;EC-END-COUNTTY</strong></td>
<td>*Differentiate the code&lt;br&gt;Whether the school is a city school&lt;br&gt;Whether the school is a county school</td>
</tr>
<tr>
<td>EC: Demographics&lt;br&gt;In county, school personnel</td>
<td><strong>EC-DEM</strong></td>
<td></td>
</tr>
</tbody>
</table>

271
<table>
<thead>
<tr>
<th>CHARACTERISTICS OF ASSESSMENT TASKS</th>
<th>ASS-CRI</th>
</tr>
</thead>
<tbody>
<tr>
<td>AS-CRI: OBJECTIVES</td>
<td>ASS-OBJ</td>
</tr>
<tr>
<td>What are the perceived purposes</td>
<td></td>
</tr>
<tr>
<td>of assessment task?</td>
<td></td>
</tr>
<tr>
<td>AS-CRI: SCAA TASKS</td>
<td>ASS-SCAA</td>
</tr>
<tr>
<td>How the SCAA tasks are used?</td>
<td></td>
</tr>
<tr>
<td>*Differentiate the code</td>
<td></td>
</tr>
<tr>
<td>AS-CRI: CONTENT</td>
<td>ASS-CNT</td>
</tr>
<tr>
<td>How important is the content of</td>
<td></td>
</tr>
<tr>
<td>assessment tasks in planning</td>
<td></td>
</tr>
<tr>
<td>stage?</td>
<td></td>
</tr>
<tr>
<td>AS-GROUPWORK ASSESSMENT</td>
<td>ASS-GRP</td>
</tr>
<tr>
<td>Do teachers plan group work</td>
<td></td>
</tr>
<tr>
<td>when they plan or devise</td>
<td></td>
</tr>
<tr>
<td>assessment tasks?</td>
<td></td>
</tr>
<tr>
<td>AS-PROGRESSION</td>
<td>ASS-PRG</td>
</tr>
<tr>
<td>How do teachers plan and</td>
<td></td>
</tr>
<tr>
<td>promote progression in pupils’</td>
<td></td>
</tr>
<tr>
<td>geographical learning?</td>
<td></td>
</tr>
<tr>
<td>AS-FORM (STYLE)</td>
<td>ASS-STY</td>
</tr>
<tr>
<td>What are the most mentioned</td>
<td></td>
</tr>
<tr>
<td>assessment task formats?</td>
<td></td>
</tr>
<tr>
<td>AS-CRI: DIFFERENTIATION</td>
<td>ASS-DIF</td>
</tr>
<tr>
<td>How do teachers use differentiation?</td>
<td></td>
</tr>
<tr>
<td>*Differentiate the code on the</td>
<td></td>
</tr>
<tr>
<td>basis of different tools that</td>
<td></td>
</tr>
<tr>
<td>enable differentiation.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>THE RECORDING PROCEDURE</th>
<th>REC</th>
</tr>
</thead>
<tbody>
<tr>
<td>4, 4.1, 4.2</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RECORDING THE EVIDENCE OF LEARNING</th>
<th>REC</th>
</tr>
</thead>
<tbody>
<tr>
<td>REC: RECORDING BY TEACHERS</td>
<td>REC-TCH</td>
</tr>
<tr>
<td>What tools do teachers use to</td>
<td></td>
</tr>
<tr>
<td>record assessment information?</td>
<td></td>
</tr>
<tr>
<td>What is the form of recorded data?</td>
<td></td>
</tr>
<tr>
<td>*Differentiate the code for</td>
<td></td>
</tr>
<tr>
<td>different forms of assessment</td>
<td></td>
</tr>
<tr>
<td>evidence:</td>
<td></td>
</tr>
<tr>
<td>quantitative marks;</td>
<td></td>
</tr>
<tr>
<td>qualitative comments.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>REC: THE CHARACTERISTICS OF RECORDING BY TEACHER</th>
</tr>
</thead>
<tbody>
<tr>
<td>REC-TCH-CHR</td>
</tr>
<tr>
<td>What kind of assessment information is</td>
</tr>
<tr>
<td>recorded by teachers?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>REC: THE CHARACTERISTICS OF STUDENTS RECORDING</th>
</tr>
</thead>
<tbody>
<tr>
<td>REC-STD-CHR</td>
</tr>
<tr>
<td>What kind of assessment information is</td>
</tr>
<tr>
<td>recorded by pupils?</td>
</tr>
</tbody>
</table>
| **REC: OBJECTIVES** | **REC-OBJ** | What are the purposes of recording? 
Try to make links to understanding formative assessment process. |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>REC: AUDIENCE</strong></td>
<td><strong>REC-AUD</strong></td>
<td>Who is the intended audience of recorded assessment information?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>ISSUES RELATED TO FEEDBACK</strong></th>
<th><strong>FDB</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FDB: THE FORM OF FEEDBACK</strong></td>
<td><strong>FDB-FORM</strong></td>
</tr>
<tr>
<td><strong>FDB: RELATIONSHIPS BETWEEN LEVEL DESCRIPTIONS</strong></td>
<td><strong>FDB-LVD</strong></td>
</tr>
</tbody>
</table>
| **FDB: THE QUALITIES OF FEEDBACK** | **FDB-QUAL** | What is the nature of feedback? 
*Differentiate this code for evaluative and descriptive feedback. 
*Differentiate this code on the basis of pupils’ involvement in the construction of feedback. |

