

***Barriers to formative assessment in school  
science:  
A Critical Realist perspective***

**By**

**Birendra Singh**

**UCL Institute of Education**

**Thesis submitted for the Degree of Doctor of Philosophy**

**I, Birendra Singh confirm that the work presented in this thesis is my own.  
Where information has been derived from other sources, I confirm that this has  
been indicated in the thesis.**

## Abstract

There is strong evidence in literature that formative assessment (FA) promotes learning, has a positive impact on teacher development and improves pupil outcomes. It nevertheless remains at the periphery of practice in schools in England. This study investigates why this is so, through an observational study of classroom practice in science in Year 8 (ages 12-13) in three secondary schools in an outer London local authority.

The study is set in the context of tensions and anxiety in schools emanating from the government's 'standards agenda', namely, the imposition of National Tests, targets, league tables of results and inspections by the Office for Standards in Education (Ofsted).

Through lesson observations, scrutiny of pupils' books, interviews with teachers and pupils and discussions with school, local, and national managers, this study confirms the dominance of tests and examinations. It confirms too the practice of 'triage' for rationing education and discovers the existence of the 'tripartite' system and the 'factory-model' of schooling within the walls of comprehensive schools. All of these mitigate against the effective use of FA to promote learning in the classroom.

The findings include teachers' lack of subject knowledge, an acute shortage of physics teachers and the disabling of teachers' agency through 'double triage' - leading to demoralisation. The government and school managers continue to use blunt instruments for monitoring narrowly defined 'standards' whilst the professional learning of science teachers is neglected. Schools are 'marking up' or even 'making up' test and coursework results through dishonest practices, masking the fact that they are failing many of their pupils.

Finally, Bhaskar's theory of learning drawn from his theory of creativity and his notion of 'unfolding the enfolded' are counter-posed against the theory of *innate ability* and used to propose a new, improved theory of assessment. This theory is *Enabling Assessment*, which builds on formative assessment and assessment for learning and provides a model of assessment with transformative potential.

## **Acknowledgements**

This is to acknowledge the debt I own to my supervisors: late Professor Roy Bhaskar, Professor Michael Barnett, Professor Shirley Simon and the late Dr Sheila Turner. I am grateful for the support of my colleagues at work, Roger Luxton and Justin Donovan, for encouraging me and facilitating funding for my part-time study, and to my colleagues in the science team, Kath Twin, Liz Lawrence and Jason Hatherill, for supporting me during periods of high pressure. My sincere thanks also go to the head teachers, teachers and pupils of the case study schools for their co-operation and openness, without which the research would not have been possible.

My heart-felt appreciation goes to my wife Carol for standing by me through thick and thin, and to my children Jenny and Asha for encouraging me whilst patiently feeling my 'absence' during this unusually long journey of learning and discovery.

I would also like to thank my brothers, Devendra and Nagendra, for doing my share of the house-hold chores during their stay with us in England.

## **Dedication**

*To the loving and grateful memory of my mother, Tetara Debi, and  
my father, Mahanth Singh, who sacrificed much so that I could have  
education, something they did not have but valued most.*

## **Contents:**

Abbreviations	17
List of Appendices	18
<b>Chapter 1 Introduction</b>	<b>19</b>
1.1 The background	19
1.2 Tensions over issues relating to assessment	27
1.2.1 Teachers' professional learning	27
1.2.2 Low expectations	31
1.2.3 The tripartite system of education	33
1.2.4 The notion of 'positive achievement'	34
1.3 The rationale for this study	37
1.4 Research Questions and Research Design	41
1.4.1 Main study Phase 1: 2005 to 2007 – National Tests in place	42
1.4.2 Main study Phase 2: 2007 to 2008 – National Tests in place	43
1.4.3 Main study Phase 3: 2009 to 2011 – post National Tests	44
1.4.4 Main study Phase 4: 2011 to 2012 – post National Tests	44
1.5 Methods of data analysis	45
1.5.1 Lenses used for analysis	46
1.5.2 The dialectic of control	47

1.5.3	The power of discourse	48
1.6	Consideration of ethical issues	49
1.7	What the study hopes to achieve	50
1.8	The structure of the thesis	50
1.9	Summary of Chapter 1	51

<b>Chapter 2</b>	<b>Summative assessment</b>	<b>52</b>
------------------	-----------------------------	-----------

2.1	Introduction	52
2.2	A brief history	53
2.2.1	From ‘patronage’ to ‘competition’	53
2.2.2	Examinations to ensure quality and report progress	54
2.2.3	The development of schooling in the USA	56
2.2.4	The U.S. Army’s Alpha and Beta tests	57
2.2.5	A ‘gate-keeping’ role for assessment	58
2.2.6	The ‘gate-keeping’ role challenged	59
2.2.7	The influence of Galtonian ideas in England	59
2.2.8	The GCSE and the idea of ‘positive achievement’	60
2.3	The National Tests	63

2.3.1	Advantages and disadvantages of the National Tests	64
2.3.2	Accountability procedures – targets and league tables	66
2.4	A time-line (1850 – 2012)	72
2.4.1	A short-lived optimistic turn	75
2.5	Assessment in secondary schools – a complex picture	77
2.5.1	A crowded assessment space	78
2.5.2	Issues with the current state of summative assessment	79
2.5.3	Negative impacts of tests	81
2.5.4	Summary of Chapter 2	81
 <b>Chapter 3 FA/AfL</b>		 83
3.1	Introduction	83
3.2	What is formative assessment?	84
3.3	Formative assessment or Assessment for Learning, which?	89
3.3.1	FA and AfL – similarities, differences and overlap	95
3.4	The notion of ‘criteria’	98
3.4.1	Mastery learning	98
3.4.2	The TGAT model	99
3.4.3	The Clarke model	100



3.4.4	Ramprasad's notion of the 'gap'	101
3.4.5	Feedback	102
3.4.6	Connecting FA with pedagogy	104
3.5	The case for FA and AfL	105
3.5.1	Evidence of success of FA/AfL	107
3.5.2	Evidence of success of FA/AfL from Ofsted	107
3.6	Efforts to support teachers - CPD programmes	108
3.6.1	The KMOFAP model of CPD	110
3.6.2	The National Strategies	111
3.7	A lack of FA in secondary school science	115
3.8	Some theoretical issues	118
3.8.1	SA, FA and AfL and AfT	120
3.9	Towards a new theory	126
3.10	Summary of Chapter 3	126

## **Chapter 4 Methodology**

4.1	Introduction	129
4.2	The Exploratory studies	132
4.2.1	Phase 1: Interviews with teachers	132
4.2.2	Phase 2: scrutiny of marking	134

4.3	A developing methodology	136
4.3.1	Learning from exploratory studies	140
4.4	Research questions and the Main Study	143
4.5	Outline of the fieldwork programme	145
4.5.1	Main study Phase 1	145
4.5.2	Main study Phase 2	146
4.5.3	Main study Phase 3	146
4.5.4	Main study Phase 4a	147
4.5.5	Main study Phase 4b	147
4.6	Field work at six levels	147
4.6.1	Classroom contexts	147
4.6.2	Departmental contexts	148
4.6.3	School contexts	148
4.6.4	Local Authority context	148
4.6.5	National context	148
4.6.6	International context	149
4.7	Amendments - theoretical and practical considerations	149
4.7.1	Informed by experience	150
4.7.2	Informed by Critical Realism	151
4.8	Organisation of Data Collection	152
4.9	Description of the case study schools	154
4.9.1	Curie High School (CHS)	154
4.9.2	Downton Community School (DCS)	156
4.9.3	Westfield Comprehensive School (WCS)	157

4.10	Summary of Chapter 4	158
 <b>Chapter 5 A Theoretical Framework</b>		 160
5.1	Introduction	160
5.2	A critique of the ideology of innate ability	160
5.2.1	Ability, attainment and performance	160
5.2.2	The Potential – Performance spectrum	168
5.3	Structural barriers to FA/AfL	172
5.3.1	Teachers as agents of change	174
5.3.2	The Transformational Model of Social Activity	175
5.3.3	Open systems and the idea of a laminated system	178
5.3.4	Educational Triage	182
5.3.5	Assessment ‘shapes’ people	183
5.3.6	Applying TMSA and the double articulation of complexity	184
5.3.7	Global/International influences	185
5.3.8	Mega level	187
5.3.9	Macro level	188
5.3.10	Meso level	189
5.3.10 (a)	The School	189
5.3.10 (b)	The Neighbourhood	198
5.3.11	Micro level	198
5.3.12 (a)	Individual Level	200
5.3.12 (b)	Teachers as individuals	203
5.3.13	Sub-individual level	204

5.4	Education and two concepts of power	205
5.5	Summary of Chapter 5	207
<b>Chapter 6</b>	<b>A Critical Realist Theory – Enabling Assessment</b>	<b>209</b>
6.1	Introduction	209
6.1.1	Shortcomings of the concepts of FA and AfL	210
6.2	Enabling Assessment	214
6.2.1	Enabling Assessment - a formula	216
6.2.2	The EA Triangle	218
6.3	‘Unfolding the Enfolded’	218
6.4	EA, FA and AfL – a comparison	225
6.5	What needs to happen for EA to occur?	226
6.6	Towards an explanatory critique	227
6.6.1	The ideology of innate ability	227
6.6.2	The Factory-Model of Schooling	228
6.6.3	Agency: teachers	229
6.6.4	Agency: schools’ senior managers	230
6.6.5	Agency: ‘Double Triage’	231
6.7	A theory of transition	232
6.7.1	The notion of a ‘good’ teacher	232
6.7.2	The notion of teachers’ mindset	232
6.8	Compensation or Transformation?	234
6.8.1	A possible resolution: a Critical Realist synthesis	234
6.8.2	Education and Power 1 and Power 2	235

6.9	Summary of Chapter 6	236
<b>Chapter 7 Data and Analysis</b>		<b>238</b>
7.1	Introduction	238
7.2	The case study schools	238
7.3	Pupil groupings: banding and setting	241
7.3.1	Pupil grouping at CHS	241
7.3.2	Pupil grouping at DCS	254
7.3.3	Pupil grouping at WCS	256
7.4	Assessment in practice in science	258
7.5	The nature and quality of feedback to pupils	264
7.5.1	Marking	266
7.5.2	Feedback through discussion in the classroom	272
7.5.3	Feedback through reports to pupils and parents	275
7.6	The quality of teaching in science	275
7.6.1	The quality of teaching in CHS	277
7.6.2	The quality of teaching DCS	289
7.6.3	The quality of teaching at WCS	290
7.7	Restricted access to the science National Curriculum	291

7.7.1	The emergence of ‘How Science Works’	294
7.7.2	The National Curriculum – an entitlement?	295
7.8	Science teachers’ workload and the nature of CPD	297
7.8.1	Teachers’ workload	298
7.8.2	CPD	299
7.9	Summary of Chapter 7	302
7.9.1	Poor quality of marking and feedback	303
7.9.2	Poor quality of teaching	303
7.9.3	Poor coverage of the National Curriculum	304
7.9.4	A lack of CPD culture	304
7.9.5	Key issues	305
<b>Chapter 8</b>	<b>Data and Analysis continued</b>	<b>306</b>
8.1	Introduction	306
8.2	Misplaced gaze - historical	308
8.3	Misplaced gaze: the Standards Agenda	309
8.3.1	Monitoring by the schools’ senior managers	310
8.3.2	Monitoring at DCS	311
8.3.3	Monitoring at WCS	312

8.3.4	Monitoring at CHS	312
8.3.5	Monitoring by Ofsted	314
8.4	What is absent in the process of monitoring?	315
8.4.1	Excessive importance given to the leadership	315
8.5	Impact on pupils and teachers of the absences	316
8.5.1	Impact on pupils at CHS	318
8.5.2	Impact on pupils at DCS	327
8.5.3	Impact on pupils at WCS	331
8.6	The Factory-Model of schooling	334
8.6.1	The Factory-Model - diagram	334
8.6.2	The Vicious Cycle	335
8.7	Summary of Chapter 8	335
<b>Chapter 9 Conclusion</b>		337
9.1	Introduction	337
9.2	The findings	339
9.3	A possible way forward	342
9.3.1	Unfolding the Enfolded	342
9.3.2	Towards high expectations	343

9.4	A new vision: Enabling Assessment	343
9.4.1	EA and the EA cycle	343
9.4.2	The five circles of creation	345
9.4.3	The five moments of the dialectic of creativity	346
9.5	Science in society	347
9.5.1	The ‘learner-centred model’	348
9.6	A new vision for teachers	349
9.6.1	A new professionalism	350
9.7	The crisis in education: risking ‘Half our Future’	350
9.7.1	A continuing danger to ‘Half our Future’	351
9.8	Every child matters, every teacher matters	352
9.8.1	A fight back for a truly comprehensive school system	353
9.9	Summary of Chapter 9	354
9.10	Ideas for further research	355
	 Bibliography	 359
	 Appendices	 369



### **Abbreviations:**

**KMOFAP:** King's Medway Oxford Formative Assessment Project

**DES:** Department of Education and Science

**DfES:** Department for Education and Science

**DCSF:** Department for Children, Schools and Families

(between 2007 and 2010)

**DFE:** Department for Education (2010 – to date)

**BERA:** British Educational Research Association

**GCE:** General Certificate of Education

**CSE:** Certificate of Secondary Education

**GCSE:** General Certificate of Secondary Education

**ERA:** Education Reform Act

**Ofsted:** Office for Standards in Education

**PANDA:** Pupil Attainment Data

**TAPAS:** Teachers Assessing Pupils – Active Support

**SSEC:** Secondary School Examination Council

**CPD:** Continuing Professional Development

## List of Appendices

A0 i: Email requesting a science teacher to participate in the research	369
A0 ii: Letter to pupil participants	370
A0 iii: Letter to the headteacher	371
A0 iv: Request to Head of Department	373
A0 v: Initial interview questions	375
A1: Marking Pro-forma A	376
A1: Marking Pro-forma B	377
A2: Extracts from the Autumn Package	378
i) 2001 KS3 National Value Added information	378
ii) 2001 KS3 National Benchmark information	379
A3: Extract from DCS Staff Handbook, pp40-41	380
A4: CHS multiple choice AfL Assessment sheet	381
A5: Example of teacher's marking	382
A6: Meeting with DHT (WCS)	383
A7: Example of book audit by Head of Science DCS	384
A8: Learning and teaching CPD forum – CHS	385
A9: Sample page from Y8 pupil's book (P <sub>32</sub> )	386
A10: Sample pupil progress card	387

## Chapter 1 Introduction

### 1.1 The Background

**“ – formative, so that the positive achievements of a pupil may be recognised and discussed and appropriate next steps may be planned;”**

- Task Group on Assessment and Testing (TGAT), 1987, Section 23.

At the January (1989) conference of the Association for Science Education (ASE) at the University of Birmingham the big hall was packed with science teachers anxious to hear Professor Paul Black, the Chairman of the Task Group on Assessment and Testing. Although there was considerable opposition to the imposition of the National Curriculum (NC) amongst teachers (Simon, 1988), science teachers were well disposed towards Professor Black and felt re-assured that he was chairing such an important Group. I was in the audience and Professor Black's speech was received with thunderous applause.

The TGAT was set up in July 1987 and reported to Secretary of State for Education and Science, the Rt. Hon Kenneth Baker MP, on 24 December 1987. The report was published soon after. One of its key recommendations was that:

“... the basis of the national assessment system be essentially formative, but designed to indicate where there is need for more detailed diagnostic assessment. At age 16, however, it should incorporate assessment with summative functions.” (Section 27)

This recommendation was rejected by the Government which imposed summative National Tests for children ages 7, 11 and 14 phased in from 1991. This was followed by the imposition of a series of monitoring steps which became known collectively as the ‘standards agenda’.

Nearly a decade later the publication of Black and Wiliam Review (Black and Wiliam, 1998a) of international literature on assessment and classroom learning highlighted the positive impact of formative assessment and re-kindled the formative assessment (FA) and summative assessment (SA) debate. This set the scene, at the start of the 21<sup>st</sup> century, for a prolonged period of tension and discussion between the educators, who wished to promote FA, and governments, which continued steadfastly to implement the ‘standards agenda’.

Tests and examinations have been the tools used for judging the attainment of pupils perhaps ever since schooling began. The term ‘assessment’ itself as used in education is of relatively recent origin. It is derived from the Latin *assidere*, meaning ‘to sit beside or with’ (Earl, 2003). The meaning of assessment given in the Chambers Twentieth Century Dictionary (revised edition, 1959) was related to the evaluation of land for the purpose of taxation. By 1998 the definition in the Collins English Dictionary, Millennium Edition (Collins, 1998) was given as ‘the act of assessing, especially (in Britain) the evaluation of a student’s achievement on a course.’ In the Education Reform Act 1988, the definition of the word ‘assess’ includes to examine and to test.

Broadfoot explains that the term ‘assessment’ may be taken to mean ‘the deliberate and overt measurement of educational performance in order to provide information beyond the interactive learning situation’ (Broadfoot, 1996).

Satterley (Satterley, 1991) introduces the term ‘evaluation’ in addition to assessment. According to him ‘assessment’ means the actual process of measurement and ‘evaluation’ means the interpretation of such measurement against ‘particular norms of performance’. In general, however, the term assessment is usually taken to include both. The terms ‘educational assessment’, ‘classroom assessment’ and ‘assessment’ appear to be used interchangeably.

In this study my use of the term assessment includes references to processes and procedures such as tests, examinations and any other measurement of educational performance of a pupil by a teacher or an institution. I have used the terms summative assessment (SA) and formative assessment (FA) in the senses defined by TGAT. Summative describes those assessment activities which occur at the end of a module or a unit or a course of instruction. Its purpose is to accredit the student with a pass or fail grade or level and/or place him/her in a rank order. Formative assessment, on the other hand, describes those activities which recognise and confirm pupils’ positive achievements and indicate the next steps in their learning.

In 1999, the Assessment Reform Group (ARG) defined summative assessment as Assessment **of** Learning (AoL) and formative assessment as Assessment **for**

Learning (AfL). Since then, FA and AfL have often been used interchangeably. However, in recent years important differences in the interpretation of the terms have emerged and these are discussed in Chapter 3, section 3. Another significant development for science education at this juncture of time was the publication of the report *Beyond 2000* (Millar and Osborne, 1998) which introduced the debate on ‘science for scientists and science for citizens’. This began a move away from the Double Award science which most schools were offering to most (80%) of their pupils and towards Triple Award science to the most able, Core and Applied science to some and core or BTEC science to the rest.

My study is set against the above backdrop. The focus of the thesis is to investigate barriers to the development of FA/AfL in science through a study of assessment practices in Year 8 in three secondary schools located in an outer London local authority. The aim is to understand the challenges and difficulties which continue to exist in putting into practice the ideas of FA and AfL in science and to develop further the conceptual and theoretical understanding of FA/AfL. This is critical to education policy and local practices alike.

My own experience as a teacher, a teacher educator, and a local authority and Ofsted inspector provides a background to the study, which originates with the premise that summative assessment is embedded in the school system in England. Formative assessment and assessment for learning, on the other hand, whilst they have been discussed and theorised with increasing vigour in the UK over the last three decades, still remain essentially at the periphery of practice in spite of clear, strong evidence

of their positive impact on pupils' learning. At the same time the UK government's 'standards agenda' has all but drowned out the summative versus formative assessment debate by the intensity with which it has been pursued.

In setting off on an investigation of the barriers to effective FA/AfL it has seemed logical to trace the origins of the idea of assessment. The dominant tradition in assessment has been summative but formative assessment has probably also existed ever since schooling began. Nevertheless, for over a century external examinations and tests have dominated as the main forms of assessment. This summative mode of assessment has prospered because it has been perceived by stakeholders to be a reliable, and relatively cheap, measure of the outcomes of schooling. Examinations have also promoted systematic selection on 'merit' and have facilitated the division of labour by grading pupils in hierarchical categories (Broadfoot, 1986, p. 59).

Formative assessment, on the other hand, has been typically considered as a part of good teaching and has remained informal. H. D. Black (Black, 1986, p. 12) made this point more than two decades ago when he wrote that teachers had always carried out formative assessment. For example, they recognised learning difficulties and corrected spellings and other errors as they moved around the classroom and through marking pupils' work. He suggested, therefore, that formative assessment had always been part of the repertoire of a good teacher, but what was lacking was a systematic approach to it. It has received much less time and attention from teachers, school managers and policy makers compared to summative assessment which was formal and systematic. These observations about the lack of attention to formative

assessment were confirmed by other authors in subsequent years. For example, Paul Black, 1993 and McCallum et al., 1993, reported the existence of some informal formative assessment but confirmed the dominance of summative assessment. In 1995 Daugherty (Daugherty, 1995) noted that the level of resources given to supporting formative assessment in the UK since 1988 had been ‘negligible’.

Since these observations were made there have been further theorising and empirical studies on formative assessment, and since 1999, on assessment for learning (AfL). Authors such as Lambert and Lines, 2000; Earl, 2003; Black and Harrison, 2004; Clarke, 2005; Gardiner, 2006; Broadfoot, 2007; Stobart, 2008; Spendlove, 2009 and Wiliam, 2011 and 2012, have all provided further expositions and explanations of the two concepts, FA and AfL. In addition, the National Strategies in England and Wales (DfES, 2004) have provided professional development programmes on AfL to many schools and teachers. There have also been reviews of research on summative assessment practices covering the period 1999 to 2013 (Moss, 2013).

Along with the above developments a debate about standards reached by British school children compared to their counterparts in competitor nations such as Germany has also been taking place. This debate, often described as the ‘standards debate’, was initially triggered by the publication of the ‘Black Papers’ (Gillard, 2011), then taken on by Prime Minister, James Callaghan, in his Ruskin College speech in 1976. Nearly a decade later, the concern about standards led to the Education Reform Act (ERA) of 1988. The National Curriculum and the associated



assessment arrangements were imposed in pursuance of this Act by the Conservative government of the day in the face of strong opposition from teachers (Simon, 1988).

During the same period in which the above developments were taking place, the publication of the TGAT report (DES, 1987) and the Black and Wiliam review (Black, 1998) of international literature on the success of FA, brought into a sharper focus issues relating to assessment and the best ways of promoting learning and thus raising standards. The TGAT recommendation of FA as the basis for NC assessment system was made because the Group saw this as more likely than the traditional SA to promote learning and thereby raise standards.

The TGAT's recommendations were, in my view, a radical departure from the status quo. Previous government reports such as the Consultative Committee report of 1911, the Norwood report of 1943, the Beloe report of 1960, the Newsom report of 1963, and the Waddell report of 1978, had addressed issues relating to widening access to examinations for pupils, through better organisation and fine-tuning of the examination system, thus further improving the summative assessment procedures and processes. TGAT on the other hand, argued for a more radical change: formative before age 16 and summative at 16, with interim summative assessments made, when required, by aggregating the ongoing formative assessments.

However, as indicated earlier, the government favoured tests and the National Tests in English and mathematics at age 7, and in English, mathematics and science at ages

11 and 14, were introduced from 1991. In addition to these externally marked tests, teachers were required to assess pupils regularly using the National Curriculum Levels, [initially ten levels, reduced to eight plus exceptional performance in 1999 (DFEE/QCA, 1999b)], and produce a 'Teacher Assessment' (TA) for each pupil in their class at the end of each Key Stage, annually for reporting to parents and also whenever a parent requested it, at two weeks' notice. Following the imposition of the National Curriculum, a range of monitoring and accountability procedures were also put in place. The Office for Standards in Education (Ofsted) inspections of schools began in 1995 and the first League Table of results was published in 1996. Targets for schools and for Local Authorities were introduced in 2001 and the government's monitoring of school targets, through Local Authorities (LAs), began in 2002. Local Authority targets were monitored directly by the Department of Education and Science through 'Stocktake' meetings with LA officers. I participated in these meetings in my capacity as the science inspector for a Local Authority. By the end of 2002 the framework and procedures for the 'standards agenda' were fully in place.

The debates and the contexts described above have produced unease and tensions at almost every level of the school system. School managers as well as teachers have been engaged in managing processes emanating from national policies. Some of these tensions arise from the way in which the two traditions of assessment have developed historically and some are due to these more recent changes. These will be discussed in chapters 2 and 3. In the next section I consider the 'tensions'.

## **1.2 Tensions over issues relating to assessment**

The TGAT report created concerns both in Government circles and among the teacher Unions, but for different reasons. The government favoured tests. The teacher Unions feared the increased workload the TGAT recommendations could create for their members. The Teacher Assessment (TA) component of the national curriculum assessment was also a source of tension and confusion. Teachers lacked confidence in their ability to produce reliable TA and doubted whether TA would be accepted by the public at large as either valid or reliable (Fairbrother, Black and Gill, 1995). Teachers did not trust their own or their colleagues' assessment results. Furthermore, Torrance and Pryor commented in 1998 (Torrance and Pryor, 1998) that TA, rather than complementing the tests as they were supposed to do, 'simply replicated them'.

### **1.2.1 Teachers' professional learning**

The question of teachers' competence in carrying out assessment has been a long-standing issue and a key source of tension, linked to the need for professional learning (in-service training) for teachers. This issue has been highlighted from time to time by government reports and other authors and researchers. The Norwood report (1943) had argued for a seven year period of in-service training for teachers of secondary Grammar Schools so that they could be enabled to assess the 'whole child', i.e. the overall achievement of a pupil whilst she/he is at school. This, the report suggested, would remove the need for external examinations which, it said, assessed only a narrow range of achievements.