<table>
<thead>
<tr>
<th><strong>TARGET SETTING PROCESS</strong></th>
<th><strong>TRG</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TRG: TARGET SETTING BY TEACHERS</strong></td>
<td><strong>TRG-TCH</strong></td>
</tr>
</tbody>
</table>
| **TRG-TCH-GN** | How doe teachers set targets to develop pupils’ learning? 
*Differentiate this code: Type of target: Generic targets; Subject specific. |
| **TRG-TCH-SB** | |
| **TRG: TARGET SETTING BY PUPILS** | **TRG-STD** |
| **TRG-STD-GN** | Do pupils set targets for themselves? 
*Differentiate this code: Generic target; Subject specific. |
<p>| <strong>TRG-STD-SB</strong> | |
| <strong>TRG: BY EXTERNAL BODIES</strong> | <strong>TRG-EXT</strong> |
| By the LEA | Do teachers mention the effect of the LEAs or the governments’ quantitative target setting on their “qualitative target setting” practices? |
| By the Government | <strong>TRG-LEA</strong> |
| <strong>TRG-GOV</strong> | |</p>
<table>
<thead>
<tr>
<th>TRG: EFFECT OF TARGET SETTING ON FORMATIVE ASSESSMENT</th>
<th>TRG-EF-FA</th>
<th>If yes, find out the ways in which this affects their target setting practice.</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is the place of target setting process in teachers’ overall formative assessment practices?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| TEACHERS’ VIEWS ABOUT DIFFERENT APPROACHES TO ASSESSMENT |
|---|---|---|
| **TEACHERS DEFINITIONS OF RELEVANT CONCEPTS** | **DEF** | **How do teachers conceptualise their** **formative assessment practices?** |
| DEF: FORMATIVE ASSESSMENT | DEF-FOR | How do teachers conceptualise their formative assessment practices? |
| DEF: SUMATIVE ASSESSMENT | DEF-SUM | How do teachers conceptualise their summative assessment practices? |
| OPN: VIEWS ABOUT THE RELATIONSHIP BETWEEN FORMATIVE AND SUMMATIVE ASS. | OPN-SUM&FOR | How do teachers use formative assessment evidence to make summative judgements of pupils’ progress? Do they see these two assessments as compatible? |
| DEF: GEOGRAPHICAL SKILLS | DEF-SKL | What are the most mentioned skills? |
| OPN: RELATIONSHIP BETWEEN ASSESSMENT AND TEACHING AND LEARNING | OPN-REL | How do teachers link their teaching to assessment? Is assessment addressed separately in this context? |
### ASPECTS OF TEACHERS ONGOING DAY-TO-DAY ASSESSMENTS

<table>
<thead>
<tr>
<th>ELEMENTS OF ONGOING TEACHER ASSESSMENT</th>
<th>ASS-ONG</th>
<th>How do teachers describe their informal assessment as a part of day-to-day teaching and learning?</th>
</tr>
</thead>
<tbody>
<tr>
<td>FREQUENCY</td>
<td>ONG-FRQ</td>
<td>How often do teachers collect annotated pupils' work during one assessment task?</td>
</tr>
<tr>
<td>TEACHERS' MENTIONED WAYS OF ELICITING ASSESSMENT EVIDENCE</td>
<td>ONG-ELI-EVI</td>
<td>How do teachers elicit assessment evidence? *Differentiate this code: Which assessment evidence do they rely on more?</td>
</tr>
<tr>
<td>THE USE AND FUNCTION OF ASSESSMENT CHECKLISTS AND RECORD SHEETS</td>
<td>ONG-CHK</td>
<td>What is the role of checklists in teachers' overall assessment practices?</td>
</tr>
</tbody>
</table>

### ISSUES RELATED TO PUPIL SELF-ASSESSMENT

<table>
<thead>
<tr>
<th>TOOLS THAT FACILITATED PUPIL SELF ASSESSMENT</th>
<th>SSA-TLS</th>
<th>What tools do teachers use to facilitate pupil self-assessment?</th>
</tr>
</thead>
<tbody>
<tr>
<td>TIMING</td>
<td>SSA-TIME</td>
<td>At what stage of the task it is used?</td>
</tr>
<tr>
<td>TYPES OF TASKS WITHIN WHICH IT APPEARED</td>
<td>SSA-TASKS</td>
<td>Is there a particular relationship between the nature of assessment task and its capacity to absorb pupil self-assessment?</td>
</tr>
<tr>
<td>ASSESSMENT CRITERIA</td>
<td>SSA-CRI</td>
<td>What is the nature of assessment criteria that are employed in such activities?</td>
</tr>
<tr>
<td>PERCEIVED VALUE</td>
<td>SSA-POS</td>
<td>What is the potential value of pupil self-assessment on pupils' learning?</td>
</tr>
<tr>
<td>PERCEIVED PROBLEMS</td>
<td>SSA-NEG</td>
<td>What are the problematic aspects of implementing these activities?</td>
</tr>
</tbody>
</table>
**ISSUES RELATED TO PUPIL PEER-ASSESSMENT**

<table>
<thead>
<tr>
<th>TOOLS THAT FACILITATED PUPIL SELF ASSESSMENT</th>
<th>PPA-TLS</th>
<th>What tools do teachers use to facilitate pupil peer-assessment?</th>
</tr>
</thead>
<tbody>
<tr>
<td>TIMING</td>
<td>PPA-TIME</td>
<td>At what stage of the task it is used?</td>
</tr>
<tr>
<td>TYPES OF TASKS WITHIN WHICH IT APPEARED</td>
<td>PPA-TASKS</td>
<td>Is there a particular relationship between the nature of assessment task and its capacity to absorb pupil peer-assessment?</td>
</tr>
<tr>
<td>ASSESSMENT CRITERIA</td>
<td>PPA-CRI</td>
<td>What is the nature of assessment criteria that are employed in such activities?</td>
</tr>
<tr>
<td>PERCEIVED VALUE</td>
<td>PPA-POS</td>
<td>What is the potential value of pupil peer-assessment on pupils’ learning?</td>
</tr>
<tr>
<td>PERCEIVED PROBLEMS</td>
<td>PPA-NEG</td>
<td>What are the problematic aspects of implementing these activities?</td>
</tr>
</tbody>
</table>

**HOW TEACHERS AGGREGATE ASSESSMENT EVIDENCE**

<table>
<thead>
<tr>
<th>MANAGING ASSESSMENT DATA WHICH IS DIFFERENT IN ITS NATURE</th>
<th>DEAL-FNL</th>
<th>How do teachers deal with the different forms of assessment data? *Differentiate this code on the basis of: marks, grades, levels and other sources of assessment evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>USE OF PORTFOLIOS</td>
<td>DEAL-PORT</td>
<td>Do teachers use portfolios to aggregate assessment data? *Differentiate this code if portfolios are used for other purposes.</td>
</tr>
</tbody>
</table>
HOW TEACHERS MAKE A FINAL LEVEL DECISIONS BY USING THE LEVEL DESCRIPTORS

<table>
<thead>
<tr>
<th>FINAL LEVEL DECISIONS</th>
<th>FNL-BFT</th>
<th>By Making General 'Best Fit' Judgements 5.2, 5.3</th>
</tr>
</thead>
<tbody>
<tr>
<td>FINAL LEVEL DECISIONS</td>
<td>FNL-PRT</td>
<td>By using 'best fit' judgements in relation to children's portfolios</td>
</tr>
<tr>
<td>FINAL LEVEL DECISIONS</td>
<td>FNL-SPT</td>
<td>5.3 By splitting the level descriptions (e.g. by creating separate statements and counting half or more as attaining level).</td>
</tr>
<tr>
<td>FINAL LEVEL DECISIONS</td>
<td>FNL-IDF</td>
<td>By identifying key aspects of level descriptions</td>
</tr>
</tbody>
</table>

HOW TEACHERS INTERPRET ACHIEVEMENT IN RELATION TO ONE LEVEL

<table>
<thead>
<tr>
<th>REWARDING A LEVEL</th>
<th>FNL-ABV</th>
<th>The level description which overall describes the child’s attainment better than the one above or below</th>
</tr>
</thead>
<tbody>
<tr>
<td>REWARDING A LEVEL</td>
<td>FNL-75</td>
<td>Students must achieve 75% or more of the statements in the level description</td>
</tr>
<tr>
<td>REWARDING A LEVEL</td>
<td>FNL-IMP</td>
<td>Students must achieve important aspects of a level description</td>
</tr>
<tr>
<td>REWARDING A LEVEL</td>
<td>FNL-INT</td>
<td>Intuition</td>
</tr>
<tr>
<td>REWARDING A LEVEL</td>
<td>FNL-100</td>
<td>Students must achieve almost 100% or more of the statements in the level description</td>
</tr>
<tr>
<td>REWARDING A LEVEL</td>
<td>FNL-50</td>
<td>Students must achieve 50% or more of the statements in the level description</td>
</tr>
<tr>
<td>REWARDING A LEVEL</td>
<td>FNL-OTH</td>
<td>Or any other approaches</td>
</tr>
</tbody>
</table>