Black and Wiliam, in their 1998 review of international literature on formative assessment, stated that there was a lack of clarity in much of the literature about the distinction between formative and summative assessments and that this had hindered teachers' understanding of issues relating to formative assessment. The Assessment Reform Group's (ARG, 1999) section (1.1) redefinition of formative and summative assessments in 1999, as assessment **for** learning and assessment **of** learning respectively, was intended to clarify the distinction between SA and FA.

Unintentionally however, it created overlap and confusion between FA and AfL.

This is discussed in section 3.3.1. In the last decade, two significant projects have further drawn attention to the need for teachers' professional learning in order for them to take up the challenges of developing FA/AfL; the Kings Medway Oxford Formative Assessment Project (KMOFAP) (2002) and the Learning to Learn project (2006). More specifically to science, the 'Professional Knowledge Base of Science Teaching', edited by Corrigan et al., (Corrigan, Dillon and Gunstone, 2011), discusses issues relating to science teachers' professional knowledge needs in great detail.

The description of the KMOFAP project by Harrison (Harrison, 2005) indicates that teachers and researchers interactions were helpful in promoting teachers' learning. The project teachers also benefitted from interacting with each other and reflecting on their practice. Drawing on Schön's (Schön, 1991) ideas of 'a reflective practitioner' and Wenger's (Wenger, 1998) notion of 'communities of practitioners', Harrison suggests that the key to developing reflective formative assessment practices "lies in promoting 'communities of practitioners' that can function as a machinery to turn ideas into classroom events". The question of how to develop or

create ‘communities of practitioners’, beyond the KMOFAP experience, is not fully explored. It is not clear how such communities would be created and, if created, how they would function as ‘machinery to turn ideas into classroom events’ without needing the support of an enabling school structure which promoted such a goal or an external stimulus, as was the case with the KMOFAP teachers.

The Learning to Learn Project (James, 2007) emphasised the importance of teachers’ professional learning and suggested that it was a ‘condition’ for the development of Assessment for Learning. Other authors, notably Lambert and Lines (2000) and James and Pedder (James and Pedder, 2006) have also emphasised the need for teachers to engage in professional learning in order to be able to develop FA and AfL.

In relation to science, subject knowledge of science teachers has been raised in the literature as an issue which could impede formative assessment. A King’s College study of the professional development needs of science teachers in 2000 (Dhillon, 2000) noted that three quarters of the secondary school headteachers surveyed for the study thought that subject knowledge was the main weakness among the science teachers in their schools. Newly qualified teachers agreed that they were less confident outside their area of specialism and in teaching the ‘Scientific enquiry’ (Attainment Target 1) aspect of the science national curriculum (DES/WO, 1991). The Royal Society’s ‘state of the nation’ report (RS, 2007) gave the following percentages of subject specialism of science teachers in England, quoting from NFER’s survey-based projection, (NFER 2006, p.113):

<b>Subject</b>	<b>Percentage of teachers</b>
Biology	47.6
Chemistry	27.5
Physics	20.1

Teachers with other science qualifications accounted for 4.8 percent.

Although the above figures are nearly seven years old, they do provide a baseline for comparison of teachers' subject specialism in the case study schools. Whilst these percentages do not directly support what the headteachers had said about weak subject knowledge of science teachers, the imbalance of subject expertise shown by the data does lend credence to their view. What degree of confidence do the 52.4% of teachers without a physics or chemistry qualification enjoy while teaching the physics and chemistry aspects of the National Curriculum?

Paul Black et al. (Black et al., 2002, pp. 15-16) argue that 'choosing a good question', for example, would require a thorough knowledge of the subject. They also remark that such knowledge is not necessarily gained from advanced study in the subject but rather gained "through understanding of the fundamental principles of the subject, an understanding of the kinds of difficulties that pupils might have, ..."

They describe this as ‘pedagogical subject knowledge’ (PSK), and this too would require teachers to engage in professional learning/development. More recently, Jones and Cowie (Jones and Cowie, 2011) have drawn attention to the interplay of ‘knowledge and action’ stating that teachers need ‘deep understanding of the subject domain’ as well as ‘student learning pathways’ and ‘pedagogical practices’ conducive to learning, in order to ‘undertake assessment for learning’. Thus the questions of appropriate CPD for teachers and their active involvement with it remain **live issues** and are key focuses of this study.

This thesis argues that tensions exist at a fundamental conceptual level between formative and summative assessments. The idea of formative assessment is that of an assessment essentially focused on the needs of the learner, situated in the ‘child-centred’ approaches of the Victorian pioneers, Mason and Parkhurst (see section 3.2); summative assessment, which came into prominence in Europe in the 19<sup>th</sup> century, has a ‘grading-centred’ approach for grading labour power which eventually facilitates the placing of pupils into positions in the division of labour: unskilled working class, then semi-skilled and so on. This social stratification-oriented function has been dominant and the placing of pupils in rank order in schools remains prominent. These differences which existed at the inception of the two traditions of assessment are still very visible today.

### **1.2.2 Low expectations**

Teachers’ beliefs about learners and learning and their perspectives on AfL are central to this study and so are the ethos and values which form the bedrock of the

education system and provide the framework for school structures. In my experience of teaching science in four different London comprehensive schools in 1970s and 1980s, I found that the main function of assessment was to facilitate the ‘banding’ and ‘setting’ procedures. All four schools used a banding system based on ability in literacy. Three out of the four science departments in which I taught were able to create ‘mixed ability’ sets within the banding system, the purpose of which was social integration. Their pedagogy and expectations remained ability related. Three of the schools were in working class neighbourhoods whilst the fourth was in a prosperous middle class suburb. In the school in the middle class suburb, parents’ expectations were generally high. Whilst all four schools had similar summative assessment procedures, in the middle class school there was a greater emphasis on good record keeping which was intended to inform teachers, pupils and parents of academic achievement and effort levels. However, in practice, the records were mainly used for reporting to parents. Teachers were not able to make effective use of the records because the process of collecting marks and grades and recording them had become too time-consuming and cumbersome. Academic progress was reported but was often inaccurate. Many parents learnt about the true progress made by their children only when the time came to enter pupils for the external examinations, General Certificate of Examination (GCE) or Certificate for Secondary Education (CSE). The CSE was intended for those who could not qualify for the higher status GCE ‘O’ levels. These two were replaced by a single 16+ examinations, the General Certificate of Secondary Education (GCSE) in 1985. In the schools in working class areas parental expectations were low and parents seemed to take education on trust, a point which was noted a century earlier by the Acland report (1911) which stated



that there was a thriving market in education but most parents did not know what to look for in good quality education and were often let down by the education **market**.

In my experience of teaching in the four secondary schools, mentioned above, pupils' failure to achieve was explained by teachers and by the schools in terms of pupils' lack of ability. Pupils' failure was rarely attributed to low expectations or a lack of good teaching or a lack of good pastoral care or a combination of all three.

### **1.2.3 The tri-partite system of education**

Following the 1944 (Butler) Education Act the tri-partite system of schooling in England and Wales was introduced although the Act itself did not require it. [In fact, the system was never truly tripartite - technical schools were expensive to build and maintain so very few were ever opened (Gillard, 2011)]. It was the Spens report of 1938 which had recommended that there should be three types of secondary school:

- grammar schools for the academically able;
- technical schools for those with a practical bent;
- new 'modern' secondary schools for the rest;

The Norwood report of 1943 had strongly endorsed this recommendation. However, the origins of the tri-partite system lie in the works of Robert Morant at the Board of Education (1863-1920) and Sir Cyril Burt (1873–1971) who worked as an educational psychologist for the London County Council (LCC). Both Morant and Burt were influenced by Francis Galton's work on eugenics and they were pivotal in

shaping the state education system in England and Wales in the 20<sup>th</sup> Century (White, 2006, p. 18).

#### **1.2.4 The notion of ‘positive achievement’**

When GCSE was introduced in 1985 I was head of science at a school in a middle class area. With the GCSE came the notion of ‘positive achievement’. In his Sheffield speech at North of England Conference (Joseph, 1984) the then Secretary of State for Education, Sir Keith Joseph, had outlined the notion of an examination at 16+ which would merge the existing examinations, GCE ‘O’ level and CSE to produce a common examination which would use criterion referencing rather than the existing norm referencing – assessing what a pupil knows and can do as opposed to what the pupil does not know and cannot do. However, this notion did not extend to the marking of pupils’ work on an ongoing basis, recognising what they knew and identifying what they needed to learn next. For most teachers the notion of ‘positive achievement’ meant ‘positive’ comments aimed at improving motivation. These were helpful but fell short of constructive feedback. In-service training (INSET) courses were provided by the Examination Boards for implementing GCSE, the thrust of which was on managing the coursework element of the GCSE assessment. Four years after the introduction of the GCSE, the National Curriculum was imposed (1989), along with the assessment arrangements associated with it. The notion of ‘positive achievement’ was somehow subsumed in the notion of Teacher Assessment (a component of the NC Assessment in which teachers assess pupils and provide a National Curriculum level for each pupil) but the tension between the tests and

examinations which grade pupils and an assessment process that recognises and confirms ‘positive achievement’ of pupils remained unresolved.

The main difficulties and tensions, at the level of teachers’ practices and perspectives concerning assessment emerging from our discussions so far, are summarised below:

- The summative tradition is embedded in the English school system
- Teacher Assessment (TA) mimics national tests
- The government’s ‘standards agenda’ is coercive in nature
- The tripartite model of schooling has shaped educational thinking, institutional arrangements and policy and practice
- Teachers lack engagement with professional learning

There has been criticism of summative assessment for nearly 100 years, and since 1976 when Bloom and his colleagues used formative assessment in the context of ‘mastery learning’, the overwhelming weight of intellectual opinion has been in favour of formative assessment and opposed to the excessive use of summative assessment. So, the case for FA and AfL appears to have been made in terms of their efficacy for learning.

Indeed, there were criticisms of external examinations almost from the very beginning of their existence. For example, the Acland Committee of 1911 raised many concerns about the dominance of external examinations, and Norwood in 1943

discussed the ill effects of external examinations on teachers and pupils. Many educationists have argued against examinations and tests on the basis that they encourage 'shallow regurgitation' and do little to promote reasoned, disciplined thought (David, 1981). In addition, the ideology of selection on 'merit' also implied school 'failures', which suited the needs of industry for 'low-paid unskilled workers' (Broadfoot, 1986, p. 57). In short, the purpose of summative assessment was to facilitate selection of this kind, which was unacceptable to most educationists (Black and Broadfoot, 1982). They argued that there was a fundamental contradiction between formative assessment which must be a central part of teaching and learning and the main purpose of summative assessment which was to provide for selection.

In addition to the afore-mentioned advocacy by authors and researchers and the strong arguments they have made, what has motivated me most is the knowledge that learning depends upon high expectations, encouragement, and a belief in pupils' capacity to learn if they are given what Bloom described, i.e. good, sensitive teaching (Bloom 1976). The role of assessment cannot continue to be that of sorting and classifying pupils on the basis of spurious scientific and bureaucratic procedures, particularly given that so because I have witnessed in my professional career the opportunities denied to pupils to learn good science using such sorting procedures.

Many teachers had welcomed the TGAT idea of formative assessment forming the basis for assessment in the National Curriculum and, like me, were enthusiastic about implementing it. However, more than two decades after the publication of the TGAT report it is clear that despite the rhetoric of formative assessment and

assessment for learning, actual assessment practice has remained resistant to change. Why is this so? One answer no doubt would be a lack of knowledge about FA. So, is CPD the main answer or are there other issues which need to be explored? These questions contribute to the rationale for my study.

### **1.3 The rationale for this study**

Whilst the above mentioned discourses were taking place with increasing intensity the voice of teachers in this debate was conspicuous by its absence. The discourse amongst academics appeared to have addressed the policy makers, but teachers and teacher unions had been less involved. The opposition by teacher unions to the TGAT proposals, citing increased workload, was one of the reasons for the ease with which the Government of the day had dismissed the TGAT recommendations on FA. Since then the teacher unions have been relatively quiet on the subject. So the question arises as to how have teachers engaged with the debate and how have they interpreted the discourses on FA and AfL?

Another issue is the way science teachers and schools have routinely used the results of assessment to create ability groups. Earlier studies, notably of Ferriman et al (Ferriman, Lock and Soares, 1994, p. 17) and Swain (Swain, 1995, p. 126), had reported the dominance of end of module tests in science in Key Stage 3 (ages 11 to 14). Daws and Singh (Daws and Singh, 1996, p. 97) found that the main use of the results of the tests was for setting according to ability.

Thus, although the notion of inherited intelligence has not often featured in discussions about assessment in schools, the notion of ‘**ability**’ has. The question therefore is how does this notion of ‘ability’ operate in the classroom and at the whole school level, and with what consequences?

The third issue emerging from literature on FA and AfL is the need for teachers to engage in appropriate **professional learning**. Government reports and academic research such as Norwood (1943), Dhillon (Dhillon, 2000), KMOFAP (Harrison, 2001) all highlighted this. The KMOFAP findings (2001) suggested that, given a supportive environment including professional development opportunities, teachers are able to engage in developing assessment practices which support learning. My study investigates the extent to which teachers have engaged with professional learning, and how it is initiated and organised at school level. Successful research projects in the field of school assessment tell us something about this. The KMOFAP (2001) project highlighted important issues, particularly in relation to teachers’ engagement with CPD programmes. KMOFAP was planned and implemented in collaboration with teachers and researchers from King’s College, London. This was a CPD model in which research evidence and contemporary practice met face to face and teachers tried to implement the research findings on various aspects of formative assessment. The researchers found that for the project teachers this experience proved to be ‘transformative’.

The National Strategy had provided CPD programmes in secondary science from 2002 to 2008. The Strategy made extensive use of research evidence to provide the latest information on Assessment for Learning, coming up with a model of CPD that was different from that of KMOFAP. Here, a cascade model led to CPD provided by local authority consultants to teachers, who opted in or were nominated by their schools to take part. Not all schools and not all teachers of the participating schools took part however, and this was a significant weakness in the Strategy's approach. Other absences included any reference to FA in the Strategy's CPD programme and having produced a CPD module on 'Assessment for Teaching' (AfT) during the 'pilot' phase of the science programme, it excluded it from the national rollout.

Thus despite the overwhelming educational case in favour of FA and AfL and despite the efforts of the National Strategy to help develop AfL, there is overwhelming evidence, discussed in chapter 3, that FA and AfL remain at the periphery of practice.

In view of these findings, and bearing in mind HD Black's observations (Nuttall, 1986) nearly three decades ago that there was a dearth of research evidence of what teachers actually did in relation to assessment and how they used the results of any assessment, there was a need for observational research on classroom practice and the issues which impact upon it. My study is a response to this need. Previous studies (Ferriman, 1994; Fairbrother, 1995) conducted questionnaire surveys of assessment practices and teachers' attitudes to assessment. These studies had partially answered Black's quest for what assessments teachers carried out and what use they made of

the results, but did not claim to have observed actual classroom practices. KMOFAP was a collaborative project between teachers and researchers aimed at improving practice by engaging teachers in dialogue at the interface of their existing practice and research evidence of how learning is enhanced through AfL. KMOFAP was not an observational study.

As an Ofsted inspector until December 2005, I saw only what was available during the inspection week of a school inspection. I could therefore not claim to have a good knowledge of what pupils actually received on a regular basis with regard to marking and feedback, but only a snapshot gained by looking at lessons or scrutinising pupils' books during the week. Also, inspections took place in the context of a superficial setting in which the inspectors saw only a partial picture of what was really happening in the day- to- day life of the school. At this point in time a decade had elapsed since the studies of Ferriman (1994) and Fairbrother (1995) and considerable changes had taken place. However, the need for an observational study of classroom practices focusing particularly on the nature of assessment and the quality of feedback had remained unmet. How had these changes impacted on the classroom and how had teachers responded to the constraints and support they experienced? In other words, how had the **tensions** described earlier played out in real schools and real classrooms?

My research thus aimed to explore the tensions relating to the implementation of formative assessment by making the classroom, the teachers, the pupils and the school managers the subjects of my study. Observations were focused on actual



classroom practices in real school settings. The practice of science teachers was observed as they went about their normal work on teaching and assessment. What they did in relation to assessment, how they interpreted what they did, the use they made of the results of the assessment and the extent to which they engaged with the debate on formative assessment/assessment for learning became the focuses of my study.

The study also examined the expectations placed on teachers by school managers in relation to FA and AfL and investigated how they responded to these expectations. In this respect, I have examined especially how the dynamics of agency and structure operated in the schools in relation to the desirability of changing the embedded culture of summative assessment in favour of a newer and superior approach, i.e. FA/AfL, but retaining the **essential** functions of SA for reporting and certification. Also examined were the prevailing discourses in the school to discover what teachers and school managers valued and promoted.

#### 1.4 Research Questions and Research Design

Drawing on the discussion in the previous sections the principal research question for the main study is:

*‘Why has the use of FA in secondary school science remained at the periphery of classroom practice?’*

The subsidiary questions are given in chapter 4, (Methodology), section 4.4. For example,

- What assessments do teachers carry out and why?

- How do these assessments impact on pupils?
- What do school managers expect teachers to do in relation to assessment?

I have adopted a qualitative approach for addressing these questions. The main study has been informed by insights gained through an exploratory study which used interviews with teachers as the main method of data collection in the first phase, and book scrutiny of marking and written feedback in the second phase. Details are given in Chapter 4.

The research design is intended to provide both a descriptive picture of assessment in practice and an explanation for the lack of FA/AfL. The analytical lenses utilised in the analysis and developed in chapter 4 were also taken into account in the design of the study.

The main study is divided into four phases:

#### **1.4.1 Main study Phase 1: 2005 to 2007 – National Tests at age 14 in place**

This was conducted in one science department at Curie High School (CHS), a large comprehensive school housed in buildings over half a century old. The school is graded by Ofsted as ‘good’. The focus of my fieldwork was Key Stage 3 (Years 7, 8 and 9)

#### **Methods and participants:**

Interviews with:

- Two science teachers
- Heads of Science Department
- Deputy Headteachers

Scrutiny of:

- Pupils' books
- School and departmental policy documents

Lesson observations

#### **1.4.2 Main study Phase 2: 2007 to 2008 – National Tests at age 14 in place**

In this phase my fieldwork involved two science departments, the science department at Curie High School and that at Downton Community School (DCS). It was focused on Year 8 only.

#### **Methods and participants:**

Interviews with:

- Two science teachers from each school
- Heads of Science Department
- Deputy Headteachers

Scrutiny of:

- Pupils' books
- School and departmental policy documents

Lesson observations

### **1.4.3 Main Study Phase 3: 2009 to 2011 – post National Tests at age 14**

This phase of the study involved three science departments, with the science department at Westfield Comprehensive School (WCS) being added to the sample. WCS is a large comprehensive school with similar age buildings to CHS. It had been graded as ‘satisfactory’ by Ofsted, just one grade above ‘inadequate’ which makes the school subject to severe scrutiny and could lead to ‘special measures’, and in extreme cases the eventual dismissal of the headteacher and even the closure of the school. The focus of fieldwork in this phase was Year 8. However, by this point the National Tests at age 14 had been withdrawn and consequently the attention of the school managers had shifted to Key Stage 4 (age 14 - 16).

#### **Methods and participants:**

Interviews with:

- Two Year 8 teachers from each school
- Three pupils from each of their Year 8 classes
- Heads of department
- Deputy Headteachers

Scrutiny of:

- Pupils’ books (Year 8)
- Policy documents
- Lesson observations

### **1.4.4 Main Study phase 4: 2011 to 2012 – post National Tests at age 14**

**The main focus in this phase was CHS**

Interviews with:

- three teachers who were under ‘extreme’ scrutiny by school managers
- the Deputy Headteacher
- the Head of Department
- the BTEC science co-ordinator
- the Key Stage 4 GCSE co-ordinator

### **1.5 Methods of data analysis**

In discussing theoretical frameworks for analysis I have used Bhaskar’s ‘transformational model of social activity’ to situate and explain the relation of structure and agency and their dynamics, in the context of understanding the reasons for the embeddedness of the culture of summative assessment, and the possibilities of transforming it in the direction of FA/AfL. I have also used the notion of a ‘laminated system’, especially as developed by Gordon Brown (Brown, 2009), to consider the multiplicity of levels of determination accounting for this embeddedness of SA. Foucault’s notion of ‘gaze’ is also utilised to analyse how policy is misdirected. This is qualified by considerations of dialectic control between structure and agency drawn from Giddens’ theory of structuration (Giddens, 1984/2003).

I have drawn on Gillborn and Youdell’s (Gillborn and Youdell, 2000) notion of ‘Educational Triage’ to analyse how teachers might be involved in this process. They explain that in a medical emergency a ‘triage’ is used to help save those who have

the best chance of survival, leaving the ‘less hopeful cases’ potentially to die. They state:

“In school, educational triage is acting systematically to neglect certain pupils while directing additional resources to those deemed most likely to benefit (in terms of the externally judged standards). These strategies seek to maximise the effectiveness of scarce resources but their effect, in practice, is to privilege particular groups of pupils marked especially by social class and ‘race’ (Gillborn and Youdell, 2000, p. 134).

Finally, I have drawn on Stobart’s notion of ‘assessment shaping how we see ourselves and how we learn’ (Stobart, 2008, p. 1) to analyse the impact of ‘ability’ grouping on pupils.

Written feedback in the form of marking is analysed using a pro-forma drawn from the literature (Appendix A1). Interview data and observation notes are analysed using coding procedures. Episodes of dialogue in lessons have been analysed using categories drawn from my own experience and also from literature on ‘questioning’ and ‘dialogue’. The analysis of interviews with pupils focuses on their perceptions of the feedback they have received.

### **1.5.1 Lenses used for analysis**

As stated, I have used Bhaskar’s notion of ‘transformational model of social activity’ (TMSA) to analyse the potential of teachers to bring about the transformation of assessment from the embedded summative assessment culture to a newer FA/AfL culture whilst retaining the use of SA for the purposes of reporting and certification.

This is developed in chapter five. One of the aims of this study is to listen to the teachers' voice on the topic of assessment, including the debate on FA and AfL. At the national level the 'standards' agenda has been dominant. This has been influenced by global economic competition as Ball and Scott point out (Ball, 2008; Scott, 2011). At the school level, schools are subject to national requirements and at the classroom level teachers are subject to stringent accountability procedures. Yet it is inconceivable that any meaningful change in classroom practice can occur without the **active agency** of teachers. Thus, the ways in which schools' structures, considered in relation to an expanded concept of the learning environment, facilitate or hinder teachers' actions and responses and the ways in which teachers as agents interact with the school structure have become part of the study.

It is often suggested that given the strong accountability culture of the last two decades teachers have become compliant with the pattern of expectation set by this culture. The nature of this compliance and how the agency/structure dynamic influences this is studied using the transformational model of social activity and the notion of duality of structure. Also considered are the prevailing discourses in the school.

### **1.5.2 The dialectic of control**

According to both the TMSA and Giddens' structuration theory, human social activities are 'recursive' (Giddens, 1986 p.3). It is argued that to be a human being is to be a 'purposive agent' who has reasons for the activities he or she engages in and is able to 'elaborate discursively' about these reasons.

Giddens' theory further states (ibid, p. 16) that power within an established social system assumes 'regularised' relations of 'autonomy' and 'dependence' between agents or collectives in the context of social interaction. The theory claims that all forms of dependence offer some ways in which those who are 'subordinate' can influence the 'activities of their superiors' so that there is in principle always a two-way interaction in any social system. This is called 'the dialectic of control'. Critical realism would agree with this conclusion. It would also argue that dependence does not necessarily mean subordination.

### 1.5.3 The power of discourse

Foucault's notion of the power of discourse is utilised to study the discourses prevalent in the school and how these influence the mechanism of structure/agency, in particular so as to help/hinder the progress of any pedagogical change/improvement in the classroom. Foucault argues that power is exercised through the development of discourse/s, and equally that discourses contribute to power. Critical Realism sharply distinguishes between **power<sub>1</sub>**, i.e. power as transformative capacity and **power<sub>2</sub>**, i.e. power as oppression. In a relationship of oppression the oppressed will always have some transformative capacity - this is the basis of the 'dialectic of control'. Ideally, the oppressed can augment their transformative capacity until they can throw off the relation of oppression. This notion is utilised in the discussion on the capacity and the potential of teachers as agents of transformation to largely transform SA to FA.



## **1.6 Consideration of ethical issues**

In conducting the fieldwork for this study I have observed the BERA guidelines on ethical issues (BERA, 2004). A key consideration was my relationship with the case study schools. As the local authority's science inspector for more than ten years I have known the schools in that capacity and I am conscious of this role being in possible conflict with my role as a researcher. I have been very alert to this potential conflict throughout my study and have endeavoured to ensure that the two roles are clearly explained to participating teachers in the case study schools both orally and in writing. I have emphasised at every meeting that participation of the teachers, pupils and indeed the school was entirely voluntary and they could withdraw from the research at any time without giving me reasons.

I have continued to emphasise that the conversations, the interviews, the book scrutiny and the lesson observations and the results would be used solely for my MPhil/PhD research purposes and publications arising from it. Names of participants or schools will remain confidential and, if needed, only used with prior written permission.

I have also explained, in relation to lesson observations, that if there were issues pertaining to child protection observed by me then I would be obliged to inform the relevant authority after discussing with the teacher involved.

My interviews with pupils have been group interviews and this has facilitated collection of good data as pupils have responded to questions in a cooperative style,

listening to each other and considering their responses thoughtfully. The group dynamic has also encouraged more confident responses avoiding any tensions which could arise in a one-to-one interview.

Permission to conduct research has been obtained from the head teachers and written requests for participation in the research have been made to all participants. Prior agreements have been reached with the teachers. For interviewing pupils, prior agreements have been obtained from them, their teachers, and their parents (Appendix A0).

### **1.7 What the study hopes to achieve**

By studying current practice at the classroom level, school level and national policy level it is hoped that the thesis will throw new light on the nature of the barriers to FA and AfL and suggest ways to overcome them. I thus hope to extend our understanding of the challenges and difficulties which continue to exist in developing and embedding Formative Assessment in school science.

### **1.8 The structure of the thesis**

In Chapter two I discuss summative assessment. In Chapter three I discuss formative assessment/Assessment for Learning and explore the tensions that arise in trying to implement them in the context of an embedded culture of summative assessment. In chapter four I discuss the methodology which underpins this research. Bhaskar's notion of TMSA is discussed in chapter five, along with Giddens' notion of 'structuration', Foucault's notion of dominant discourse, the notion of 'educational triage' and Stobart's notion of 'shaping' the learner. Chapter six presents a description of a new theory of assessment which attempts to build on the ideas of

FA, diagnostic assessment and scaffolding. The data and its analysis are given in Chapter seven. Chapter eight presents the findings and Chapter nine discusses conclusions. This is followed by the bibliography and appendices.

## **1.9 Summary of Chapter 1**

This chapter opens with a review of the summative/formative assessment debate in the midst of a progressively coercive ‘standards agenda’. The first phase of this agenda was ushered in by the ‘black paper writers’. The second phase was started by Labour Prime Minister James Callaghan with his Ruskin College speech in 1976, and by 2002 all aspects of the ‘standards agenda’ were in place. The severity with which this agenda was pursued all but drowned out the pedagogical debate about SA and FA. However, initially the TGAT report (1987) then Black and Wiliam’s review (Black, 1998) and then in 1999 the publication of *Assessment for Learning* (ARG, 1999) kept the debate alive.

Given the clear evidence that FA/AfL promotes learning, the chapter introduces the aims of this thesis which are to understand the nature of the challenges that still exist in attempting to develop and use FA in science classrooms. Teachers are well disposed to the ideas of FA but SA remains dominant with FA still at the periphery of practice. What teachers do in the name of assessment, how they use the results of these assessments, the extent to which teachers engage in professional learning and the extent to which the notion of ability influences classroom activities are the key questions explored as the thesis investigates the impediments to the use of FA/AfL through an observational study of classroom practice.

## Chapter 2 Summative Assessment

### 2.1 Introduction

In chapter one I have described external examinations and tests as the main instruments of summative assessment and argued that examinations are embedded in the school system in England. The argument in chapter one is that the tests and examinations have prospered over many years because they are perceived by stakeholders to be reliable and relatively cheap measures of the outcomes of schooling (section 1.1). In this chapter I describe the origins of examinations and discuss how, from an arguably emancipatory beginning, they have become parts of a variety of instruments for rationing education. An alternative to external examinations, i.e. assessment by teachers of their pupils' 'whole achievement' during the period of compulsory schooling has been discussed at least since 1911 (Acland, 1911) but remains problematic because teachers lack the skills and competencies required to produce, accurate and reliable assessment which can command public confidence. Finally I discuss issues, both weaknesses and potential, for teachers' summative assessment practices.

Looking at the history of the development of assessment in England I describe here how examinations came into being to facilitate '**selection on merit**' and were campaigned for and even demanded in order to promote fair competition and to give parents an accurate picture of the progress their children were making.

## **2.2 A brief history**

### **2.2.1 From ‘patronage’ to ‘competition’**

Historically, examinations have been seen as fair means of selecting candidates for high public office. In China ‘written public examinations were used to select candidates for public service as far back as the beginning of our Christian era’ (Beloe, 1960). Here in England the report of the Northcote-Trevelyan Commission of 1853 on entry to the British civil service ushered in the age of ‘selection on merit’. This report by Sir Stafford H. Northcote and Sir Charles E. Trevelyan proposed a Central Board of Examiners to conduct competitive examinations for selection to the civil service. It also recommended the abolition of ‘nomination and patronage’ which had been the main means of securing entry to the civil service or to an officer grade in the armed forces. The report was accepted by the government of the day, and led to the introduction of a competitive examination for entry to the civil service.

At about the same time universities set up examination boards to conduct examinations, the first of which was the Oxford Delegacy of Local Examinations, established in 1858. Cambridge followed suit soon after and the University of London introduced entrance examinations open to men in 1858; women were not admitted until 1878. The military colleges of Sandhurst and Woolwich introduced entrance examinations in the 1870s. These changes had their effect on schools. The curriculum of many schools began to be influenced by the requirements of the civil service and the army examinations, and the London University’s ‘Matriculation Examinations’. The purposes of these examinations were both social and educational; they began to influence the school syllabus and they helped regulate

competition and selection based on merit, thus facilitating a shift away from the old system of patronage.

### **2.2.2 Examinations to ensure quality and report progress**

During the middle of the nineteenth century there was also a growing demand for external examinations linked to the quality assurance of the educational provision. The number of private schools had increased rapidly to meet the needs of the newly emerging middle class (Gipps, 1990) and there were also individuals who were offering education. Parents, however, were often unaware of what to look for and what was of worth. J. E. Rogers, writing in 'Education in Oxford' (Acland, 1911) makes the point that many of the advertisements about the quality of education being offered by individuals and by institutions were deluding people. He argued that it was difficult for a lay person to judge the quality of education on offer. They had to take the education available on trust, and in so doing they were frequently deceived. In the early 19<sup>th</sup> century the founders of the Local Examination movement argued that there was a need for public examinations which would enable parents to judge the merit of the education they paid for. This argument provided a strong rationale for external examinations and the first public examination which was offered by the College of Preceptor's in 1853 had the following laudable purposes (Acland, 1911):

“...to test progress, and afford at once, to the teacher and the pupils, a satisfactory criterion of the value of the instruction they receive.”

There is evidence too that the elementary school curriculum was also moving in an encouraging direction (White, 2006, p. 7). White describes how the 'post-1870

system was rapidly moving beyond the basics'. Pupils were 'able to study more advanced subjects' including vocational and practical subjects in "higher grade schools". He argues that if this trend had continued, 'this pressure' from the school level would probably have influenced the policy makers and as a result all pupils could have had opportunities to learn 'as far as they wished'. White states,

"All young people could have been encouraged to go on learning as far as they wished, whether in practical or vocational directions or for more intrinsic reasons. Once one level of learning had been attained, the way could have been open to move on to another." (White, 2006, page 7).

Thus it was theoretically possible that the English education system could have moved in this direction after 1900. White explains that during the 20<sup>th</sup> century 'short-lived attempts' were made to move the system in this direction but nothing of significance came of it because, as John White puts it:

"But such a policy required governments committed to furthering the education of every child and the commitment to fund it. These were rare" (ibid p 7).

Instead, mass schooling took a very different direction. First in the U.S (Earl, 2003, p.3) and then in England, the organisation of schools was heavily influenced by the industrial model of production, and assessment and teaching were influenced by Francis Galton's (Galton, 1869) idea that '**intellectual ability is inherited**'.

### 2.2.3 The development of schooling in the USA

The development of mass schooling in the US was modelled on a factory assembly line. Earl (Earl, 2003, p. 3) puts it as follows:

“For most of the 20<sup>th</sup> century, and even now, factories have dominated the economy, and schools organised around a factory model have been consistent with the world around them. Kindergarten-sized units of raw material are put into the first bench of the “plant” and sequentially moved through the “stations” (grades) on the assembly line. They spend a fixed amount of time at each one (a school year). If, at the end of the allotted time, they are not “done”, they are sorted into “streams” or “tracks” and moved to other parts of the building.”

Earl argues (Earl, 2013, p. 12) that this ‘metaphor for schools has largely fit the times’ and it has continued to be so because society has accepted that in the past the model had been successful for most pupils. Many left school at an early age but this was not seen as a problem. They had jobs to go to in ‘factories, mines, farming, fishing and the like’. Society had also accepted in the past that “education, beyond the basics, was a scarce resource, *necessary only for a few*” (ibid. P.12). (My emphasis).

There was opposition to this model of schooling by many thinkers and educators. Earl points out that Dewey argued as early as 1916 (Dewey, 1916) that there was a need for education to serve all pupils, pointing out that “education was necessary in order for society to continue to grow in democratic, social, and moral ways” (ibid). However, the dominant model of schooling continued to mimic the industrial production line.



Thus it was that teachers became quality control agents whose job it was to grade and sort pupils in order to decide who goes on to higher levels of schooling. This kind of sorting and selecting became a function of schooling (Stiggins, 2001) and in order to facilitate this process ‘**intelligence tests**’ were sought after and developed. The use of such tests by the U.S. Army, in selecting candidates for officer grades encouraged the development of a whole range of mental measurements in their image.

#### **2.2.4 The U.S. Army’s Alpha and Beta tests**

During the First World War, the U. S. Army felt the need for ‘a quick-to-administer’ intelligence test to select recruits for advanced training. Two prominent psychologists Lewis Terman and Robert Yerkes, in collaboration with others, developed the Alpha and Beta tests. The Alpha tests emphasised verbal abilities and were given to all recruits. The Beta tests emphasised non-verbal ability and were for those who failed the Alpha test.

Terman was a member of the Advisory Council of the American Eugenics Society from 1922 to 1935, and strongly supported the classification of school children according to ability. His views are vividly expressed in the following quote from his 1922 paper, quoted by White (2006, p. 24):

“... all the available facts that science has to offer support the Galtonian theory that mental abilities are chiefly a matter of original endowment ... It is to the highest 25 percent of our population, and more especially to the top 5 percent, that we must look for the production of leaders who will advance science, art, government, education, and social welfare generally ... The least intelligent 15 to 20 percent of our population ... are democracy’s ballast, not always useless but always a potential liability ...”

The spurious objectivity offered by the ‘intelligence tests’ encouraged the development of Scholastic Aptitude Tests (SATs), first introduced in 1926 (and later changed to Scholastic Assessment Test (SATs)) in the United States. These SATs were justified, in the words of Henry Chauncey, the first president of the Educational Testing Service (ETS), as follows:

“... to dispose of the existing, undemocratic American elite and replace it with a new one made up of brainy, elaborately trained, public- spirited people drawn from every section and every background” Earl (2003, p.7; 2013, p.13).

### **2.2.5 A ‘gate-keeping’ role for assessment**

The SATs were external tests and became the means by which the ‘gate-keepers’ exercised power and provided or denied access to higher education and ‘... gave the illusion, at least, of objective measurement’ (ibid, p.13). Teachers anxious to help their pupils gain access to higher education were soon drawn into this role of examinations and tests. It was not enough to use tests and examinations to report the progress pupils were making and for reporting achievement to parents, other teachers

and the pupils themselves; the role of assessment in sorting students in different ‘tracks’ became important.

### **2.2.6 The ‘gate-keeping’ role challenged**

This role remained dominant for over four decades but was challenged in a serious manner in 1971. Bloom, Hastings, and Madaus wrote a book, *‘Handbook on the formative and summative evaluation of student learning’* (Bloom, Hastings and Madaus, 1971) in which they challenged the prevailing practice and argued that the ‘primary purpose of schooling was to develop the individual’. They emphasised that teachers could play a very important role in using evaluation to ‘improve and extend students’ learning’.

My argument here is that there have been at least two strands of thinking from the very beginning of schooling: the strand which argues that assessment should focus on improving and extending pupils’ learning and the strand which emphasises the ‘sorting’ and ‘selecting’ role of assessment. Galtonian ideas have been very influential in promoting the latter strand of thinking.

### **2.2.7 The influence of Galtonian ideas in England**

As I have argued in the previous chapter, section 1.2.3, one of the persistent threads of thinking running through the English education system is the notion of *innate ability* - an idea which appears to have become ingrained in educational thinking (Broadfoot, 1996; Lambert and Lines, 2000; Stobart, 2008). Despite the

establishment of comprehensive schools in 1967 (Circular 10/65) the notion of ‘ability’, enshrined in the tri-partite system of schooling, still prevails. Government spokespersons and documents continue to use the word ‘ability’ even when it is redundant. The Secretary of State for Education, Sir Keith Joseph, for example, referred to the ‘whole ability range’ in his Sheffield speech in January 1984 when talking about ‘positive achievement’. The words ‘all pupils’ instead of ‘whole ability range’ would have sufficed. The idea of ‘positive achievement’ also remained at a purely rhetorical level.

#### **2.2.8 The GCSE and the idea of ‘positive achievement’**

In June 1984 the Government announced its decision to introduce the GCSE (DES and WO, 1985). The objectives for these examinations were set out in the Government White paper ‘Better Schools’ (DES, 1988a), and the intention was to raise standards by recognising the ‘positive achievement’ of pupils. As mentioned earlier, Sir Keith Joseph, in his speech in Sheffield to the Northern Education Association in January 1984, talked about the need to raise standards and levels of achievement in secondary schools and the role of examinations in achieving this. He stated that the GCSE examination was to be open to the ‘whole ability’ range so that candidates were to achieve their grades by ‘positive achievement’ (Kingdon and Stobart, 1988). This contrasted sharply with the GCE and CSE where pupils gained their grades by ‘failing to achieve a higher grade’.

It was expected that the GCSE would help raise standards by using a variety of different methods of assessment, including oral assessment in English and Modern Foreign Languages and assessment of practical work in science, design technology and home economics. GCSE was thus intended to be a clear departure from the previous examinations and was a culmination of work done over a period of twenty years by many in the field of education. As the HMI report (DES, 1988b) stated in section 6;

“The GCSE is much more than just a new public examination. It embodies the aspirations, developed over two decades, of many educationalists ...”

Despite the enthusiasm for positive achievement in the above description, the GCSE failed to deliver these laudable aims. Kingdon and Stobart (Kingdon and Stobart, 1988 ) note that

“... those who actually developed the syllabuses and specimen papers took the more restricted approach of addressing the GCSE to the existing GCE and CSE candidature”.

This meant the GCSE, in practice, catered for the top 60% of the pupil population as was the case with the examinations it replaced. Soon, only grades A to C, the old ‘O’ level grades were considered as pass grades. The idea that most pupils (80 to 90%) would achieve ‘average’ standards has not happened in science.

My own enthusiasm for the GCSE was based on a belief that the new examination would remove the superior and inferior labels attached to the GCE ‘O’ level and CSE examinations respectively; and also, that the notion of ‘positive achievement’

would raise expectations of pupils in the minds of teachers and in the minds of the pupils themselves. As mentioned previously, when I began my teaching career it was in a mixed comprehensive school in a white working class suburb of London. Expectations were low. Teachers did not expect many pupils to gain GCE 'O' level passes. My first examination class was a 5<sup>th</sup> year (now called Year 11) physics group consisting of 14 fifteen year-old boys. I was told that they had not had a physics teacher in the 4<sup>th</sup> year (now called Year 10) and my task was to teach them some physics in the remaining two and a half terms before their GCE or CSE examinations. The pupils responded well and were eager to learn but were not sure that they would succeed despite my assurances. I too was not sure. At the end of the two and half terms of physics I entered them for the CSE examinations. They were not expecting to be entered for GCE 'O' levels and I was not confident enough to enter any for the GCE - as an untrained graduate in my first year of teaching I had only a limited knowledge of the GCE syllabus.

Seven pupils out of the 14 gained grade 1 (equivalent to GCE 'O' level grade C); the other seven gained grade 2. This was hailed as the best physics result in the history of the school. I was pleased with the results and enjoyed the recognition accorded to me as a good teacher. Years later when I reflected on the results I could not help feeling that the fact that the results were hailed as the best ever for the school was a symptom of the low expectations of the children who attended the school. The boys were expected to fail but they did not. I am sure that if they had attended a school with higher expectations and had been taught physics for two years, many would have achieved 'O' level pass grades.

When the National Curriculum (NC) in science was introduced in 1989, I opposed it, considering it to be inimical to teachers' professionalism. In time however, I welcomed it. I reasoned that the advent of the national curriculum was an opportunity for pupils to receive an 'entitlement' curriculum in which aspects of physics, chemistry, biology and earth science would be taught to all pupils, which was not the case prior to this point. Such a 'balanced science' curriculum was considered good practice but not all schools provided it prior to the introduction of the NC. I had assumed that adequate in-service training would be provided to enable teachers to teach all the four areas of science. However this did not happen, and at the same time the imposition of the national tests and associated accountability measures changed the dynamics of the discourse and, as mentioned earlier, set the scene at the start of the 21<sup>st</sup> century for a period of intense focus on external tests.

### **2.3 The National Tests**

The Education Reform Act (ERA) 1988 imposed a National Curriculum (NC) and associated assessment arrangements, in the teeth of strong opposition from teachers (Simon, 1988). After trials, National Tests were imposed and were phased in, between 1991 and 1993. As mentioned above I had begun to view the National Curriculum as the minimum 'entitlement curriculum' for all. Similarly, I began to see national tests at ages 11 and 14 as opportunities for providing accurate information about pupils' progress at these crucial ages which was not available prior to the introduction of the tests. Teachers' assessment of pupils in the early years of secondary school was notoriously unreliable. Most parents discovered the

lack of progress their sons and daughters had made when it came to entering them for external examinations at age 16 – only a minority were allowed to do the higher status GCE ‘O’ levels, the rest were entered for the lower status CSE (Certificate of Secondary Education). Yet, in the years before age 16, many parents had been told that their sons/daughters were making satisfactory/good progress.

I found the ten-point assessment scale – NC levels 1 to 10 - which TGAT had devised helpful. The ten levels were later revised down to eight plus an exceptional performance level (DFEE/QCA, 1999b). I considered them to be a good resource which could provide useful criteria for lesson planning, writing schemes of work and assessing pupils’ progress. However, I also saw the need for a sustained period of appropriate CPD for teachers in the use of the NC levels for constructing lesson criteria and for assessment. Training was provided with good success to primary school teachers, though in secondary schools training was limited.

### **2.3.1 Advantages and disadvantages of the national tests**

In my view, The National Tests in science at Key Stages 2 and 3 made a significant contribution to the improved coverage of science in both primary and secondary schools. The test papers covered all four Attainment Targets of the Science National Curriculum (DES/WO, 1991): Scientific investigation (Sc1), Life and living processes (Sc2), Materials and their properties (Sc3), and Physical processes (Sc4), roughly equivalent to practical laboratory work, biology, chemistry and physics respectively. This was a fuller and more comprehensive coverage of the science



curriculum at ages 11 and 14 than there had ever been in my twenty year experience of teaching science in London schools. The descriptions of Sc1 and Sc2 were later changed to 'Scientific enquiry' and 'Life processes and living things' (DFEE/QCA, 1999b).

The tests had also ushered in a period of intense in-service training (CPD) for primary school teachers through 20-day and 10-day courses, offered nationally. As a result the quality of teaching in science in primary schools improved considerably. There was no such intensity or demand in the secondary schools. There was no ring-fenced government funding for 20 or 10 day type courses for secondary school science teachers and there was a lack of CPD culture in secondary schools. A small minority of teachers were interested in CPD and some science teachers attended, for example, the Association of Science Education (ASE) annual and regional conferences to improve their knowledge and skills and discover new resources. However, most secondary school science teachers were ill disposed towards CPD or remained indifferent to it.

One of the disadvantages of the national tests was that they encouraged a dependency culture. Teachers relied on the tests in order to produce summative judgements and opportunities for improving skills in assessment, SA, FA or AfL, were missed or not taken.

The Key Stage 3 Science National Strategy launched in 2002 (later becoming part of the Secondary National Strategy in 2005) did provide some CPD to teachers.

However, whilst the contents of the CPD units were good (I was involved in teaching them) the programme was patchy. The training was available to local authorities and schools but the dissemination was based on a cascade model which did not happen in most schools in the Local Authority I was working for. This was because the teachers who attended the training were unable to cascade it to their colleagues because the schools had made no provision for this in their school timetable and were ill-prepared for it. The Strategy proved to be too short lived and too limited in scope to make a significant difference to the quality of teaching and assessment.

### **2.3.2 Accountability procedures – targets and league table of results**

Moreover, the accountability and monitoring instruments of the standards agenda, namely, the league tables of results, the target-setting programme, and Ofsted inspections, had taken hold of the education agenda and were fully in place by 2002. The collection and analysis of data (pupils' performance data in the national tests) had become a key priority for schools, as their effectiveness was measured mainly on the basis of the results of the tests. This further consolidated the already embedded summative assessment. Teachers became even more focused on producing 'test results'. There was no expectation that teachers would update their knowledge and skills continually, neither was much effort or money put in to facilitate this.

As mentioned above, target-setting for schools began in England and Wales in 2001. Teachers, schools and Local Education Authorities (LEAs) were put under a great deal of pressure to show year on year improvement in pupils' performance in the

national tests. One example was a circular (NAHT, 2001) from the National Association for Head Teachers (NAHT) to its members stating that the ‘Government’s unrelenting emphasis on target setting’ continued to ‘add to pressures on NAHT members’. The Government had proposed national targets for 2004 in English and mathematics at age 11 and for English, mathematics and science at age 14 as follows:

### **National Targets for 2004 in English, mathematics and science**

(Source: Department for Education and Skills, Standards and Effectiveness Unit, 18<sup>th</sup> September 2001 letter to Chief Education Officers)

<b>Age</b>	<b>EN</b>	<b>Ma</b>	<b>Sc</b>	<b>EN</b>	<b>Ma</b>	<b>Sc</b>
	<b>Level 4+</b>	<b>Level 4+</b>	<b>Level 4+</b>	<b>Level 5+</b>	<b>Level 5+</b>	<b>Level 5+</b>
11	85%	85%	none	35%	35%	none
14				75%	75%	70%

The Government’s Standards and Effectiveness Unit (SEU) attached great importance to ‘the process of target setting’. The following excerpt from the Department for Education and Skills (DfES) consultation paper (September, 2001, Re: DfES 0541/2001) on targets for English and mathematics at Key Stage 2 in 2004 made clear:

“As part of the Education Development Plan over the past three years, targets have been agreed between the department and all LEAs, and these in turn have guided LEAs in their discussions with individual schools...”

It continued,

“We believe that (this) process has generally worked well and should continue to lie at the heart of target setting for 2003 and 2004. Through our Education Advisers and Regional Directors, we will be discussing appropriate LEA level targets for 2004 with authorities in the first half of the Autumn Term 2001, as a basis for LEAs’ discussions with their schools.”

Accordingly, the Chief Education Officers of each LEA met his/her SEU Adviser and the Literacy, Numeracy and Science Regional Directors, in early October. Following these meetings LEAs were asked to “submit their provisional targets for 2003/4 for Key Stages 2, 3 and 4 by October 26<sup>th</sup> 2001 on the pro-forma provided.” (DfES Communications with LEAs, 18 September 2001).

Schools’ national test results, the national average scores and the national test results of schools with similar characteristics were sent to each school in the Autumn Term, in a document known as the Autumn Package. This was a DfES produced document containing ‘National Summary Results’ for each Key Stage. It also contained ‘National Value Added Information’ and ‘National Benchmark Information’ (Appendix A2). The ‘National Summary Results’ provided information which was intended to enable schools to compare their own performance with national results. The section on ‘National Value Added Information’ provided two tools: value-added line graphs and ‘Progress Charts’ in the form of bar graphs (Appendix A2ai). The

bar graphs were intended to ‘assist in setting realistic but challenging targets for individual pupils, groups or classes and the whole school’. These tools were intended to enable schools “to compare the progress of [their pupils] with progress achieved nationally taking into account prior performance” (DfES, 2001).

The National Benchmark Information was intended to enable schools “to make more detailed comparison of [their school’s] performance against national performance results and with other schools grouped according to one measure of pupil disadvantage – free school meals” (DfES, 2001).

Another document known as the PANDA (Performance and Assessment) report was sent to each school in late October of the same year. The Autumn Package was a national public document and could be obtained from the DfES website, [www.dfes.gov.uk](http://www.dfes.gov.uk). The PANDA was specific to a school, not a public document, and was Ofsted’s Performance and Assessment Report. It provided schools with information about the performance of their own pupils in relation to pupils in other ‘similar’ schools: schools that had similar proportions of pupils with free school meals and had a similar prior attainment pupil profile. It provided schools with appropriate benchmark tables and also provided “a broader overview of the school’s context, including information that is not used for benchmarking purposes (such as the proportion of pupils with special educational needs), but which may influence pupil performance.” The PANDA showed performance trends over time and comparisons of the performance of boys and girls separately, all against national indicators.

Both the Autumn Package and the PANDA were incorporated into a web-based database called RAISEonline ([www.raiseonline.org](http://www.raiseonline.org)) in 2006. RAISE stands for Reporting and Analysis for Improvement through School Self-Evaluation.

The information was intended to be used by schools to review the progress of their pupils, based on the results of the national tests during the previous year, and to set appropriate targets for the current year and the year after. Local Education Authorities had the responsibility of overseeing this process and ensuring that schools set sufficiently challenging targets, supposedly based on the prior attainment of pupils and national expectations. The average Level at the end of Year 6 (age 11) was set at Level 4. The average at the end of Key Stage 3 (age 14) was Level 5. A child achieving L4 at age 11 was expected to reach L5 if the progress he/she made was average. Moving up to Level 6 from Level 4 would be considered good progress by Ofsted Inspectors and was considered to be associated with good teaching.

Head teachers were expected to use the information from RAISEonline to analyse the performance of pupils in the national tests in English, mathematics and science and in GCSE 16 plus examinations. They were expected to evaluate the school's overall performance in relation to national averages and also in relation to the performance of schools with similar characteristics, both of which comparators were available through RAISEonline. In the light of their analysis targets were then agreed with the Local Authority (LA). Local Education Authorities were subsequently merged with Social Services following the Children Act of 2004, their statutory

responsibilities becoming part of the brief of every Local Authority Children's Services.

From 2007 schools were required to agree targets with their School Improvement Partner (SIP) who was considered to be an independent adviser but was employed and monitored by the Local Authority.

As described above, since 2002 every school had been required to set targets in English, mathematics and science for national tests at age 14. However, these tests (at age 14) were withdrawn in 2009 following the failure of the test agency to provide the 2008 test results on time.

National tests at age 14 were replaced by Teacher Assessment (TA) in 2009. However, there was a serious flaw in this move. The increasing dependency on the tests had meant that there had been little or no improvement in teachers' skills in developing summative assessment (TA) or formative assessment or assessment for learning over a period of ten years. Teachers had become dependent on the national tests for assessing or reporting pupils' progress. The already secure summative assessment culture in schools had been further strengthened as many teachers had been markers of national tests and had acquired expertise in constructing tests papers, not skills in TA. An unintended consequence of the withdrawal of the national tests was that once the tests were withdrawn schools began to **'mark up'** and inflate the attainment levels at age 14, indeed, turning the clock back to pre-national curriculum/national tests days, when neither the pupil, nor the parents nor

the teachers really knew what progress pupils were making in the four years prior to Year 11, the GCE/CSE year.

## 2.4 A timeline of key dates and developments since 1850:

**Table 2**

1853	-	Northcote-Trevelyan Commission
1858	-	Oxford Delegacy of Local Examinations
1911	-	Acland Committee Report
1938	-	Spens Report
1943	-	The Norwood Report
1944	-	The Butler Education Act
1960	-	The Beloe Report
1963	-	The Newsom Report
1965	-	Circular 10/65 Comprehensive Schools
1970	-	Circular 10/70 - Mrs Thatcher rescinds Circular 10/65
1976	-	James Callaghan's Ruskin College speech
1984	-	Sir Keith Joseph's Sheffield speech
1986	-	GCSE introduced
1987	-	TGAT report
1988	-	Education Reform Act (ERA)
1989	-	National Curriculum imposed
1991	-	National Curriculum revised
1991	-	National Record of Achievement introduced



- 1993 - National tests at 7, 11 and 14 introduced
- 1994 - The Dearing Review of NC and  
Assessment arrangements
- 1995 - Ofsted Inspections of schools begin
- 1996 - First National League table of results  
published
- 1999 - PANDA report produced
- 1999 - National Literacy Strategy introduced
- 2000 - National Numeracy Strategy introduced
- 2000 - The Autumn Package introduced
- 2001 - Targets introduced.
- 2002 - National Key Stage 3 Science Strategy  
launched
- 2002 - DfES monitoring of targets through begins
- 2005 - National Strategy extended to Key Stage 4
- 2006 - RAISEonline replaces the PANDA and the  
Autumn Package
- 2006 - A new Ofsted framework, short inspections,  
introduced
- 2008 - National Tests ‘fiasco’
- 2009 - National Tests at Key Stage 3 abandoned in  
England (13<sup>th</sup> January 2009)

- 2010 - New Con-Lib Coalition Government formed
- 2011 - National Strategy wound-up
- 2012 - New Chief Inspector of Schools, Michael  
Wilshaw, appointed
- 2012 - Circular 2012/03 School Development  
Planning and Target Setting

The time-line shows, over 150 years of developments in tests, examinations, curriculum and monitoring procedures.

As discussed in section 2.2.1 the Northcote-Trevelyan Commission of 1853 was a turning point in the concept of external examinations as guarantors of fair competition. The Universities soon established benchmarks for the standards expected at the end of secondary schooling. However, these were intended for the 'elite'. There was no systematic provision of secondary school education for the masses.

The Acland Committee report (1911) was a response to the demands of the middle classes for 'authentic' education amidst a 'chaotic' market in education. Despite its acknowledgement of the adverse effects of external examinations on teachers and pupils it recommended the School Certificate (SC) and Higher School Certificate (HSC) which helped bring order to the market. Norwood (1943) was vehemently against external examinations but considered that the teachers and schools were not skilled enough to provide reliable assessment of all the qualities of a pupil, the

‘whole child’, and recommended a seven year programme of in-service training for grammar school teachers and, in the interim, recommended the General Certificate of Education (GCE) Ordinary (O) level and Advanced (A) levels examinations.

However, the post-war government had a number of priorities and providing seven years of in-service training to teachers was not one of them. As mentioned in Chapter one the tripartite system of school education was set up following the Butler Education Act of 1944. This system of school education cemented both the class nature of the provision and the ‘inherited ability’ notion of access to it. The 11+ examinations ensured both.

The Beloe report (1960) was aimed at extending access to an external examination to those who did not have the ‘*ability*’ to do GCE ‘O’ level and recommended a Certificate of Secondary Education (CSE). It soon became clear that the CSE was an inferior examination and many parents began to reject it for their children. The Newsom report (Newsom, 1963) highlighted the plight of pupils in the middle and bottom bands in secondary schools. They were simply ‘neglected’.

#### **2.4. 1. A short-lived optimistic turn**

Circular 10/65 (DES, 1965) represented an optimistic turn and had come as a result of people’s campaign against the 11+. Full comprehensive education for all demanded by the campaigners for comprehensive schools, with Caroline Benn as one of the leading figures, was at last in sight. However, Thatcher’s Circular 10/70

(DES, 1970) put an end to this dream. It rescinded the requirement of Circular 10/65 that all local authorities were to 'plan' for conversion to comprehensive schools and also rescinded Circular 10/66 which had said that no funding would be given for new non-comprehensive schools.

James Callaghan's Ruskin College speech (1976) brought in the second phase of the 'standards agenda', the first phase being the 'Black Papers' which attacked comprehensive schools accusing them of 'lowering standards'. The ERA (1988) ushered in a more stringent phase, the third phase, with the imposition of the National Curriculum and national tests.

This time between 1989 and 2009 represents a period of consolidation of the instruments of the 'standards agenda' with a very repressive monitoring structure in place, paying only lip-service to improving the skills of teachers and improving the quality of teaching.

In the period since 2010, the actions of the new coalition government represent a further turn for the worse with the grip on the scrutiny of teachers tightening further, with coercive monitoring leading to 'bullying' and 'triage' of teachers, the marking up or making up of Key Stage 3 and GCSE results and with teachers' morale at rock bottom.

The time-line above is not exhaustive and is intended to give a brief picture of the key dates in national development of school education. It also gives markers to what

I have described in this thesis as the ‘misplaced gaze’ of successive policy makers which in my view has propelled school science into a cul-de-sac from which it needs to be pulled out. Arguably, similar issues bedevil the whole of school education but my thesis has only focused on school science.

## **2.5 Assessment in secondary schools – a complex picture**

In the midst of all this, assessment in secondary schools has become very complex. Elaborate arrangements are made in order to conduct examinations and complete Teacher Assessments (TA). Recording, analysing and disseminating the results of these assessments takes up a considerable amount of teachers’ time. As mentioned earlier, by 2002 all components of the standards agenda were in place. Tests and examinations began to take up a huge amount of pupils’ time. Writing in the Times Educational Supplement (Henry, 2002), Julie Henry stated that:

“pupils are losing more than a year of secondary teaching because of the time taken to prepare for and sit exams. ... [for example]... at William Farr School in Lincoln students are out of lessons for up to 46 weeks in the course of their seven-year secondary careers. The school year is 38 weeks.”

What I discovered during the course of my fieldwork was that the testing programme had intensified and schools were collecting larger and larger amounts of pupil performance data.

### 2.5.1 A crowded assessment space:

Can FA/AfL be accommodated in the already crowded space of assessment in schools? If it is to be so, what needs to happen? What needs to change?

Many secondary schools use Cognitive Assessment Tests (CATs) to assess pupils on entry to the secondary school which are a form of intelligence test (Gillborn and Youdell, 2000, p. 60). Some used the results of the national tests at Key Stage 2 (age 11) in English and mathematics. Others used both. Since the abolition of the National Tests the use of CATs has increased.

Many teachers have been markers for the national tests and have acquired considerable skills in their operation. They are able to transfer these skills to constructing and marking tests for their pupils. As Torrance (1993a) indicates, teachers use these skills to set more frequent tests and place pupils into even more **finely** graded categories *without changing their approach to teaching*.

Given the above context, how have teachers and schools responded to the views, expressed in the original Ofsted Handbook for inspecting secondary schools (Ofsted, 1999 p.44), and which has continually been encouraged, that teachers should ‘...assess pupil’s work thoroughly and use assessments to help and encourage pupils to overcome difficulties’?

Have teachers and schools the capacity to do so? What are the barriers that might prevent this from happening?

Lambert and Lines (2000) suggest the following barriers:

- perhaps the teachers' lack of knowledge of the evidence of success of FA provided by the Black and Wiliam Review;
- perhaps FA appears to be a daunting prospect;
- the external examination system's capacity for distorting teaching and learning;
- providing effective feedback is very demanding;
- teaching conceptualised as 'covering' the syllabus;
- increasing accountability of teachers and schools has encouraged teachers to rely on formal tests.

These points appear to be valid and summarise some of the issues I have raised in the earlier pages. I will return to them after the discussion of my fieldwork data in chapter 7.

The above discussion shows that the tension between the summative and formative purposes of assessment remains prominent in the work of schools. However, summative assessment too has come under scrutiny as it becomes clearer that teachers lack skills of developing and carrying out accurate and reliable summative assessment (Black, 2013) despite the fact that many teachers have been markers of the national tests.

### **2.5.2 Issues with the current state of summative assessment**

Two reviews of literature on summative assessment reveal ‘areas of need’ and also ‘areas of promise’ with regard to teachers’ summative assessment practices. Harlen’s review (Harlen, 2004) and a relatively recent review of research on summative assessment practices in classroom covering the period a twelve year period (1999 to 2011) and gathered from an Education Resources Information Centre (ERIC) (Moss, 2013) throw light on the current weaknesses (‘areas of need’) and potential for improved practice (‘areas of promise’).

The reviews conclude that although teachers have been testing more frequently and have analysed more test results (in England, the national test results when the tests were in place) and in the USA, in general many of them are ill-prepared and are ‘insufficiently skilled’. This leads to ‘inaccurate and unreliable’ summative assessment. The reviews also reveal that despite theses weaknesses and evidence to the contrary, teachers often report ‘positive beliefs’ about and a ‘high level of confidence in their assessment skills and competence’ (Rieg, 2007) and (Black et al., 2010).

### **2.5.3 Negative impact of tests**

A review of research on assessment **of** learning (summative assessment) carried out by Harlen and Deakin-Crick (ARG, 2002) discussed the negative impact of testing on pupils’ motivation to learn. The purpose of their review was to seek out evidence linking summative assessment and testing to motivation. A comprehensive search through published research found 183 studies, 19 of which provided ‘sound and valid empirical evidence’ of the negative impact of testing on pupils’ motivation. It also found that the negative impact could be reduced by ceasing to focus teaching on



test content, i.e. not teaching to the test, and by ending the practice of ‘training’ pupils in how to pass the tests. The reviewers argued that FA or AfL could raise pupils’ attainment without the adverse effects of testing by helping pupils to learn better. This would involve not teaching to the test and not engaging in repeated test-taking or test practice.

The two reviews, one by Black and Wiliam (1998) and the other by Harlen and Deakin-Crick (2002), together with the findings of Lambert and Lines (2000), ARG (2002), Stobart (2008) and Wiliam (2011) strongly suggest that pupils will be better motivated, learn better and achieve more if schools promoted formative assessment/assessment for learning practices and use summative assessment only when it was really necessary.

## **2.4 Summary of Chapter 2**

This chapter gives a brief history of the use of external examinations and tests and discusses how they facilitated ‘selection on merit’. Drawing on the developments in the USA the chapter discusses how the ‘Factory-Model’ of schooling seemed to be a logical concept given that the society was heavily influenced by this industrial model. In the industrially advanced economies pupils without a high school education or any recognised qualifications were able to get manual, semi-skilled and skilled jobs in industries such as ship-building, mines, steel mills etc. and were not

considered ‘failures’. It was accepted that education beyond the basics was a scarce resource restricted to a few.

The Galtonian ideas of eugenics were influential in the development of intelligence tests. The development of Alpha and Beta tests by two prominent Harvard psychologists Terman and Yerkes, for the use by US army for the selection of recruits for officer grade training, encouraged the use of such tests for selecting pupils for higher education. In England, the ideas of Sir Cyril Burt about the inherited nature of intelligence dominated and determined the direction of secondary school education.

In the next chapter, chapter 3 I present a literature review on formative assessment/assessment for learning.

## Chapter 3 Formative Assessment / Assessment for Learning

### 3.1 Introduction

This chapter considers the concept of formative assessment and assessment for learning as developed in the literature. In section two, I discuss the question ‘What is formative assessment?’, section three deals with similarities and differences between FA and AfL and section four considers the notions of assessment criteria and the importance of feedback. Section five presents the case for FA/AfL and in section six the role of CPD is reviewed. Section seven discusses the lack of FA in secondary school science and section eight summarises the key theoretical issues. Section nine introduces Enabling Assessment (EA), a new theory that synthesises existing ideas and builds on these to propose a more effective approach to classroom assessment. A brief summary of the chapter is given in section ten.

As we have seen in Chapter One the Task Group on Assessment and Testing (TGAT, 1987) renewed the debate in the UK on the desirability of formative assessment, and subsequently the Black and Wiliam Review (1998) provided evidence of the success of formative assessment. Coming towards the end of the 20<sup>th</sup> century these two documents set the scene for the current debate. In 1999 the Assessment Reform Group (ARG) presented a pamphlet entitled ‘Assessment **for** Learning’ (my emphasis) at the British Educational Research Association (BERA) conference, in which formative assessment was described as assessment **for** learning and summative assessment as assessment **of** learning and the characteristics of

assessment for learning were outlined. Since then further developments both in the research field and in the policy domain have taken place.

The UK Government's Key Stage 3 Science Strategy was launched in 2002 and was influenced by the findings of the Black and Wiliam Review. It included professional development modules for teachers on Assessment for Learning and Assessment for Teaching (AfT) in its science pilot programme (DfEE, 2001). After the pilot phase, only the AfL module was rolled out nationally. Subsequently, Assessment for Learning was also developed by both the Primary and the Secondary National Strategies in Continuous Professional Development (CPD) units for teachers.

So what, then, **is** formative assessment/assessment for learning?

### **3.2 What is formative assessment?**

The origins of formative assessment are linked to the child-centred education advocated by the mid-nineteenth century pioneers Charlotte Mason, a teacher in the middle-class schools of Victorian England, and Helen Parkhurst of Dalton, Massachusetts. Their efforts to develop child-centred education required a change in the nature of assessment and H. D. Black (Black, 1986) noted that this meant, among other things, that all written work was to be corrected in order to help teachers discover what children needed to learn. This, in my view, relates to the crux of the issues surrounding FA: teachers need to know the 'gap' in pupils' learning.

H. D. Black (1986) describes another example from the Victorian era which provides an illustration of an early attempt at criterion referencing. He describes a letter written by the Reverend George Fisher, Principal of Greenwich Hospital School, and quoted by E. Chadwick (Chadwick, 1864) which introduced the notion of **criteria for assessment**. Reverend Fisher described the existence of a ‘scale book’ in the school which provided examples of work at different levels of attainment. These levels were used as ‘fixed standards’ against which individual pupils’ work was compared. The ‘scale book’ provided examples of ‘standards’ in mathematics, navigation, scripture knowledge, grammar and composition, French, history, drawing and practical science – essentially in more or less the entire curriculum.

Thus so far we see that a purposeful application of formative assessment would require a learner-centred approach, marking of pupils’ work, and the use of an agreed ‘standard’ for assessing learning. One such approach was developed in the US in the nineteen seventies by Bloom and his colleagues (Bloom, Hastings and Madaus, 1971), who applied formative evaluation in the context of their espousal of ‘mastery learning’. They borrowed the term ‘formative evaluation’ from Scriven’s use of it in the context of curriculum improvement (Scriven, 1967). They argued that formative evaluation was useful not only for curriculum construction but also for ‘instruction’ and pupils’ learning. They proposed that formative evaluations should be directed at providing information which the teacher would use to ‘alter instruction’ or review those ideas and themes on which pupils were having ‘great difficulties’. So, formative evaluations were to provide information for the benefit of teachers too. This is an important point for my thesis which I return to later in Chapter 6.

Bloom's notion of formative evaluation was in the context of mastery learning and the belief that all learners could achieve 'mastery' (in essence, 'deep' learning in the subject). This would happen if the teaching was 'systematic' and carried out 'sensitively', if pupils were helped when they experienced difficulties and given sufficient time to achieve mastery, and if there were clear criteria of what constituted mastery.

Thus, the following five components of formative assessment emerged from the pioneers of this approach:

Formative evaluation/assessment should:

- be learner-centred – identify the 'difficulties' children were experiencing
- have agreed standards or criteria for judging achievement and for feedback
- be an integral part of systematic and sensitive teaching
- lead to information for teachers' action to 'alter instruction'
- be based on a belief that all learners can in principle achieve learning goals

Bloom emphasised the importance of '*systematic*' and '*sensitive*' teaching and in his view, formative evaluation was integral to this.

Although there wasn't a programme in the UK similar to that of Bloom and his colleagues in the USA, there were authors who advocated similar ideas to those of Bloom. In 1986, Lloyd-Jones argued that formative assessment was 'integral with learning' and that the information gained was to provide feedback to both pupils and the teacher (Lloyd-Jones et al., 1986, p. 2). He suggested that assessment should

provide information on areas of both weakness and strength. For learners it was to provide encouragement and motivation and for the teacher it was to provide continual feedback on whether or not ‘teaching/learning’ objectives were being met. He envisaged that much of the teacher’s day-to-day teaching strategy would depend upon feedback, implying the notion of ‘altering instruction’. Gipps (Gipps, 1994) agreed with the importance of feedback and indeed argued that formative assessment was assessment that was used to feed back into both teaching and learning. Thus, Bloom, Lloyd-Jones and Gipps all emphasised the role of feedback and the key idea that it is to be for *both the learner and the teacher*. I have built on this ‘dual role’ of feedback in developing EA, a newer concept of assessment.

Critical to the notion of feedback are two elements: the one is the need for a yardstick against which achievement can be measured - this is variously described as ‘teaching/learning objectives’, agreed standards, criteria etc., the second is the use of this information to improve the level of achievement in relation to the yardstick. This is summed up precisely in Ramprasad’s definition of feedback (Ramprasad, 1983, pp. 4-13) as follows:

**“Feedback is information about the gap between the actual level and the reference level of a system parameter which is used to alter the gap in some way.”**

Sadler (Sadler, 1989) helpfully developed Ramprasad’s notion of feedback specifically in an educational context, insisting that it should lead to a reduction in

the ‘gap’ between pupils’ knowledge and understanding and the desired goals to be reached. Sadler’s notion of ‘desired goals’ is akin to Ramprasad’s reference level.

It was broadly in this context that Sir Keith Joseph in his Sheffield Speech, (June 1984, SEC, 1986), emphasised the need for the new 16+ examination, the GCSE, to recognise the ‘positive achievements’ of the examinee, that is, to recognise what they know, understand and can do. The TGAT report (DES 1987) built on the notion of ‘positive achievement’ developed during the period leading up to the introduction of the GCSE (1986), presenting a notion of formative assessment which involved confirming pupils’ positive achievement and discussing the next steps in the learning process. The report emphasised the need for using criteria for assessing pupils’ progress and involving pupils in the process of feedback so that they could respond to the teacher’s comments and discuss the next steps. This was an important moment in the conceptualisation of formative assessment. The speech highlighted the need for **confirmation** of pupils learning and **involving** pupils in discussion about the next steps.

By requiring pupils to be involved in the process of feedback TGAT emphasised the pupil-centred nature of FA. It did not however make explicit the need for feedback to teachers and the need to ‘alter instruction’ as clearly as Bloom and others had done, although feedback to teachers is implied in the report (TGAT section 23):

“... - diagnostic, through which learning difficulties may be scrutinised and classified so that appropriate **remedial** help and guidance can be provided;” (My emphasis on remedial).



Later, in Section 27, the report explains that it does not ‘see the boundary between the formative and diagnostic purposes as being sharp or clear’ and states:

“If an assessment designed in formative terms is well matched to the pupil, it is likely to provide some information which will help in the diagnosis of strengths and weaknesses.”

This last comment is helpful, but the association of diagnostic assessment with ‘remedial’ help is, in my view, less so because the word ‘remedial’ is often associated in the mind of teachers with ‘special educational needs’, implying that such ‘help or guidance’ will only be required in a limited number of cases. Later in this thesis I use the term diagnostic assessment much more broadly in relation to ‘scaffolding’ learning.

### **3.3 Formative Assessment or Assessment for Learning - which?**

My purpose here is to suggest that there remains an unanswered question, are FA and AfL the same or are they different? In their publication *Assessment for Learning: Beyond the Black Box* (ARG, 1999) the Assessment Reform Group defined or perhaps one might say re-defined formative assessment as assessment **for** learning, and made the distinction between this and assessment **of** learning which they used to describe summative assessment. The Group stressed the importance of distinguishing assessment for learning from ‘other current interpretations of

classroom assessment', and described the characteristics of assessment for learning on page 7 of its booklet as follows:

- 1 it is embedded in a view of teaching and learning of which it is an essential part
- 2 it involves sharing learning goals with pupils
- 3 it aims to help pupils to know and to recognise the standards they are aiming for
- 4 it involves pupils in self-assessment
- 5 it provides feedback which leads to pupils recognising their next steps and how to take them
- 6 it is underpinned by confidence that every student can improve
- 7 it involves both teacher and pupils reviewing and reflecting on assessment data.

Clearly there is overlap here with the features of formative assessment discussed above, but are the two the same thing? The ARG characteristics appear to shift the emphasis from the teacher to the learner, and raise important questions which indicate a need for further clarification. For example, point 2 refers to sharing learning goals with pupils but what is then to be done to achieve the goals is not made explicit.

Similarly, point 3 is problematic if Sadler's idea of the gap between the current knowledge and the desired level is considered. Pupils can be told of the standards they are aiming for but this might not mean anything precisely because of the possible existence of a gap between their current knowledge and what they need to learn to reach these standards. To be effective, point 4 would require, for example, an emphasis on teaching and training pupils in the skills of self-assessment and linking the process to Sadler's idea of the 'gap'. The notion of 'recognising the next steps' in point 5 requires a level of knowledge, understanding and skills which the pupil might not have at the time of receiving the feedback, and in some cases such as science it is possible that the teacher might not have the requisite knowledge to identify the next steps. Point 6 refers to 'improve' rather than achieve the desired goals, which seems to lower expectations, and finally point 7 is perplexing. What might 'reflecting on data' mean to the teacher and to the pupil, and what data is being referred to? Is it performance data obtained from the National Tests, the information gathered by the teacher's own assessment of pupils, pupils' own self-assessment or some combination of these?

Following the publication of 'Beyond the Black Box' (ARG, 1999) many practitioners and educationists accepted the ARG's re-definition of formative assessment as AfL, and subsequently the two descriptions are often used interchangeably. However, a more attentive reading of the literature throws light on some differences. Stobart (Stobart, 2008, p. 146) for example, states that assessment for learning has been introduced partly because of the many 'misunderstandings' the term 'formative' generates. One of these is the belief among some teachers that regular classroom tests used for monitoring progress are formative. These tests,

according to Stobart, are ‘mini-summative’ tests because the results are not used to modify teaching and learning. He argues that the same is true for marking class work, which is often described as formative but usually isn’t, because the purpose is to provide evidence for later ‘summative judgements’.

Drawing on a review of French literature on formative assessment by Allal and Lopez (2005) Stobart describes three kinds of formative response to assessment information:

- Interactive
- Retroactive
- Proactive

Interactive is based on the interactions of the learner with the various components of teaching resources i.e. the teacher, other pupils and the lesson content and materials. This becomes the classroom culture and pervades the day-to-day lesson activity. He argues that such a classroom environment results in the *continuing adaptation of learning* particularly through feedback and guidance. He suggests that this is the focus of those who use the term **assessment for learning**.

Retroactive is the term used for formative assessment conducted after a phase of teaching. This is often done through a test and is about addressing the learning difficulties identified by the test. According to Stobart, this is ‘test and remediate’ model of formative assessment and is dominant in the US.

Proactive response is when evidence leads to future changes in teaching. In the French context with their emphasis on ‘whole class’ teaching, this involves using differentiated activities to meet the differing learning needs of pupils. In a wider context it would mean that teachers modify their future teaching in response to the evidence from their current pupils. This approach mainly benefits the next group of pupils because the results of tests and external examinations could arrive too late to benefit the current cohort.

Stobart treats AfL as ‘a particular emphasis within formative assessment’ (Stobart, 2008, p. 195). According to him, AfL’s primary focus is ‘interactive student learning’ differing from what he sees as those interpretations of FA which focus mainly on teacher learning, intended to improve pupil learning by bringing about changes in teaching and the curriculum.

Black et al (Black et al., 2002, p. i) provide some helpful distinctions between AfL, SA and FA. Firstly they suggest that,

- AfL is any assessment whose first priority, both in its design and practice, is the promotion of learning
- It differs from assessments which are designed primarily for the purposes of selection, certification, ranking or accountability (SA).

They go on to explain that,

“An assessment activity can help learning if it provides information to be used as feedback by teachers and their pupils in assessing themselves and each other, to modify teaching and learning activities in which they are engaged.”

(Black et al., 2002, p. 1)

They state that,

“Such assessment becomes ‘formative assessment’ (FA) when the evidence is actually used to meet the learning needs” (ibid, P.1).

Stobart interprets this by stating that AfL is the purpose i.e. the intention, whereas FA is the function, i.e. the effect, of assessment (Stobart, 2008, p.195). Here, Stobart makes a distinction between the words purpose and function. However, often these words are, in most contexts, used interchangeably. So, the distinction does not appear to lead to greater clarity.

For David Spendlove (Spendlove, 2009) the concept of ‘formative assessment’ is central to his understanding of Assessment for Learning. He defines formative assessment (Spendlove 2009, p. 4) as

“... assessment for enhancing and shaping learning through modifying teaching”.

He asserts that formative assessment “shapes the learning process and is implicit in AfL” (ibid) and argues that the ‘retroactive’ approaches of FA described earlier in this section do not contribute to AfL. For formative assessment to contribute to AfL it needs to have an immediate impact on changing teaching and learning. If it is

intended for longer term adjustments then it is ‘not AfL’. In Spendlove’s view, AfL is ‘dynamic’ and requires ‘daily’ adjustments to teaching and learning informed by feedback. Spendlove’s enthusiasm for AfL is evident in the following statement:

“AfL recognises teachers as informed professional decision makers. It is the teacher who decides the balance and timing of formative and summative assessments; it is the teacher who uses this information intelligently to inform their teaching; it is the teacher who as a consequence nurtures reflexive, resilient and autonomous learners.”

However, this statement accepts teachers as informed professionals and does not make explicit the need for teachers to engage in professional learning in order to continually enhance their ‘guild-knowledge’.

### **3.3.1 FA and AfL - similarities, differences and overlap**

The discussion above points to the intertwined nature of FA and AfL; the differences and the similarities are not always explicit and there is overlap between the two concepts. Under the heading ‘feeding back’ TGAT (section 37) suggested the following:

‘... formative assessment must be in terms to which pupils can respond and which clearly indicates what has been and what remains to be achieved’.

This implied the idea of a **goal (or criteria)** to achieve.

Thus TGAT placed significant emphasis on the role of the teacher in the process. It was a teacher-led process but it was also to involve pupils in discussions on what they had achieved and what their next steps would be. Assessment for Learning, as

outlined by ARG (1999) moved the idea of pupil involvement much further towards them taking more and more responsibility for their own learning. Other authors have also suggested more active involvement of pupils with the aim of empowering them.

Thus two trends emerged:

- The first was teacher-led confirmation of achievement, diagnosis of learning needs by identifying strengths and weaknesses and teacher-pupil discussion in planning the next steps with perhaps an implicit intention to empower pupils so that they can take responsibility for their own learning.
- The second explicitly emphasised empowerment of pupils by encouraging them to take responsibility for their own learning through self-assessment and peer-assessment.

The proponents of FA have continued to argue, in my view rightly, that the aim of FA is the eventual empowerment of the learner so that the learner becomes autonomous, self-monitoring, confident, and so on. However, this argument has not always been made explicit and as Wood pointed out (Wood, 1998), in order to achieve this empowerment with most/all learners, good teaching, good feedback and scaffolding are required. This position is reinforced by Black and Wiliam's Review findings in 1998 concerning the impact of feedback on learners. They reported that feedback needed to be **sensitive** to learners' feelings and self-esteem and given in a way that built confidence and **self-belief**.



The AfL focus on empowering pupils is not contested and both Stobart and Spendlove claim that AfL is dynamic, interactive and orientated towards the here and now. It seeks the active involvement of pupils in lessons. The point is, however, that this articulation pre-supposes that good feedback and good scaffolding, i.e. good FA, are already in place.

FA emphasises ‘altering teaching’ to meet ‘learners’ needs’. It emphasises sensitive teaching and good feedback in order to enhance learning. For these to be achieved FA emphasises the need for teachers to engage in professional learning. In AfL, ‘learners’ needs’ and the need for teachers’ professional learning are implicit. Marshall and Drummond (Marshall and Drummond, 2006) make a distinction between teacher actions that demonstrate the ‘spirit’ of assessment for learning and those that only comply with the ‘letter’. They maintained that the ‘spirit’ of assessment for learning required careful selection and sequencing of tasks and especially high organisation based on ideas. Jones & Cowie (Jones and Cowie, 2011, p. 51) suggest that this implies that,

“... for the teachers to undertake assessment for learning they need a deep understanding of the subject domain, likely student learning pathways and pedagogical practices likely to move student learning forward. Thus the interplay of knowledge and action is a key issue in assessment for learning.”

A strong feature of both FA and AfL that might require further explanation is creating a ‘culture of success’ in the classroom. By this I mean a classroom in which all pupils feel valued and capable of achieving and are expected to learn and achieve.

Teachers have high expectations of them. Pupils are willing to ‘think aloud’ and get things wrong. They receive good feedback. Scaffolding is such that feedback is given sensitively, the ‘gap’ (Ramprasad, 1983; Sadler 1989) between current achievement and the ‘desired’ level is addressed taking into account pupils’ self-esteem, motivation and sense of well-being. In the process of scaffolding, teachers would need to have **good subject knowledge** and a good understanding of the criteria used for success in their subject.

### **3.4 The notion of ‘criteria’**

As we have seen, the pioneers of formative assessment suggested the use of agreed standards or criteria in a subject for assessing achievement and for identifying the next steps in learning. Various models have emerged from different authors: Bloom, for example, specifies mastery learning, for Sadler it is the learning goals, for Lambert and Lines it is a pre-determined set of explicit knowledge, understanding and skills, and for TGAT it was the ten levels of attainment outlined in the National Curriculum assessment document. Some of these notions of ‘criteria’ or desired learning goals are discussed below:

#### **3.4.1 Mastery learning**

I referred briefly to Bloom’s work in section 2.2.6. Here I expand on his notion of ‘mastery learning’. Bloom and his colleagues stated that the use of “mastery learning” techniques could help 80% of pupils in a class attain the same standards as the top 20% of pupils taught by the same teacher using conventional techniques. The

change is brought about by an emphasis on helping pupils develop positive attitudes to learning by providing good teaching based on a good understanding of what pupils already know. They describe this as a ‘cognitive entry characteristic’.

The idea of mastery learning requires the teacher to allow the pupil to master each ‘sequential unit’ of work. Consequently, each element of learning in the ‘hierarchy’ takes place when the pupil has mastered the previous element. This helps to develop more positive attitudes, what they describe as ‘affective entry characteristics’. By ensuring that the cognitive levels match the demands of the next element of learning, motivation is improved and progressively pupils engage with the learning tasks with more positive attitudes. Lambert and Lines (2000) consider the concept of mastery learning to be roughly equivalent to the idea of ‘deep’ learning used in the British context. It is also equivalent to the fourth stage in Bhaskar’s (Bhaskar, 2012, p. 105) concept of the dialectic of learning in which knowledge is effectively in-built to the learner and becomes part of him/her; and having acquired mastery of that level the learner is ready to move on, both to use the knowledge and to acquire the knowledge at the next level.

### **3.4.2 The TGAT model**

In contrast to the ‘mastery learning’ model of Bloom and colleagues, the TGAT model (DES, 1989) is much less detailed. It provides guidance on how to use the level descriptors and although it can be argued that the model illustrates the progression of key ideas in the topic being studied it does not require teachers to ensure the mastery of the key ideas in order for pupils to access the intended learning.

In England and Wales, following the TGAT model, a ten level progression ladder (later reduced to eight levels plus an exceptional level) was devised for each curriculum subject. This was intended to be used for judging the performance of pupils from age 7 to 14. The average expected level at 14 was Level 5. Level descriptors for science, as for other subjects, were included in the subject specific NC documents '(DES/WO, 1991).

### **3.4.3 The Clarke model**

Another model for using 'learning objectives and success criteria' which has popular appeal, particularly in primary schools but also in some secondary schools, is due to Shirley Clarke (Clarke, 2001). The main focus here is on linking learning objectives and success criteria for the lesson, but there are also references to 'sharing learning objectives with students long- and short-term'(Clarke, 2005). The approach is based on ARG's 'Assessment for Learning: Ten Principles' (ARG, 2002) This concept includes, in addition to the knowledge, understanding and skills of the TGAT model, the so called 'big picture' learning objectives. Both the ten principles of ARG and the 'Application of Learning Objectives' of Clarke would arguably require a considerable amount of professional learning for teachers on how to adopt such an approach.

Clarke's (Clarke, 1998) use of 'learning intentions and success criteria' in terms of WALT (We Are Learning To) and WILF (What I am Looking For) have proved popular with many UK primary school teachers and have influenced classroom practice. The Secondary National Strategy has also promoted the use of WALT and WILF for use in the classroom describing them as 'learning objectives' and 'learning

outcomes’ respectively. However, there is a danger that these acronyms end up being used without teachers understanding their meanings.

Spendlove (Spendlove, 2009) describes the differences between learning objectives, learning outcomes, success criteria, and assessment criteria.. He discusses the link between learning objectives (learning intentions) and assessment criteria and introduces the acronym OLII (our learning intention is). Such descriptions of learning objectives and outcomes also require the engagement of teachers in developing these for the subjects they are teaching and the pupils in their class.

#### **3.4.4 Ramprasad’s notion of the ‘gap’**

My preference is for learning goals towards addressing the ‘gap’ (Ramprasad, 1983) between current learning and desired learning by using the concept of ‘knowledge, understanding and skills’ as emphasised by Lambert and Lines. The TGAT model based on the National Curriculum levels appears to be flawed because the task of identifying the levels in terms of knowledge, understanding and skills is often left undone. Thus learners usually know what levels they are in terms of numbers (Level 4 or Level 5 etc.) without knowing what this means in terms of the knowledge, understanding and skills they already have and what they need to acquire. The current learning could and should be accurately assessed through observing pupils in lessons, marking their work, engaging in discussion, and including short tests (SA), if deemed appropriate. Lessons can then be planned with steps marked out for closing the gap. In other words, arriving at the learning objectives for lessons would

require the teacher to engage rigorously with formative and diagnostic assessments, keeping in mind the ‘desired goals’. Diagnostic assessment here would mean finding out what pupils need for taking the next steps towards the ‘desired goals’ which could involve accurate summative assessment. In addition, diagnostic assessment could indicate what the teacher might need to do (or learn) in order to support (scaffold) pupils’ next steps.

### 3.4.5 Feedback

There is general agreement among authors that feedback is a crucial element of formative assessment/assessment for learning. Many argue that feedback is part and parcel of good teaching and most teachers provide some feedback as part of their normal teaching. The argument here is that feedback needs to be, in the words of Bloom (1971), a part of ‘systematic’ and ‘sensitive’ teaching. It is to be *systematic*, in that the learners’ needs are identified in a systematic way, with feedback given about the areas where improvement is needed; it is to be *sensitive* to learners’ feelings and their self-esteem. My argument here is that providing feedback and adapting teaching are *iterative* processes. In my view, scaffolding requires both to work together and in order to scaffold pupils’ learning effectively teachers will need to engage fully with the intricacies of feedback and appreciate the iterative nature of scaffolding and teaching.

Crooks’ (Crooks, 1988) and Black and Wiliam (Black and Wiliam, 1998a) have provided good advice on feedback. Other authors such as Lambert and Lines (2000),

Clarke (2005), Gardner (2006), Broadfoot (2007), Stobart (2008), and Wiliam (2013) have generally endorsed Crooks' advice which consists of the following three key points.

Feedback:

- needs to be given regularly and while still relevant
- should be specific to the task.
- should highlight the skills, knowledge and attitudes which are seen to be most important in the subject a pupil is studying.

The last bullet point could mean, for example, that in a science laboratory, attitudes to safe practices and collaboration with other pupils would be observed and feedback provided during the lesson.

Black and Wiliam's advice on feedback is based on Butler's (Butler, 1988) study which involved 48 Israeli pupils, age 11, selected from 12 classes across 4 schools. Half of the pupils were in the top quartile of their class on tests of mathematics and language, the other half were in the bottom quartile. They were given two types of tasks in pairs. The tasks were not curriculum related but tested convergent and divergent thinking. Each student received one of three types of written feedback with returned work.

One third of the group were given individually composed comments on the match, or not, of their work with the criteria which had been explained to all of them beforehand. Another third of the group were given only grades based on the scores

they had achieved in the preceding session's work. The third group received both comments and grades.

Scores in the work done during the whole of the programme showed that:

- the comments only group maintained a high level of performance throughout
- the comments-with-grade group showed a significant decline in scores on both tasks, as did the 'grade' only group

The main difference between the higher attaining and the lower attaining groups was that interest was 'undermined' for the lower attaining pupils by either of the two feedback methods containing grades whilst the interest of high attaining pupils was maintained at a high level irrespective of the method of feedback. I have made use of the above 'advice' on feedback in my analysis of written feedback to pupils (see chapter 7).

A more recent study (Ruiz-Primo and Li, 2013) confirms the need for addressing the 'gap' identified by Ramprasad and Sadler. It states on page 20:

"The formative role of feedback cannot be fully understood without connecting it to the targeted learning goal and comparing the actual achieved level and the expected level as defined by the criteria (Ramprasad, 1983; Sadler, 1989)."

### **3.4.6 Connecting FA with pedagogy**

In recent years, discussions on formative assessment and summative assessment in the context of the broader issue of pedagogy have emerged and as discussed in Chapter 2, there are weaknesses in teachers' summative assessments too. Thus, a



broader pedagogical approach is needed. Black (Black, 2013) discusses these issues drawing attention to the importance of ‘dialogue’ and active involvement in the pedagogical process of both the teacher and the pupil. Having reviewed recent literature relating to the state of assessment in Australia, England, Scotland, Wales and Sweden, Black argues for a programme of ‘Assessment Literacy’ for teachers in which the approach to assessment will lead from *‘discord to harmony’ between SA and FA*. The notion of ‘assessment literacy’ for teachers is crucial because teachers would need training and support in acquiring skills and expertise in all forms of assessment, including summative assessment.

### **3.5 The case for FA and AfL**

Here, I review the case which has been made for formative assessment/assessment for learning in the literature. We know that FA and AfL focus exclusively on teaching and learning. They should enable teachers and learners to track the learner’s progress against a set of goals which the learner endeavours to reach. They are pupil-centred and thus have the potential for empowering pupils (Broadfoot, 1994, Daws & Singh, 1996). Black and Dockrell’s study (Black and Dockrell, 1984) found that despite the fact that it was difficult to discover the ‘extent and the nature’ of formative assessment practised by individual teachers, there was evidence that ‘carefully planned approaches to formative assessment could have a positive ‘impact’ on both teaching and learning’ (Black, 1986, p. 11).

Strong evidence of success for FA comes from Black & Wiliam's Review (1998). The review found evidence of success for formative assessment, showing effect sizes of between 0.4 and 0.7. (Black & Wiliam 1998a: 41).

It is also argued that formative assessment helps improve motivation and encourages pupils' desire to learn and therefore has the potential to promote life-long learning. Harlen (Harlen, 2006) suggests that assessment is one of the key factors that affect motivation, explaining that developing motivation for learning is an important purpose of modern education and suggesting that formative assessment practices can make a significant contribution to this. Other earlier sources had also supported this claim. For example, the KMOFAP project (1999 – 2001) found evidence of improved motivation both in the project teachers and the pupils.

Support for AfL also come from Ofsted inspections included in Her Majesty's Inspector for Schools' (HMCI's) annual reports which describe the use of AfL as one of the key characteristics of 'outstanding teaching'.

Another concept in relation to assessment emerges when considering the kind of learners the educational enterprise should aim to create and shape. Stobart (Stobart, 2008) and Spendlove (Spendlove, 2009) argue that assessment creates and shapes learners and the purpose of formative assessment and assessment for learning is to shape the learning 'journey' in such a way as to help learners gain autonomy and increase their ownership of the learning process.

### **3.5.1 Evidence of success of FA/AfL**

Substantial empirical evidence for the success of formative assessment was provided by the Black and Wiliam Review (Black & Wiliam 1998), published in March 1998 and covering a ten-year period, from 1988 to 1998. The impact of FA in terms of effects on achievement is given above (section 3.5), and a further significant finding was that the Review did not reveal even one report of a *negative impact* of formative assessment on learning. An example from the Review that has particular relevance to my study is that of Whiting, Van Burch and Render (1995). This study gives an account of Whiting's experience and importantly his record of success with about 7000 students over a period of 18 years of using mastery learning with his students. This involved regular testing and feedback to students, with a requirement that they score at least 90% in order to move immediately to the next task or they study the topic further till they achieve the minimum score (90%) required. Black and Wiliam stated that this study had ecological validity – it was an account of work in a real classroom of a method used by a teacher over many years. The key point for me is the teachers' belief that '*all pupils can succeed*' and the impact of this on pupils' achievement.

### **3.5.2 Evidence of success of FA/AfL from Ofsted**

Evidence of the success of aspects of formative assessment is found in the HMCI's 2009 Annual Report (Ofsted, 2009). The report states in Section 60 that the quality of teaching and learning was 'outstanding in 12 % of the secondary schools inspected between September 2008 and July 2009'. It also states in Section 67 that:

“Improvement in provision is often linked with better use of assessment as a key element of teaching and learning”

It further adds:

“Careful **marking** of pupils’ work provides clear guidance for improvement.”

(My emphasis)

The strong evidence of success of FA/AfL led to the development of programmes for improving teachers’ skills in these aspects. These are discussed in the next section, and more detail is provided on the impact of, and therefore the case for, FA/AfL.

### **3.6 Efforts to support teachers - CPD programmes in FA/AfL**

The issues illustrated by the examples in the Black and Wiliam Review (1998) present considerable challenges for teachers’ professional learning, including their belief in learners’ ability to learn and their levels of engagement with the processes which would enable them to implement and embed FA and AfL practices.

These challenges are not insignificant and the barriers are many, as Lambert and Lines (2000, p. 120) point out:

- the high stakes external examination system ... influences how teachers see the role of assessment;
- certain principles that underpin external assessment – notably the importance of reliability are carried forward and influence unduly aspects of classroom assessment;

- the public accountability of teachers and schools has also tended to raise the stakes in assessment and increase the propensity of teachers to rely on formal tests

They also point out that for teachers to ‘move fully’ towards adopting FA would require a significant ‘shift in their beliefs and perceptions’. They state that even if the ‘public accountability’ functions of the ‘external summative assessment’ was to be ‘toned down somewhat, there is, in our view, no guarantee that formative assessment would flower in its place’ (ibid, p.121). They state further that FA requires:

- the application of mental and physical energy;
- secure subject knowledge;
- a growing understanding of the pupil as a learner.

They assert (ibid, p.174) that:

“Labelling children too rigidly according to ‘ability’ is perhaps the most damaging unfairness of all, and runs counter to the goals of formative assessment.”

In the last three decades there has been a considerable amount of research and writing suggesting the importance of appropriate CPD for teachers for developing formative assessment and assessment for learning. Teacher training courses have included modules on teaching, learning and assessment. The KMOFAP project produced a successful model for CPD. However, the main source of training for serving teachers and school managers has been through the efforts of the National Strategies. Many teachers, school managers, and Local Authority officers have

undergone professional development training on assessment for learning. Here, I describe first the main elements of the Project.

### **3.6.1 The KMOFAP model of CPD**

The King's Medway Oxford Formative Assessment Project (KMOFAP), 2001, involved twelve science teachers, twelve mathematics teachers from six comprehensive schools in two Local Education Authorities (Medway and Oxfordshire) and four researchers from King's College, London's School of Education. The project began in January 1999 and lasted for 18 months. Teachers were introduced to the key findings of the Black and Wiliam Review and were asked to develop innovative ways of applying formative assessment strategies in their classrooms.

Teachers and researchers focused on the following areas:

- Questioning
- Marking
- Pupil self-assessment and peer-assessment

The project teachers experimented with self- and peer-assessment ideas in a variety of ways and with some of their classes. Black and Harrison (Harrison, 2001) reported that teachers made significant progress in *adapting marking* to maximise good feedback. More generally, the KMOFAP researchers commented that when confronted with any ideas about improving day-to-day learning, teachers would worry about how they would find time for them. However, during the project, teachers came to see that the time spent on formative assessment was worthwhile. Teachers' initial doubts about

the efficacy of formative assessment methods were overcome and they found that their time engaging in the project was well spent.

### **3.6.2 The National Strategies**

The National Strategy for Key Stage 3 was put in place by the Department for Education and Skills (DfES) and supported by a national team of trainers and educators. The initial aim of the Strategy was to raise standards in English, mathematics and science by improving teaching and learning, developing cross-curricular skills such as literacy and numeracy, and helping pupils who enter Year 7 (age 11) with attainment below level 4 in the National Curriculum to make faster progress (DfEE, 2001; Ofsted, 2004).

The Secondary Strategy started in September 2000 with a pilot phase (Ofsted, 2004). This involved 205 schools in 17 volunteer Local Education Authorities (LEAs). The pilot phase began with English and mathematics and a year later included science, then subsequently the foundation subjects. Some schools began to pilot information and communication technology (ICT) in 2001.

When it was rolled out nationally in 2001 (following the pilot phase), the National Strategy consisted of strands covering English, mathematics, science, ICT and other foundation subjects such as modern foreign languages and design and technology. From September 2004, behaviour and attendance and school improvement strands were also included.

Although the Strategy's rhetoric was of '*transformation*', its programme fell woefully short of this. The cascade model of training meant that only one or two teachers from each department of the schools attended the CPD courses, with an expectation that the teachers attending the courses would deliver the training to the rest of the department. I was involved in providing training in my capacity as a local authority adviser. There was no structure in place in schools to facilitate the cascading of training, and compared to the KMOFAP approach, where there was intensive collaboration between teachers and researchers, the cascade model of CPD presented a light touch interaction with teachers. The practical support promised by the Strategy was provided with a ratio of approximately one LEA consultant to 100 science teachers, arguably too little to make a significant impact.

The pilot phase of the National Strategy science strand began in September 2001 and was completed in July 2002. There were 17 CPD units, of which two were on assessment:

- Unit 11 on Assessment for Learning
- Unit 16 on Assessment for Teaching

Unit 11 on AfL drew on Black & Wiliam's publication, (Black, 1998) and ARG's publication, (ARG, 1999). Unit 16, however, was the Strategy's own invention, and provided a separate CPD unit that suggested that AfT was different from AfL. The AfT unit contained materials which were absent from the AfL CPD unit, including sections on 'Teachers' subject knowledge and understanding', 'Teachers' Pedagogical Knowledge' and 'Teachers' knowledge of progression in science from Key Stage 2 (age 11) to Key stage 3 (age 14)'. In my view it was a highly relevant



CPD unit. When the science strategy rolled out nationally, the AfT unit was not included in the programme. The reasons for this are not clear (possibly related to a significant cost nationally), but given the issues which existed (and still exist) around science teachers' subject knowledge it was a real loss.

In addition to the above, the Key Stage 3 National Strategy also produced whole-school, generic CPD training materials on Assessment for Learning. An A4 ring binder housed the training materials and was referred to as the 'Assessment for Learning (AfL) training folder'. The folder contained 6 CPD Units in addition to a 28 page introduction (DfES, 2004). It emphasised 'the formative use of summative assessment' drawing on Biggs' (Biggs, 1998) notion of using external summative assessment structures formatively.

Thus two important points emerge about the Strategy's conceptualisation of assessment in relation to teaching and learning; Firstly, the use of 'external summative assessment structures' formatively. This can be helpful but suggests a 'bolt-on' model for formative assessment and has the drawback that by the time such results are available it is often too late to go back and teach/learn the topic and or concepts and skills which needed to be re-taught and learnt. Secondly, even though it was not pursued, the Strategy's formulation of AfT was helpful as it pointed to what was missing from AfL, i.e. an emphasis on teachers' subject knowledge and 'pedagogical subject knowledge'.

As all this suggests the National Secondary Strategy was investing a considerable amount of money and effort in promoting Assessment for Learning. The aim,

however, was to fit AfL into the target-setting framework, a framework based on output measured by the National Tests. The whole-school approach was essentially addressed to school managers, initially in 2004 to School Key Stage 3 managers, then in 2005 to senior leaders and subject leaders. The main thrust of the Strategies, however, remained focused on maximising the outputs of the National Tests. Thus methods such as ‘catch-up programmes which could boost examination performance were promoted, irrespective of their contribution to ‘deep’ learning.

The whole-school programme was presented only to those schools which chose AfL as a whole school policy. However, the guidance provided was detailed and gave numerous strategies for promoting AfL. What it argued for was a systematic use of assessment for learning to support learners and raise standards as measured by the national tests. The Strategy’s approach appeared to be based on an assumption that summative assessment, (that is the National Tests), and a systematic AfL could co-exist and indeed complement each other. This was contrary to what many researchers and authors had argued, namely that there were clear tensions between summative assessment which was deeply embedded and formative assessment which remained at the periphery of practice. Lambert and Lines (2000) summarise these tension aptly (p.191) by stating that SA and FA:

- fulfil entirely different purposes;
- are valued unequally by the system;
- co-exist in conditions of tension.

The KMOFAP's approach to helping teachers develop formative assessment as indicated in section 3.6.1 was, by comparison to the National Strategy's CPD programme, more intensive. The collaboration between teachers and researchers was sustained over a period of 18 months and focused on fewer objectives than the National Strategy's programme. The focus of KMOFAP was on questioning, marking and self-assessment, and peer-assessment. The ratio of teachers to researchers in the KMOFAP programme was 6:1. In contrast, the Strategy did not prioritise any themes and covered 'Assessment for Learning', and 'whole-school training' over ten themes. As already mentioned, the ratio of teachers to consultant was about 100 to 1.

Despite these not inconsiderable efforts to promote AfL, the ARG (ARG, 2009, p. 22) reported that its use had remained 'patchy'. This was particularly the case in science (SNS, 2007; HMCI report, 2008; Ofsted 2008), and my field work investigated what the reasons for this might be.

### **3.7 A lack of FA in secondary school science**

The available evidence suggests that systematic formative assessment (FA) is infrequent, unsystematic, and inconsistent in secondary schools generally and in science particularly. Nearly three decades ago, H. D. Black (1986) noted that the use of formative assessment was not systematic or purposeful. Other researchers and authors such as Daws and Singh (1996), Black and Wiliam (1998), HMI (2004), Black et. al. (2007), the National Strategy (2007), the annual reports of the Chief Inspector for Schools (Ofsted, 2008, 2009), the Ofsted report on the impact of

National Strategy support on Assessment for Learning (Ofsted, 2008b), (Gadsby, 2012) and (William, 2012) have all reported a lack of systematic and purposeful use of formative assessment. The point here is that efforts to develop formative assessment or assessment for learning, though persistent, have resulted either in failure or at best in limited success.

The Ofsted (Ofsted, 1996) report on Secondary Schools stated that:

“Marking is usually conscientious but often fails to offer guidance on how work can be improved. In a significant minority of cases, marking reinforces under-achievement and under-expectation by being too generous or unfocused. Information about pupil performance received by the teacher is insufficiently used to inform subsequent work.”

And in 2004 Ofsted (Ofsted, 2004) stated:

“Assessment for Learning, including well-focused use of marking, is not yet well enough established to promote progress systematically. Relatively few schools have a thoroughgoing, rigorous approach. The Strategy is responsible for some improvement, for example, in mathematics, but there is a long way to go if practice is to become consistently better across the range of subjects. This issue will be a major focus of the Strategy from spring 2004.”

The Secondary National Strategy’s own evaluation of the state of Assessment for Learning in 2007 (SNS, 2007) found that marking and feedback did not show pupils how to improve.

“Whilst marking of pupils’ work was nearly always up-to-date, little guidance was provided to pupils on how to improve.”

“Teachers’ comments tended to be targeted on presentational aspects of the work or general exhortations to ‘work harder’ or ‘listen more carefully.’” (p. 35).

An example of teachers’ comments:

“complete all the work set in the lesson and use a ruler to draw tables” (Year 8 pupil)

An example of a pupil’s dilemma:

“I know my **target** and my **level** at the moment. I’m not really sure what to do to make my Level 4 into Level 5 ...” (Year 9 pupil) [my emphasis].

Her Majesty’s Chief Inspector of Schools’ (HMCI’s) annual report, 2008 (Ofsted, 2008a) based on the inspection of 1164 secondary schools in England between September 2007 and July 2008, states:

**“There is rather more unsatisfactory teaching in science than the other core subjects”**

“The use of assessment to track pupils’ progress, plan work, and manage whole-class discussion continues to be a common weakness.”

Ofsted (2008): impact of the Strategy – section 54, states:

“In response to initiatives from the Secondary National Strategy, there has been some improvement in teachers’ use of assessment for learning in science. In general, however, it was insufficiently diagnostic and not used well enough to plan work for groups. The emphasis has been on **summative assessment and preparation for tests and**

**examinations.** Students generally knew what their targets were in science, but lacked sufficient guidance to be certain about what they needed to do to meet them.” [My emphasis]

The HMCI (Ofsted, 2009) report states:

“The quality and use of assessment are inconsistent both within and between schools”

“Assessment frequently emerges as an area for improvement.”

More recently, in the Stewart interview with Dylan Wiliam, in the TES (Stewart, 2012, p. 24) , on the implementation of AfL, Wiliam commented that ‘most schools were doing it wrong’. Likewise, Gadsby (Gadsby, 2012, p. 1) states that, “many well-intentioned teachers are engaging with the letter of AfL rather than the spirit of it”.

The above evidence strongly suggests that formative assessment or assessment for learning has remained impoverished in secondary school science classrooms. As already indicated, my fieldwork was focused on the possible reasons for this impoverishment.

### **3.8. Some theoretical issues**

The discussion in this chapter shows that whilst the notions of formative assessment and assessment for learning share some key ideas such as feedback, the use of criteria and the involvement of pupils, there appear to be some differences. In Black and Wiliam’s view, an assessment is ‘formative’ when it is used to meet pupils’ learning needs. Spendllove states that formative assessment shapes the learning

process. Both Spendlove and Stobart assert that AfL is dynamic, interactive and here and now. However, whilst Stobart sees AfL as ‘a particular emphasis within FA’ Spendlove sees FA to be ‘implicit in AfL’. In short, there appears to be some confusion about the definitions of FA and AfL.

Despite discussions, research findings, encouragement and support to teachers, the practice of FA and AfL in science remains at best sporadic and weak. Given the fact that the summative culture is deeply embedded in schools’ pedagogy, it was to be expected that there would be challenges and barriers at every level: national, school and classroom. So what is the nature of these challenges and how might they be overcome? Chapter 5 describes a framework for studying these questions.

My overarching research question is *‘Why has the use of formative assessment or assessment for learning remained at the periphery of practice in school science?’* despite the widespread belief and extensive evidence that they promote learning and despite various efforts to promote them.

Prominent amongst the barriers identified in the literature are:

- the notion of ‘innate ability’ as fixed and limited conceived as being counter to the aims of formative assessment;
- teachers not knowing enough about their pupils’ learning needs;
- marking and grading, much of which tend to ‘lower the self-esteem’ of pupils;

- feedback often serves social and managerial purposes rather than helping pupils to learn more effectively;
- the dominance of national tests and exams;
- the ‘standards’ agenda and the ‘accountability culture’
- lack of a sustained CPD programme focused on integrating assessment, teaching and learning

In addition to the above, another possible barrier could be the uncertainty about the definition, conception and description of FA and AfL in the literature as discussed above.

### **3.8.1 SA, FA, AfL and AfT**

In the preceding pages I have described some key features of summative assessment, formative assessment, assessment for learning and assessment for teaching. In this section I consider some issues with SA and then draw together similarities, differences and overlaps between FA and AfL and discuss how a new concept of assessment, that of *Enabling Assessment (EA)*, can build on these. This involves a new definition for diagnostic assessment and draws attention to an iterative relationship between FA and scaffolding.

Firstly, some issues around summative assessment. The purpose of summative assessment as defined by TGAT (DES, 1987) is ‘for the recording of the overall



achievement of a pupil in a systematic way'. In 2004 Wynne Harlen endorsed the above purpose in the following words:

"... unlike assessments that are formative or diagnostic, the purpose of summative assessment is to determine the student's overall achievement in a specific area of learning at a particular time – a purpose that distinguishes it from all other forms of assessment (Harlen, 2004)".

In this thesis, whilst the main focus is on formative assessment and assessment for learning, the importance of accurate and reliable summative assessment for the purposes of information on the overall achievement of a pupil and for certification purposes are recognised and endorsed. Indeed, accurate, valid and reliable summative assessments are crucial for maintaining the integrity of the system of assessment and for sustaining public confidence in the system.

However, external examinations have long been subjected to severe criticisms for their effect on 'narrowing' the curriculum and treating a pupil only as an 'examinee' not as a learner and a person (Acland, 1911; Norwood, 1943). The debate has centred around the question of external examinations and tests on the one hand, and internal school assessments carried out by teachers on the other. In fact the Norwood report had recommended a seven year period of sustained 'in-service training' for grammar school teachers so that they could gain the competence needed for carrying out assessments which would assess the 'whole' child, the whole achievement of a pupil during her/his time at the school.

However, the post-war government had more pressing priorities, so funding was not available for implementing such an ambitious programme of ‘in-service’ training even though it was aimed only at grammar school teachers. Since then there has been little or no political will or ministerial enthusiasm for launching such a programme for training teachers for the task of carrying out accurate, valid and reliable assessments. Education researchers and authors, however, have continued to draw attention to the shortcomings of teachers’ classroom assessments and the possibility of a future in which teachers could be trained to assess the ‘whole’ child with competence and integrity and ensuring the validity and reliability which would merit public confidence in the process.

Two reviews of summative assessment practices, the 2002 review (Harlen and Crick, 2002) and the 2004 review (Harlen, 2004) described the state of summative assessment in schools and its impact on pupils and teachers. These reviews found that following the introduction of the national curriculum tests in England, lower attaining pupils tended to have lower self-esteem than higher attaining pupils. Lower self-esteem often led to a decrease in motivation and effort and eventually to lower academic success. What is more, the researchers found that:

“ High stake tests impacted teachers, making them more likely to choose teaching practices which transmit information during activities that are highly structured and teacher controlled” (Moss, 2013).

Evidence for this was found in my fieldwork where teachers did not answer pupils’ questions stating that they will not be asked such questions in the SATs.

A relatively recent review of research on summative assessment practices in classrooms covering the period 1999 to 2011 (but excluding the studies covered by Harlen & Crick (2002) and Harlen (2004)) confirms many of the findings of the earlier reviews and concludes as follows:

“The studies in the review ‘reveal areas of need and areas of promise regarding teachers’ summative assessment practices” (Moss, 2013).

Although teachers are testing more frequently, many lack competence in the skills required for accurate and reliable assessment. Teachers, however, continue to retain positive beliefs about and a high level of confidence in their assessment skills ‘despite evidence to the contrary gathered through observations and teachers self-report’ (Black et al., 2010; Reig, 2007).

There are many instances where teachers ‘misinterpret’ pupil achievement or ‘misestimate’ their ability (Kilday et al., 2011). What’s more, teachers arrive at their views of pupil achievement through idiosyncratic methods (Moss, 2013, p. 252). They include behaviour as a category either for rewarding good behaviour with a higher level/grade or for punishing poor behaviour by ‘adjusting’ marks down (Wyatt-Smith et al., 2010).

Mackiney et al. pointed out in 2009 (Mckinney et al., 2009) that traditional and routine practices were common across the board with low-level recall and objective tests being used frequently by teachers regardless of grade level or subject area. They

concluded that low-level testing could be found in many classrooms where it impacted both the quality of learning and the motivation of the pupils, adding that pupils had no choice but to engage in those assessments. Much earlier, McMillan & Nash (McMillan and Nash, 2000) had discovered that even when teachers recognised effective assessments practices, they often saw the realities of their classroom environment and other external factors imposed on them as too difficult to overcome.

Despite these shortcomings in teachers' practices of summative assessment many researchers still believe that teachers' summative assessment practices have the potential to positively influence students, and that teachers can do summative assessment without the negative effects associated with external examinations and tests (McMillan, 2003; Martinez et al., 2009). Moss (2013, p. 253) states that six of the 16 studies reviewed took the above positive view of possible future developments. These were: Black et al., 2010; Brookhart and Boronwicz, 2003; Brookhart and Durkin, 2003; Brookhart et al., 2006; McMillan & Nash, 2000; Wyatt-Smith et al., 2010).

I belong to the above minority group of researchers who believe that the future of assessment should lie with teachers, with moderation between colleagues ensuring quality control and providing a collegiate environment for developing teachers' skills in all forms of assessment, including summative and formative assessments. Public acceptance of a new system in which internal school assessments by teachers could be the main source of summative assessment would be hard to achieve, perhaps because examinations have been seen for decades as reliable and impartial. Nevertheless, the first steps towards gaining public acceptance need to be taken by enhancing teachers' skills in classroom assessment.

Assessment for teaching (AfT) appeared in one of the CPD modules produced by the National Strategy in Science (Unit 16) and has been discussed in section 3.6.2. It emphasised the need for developing teachers' subject knowledge, especially in the light of the fact that there was and continues to be a severe shortage of physics teachers and, to a lesser extent, chemistry teachers, in state schools in England. It also discussed the importance of teachers' pedagogical subject knowledge. Many teachers were not fully confident with the progression indicated in the science National Curriculum and the AfT module attempted to address that.

The overlap and differences between FA and AfL have been discussed in section 3.3.1. Some authors have emphasised AfL over FA (Stobart, 2008; Spendlove, 2009), others have assumed the interchangeable nature of FA and AfL (Black, 2003; Black and Wiliam, 2012).

There appears to be the emergence of two schools of thought: one promoting AfL assuming that FA is subsumed in it, the other promoting FA assuming that AfL is subsumed in it. For example, the National Strategy in science (DfEE) focused almost exclusively on AfL and the Sage Handbook of Research on Classroom Assessment (Sage, 2013) has chapters on SA and FA but none on AfL. As explained in section 3.3.1 different authors have emphasised different aspects of AfL with most agreeing on feedback as the common feature of FA and AfL. AfL appears to emphasise self- and peer-assessment and dialogue in the classroom. FA encompasses these characteristics of AfL and promotes marking and feedback. Thus, there are some key concepts which are explicit in one but not explicit in the other.

In addition, attempts to engage teachers in CPD relating to FA/AfL have met with little success.

### **3.9 Towards a new theory**

In view of the fact that there remain significant differences in the interpretations of FA and AfL and of the notion of feedback and that their use to support learning, despite their clear potential to do so, is ‘patchy’, it is appropriate to theorise and develop these concepts further. The depth of overlap or confusion is indicated in the title of a paper presented in 2003 by the King’s College, London’s Assessment for Learning Group to the AERA conference (Black, 2003). The title was “The Nature and the Value of Formative Assessment for Learning”, suggesting the authors have seen the need for combining both FA and AfL in their discussions, or indeed deemed them too intertwined to separate.

In view of the above discussion I have endeavoured to re-conceptualise these ideas of FA, AfL and elements of SA in a new theory of assessment. This captures the essential features of both FA and AfL, avoids their limitations and takes them forward, together with appropriate elements of SA, to *‘Enabling Assessment’*. This theory *is* further expanded in chapter 6.

### **3.10 Summary of Chapter 3**

In chapter three I have carried out a review of the literature on formative assessment and assessment for learning. The key components of formative assessment/

evaluation and of assessment for learning as they appear in the literature are outlined below:

Formative assessment should:

- be learner-centred and identify pupils' achievement and confirm it
- diagnose the 'difficulties' children are experiencing
- have agreed criteria for judging achievement and for feedback
- be an integral part of systematic and sensitive teaching
- lead to information for teachers' action to 'alter instruction'
- be based on a belief that all learners can, in principle, achieve learning goals

Bloom considers FA to be integral to 'systematic' and 'sensitive' teaching.

The key elements of Assessment for Learning are:

- it is embedded in a view of teaching and learning of which it is an essential part
- it involves sharing learning goals with pupils
- it aims to help pupils to know and to recognise the standards they are aiming for
- it involves pupils in self-assessment
- it provides feedback which leads to pupils recognising their next steps and how to take them
- it is underpinned by confidence that every student can improve
- it involves both teacher and pupils reviewing and reflecting on assessment data.

The role of feedback is central to the concepts of FA and AfL and teachers would need to engage with the intricacies of feedback for it to be effective in enhancing learning and self-esteem. Some authors have argued convincingly that the purpose of feedback is for both the learner and the teacher.

However, there remain some differences in the conceptualisation of FA and AfL.

Are the two concepts the same or are they different? There isn't a convincing answer to this question.

In introducing the new concept of *Enabling Assessment*, an enhanced theory of assessment, I build on the concepts of FA, AfL and SA.

In the next chapter I discuss the methodology for the research which has led to this thesis.



## **Chapter 4 Methodology**

### **4.1 Introduction**

The purpose of this chapter is to describe the methodology which underpins the research for this thesis. It is a qualitative case study with a Critical Realist perspective. The epistemological claims are based on the empirical work which has followed the norms and procedures of qualitative research: firstly the data, which has come from interviews, lesson observations, scrutiny of marking and policy documents; secondly, the procedures used to interpret and organise data through ‘coding’; thirdly, other norms and procedures which include non-statistical sampling, the writing of memos, diagramming and written and verbal reports (Strauss and Corbin, 1990). The main focus of data collection has been on two levels, the classroom (micro) and the school (meso). However, a national (macro level) picture has also been gained through interviews and the study of policy documents in relation to the Government’s ‘standards agenda’, and at an individual level repeated interviews of pupils and teachers have helped consolidate the data. I have taken care to choose methods which would best elicit information pertaining to my research questions. My fieldwork has been conducted with ethical considerations uppermost in my mind and due importance has been given to empathy in listening to pupils, teachers and managers.

In this chapter section 4.2 describes the exploratory studies and section 4.3 discusses how the methodology has developed as my research has progressed. In section 4.4 I discuss the appropriateness of a case study approach for the main study, section 4.5 outlines the fieldwork programme for the main study and section 4.6 discusses the

theoretical and practical influences which have guided the fieldwork. In section 4.7 I discuss how my study has been influenced by Critical Realism. The process of data collection is outlined in section 4.8 and section 4.9 summarises the key features of the case study schools. Section 4.10 provides a summary of the chapter.

Over a period of seven years an ethnographically informed case study has been carried out in which initially my principal case was of one science department in a secondary school, with two others subsequently added for the purposes of comparison. The study is ethnographically informed insofar as my professional relationships with the teachers in the case study schools are concerned. My knowledge of the sites was acquired over a period of more than seven years during which time my acquaintance with the teachers in my capacity as the local authority's Science Inspector placed me in a privileged position in terms of access to teachers and school managers. This provided me with the advantage of an 'insider' perspective, giving ease of access and placing me in a position of trust which encouraged the sharing of information. In turn this placed a responsibility on me to be alert to the possibility of bias in the data, to maintain strict confidentiality regarding the identity of the participants and to be sensitive in the use of the data. I have been mindful of the 'researcher and researched' relationships and have adopted a listening mode, listening with empathy to the narratives of teachers, pupils and school managers as they describe the way they have gone about their daily business.

This is an observational study and my methods have been informed by some of the approaches used by previous researchers in this field. Notably, these are:

- Questionnaire surveys such as the studies of Ferriman et. al.(1994); TAPAS (Teachers Assessing Pupils – Active Support) surveys by Fairbrother et. al. (1995), and Daws and Singh (1996),
- Project-based research such as the KMOFAP (2001) and the Learning to Learn Project (James, 2007).

The comment by H. D. Black (Black, 1986, p.11) that there was ‘a lack of observational studies on teachers’ actual practice of assessment’ has been a factor providing an incentive for undertaking this research. Another incentive for my observational study of practice at the classroom level is provided by my experience as an Ofsted inspector from 1998 to 2005. In this capacity I was able to observe lessons, examine marking and scrutinise school policies on the quality of teaching, curriculum and assessment. However, these activities were conducted within Ofsted’s restricted framework and in the context of the ‘inspection week’ which was a highly unusual week in the life of a school, in that schools were preparing often for months in advance for the inspection. Whatever the Ofsted inspectors saw was, at least in part, merely a result of this preparation. At the end of each inspection I was left with a feeling that what was seen was not what the pupils normally received and experienced as a matter of day-to-day routine. So, ‘What do teachers do outside and apart from the ‘Ofsted week’?’ had become a persistent question in my mind, and this question partly coincided with the implied questions on H.D. Black’s mind a decade and half earlier. Thus the wish to observe some key aspects of everyday classroom practice contributed to the methodology developed for my main study.

My research methods for the main study have been also influenced by the exploratory fieldwork in which there were two phases. Phase one was carried out between 1998 and 1999 and consisted of group and individual interviews with secondary school science teachers. Phase two took place from 2003 to 2004 and consisted mainly of scrutiny of pupils' books. The main study began in September 2005 and consists of four phases. The first three were 2005 to 2007 at Curie High School (CHS), 2007 to 2008 at CHS and Downton Community School (DCS), and 2008 to 2011 at CHS, DCS and Westfield Comprehensive Schools (WCS). I returned to CHS for the period 2011 to 2013 for the fourth and final phase.

## **4.2 The exploratory studies**

### **4.2.1 Phase 1: Interviews with teachers**

In the first phase of the exploratory fieldwork the research questions were limited to the impact of the national curriculum on assessment and the proportion of the assessment that was formative.

I carried out five semi-structured interviews with science teachers in three different schools, each located in a different local authority. The first of the five interviews was with a group of three science teachers with posts of responsibility, including the head of department, from school 1. The other interviews were with individual science teachers, two from each of schools 2 and 3, each including the head of department.

The interview schedule is included in Appendix A0. Permission for the interviews was obtained from the headteachers and the heads of department and the consent of the interviewees had been given. The selection of the schools was opportunistic as I had known contacts at the schools, either the head of department or the deputy headteacher or the head teacher, through my capacity as a university tutor for PGCE science students. Before the interviews I explained to the interviewees that their participation was voluntary and the information gained would be used for the purpose of my MPhil/PhD study at the Institute of Education, London. I also explained that the information would be confidential and if the name of the school or of the interviewees were needed for any publications arising from the interview data, prior written permission would be sought from the head teacher and the respondents involved.

My initial findings from the exploratory fieldwork (phase 1) were as follows:

- End of Unit/Module tests were the dominant mode of assessment;
- End of Year examinations took place in all three science departments;
- Test and examination results were used mainly for ‘ability’ grouping and for reporting to parents;
- Examinations and tests had become more formal as a result of the introduction of the National Curriculum;
- Test and examination data were collected as evidence for Ofsted inspections;
- Teachers’ discussions with pupils regarding their progress, their learning and how to improve were mainly in relation to improving GCSE coursework;
- Two out of the three schools used few or no formative assessment strategies;

- In the second school, where some FA was used, the head of department thought that the proportion of FA: SA was about 25%:75%.
- There was also evidence from the third school, a large comprehensive school, of a reductions of time spent on scientific investigations and laboratory work, as a result of the National Curriculum.

Previous researchers in this field had used questionnaires to gather information, so the intention here was to try interviewing as a technique which might provide a more flexible interaction, perhaps leading to richer information about teachers' classroom practices. Also, interviewing has long been one of the key methods of data collection (Hammersley and Gomm, 2008) when the researcher wishes to 'understand' the perspectives of the researched, and this has been one of my intentions in undertaking this study.

#### **4.2.2 Phase 2: Scrutiny of marking**

The second round of exploratory work consisted of scrutinising the marking of pupils' books and took place in the academic year 2003/2004. Two schools and six teachers (three in each school) were involved in this phase. School 1 was an established school with over 1500 pupils while School 2 was a new school with only two cohorts, years 7 and 8 (ages 11 and 12), and five hundred pupils in all. Schools 1 and 2 were later included in my main study and named CHS and DCS respectively. Both schools were in the same local authority.

All six teachers taught classes in Year 7 and Year 8, and three books from across these classes were randomly selected from each science teacher (18 books

altogether). Based on the main features of lessons and marking found in the books, I created a pro-forma (marking pro-forma A – Appendix A1) for recording comments, grades and marks etc. This pro-forma was later amended (marking pro-forma B – Appendix A1) to accommodate ideas from the literature and from the National Strategy's Assessment for Learning documents.

A preliminary analysis of the book scrutiny is indicated in table 4.2.2, where T<sub>21</sub> stands for teacher 2 at school 1 and T<sub>32</sub> stands for teacher 3 at school 2 and so on. T<sub>12</sub> is a geographer and head teacher, T<sub>21</sub> is a physicist, and T<sub>32</sub>, T<sub>22</sub>, T<sub>11</sub> and T<sub>31</sub> are biologists. T<sub>32</sub> is head of science and T<sub>22</sub> is a Key Stage 3 science Co-ordinator.

**Table: 4.2.2**

Action	Teacher	Description/comments
Spellings/literacy corrected	T <sub>11</sub> , T <sub>12</sub> , T <sub>22</sub> , T <sub>32</sub>	
Encouraging comments	All	
Use of learning objectives	T <sub>11</sub> , T <sub>21</sub> , T <sub>31</sub> ,	T <sub>22</sub> uses **WILF (success criteria) when it should be *WALT (learning objectives)
Comments on presentation	All	E.g. Use ruler, underline the date
Science corrected	T <sub>32</sub>	How to improve a line graph shown (Y7 Nico 29/09/03).
Incorrect science not corrected	All teachers	T <sub>21</sub> is worst in this group of teachers.
Incorrect science ticked or wrongly corrected	T <sub>22</sub>	'Correction' shows wrong direction for current – h/w 15/06/04
Unfinished/incomplete comments	Five out of six teachers	
Pupils' responses to T's comments	T <sub>32</sub>	One pupil responds: "Ok sir."

Criteria used	No evidence of any teacher using criteria for assessment	
Tasks set	T <sub>12</sub> , T <sub>22</sub>	
Grades given	All	

All teachers gave marks and T<sub>32</sub> used smiley faces.

**\*WALT: This stands for ‘we are learning to’** (Clarke, 2001) [Learning objective]

**\*\*WILF: This stands for ‘What I’m looking for’** (Clarke, 2001) [Success criteria]

Preliminary analysis of marking showed that it focused mainly on the presentation of work, e.g. the use of rulers and emphasis on neatness, and the correction of spellings. Errors in science were generally not corrected and incorrect science was often ticked. Some corrections, especially of physics, were actually wrong. There was only one clear instance of a teacher actually evidencing the correct science in my sample (AT/NICO/7J2/29Sept). Many opportunities to correct pupils’ incorrect science and to show how to improve were missed. Very little evidence of practical work was found. Work covered was generally at the primary school level, one to two years below the National Curriculum expectations for the age of the pupils.

Lessons learnt from the two exploratory studies have informed the design of the main study. In section 4.3 below I describe how I arrived at the methodology for this study, which has undergone changes and adjustments as my fieldwork has progressed and as I have discovered that the process of researching is less neat and more ‘messy’ than the appearance of a final thesis might suggest. As Bryman and Burgess (Bryman, 1994) put it:



“... the research process then, is not a clear cut sequence of procedures following a neat pattern, but a messy interaction between conceptual and empirical worlds, deduction and induction occurring possibly at the same time”.

### **4.3 A developing methodology**

My study was initially informed by at least five influences:

- The TGAT report (1987);
- HD Black’s observations nearly 20 years ago that there was a dearth of research evidence about what teachers actually did in relation to assessment and how they used the results of such assessment;
- Information obtained from the questionnaire-based studies and the project-based studies mentioned earlier;
- My experiential knowledge of assessment in science education;
- The King’s College (London) study of the professional development needs of science teachers (2000).

The questionnaire surveys by Ferriman et al. (1994), Daws and Singh (1996) and Fairbrother et al. (1995) produced valuable findings in relation to the dominant modes of assessment and the main uses of the results of assessments. However, they also revealed the need for an in-depth study of day-to-day practice in order to observe what teachers actually did and to discover the reasons they gave for the things they did.

Interviews associated with survey research have tended to be structured. In such interviews the interviewer has control over a list of questions (the interview schedule) that has been prepared beforehand. It is assumed that the interviewer will be able to manipulate the situation and get answers to the pre-planned questions. It has been argued that such an approach gets the answers but does not allow scope for the questions to be 'considered'. Authors such as Corbin (1997), Oakley (1981), and Stenhouse (1984), have used interviews in their empirical work which mark a departure from pre-planned questions and indicate a movement towards more informal, unstructured interviews, enabling the interviewee to 'consider' the questions and their answers.

In a survey style interview the interviewer does not need to know the research setting well or to have a long-term relationship with the researched. Many authors have criticised this model stating that in this approach to interviewing the interviewer is in an 'unnatural relationship' with those who are interviewed.

Generally, most field researchers have preferred semi-structured or unstructured approaches to interviewing (Burgess, 1984, p. 102). My initial interviews were semi-structured but one of the lessons learnt from the interviews during the exploratory phase related to the semi-structured interview as an instrument for data collection. During the group interviews it became increasingly obvious that when I deviated from the interview schedule the conversations with the teachers seemed to become more informative and appeared more genuine. As more interviews took place they became more unstructured or *conversational* as the fieldwork proceeded. Burgess'

(1984) refers to such social science research interviews as ‘conversation with a purpose’. Charmaz (Charmaz, 2006, p. 25) has referred to ‘intensive interviewing’. She has drawn on Loftland’s (Lofland and Lofland, 1995) description of an interview as a *‘directed conversation’*. In the light of this my interviews for the current study fall into this category, and I have employed a more conversational approach.

In a conversational style interview, the need to develop trust and rapport between the researcher and the interviewee has been emphasised. There have also been numerous warnings about the ‘dangers of over-rapport’ because, it has been argued (Burgess, 1984), this could introduce ‘bias’ in the data. It is advised that the relationship should be friendly but not over-friendly. In order to avoid such bias, development of ‘rules’ and ‘proper’ interview behaviour has to take place. I have endeavoured to develop such rules and proper behaviour during my fieldwork. One technique I have used, which has also demonstrated an ethical approach, is to state the purpose, the confidentiality and the voluntary nature of participation before each interview, each lesson observation and each book/document scrutiny.

In conducting my interviews I have paid heed to the advice offered as long ago as 1948, by Zweig (Zweig, 1948) who described the ‘key characteristics’ of the interviewer who engaged in ‘unstructured’ interviews. He describes the researcher as a friend who shows interest in the work and life of the person with whom ‘a conversation’ occurs. He emphasised the need for the interviewer to have a detailed knowledge of the empirical setting and to have prepared well for the interview. In my approach I have focused on listening, enabling respondents to express their

thoughts fully in their own style. This has run the risk that occasionally the conversation has not been as directed as was intended and has gone on for too long. However, the information so obtained has been meaningful and provided a deeper insight into the thinking of the respondents and thus, arguably, into the nature of the actual practice.

The interview as a method of data collection has come under strict scrutiny by many critics, who throw doubt (Atkinson, Coffey and Delmont, 2003); Potter and Hepburn 2005) on the idea that interviews can ‘tap stable attitudes or perspectives that govern people’s behaviour beyond the interview situation’ (Hammersley and Gomm, 2008). Accordingly, in the main study, care has been taken to interview the same respondents at regular intervals several times in order to increase the stability of the interview data.

#### **4.3.1 Learning from exploratory studies**

In my exploratory study most of the interviewees showed little awareness of the literature on formative assessment. Their views ranged from openly not knowing what formative assessment was to considering regular testing as formative assessment. There was a general lack of awareness of the debate around formative versus summative assessments.

The group interview with the science teachers from School one and the interview with the head of science department of School two (exploratory fieldwork) indicated

that some of the responses were prompted by what they thought I would want to hear. The interviews with the two heads of department may have been influenced in this way. For one head of department it may have been her knowledge that I was a friend of the head teacher. As for the head of department in the group interview, I had known him in my capacity as the university tutor for PGCE students placed in his department for teaching practice. His responses appear to indicate this knowledge and I felt that he was saying things to create a good impression of his department. Or, could it be that I was reading this into his responses? It is difficult to know. This experience highlighted issues relating to the validity of the interview data and the possibility that the way the researcher's role is perceived by the respondents could influence responses and thus could contaminate the data. Equally, the researcher might read into the interview data motives which may or may not be there. This made me aware of the pitfalls inherent in this method and, consequently, I became very alert to these possibilities. The awareness raised during the initial exploratory interviews has thus served as a sensitising experience and has been pertinent to the main study.

The schools selected for the main study are in the same local education authority, in which I have worked as the science inspector. My roles as the science inspector and as a researcher could present a conflict of interest on my part to which I have been alert. The responses of teachers to my questions are likely to be influenced by their knowledge of my role as the LA inspector. This has required caution in interpreting what teachers have said to me in response to my enquiries and questions. In this context I am mindful of the ethical issues involved and complied fully with the BERA guidance on ethical research, (BERA, 2004) and ethical guidance for

educational research (BERA, 2011). I have obtained permission from headteachers, heads of department and the individual teachers I have worked with. Every time I have met teachers in the course of my fieldwork I have taken care to explain my role as a researcher and assured them that they could withdraw from the research if they wished without feeling any obligation to me. I have also assured them of the complete confidentiality of the process.

The ethics of interviewing pupils have also been rigorously observed. Permission for interviewing has been obtained from the head teacher, the class teacher and parents when pupils have volunteered to take part in interviews. At every meeting it has been explained to pupils that: they are volunteers and as such they can withdraw from participating without giving reasons, that all interviews would be confidential, and that their prior written permission would be obtained if they were to be named in any publication arising from the research. All interviews with pupils have been conducted in places where more than two people were present.

Further learning from the exploratory field work has been the appreciation that the classroom practice of assessment is influenced heavily by the schools' expectations, ethos and structures, by teachers' expectations of pupils and also by national policies. Thus while observing classroom practice it was important to bear in mind the multifaceted nature of the factors which influence it.

#### 4.4 Research Questions and the Main Study

As already outlined, H.D. Black (Black, 1986, p.11) alluded to questions around what teachers actually did in the name of assessment and in relation to formative assessment he thought that it was informal and unsystematic. My exploratory study and evidence from the literature indicated that nearly a quarter of a century after the questions raised by H.D. Black, formative assessment in secondary schools remains ‘informal and unsystematic’.

The principal research question for the main study therefore is:

*‘Given that the intellectual argument in favour of formative assessment (or assessment for learning) has been won, why has the use of FA in secondary school science remained at the periphery of classroom practice?’*

In order to address the main question, the following subsidiary questions have been researched:

- What assessments do teachers carry out and why?
- How do these assessments impact on pupils?
- What do school managers expect teachers to do in relation to assessment?
- To what extent have teachers and managers engaged with the FA/ AfL debate?
- How have the teachers and the school managers engaged with professional learning in order to implement and develop FA/AfL?
- What kinds of assessment discourses are dominant at the school?

- To what extent does the ideology of ‘innate ability’ act as a barrier to FA in the classroom?

An associate question which might aid the analysis would be:

- *How do the dominant discourses impact upon agency in the school in relation to implementing the structures of FA/AfL?*

As indicated above, a case study approach was adopted, beginning with one science department and then extending to two more science departments (two more schools) with a view to carrying out a comparative study. The three case study schools were in the Ofsted categories of ‘outstanding’, ‘good’ and ‘satisfactory’.

A Case Study approach is thought to be best suited for the kind of enquiry I have undertaken. As Robson (Robson, 2002) describes it, a case study is an established research strategy where the focus is on a specific case. The case is interpreted very widely to ‘include the study of an individual person, a group, an organisation etc. in its own right, and taking its context into account’.

Gillham (Gillham, 2000) suggests that ‘a case’ can be defined in a variety of ways - an individual or a group, it can be a family or an institution or part of an institution. It can be a school, a department in a school, or a class. It can even be a whole community, a town, a profession or an industry. All these are single case studies, but one can study ‘multiple cases’. He suggests that it all depends upon ‘what you want to find out’.



He defines the word ‘case’(Gillham, 2000, p. 1) as follows:

- “a unit of human activity embedded in the real world;
- which can only be studied or understood in context;
- which exists in the here and now;
- that merges in with its context so that precise boundaries are difficult to draw.”

My study of assessment practices in three science departments meets these criteria. It is an example of ‘multiple cases’ in which Curie High School (CHS) is the principal case and the two companion cases are Downton Community School (DCS) and Westfield Comprehensive School (WCS).

It is also a study of human actions, particularly teacher-pupil interactions, in the ‘real world’ setting of science classrooms and science departments. The school context is also important and the interface between actions taken by teachers in the classroom and those taken by the schools’ senior managers to support learning will be key features of this study.

## **4.5 Outline of the fieldwork programme**

### **4.5.1 Main Study phase 1:**

September 2005 to July 2007 (National Tests at age 14 in place)

School: CHS, focused on Key Stage 3:

- Interviews with two teachers
- Scrutiny of marking

- Lesson observations
- Interview with head of department (HOD)
- Interview with deputy headteacher (DHT)
- Scrutiny of policy documents

#### **4.5.2 Main Study phase 2:**

Sept. 2007 to Dec. 2008 (National Tests at age 14 still in place)

Schools: CHS and DCS, focused on Year 8.

Two science teachers per school, one Year 8 class per teacher

- Interviews with two teachers from each school
- Interviews with both HODs
- Interviews with both DHTs
- Book scrutiny
- Lesson observations
- Review of policy documents

#### **4.5.3 Main Study phase 3:**

January 2009 to July 2011 (post- National Tests)

Schools: CHS, DCS and WCS, focused on Year 8

Two teachers per school, one class per teacher and three pupils from each class and Fieldwork as in Main study phase 2, plus interviews with 3 pupils per Y8 class.

#### **4.5.4 Main Study phase 4 (a):**

2011 to 2013 – (a ‘coercive’ period of monitoring and accountability)

School: CHS, focus on Y8

Interviews with:

- three teachers who were under ‘extreme’ scrutiny by school managers
- the Deputy Headteacher
- the Head of Department
- the BTEC science co-ordinator
- the Key Stage 4 GCSE co-ordinator

#### **4.5.5 Main Study phase 4 (b): 2011 to 2012**

**Schools: CHS, DCS, WCS Focus on Year 8**

- From CHS scrutiny of books from a top and a bottom sets taught by each teacher (4 classes in total).
- From DCS interview with a teacher who taught both a top set and bottom set
- From WCS top set from one teacher and bottom set from another

#### **4.6 Fieldwork at six levels: Classroom, Department, School, LA, National, International**

##### **4.6.1 Classroom contexts:**

1. Focusing on Year 8
2. The nature of written feedback - marking through book scrutiny
3. The nature of oral feedback – episodes of classroom dialogue

4. Discussion with class teachers – two from each school (6 teachers)
5. Discussion with pupils – three from each Year 8 class (18 pupils)

#### **4.6.2 Departmental contexts:**

6. Policy documents
7. Interviews with the heads of department

#### **4.6.3 School contexts:**

8. Characteristics of the schools
9. Schools' procedures and policies
10. Interviews with headteacher/deputy headteacher

#### **4.6.4 Local Authority context:**

11. The role of the LA
12. Interviews with LA officers

#### **4.6.5 National context:**

13. Government policy
14. Ofsted inspection framework
15. League Tables of results
16. Interview with the National Strategy' Science Strand Director
17. Interview with the Regional Science Strand Director

#### **4.6.6 International context:**

18. PISA studies and International comparison

#### **4.7 Amendments informed by theoretical and practical considerations**

The book scrutiny in 2003/2004 (exploratory phase) provided a benchmark and a framework for studying written feedback during my main study. A pro-forma for the analysis of written feedback (marking) informed by the literature is shown in Appendix A1. Following an HMI report (Ofsted, 2004) on marking, the sampling frequency of books for evaluating written feedback was postponed for a period of time.

In common with the findings from my exploratory study, the HMI report on marking suggested that feedback through marking was weak and as a consequence the National Strategies would begin their work of supporting and improving this aspect of teaching from the spring term 2004. In view of this the sampling for the book scrutiny for the main study was done from 2006 onwards at regular intervals on a rolling programme, to see if the National Strategy's support led to improved marking and if so, how. A small sampling of marking took place in 2006 and another in 2007. Larger sampling was carried out during the autumn term of 2008 and the spring term of 2009. This continued at regular yearly intervals till the end of the summer term in 2012.

Pupils' marked work was analysed to explore the nature of written feedback. The analysis drew on the notion of criterion referencing (TGAT, 1987) and 'comments-only' marking, based on Butler's study (Butler, 1988).

The sudden withdrawal in 2008 of the National Tests at age 14, threw the case study schools into a period of confusion and this prompted another adjustment to my schedule of data collection in order to capture the impact of this new and important development. With the abandonment of the national tests schools and teachers found themselves unsure of what was expected of them in the aftermath. My fieldwork continued in the months which followed and I have described this as phase three of the main study. This was originally intended as the last phase of my study, but the discovery of a ‘coercive’ monitoring process at one of the case study schools prompted further fieldwork in order to capture the way the monitoring of teachers was working. This necessitated a fourth, and final phase.

#### **4.7.1 Informed by experience**

My experience in the field, professional knowledge of the empirical settings and familiarity with the systems relating to the theoretical field meant that I entered the research site not as a novice but as an experienced observer of the schooling process. The research approach employed in my initial exploratory fieldwork was led by data. This is evidenced in the choice of next steps in the fieldwork and in the conceptualisation of the ‘problem’. This data-led approach chimed with and preceded my reading of the ‘grounded theory’ literature. The process at that point could be characterised as step 1(a) of the ‘discovery context’, drawing on the distinction in the philosophy of science between the *context of discovery or enquiry* and the *context of presentation* (articulation, elaboration) or *justification*. My reading of the principles, procedures and techniques of grounded theory as espoused by Strauss and Corbin (Strauss and Corbin, 1990) then informed my subsequent main

study fieldwork and analysis. This could be characterised as step 1(b) of the context of discovery.

After reaching this point in the summer of 2011, I started to become more familiar with the conceptual framework of Critical Realism and soon afterwards Professor Roy Bhaskar became my supervisor. Thus Critical Realism came to inform step 2 of the context of discovery or enquiry, and then subsequently to inform the writing of the final draft of my thesis - which then effectively defines the context of presentation and theoretical elaboration (or justification). Thus, as I discuss in section 4.7.2, Critical Realism has furnished me with a conceptual framework in which I can articulate my findings and pose relevant new questions.

#### **4.7.2 Informed by Critical Realism**

In chapter five I have described the conceptual framework provided by Critical Realism and especially the transformational model of social activity (TMSA). The TMSA is further developed in the idea of the four planar social being, as discussed in chapter 2, section nine of *'Dialectic: The Pulse of Freedom'* (Bhaskar, 2008a) and the conceptualisation of social being as involving up to seven levels of scale (Bhaskar et al., 2010). This last model provides a set of social categories in which to relate and discuss the global, national, macro, meso, micro, individual and sub-individual elements (structures and mechanisms, institutions, practices) which combine to act as barriers to formative assessment/assessment for learning. It

provides a framework which enables me to show how deeply (and frequently unconsciously) embedded examinations and tests are as classroom practices.

In addition to the TMSA and the seven levels of scale, I have also drawn on the notion of ‘the learning environment’(Brown, 2009) ’ in order in particular to develop the ontological framework for considering the extent to which the explanatory structures and mechanisms which facilitate or constrain the development of FA/AfL are operating at different ontological levels.

In studying the barriers to formative assessment/assessment for learning which are causally generated and are efficacious at these seven levels of scale I hope to isolate and display the mechanisms or structures which are inhibiting the growth of FA/AfL.

Furthermore, I draw on Bhaskar’s notion of ‘Unfolding the Enfolded’, as expounded in the Philosophy of Meta Reality (Bhaskar, 2012, p. 127) and ‘*From Science to Emancipation*’, chapter 11 (Bhaskar, 2011a), to elaborate my synthesis of the ideas of formative and diagnostic assessments with the notion of scaffolding to explore a new conceptualisation of FA and AfL. The intention is to ‘under-labour’ (Bhaskar, 1975/2008, p. 10) for education and to help teachers to create a ‘classroom culture of success’ (Black and Wiliam, 1998a).

#### **4.8 Organisation of Data Collection**

The data collected was organised as follows, for the reasons given:



- Key characteristics of the schools, to provide background information including socioeconomic factors.
- Interviews with teachers, pupils, school managers, Local Authority managers and National Strategy managers in order to evaluate their priorities in assessment, the nature of assessment being undertaken, the uses made of the assessment and their views on the merits and uses of FA and AfL.
- Interviews with pupils to learn how assessment impacted on them and their learning
- Book scrutiny in order to study the nature of marking and written feedback
- Lesson observations for studying the nature of teacher-pupil and pupil-pupil dialogue
- The study of policy documents and procedures to evaluate the schools' ethos and expectations

In the first phase of the main study focus was on assessment in the science department of Curie High School (CHS), the principal case in this study. In the second and third phases, the science departments of Downton Community School (DCS) and Westfield Comprehensive School (WCS) respectively were included. In the fourth and final phase, I returned to the principal case, CHS where there were considerable activities taking place in relation to CPD and monitoring of teachers. In a new development all teachers in the school were being required to participate in CPD on AfL, and then implement the key elements into their day-to-day teaching. Implementation was being rigorously monitored by the senior managers, in some

cases leading to disciplinary action against teachers who were judged not to be using the techniques discussed in the CPD sessions.

## **4.9 Description of the case study schools**

### **4.9.1 Curie High School (CHS) – the principal case**

As mentioned above, Curie High School is the principal case for my study, and its features make it an especially educative case. At the time of the study historically the standards had been low and consistently below national averages. The school's senior managers had been working hard to improve standards and the school moved to the Ofsted grade 'good' in 2008 from 'satisfactory' in 2005.

The school has a welcoming attitude to pupils and parents, with open access to the school office. It has an unusual approach to behaviour management as it employs 'Green Coats' (non-teaching staff who wear green uniforms) to manage behaviour in the corridors and in the playground, thus removing this perceived burden from the teachers.

The school's neighbourhood is relatively poor and its ethnic mix has changed considerably in the last decade with a significant increase in eastern European and African families. The Ofsted reports (2005, 2008) state that almost all students are from the local neighbourhood which has high levels of deprivation. The proportion of pupils entitled to free school meals, an index of poverty, is more than twice the national average.

The sixth form is relatively small (see Chapter 7), perhaps in resonance with the school's 'reputation' for low standards and low expectations. It has a traditional approach to pedagogy and assessment which follows closely Sir Cyril Burt's idea of a 'treble-track' system for organising teaching groups. Burt outlined his 'treble-track' approach over 80 years ago in a memorandum on 'backward children' written in 1925 (White, 2006, p. 9), and discussed in section 5.1.

Teachers have been subjected to severe scrutiny through mini-inspections of their lessons every term by the school's senior managers. As a consequence, teachers' morale has been low and some have lost their jobs through this process.

The science department is well resourced with 11 science laboratories, 3 preparation rooms and 4 smaller rooms/chemical stores attached to the larger laboratories. There are good IT facilities in each lab which include projectors and white boards. There are three full-time laboratory technicians and there is a good supply of books, chemicals and science apparatus.

Standards, as measured by the results of the national tests at age 14 and GCSEs at age 16, remained below national averages up to 2008, then dramatic increases began to happen in the results of Teacher Assessment at age 14 and the GCSE passes at age 16. This is discussed in Chapter 7.

#### **4.9.2 Downton Community School (DCS) – an ‘ideal’ case**

Downton Community School, in contrast, is a relatively new school opened in 2002. Its special features make it an ideal companion case. It is housed in an impressive new, purpose built building. It has a welcoming and expansive ‘façade’ with open access to the school reception. There is a very spacious front foyer and a cafe open to pupils, parents and the public.

The school enjoyed an ‘outstanding’ Ofsted grade during the period of my fieldwork and prided itself in having a ‘pupil-friendly’ ethos. It also had a whole-school policy on assessment for learning (AfL) and considered itself to be at the fore-front of pedagogical innovation.

In 2003 there were only 359 students (Ofsted report 2003), 180 in Year 7 and 179 in Year 8. Most students were from a White British background. The science department was well-resourced. With only 359 pupils and at that time ‘hand-picked’ staff, the school provided an ideal setting for studying marking and feedback. I therefore included the school in my exploratory study and later in my main study.

By 2008 the school had 1316 pupils in total, with 156 (11.9%) in the sixth form. Standards, as measured by national tests at age 14 and GCSE at age 16 remained below national average throughout the period of my fieldwork.

#### **4.9.3 Westfield Comprehensive School (WCS) – a ‘satisfactory’ school**

Westfield Comprehensive School provides an appropriate case for comparison with CHS and DCS. It is in Ofsted’s ‘satisfactory’ category and was on two sites about a mile apart at the time of the Ofsted inspection in 2002 when it was graded as ‘satisfactory’. At the time it had 1622 pupils aged 11-18 with only 187 (11.5%) in the sixth form, much lower than the national average. Twenty nine per cent of the pupils were eligible for free school meals, which was above the national average, indicating the degree of economic deprivation of the neighbourhood. Twenty three per cent of the pupils had English as an additional language. The Ofsted report had noted that “provision for pupils with English as an additional language was unsatisfactory” and provision was ‘barely satisfactory’ for those pupils with special educational needs.

In 2006 the school moved to a single site, and the 2007 the Ofsted report noted that nearly half the pupils were from minority ethnic backgrounds and that English was not the first language of one third of these pupils. These proportions were well above average and growing. This was also true of the proportion of students entitled to free school meals. Like Curie High Westfield Comprehensive had historically low standards and low expectations. The school ethos was less academic and it remained in Ofsted category, ‘satisfactory’ throughout the period my study.

The school is more security conscious than the other two schools, with access to the school office restricted by the use of electronically operated gates. A seven foot high spiked iron fence surrounds the school making it less welcoming to pupils, parents and visitors. The school, unlike the others, has a strong focus on pupils’ ‘behaviour’.

This has emerged (as we will see in chapter 8) as an instance of significant **misplaced gaze** by the school.

Standards, as measured by national tests and GCSE results, remained below average throughout the period of the study.

These characteristics of Westfield Comprehensive School made it an appropriate case for study in comparison with the other two schools.

#### **4.10 Summary of Chapter 4**

Chapter 4 describes the methodology and draws attention to the ‘twists and turns’ in the process of my fieldwork. My fieldwork began with an ethnographical approach intended to capture the teachers’ voice in the SA/FA debate. It was initially influenced by one of the Grounded Theory ideas, i.e. of letting data inform the journey of observation, recording and note taking etc. and moving to the next phase of observations guided by the analysis of the data already obtained. For example, my main study is informed by the findings of my exploratory studies on marking pupils’ work and in-depth interviews with teachers, head of departments, deputy head teachers and head teachers. This threw up the need for interviewing pupils in order to get their views of what the school is offering them and what kind of feedback they are receiving from their teachers.

After Phase 1 of my main study my focus changed to a more manageable workload by focusing only on Year 8, the rationale being that they were well into their secondary schooling and a year away from the pressures of the National tests. In phase two there were two schools involved, CHS and DCS. In phase 3 one more

school, WCS, was added and pupils were interviewed. Phase 4 of the fieldwork became more complex than I had anticipated. The discovery of the ‘double triage’ and the discovery that the books were not being marked at all for pupils in the bottom set required further research.

As indicated in 4.7.1 when I became aware of Critical Realism my fieldwork was at the final phase (Phase 4). However, my approach to interviewing had many features of a critical realist approach (Smith and Elger, 2014) . For example, empathy with teachers, repeated interviews allowing the scope for in-depth discussion which teachers felt enhanced their understanding of the issues.

## **Chapter 5 A Theoretical Frame work:**

### **A critique of the ideology of fixed innate ability and an analysis of structural barriers to the development of formative assessment/assessment for learning**

#### **5.1 Introduction**

This theoretical framework, as I have intimated in chapter four, is largely informed by the philosophy of critical realism, in particular three strands: (a) from the philosophy of social science, I have taken the transformational model of social activity (TMSA) and the concept of ideology; (b) from the philosophy of science and applied critical realism, I have taken the idea of open systems and the device of constructing laminated systems, of two kinds, the first defined by different ontological levels and the second by different levels of scale; (c) from the philosophy of meta-reality I have taken the model of learning as involving ‘unfolding the enfolded’.

In this chapter I present a critique of the ideology of *innate ability* and then tell a dual story which concerns the persistence of the structural barriers to the development of formative assessment/assessment for learning and the consequences of the intensification of the ‘standards agenda’.

#### **5.2 A critique of the ideology of fixed innate ability**

The ideology of ‘innate ability’ which has antecedents in the dim and distant past in the ideas of Manu (Saksena, 2000), the Hindu law-giver, and in more modern times



in the eugenic theories of Galton (1865), was transferred into the basis of schooling in England by Sir Cyril Burt (1925) and remains embedded today in state institutions and the practices in schools. Gillborn and Youdell (2000) explain how the ideology of innate ability is used to operate 'educational triage'. Stobart (2008) observes that the word 'ability' has replaced the acronym, 'IQ'.

As seen in chapter 1, what could be described as the Burt and Morant legacy – the ideology of '*innate ability*' - has been enshrined in the English school system since 1927 and has provided a pseudo-scientific legitimacy (Blum, 1978; Kamin, 1974) for a fundamentally class-based system of education, arguably geared more towards managing social stratification and the division of labour than to providing good quality education for **all**. This has provided a way of managing social inequality so that there are clear, relatively fixed, categories and ratios of people who can be fitted into allocated places within the social structure.

However, the current economic crisis, in the context of the chronic de-industrialisation of the UK, has all but displaced received notions of the division of labour. There is a significant minority of pupils in the UK who are unlikely to ever get on the employment register. It is now even more likely that those who are failed by the school system will end up on the unemployment heap with all the unpleasant consequences that this entails. Marshall (Marshall, 2013, p. 3) makes this point as follows:

“... we have a good idea of what happens to the 40% who do not achieve five good GCSEs. If you are part of this group then there is a greater than one in four chance that two years later you will be NEET, that is, not in any kind of employment, education or training. If you were one of those (nearly 4%) who had gained no GCSEs at all, there was a greater than one in two chance that you would be NEET”.

The classic work of Marsden (Marsden, 1987) *‘Unequal educational provision in England and Wales’* gives a vivid account of the inferior provision for children from poorer socio-economic backgrounds covering the period 1932 to 1987. Many ethnographic studies such as Willis’ *‘Learning to labour’* (Willis, 1977) show how a culture of lower expectations, of not being good at ‘intellectual things’, of not being ‘clever’, is inculcated in pupils from lower socio-economic class backgrounds at school. More recently Jones (Jones, 2011, p. 90) discusses the ‘demonization of the working class’ and particularly those labelled the ‘non aspiring, working class’ by New Labour and the ‘non-deserving poor’ by the Conservatives. These divisions contribute to a sense of justification for restricting educational resources for the poorest children. Within the school system, however, the same result is achieved through the mechanism of ‘ability grouping’.

Stobart (Stobart, 2008) in his book *‘Testing times: uses and abuses of assessment’* explains how the notion of ability shapes our perceptions of who we are and how, for many pupils, this perception ‘marks’ them for life. Many pupils see themselves as ‘thick’ (Allen and Ainley, 2007). The notion of ability, which has largely replaced the notion of intelligence in national, local and school discourses, continues to

display the characteristics of ‘intelligence’ as described by Cyril Burt in his 1925 Memorandum. This was included in the 1927 Board of Education Handbook for Elementary school teachers which recommended a ‘treble-track’ system:

“The ideal plan would perhaps comprise a ‘treble-track system – a series of backward classes for slow children, a series of advanced classes for quick children, both parallel to the ordinary series of standards for children of ordinary average ability.” (Board of Education 1929:422)

As described earlier, Burt’s thread of thinking continues to hold sway in national thought and ideology. This is despite the fact that this thread of thinking is at least 150 years old: its origin in the modern period lies in Galton’s article in Macmillan’s Magazine in 1865 (Galton, 1865). However, arguably, Galton’s and Burt’s ideas have roots in much more archaic thought. The Hindu philosopher Manu (Saksena, 2000), a couple of thousand years before Galton, had developed similar ideas. The difference was that he did not wish to eliminate the Sudras (the lowest caste) as Galton did for his ‘caste B’. Rather, Manu wanted ‘Sudras’ to remain subservient to and serve the Brahmins and the other two castes throughout their lives (Saksena, 2000). The ‘Manu Code’ (ibid) describes the roles and duties of the four castes: Brahmins should read and teach the Vedas (Hindu Holy Scriptures); Chhatriyas and Vaishyas should study the Vedas and be taught by Brahmins; while the Sudras, were not even expected to read the Vedas. It would be comforting to think that these ideas were quaint historical curiosities. However this is not so, they have powerful influence even today. Most working class pupils are not even expected to read at the level at which, say, the Guardian is written.

Adolf Hitler in his book, *Mein Kampf* (Hitler, 1925) referred to the idea of marriages between the so called Aryan races and others and asserted that in such a union the quality of the Aryan stock declines. Such suppositions have of course been shown to be scientifically baseless as is the whole idea that the capacity to read or engage in an intellectual pursuit depends upon the wealth or caste or ethnicity or social status of parents.

I am not arguing here that the system of schooling in England is imbued with ideas as primitive as those of Manu or as sinister as those of Hitler, but I am arguing that the thread of thinking associated with the notion of *innate ability* is connected with genetically determinist ideas; and that this notion continues to play an important role in the English education system, using practices such as banding and setting.

Many policy and government documents use the notion of ability when the word is not needed to express what is intended. In order to illustrate this point I present here five quotations, one from the Education Reform Act (ERA 1988), quoted in the 'Foreword' of 'Science for ages 5 to 16' (DES/WO, 1988), one from 'GCSE: A General Introduction' (DES and WO, 1985) and three from the GCSE guidance to science teachers (SEC, 1986, p. 16) (my emphasis in bold):

The ERA defines the programmes of study as:

'... the matters, skills and processes which are required to be taught to pupils of different **abilities** and maturity during each key stage'. **Quote 1 (ERA)**

The Act only needed to say **all** pupils instead of pupils of different abilities.

Again,

“The Government’s fundamental objectives, which ... are to improve the quality of education and to raise standards of attainment by stretching and stimulating pupils **throughout the ability range**”.

**Quote 2** (from GCSE: A

General Introduction, Section 5)

Here, ‘**pupils throughout the ability range**’ could be simply replaced by **all** pupils.

“... to raise standards across the **whole ability range**, to bring 80 – 90 per cent of all pupils up to and beyond the standard of performance then associated with the ‘average’ and ‘measure achievement of what pupils know, understand and can do. ...”

**Quote 3 (SEC, 1986)**

This statement could have said,

“... to raise standards *for all pupils*, to bring ...”

And the statement,

“... the GCSE offers to pupils of **all abilities**, to give credit for oral and practical work, and to reward worthwhile and positive achievement, even if for some pupils this is at a modest level.”

**Quote 4**

**(SEC, 1986)**

In this quote, if we replace ‘pupils of all abilities’ by all pupils, the quote would read:

“...the GCSE offers to *all pupils*, to give credit to oral and practical work, and to reward worthwhile and positive achievement, even if for some pupils it is at a modest level.”

And finally from SEC, 1986:

“All examinations must be designed in such a way as to ensure proper discrimination so that candidates **across the ability range** are given opportunities to demonstrate their knowledge, abilities and achievements: that is, show what they know, understand and can do.” **Quote 5 (SEC, 1986)**

In this quote, replacing the words ‘candidates across the ability range’ by *all candidates*, would read:

“All examinations must be designed in such a way as to ensure proper discrimination so that *all candidates* are given opportunities to demonstrate their knowledge, abilities and achievements: that is, show what they know, understand and can do.”

As can be seen, the references to ‘ability’ are not needed to make the points intended in the statements. In fact the words relating to ability are redundant. It can be argued that the references to ‘ability’ in the above documents are intended as reminders for including all children. If so, why not use the word, **all**, instead? The continued use of ‘ability’ simply re-enforces the idea of this notion as a determining factor in underpinning the possibilities and limitations on what pupils can achieve.

My argument is not that pupils do not have different abilities but that revealed ability is not an innate, natural or essential characteristic of human beings. Revealed ability

is contingent upon the many social, cultural, economic and psychological factors which determine it at a given moment. Moreover, although humans have an innate capacity to learn and do many other things like love and quarrel, magnifying and reifying this notion of ability and giving it a central place in the organisation of education is not only irrational but also potentially harmful because it tends to hinder learning for many by denting their self-esteem (Ireson, 2000). In this way, it lowers the expectations pupils have of their own capacities. It also lowers the expectations others have of them, engrossing them in a **vicious** cycle of low expectations, low achievement and continued poverty which for Bourdieu, for example, amounts to continuing material and symbolic violence.

### **5.2.1 Ability, attainment and performance**

In the literature on assessment concepts such as potential, ability, attainment and performance abound and although they mean different things, they are often confused or used synonymously. In addition, the term ability is also often used to mean innate ability or fixed ability. Of course, in terms of the potential of a newly born child, undoubtedly, there will be innate potential and when one considers ability, innate ability. Even in an ideal world children will differ in levels of attainment generally, and in specific subjects or disciplines. So, for example, Kant had a great ability in philosophy, Mozart in music, Leonardo da Vinci in both art and science, Einstein in physics and so on.

### 5.2.2 The Potential – Performance spectrum

There are a number of concepts which need to be considered here, including potential conceived as abstract ability, then (concrete, developed) ability, capacity, capability, attainment and performance. In fact one could call this a potential – performance spectrum. On this spectrum all the terms may be assumed to have a genetic component, but as one moves from pure potential on the left to revealed performance on the right the significance of the genetic component will reduce and the significance of the social (or more generally non-genetically determined, i.e. acquired or formed, whether by nature or society) will increase.

The distribution of the genetic and the social components are critical areas for considering. Little is known about the distribution of the genetic component (it may or may not prove to be very uneven and unequal), but we do know that the distribution of the social component is hugely uneven. Measuring the components of the spectrum is, of course problematic. For example, we can't say that performance is necessarily a good measure of attainment because performance depends upon many other factors (some of which have been addressed by David Scott in his critique of the PISA performance tables (Scott, 2011)). However, we can sometimes say that a measure is inappropriate, for example, the use by the case study schools of performance in *English* for banding in *science*.

This thesis holds that from the standpoint of the Bhaskarian concept of learning as 'unfolding the enfolded' everyone has, effectively, infinite potential as a learner. It also argues that ideologically there is a persistent tendency to reduce terms in the potential – performance spectrum to the left hand side, especially to a portmanteau



(‘catch all’) concept of ability, and to reduce the socially formed component to an assumed genetic component; so that all degrees of manifest attainment and performance are reduced to and explained in terms of assumed *innate ability*.

It is important to appreciate that attainment and performance are different concepts; attainment is what has been arrived at in terms of knowledge and skills, and, performance is what one shows or demonstrates on a particular occasion. A newly born baby has enormous potential but one wouldn’t say that it has ability except in relation to ability to learn. So, to say that someone has ability normally presupposes that the person must be capable of exercising their potential in some way. The term capacity is also used to mean ability and to mean potential. Another term, capability, is also found in the literature and means more than capacity. For example, one might have the capacity to learn Mandarin but might not have the capability of doing so because capability means that you also have the time, the money and a tutor and so on.

A big issue is of course the question of the extent to which properties such as potential, ability, attainment and performance are derived from an innate or effectively genetic component. It would be wrong to suppose that there is no genetic component in say potential or ability or even attainment and performance. As one moves along the spectrum we assume a greater degree of actualisation of the potential, of the infinite possibility. When one talks about attainment or knowledge, one is talking about what has been arrived at.

Summative assessment mainly measures performance but it is designed to capture attainment. However, as mentioned earlier, the genetic aspect is not the only component of attainment. There is a significant component which is socially

acquired. The nature versus nurture debate is of course an old one, which I do not intend to enter into in this thesis. However, what I am concerned with is the social component which makes a huge difference in education. It is therefore significant that the social conditions or states which are beneficial for learning and in education in general are inequitably distributed.

For the Critical Realist tradition, ‘ability’ is always at least partially a social product, dependent on anterior social products. Likewise, performance is even more obviously a product. Furthermore, by Bhaskar’s conception of *unfolding the enfolded* (Bhaskar, 2011a, chapter 11) every learner always has immense potentialities, no matter how the learner has been hitherto formed and whatever his or her genetic endowment or social circumstances. Everyone is a ‘concrete singular’ and they have infinite potential. The job of the school and the teacher is to bring out the ‘genius’ in everyone.

The argument of this thesis is that in the English education system the concept of ability has been conflated with *innate ability* so that the socially determined part of ability has been neglected or undermined. The conclusion of this argument has been that the measures used to select pupils for the tri-partite system of education did not reflect their potential, ability or attainment rather they largely reflected their access to resources and differences in their socio-economic status. Many authors have drawn attention to this seriously debilitating concept and the way it has been used to ‘ration’ education (Gilborn and Yeodell 2010, Benn & Chitty, 2011). In this thesis I have researched the role played by the notion of ability interpreted or assumed as

*innate ability*; its influence on official discourses, school structures, school discussions and classroom practice. I have made references to the continued use of the term ability, generally taken to mean assumed *innate ability*, in government publications which lend credence to the tripartite system of schooling and notions of streaming and setting. I have focused mainly on how this concept operates at school and classroom levels, particularly in relation to formative assessment, expectations and pupils' sense of self-worth.

The ideology on which the tripartite system of education was based had asserted that children could be divided into three ability bands; able, with some ability and less able. This was Sir Cyril Burt's formulation in his article on 'backward children'. This ideological structure has been reproduced over decades and although tinkered with it still exists within the so-called 'comprehensive' schools. It has been argued, for example by the DHT of Curie High School, that dividing pupils into sets on the basis of roughly equal attainment levels makes pedagogical sense in relation to the pitch of lessons and the curriculum to be covered. It also makes sense according to the DHT because it is perceived as an easier task for teachers to teach pupils organised into such sets. However, what the DHT did not reveal is that, in practice, for the lower sets, it restricts access to specialist teachers, good books, a balanced science curriculum and, importantly, high expectations. This is in practice a legacy of the tripartite system of schooling which was grounded in an ideology that you could divide children into three distinct groups according to ability. The 11+ examination was used to select the most 'able', around 20%. In general, the exact interplay of genetic and social components of ability was never discussed but it was assumed that balance reflected *innate ability* more or less.

Banding in a secondary school plays a crucial part in allowing or restricting access to educational resources. Thus, groups are more likely to reflect social class in which the single most important factor is some combination of variables relating to parental wealth. Once the banding has been decided, as soon as the pupil enters secondary school, her/his fate is practically sealed in relation to access to resources and the curriculum.

### **5.3 Structural barriers to FA/AfL and the consequences of the ‘standards agenda’**

This section tells a dual story in considering the structural barriers and their role in the reproduction of what I will call the factory-model of schooling and the intensification of the pressures for summative assessment under the regime of the ‘standards agenda’. As previously discussed there are three main moments of the ‘standards agenda’; the first moment beginning with the ‘black paper’ writers, second beginning with Callaghan’s 1976 Ruskin College speech; and the third moment beginning in the 1990s with the imposition of the National Curriculum, coinciding with an intensification of the process of globalisation, resulting in a regime characterised by what I propose to call ‘double triage’ and an acute crisis in morale and motivation of school science teachers.

At least eight interconnected issues relating to barriers to the implementation of formative assessment/assessment for learning in school science have emerged from the discussions of the literature on assessment:

- the ideology of ‘innate ability’, the **vicious** circle of low expectations and the process of ‘triage’;
- the ‘standards agenda’ and the misplaced gaze that results from it
- the great extent to which summative assessment is deeply embedded in the social organisation and culture of schools;
- the shortcomings of the concepts of FA and AfL as articulated in literature;
- inadequacies in science teaching, including the re-emergence of a ‘two cultures’ education - ‘science for the scientist and science for the citizen’;
- the limited extent to which teachers have engaged with professional learning;
- the complicity of the agency of teachers in the process of ‘triage’, and the resultant process of ‘double triage’.

It is generally agreed that teachers’ involvement in the debate on FA/AfL and SA is essential and that teachers’ role is central to the uptake and the development of FA/AfL. However, the challenges facing teachers in this regard are complex, as Black and Wiliam pointed out, in *Inside the Black Box* (Black and Wiliam, 1998, p. 15):

“...the improvement of formative assessment cannot be a simple matter. There is no ‘quick fix’ that can be added to existing practice with the promise of rapid reward”.

They asserted that improvement would only come about if each teacher found his/her own ways of incorporating the ideas of FA in their classroom work. However this assertion, in my view, risks individualising a problem which is multi-layered and involves deeply rooted practices and structures. I agree with authors such as Black

et. al (Black et al., 2003), Spendlove (2009) and Lambert and Lines (2000) among others, that formative assessment has to be within the control of the individual teacher. Nevertheless, given that SA is deeply embedded, it has to be asked whether teachers, acting on their own as agents of change, can bring about a transformation of summative assessment in the direction of formative assessment.

For the convenience of the reader I repeat here the quote from Johnston et al (Johnston et al., 1995, p. 361) referred to above:

“Most of the teachers in this study [*Johnston et al. 's study*] were caught in conflicts among belief systems and institutional structures, agendas, and values. The point of friction among these conflicts was assessment, which was associated with very powerful feelings of being overwhelmed, and of insecurity, guilt, frustration, and anger ... [*our*] study suggests that assessment, as it occurs in schools, is far from a merely technical problem. Rather, it is deeply social and personal”.

### **5.3.1 Teachers as agents of change**

In considering teachers as agents of change some important questions arise:

- Can teachers be effective agents of change? More specifically, to what extent can they bring about a transformation from summative assessment to formative assessment through their individual (or collective) praxis?
- To what extent does their praxis (including their potentially transformed praxis) need to be complemented by changes at other educational and social levels?
- What are the issues involved and how might the agency/structure dynamic operate in this context?
- What transformations would this require in teachers' own praxis?
- To what extent and how can teachers become ‘transformed transformative’ agents?

The work of Gillborn and Youdell (2000) casts doubts on teachers' capacity to transform the assessment system which underpins schooling in England, because they are themselves implicated and are complicit in sustaining the practice of 'educational triage' - a system of schooling which is repressive, unequal and unfair. I discuss 'educational triage' in section 5.3.4 below.

Gillborn & Youdell imply that teachers' complicity is forced upon them through coercive accountability procedures dictated by international competition. This is true particularly in the third moment of the standards agenda. However, teachers had low expectations of working class pupils and used ability groupings for 'rationing' access to educational resources much before the 'standards agenda' came into effect. So, based on the evidence from practice, teachers' complicity, willingly or unwillingly, is evident. Teachers will need to reflect on this if they are to become 'transformed transformative agents' of change.

### **5.3.2 The Transformational Model of Social Activity (TMSA)**

In the section 'On the Society/Person Connection' in *The Possibility of Naturalism* (Bhaskar, 1979/98) Bhaskar develops the transformational model of social activity. He argues that as society pre-dates any individual and any round of human agency, it is not true to say that human agents 'create' society. One must rather say that they 'reproduce or transform it'. Society effectively 'pre-exists the individual', and all human activity presupposes 'the prior existence of social forms'. "If we consider 'saying, making and doing as characteristic modalities of human agency' then all of

these require making use of existing social forms.” Thus, he argues that ‘the necessary pre-existence of social forms’ suggests ‘a radically different conception of social activity from that which typically informs discussion of the society/person connection’. He suggests an Aristotelian conception of social activity ‘in which the paradigm is that of a sculptress at work, fashioning a product out of the material and with the tools available to her’. He calls this ‘the transformational model of social activity’.

Bhaskar emphasises the importance of ‘distinguishing categorically between people and societies, and correspondingly between human actions and changes in the social structure’ because although social structures depend upon human activity, they characteristically have very different ‘properties’ from those possessed by individuals. Thus human actions may be characterised by the properties or qualities of ‘purposefulness’, ‘intentionality’ and ‘self-consciousness’ but these are not characteristics of social structures, nor are human actions necessarily immediately manifest as changes in social structures.

The conception of social activity which Bhaskar proposes is that ‘people, in their conscious activity, for the most part unconsciously reproduce and occasionally transform the structures governing their substantive activities of production’. In other words, we as human beings engage in day-to-day ‘first order’ activities and in doing so, through our praxis or activity, reproduce or transform social structures (at a second order level). Of course, saying that, for example, teaching is indeed a ‘first order’ activity for the teachers is consistent with both the teacher just viewing the job



instrumentally, simply as a means to an end (where the end is for example, the salary that comes with the job, sustaining the biological and material well-being of the teacher) and seeing it as possessing intrinsic value, as a job worth doing for its own sake.

We should note that in much of the literature on formative assessment there is an emphasis on influencing and persuading teachers to adopt FA/AfL in their practice. Thus Black et al. address their appeal to teachers and policy makers. However, they do not fully explore the structures and other aspects of the social world which govern and maintain the embeddedness of summative assessment and inhibit the development of formative assessment.

Spendlove (Spendlove, 2009) sees teachers as ‘informed professional decision makers’. Wiliam (Wiliam, 2011) asserts that only ‘good’ teachers can bring about this transformation and argues that in order to be good or better they will need to engage in professional learning. Most other authors concur with the argument that teachers will need to engage fully with professional learning in order to develop and use FA and/or AfL (Lambert & Lines, 2000, James & Pedder, 2006). However, will professional learning alone suffice? Will not the teachers also need to confront their ‘mindset’ (Dweck, 2007) and reconsider their ‘beliefs and perceptions’ in relation to their pupils?

I would argue that even with an increase in their professional learning teachers acting *alone* cannot implement a shift from summative to formative assessment. It can also be maintained that the legacy of Burt’s 1925 paper on the ‘backward child’ continues to haunt educational thinking and educational structures in 21<sup>st</sup> century

England – a legacy which remains a barrier to promoting learning for **all**, and which will need to be challenged and overcome.

Whilst teachers acting alone, as individuals, cannot transform the practice from SA to FA collectively they can initiate many changes in their own classrooms. Indeed in principle, if they were sufficiently determined, they could transform the practice because everything else depends on them – they are effectively like the element of the necessary deep infrastructure composed of qualities such as trust in the social field of education. One cannot envisage school managers teaching all the children, they would still need to use teachers.

Whilst it is indisputable that teachers will need to be actively involved in bringing about such a transformation, it must be asked whether this approach overestimates the capacity of teachers as agents of change to overcome all of the constraints placed on them by existing educational and social structures. In order to explore the constraints facing teachers I discuss below ‘open systems’ and the idea of ‘a laminated system’

### **5.3.3 Open systems and the idea of a laminated system**

Schools are open systems and school assessment is a complex process (see section 2.5). Barriers to formative assessment/assessment for learning are multiple, complex and related in various ways. In open systems, phenomena are generated by a multiplicity of mechanisms, often at radically different ontological levels (Bhaskar et al., 2010). This means that more than one mechanism will almost always be needed

to explain the phenomena. Thus, in order to study the phenomena of the barriers to FA/AfL I have made use of Bhaskar's notion of a laminated system (Bhaskar et al., 2010). Bhaskar developed the concept of a laminated system as a device to fix the idea of the irreducibility of different mechanisms in the formation of an open systemic phenomena or event, bearing in mind the constant tendency of empiricist and empiricist-informed work to revert to an actualist and, therefore, mono-disciplinary reduction (Bhaskar and Danermark, 2006).

There are at least four main types of laminated systems.

- (i) The first type of laminated system is the one composed of different ontological levels as elaborated by Bhaskar and Danermark (Bhaskar and Danermark, 2006) and Gordon Brown (Brown, 2009). It was initially introduced as a device especially created to take account of the complexity of open systems and the emergent character of mechanisms at some of the ontological levels involved in their explanation.

- (ii) The second type of laminated system is that of the four planar social being.

This is the idea that every social event occurs simultaneously on each of four planes: (a) material transactions with nature; (b) social interactions (of those communicative and other e.g. material kind) between agents; (c) social structure, conceived as distinct from agency but which agency reproduces or transforms – this will include everything from language

and knowledge to economic and political systems; and (d) the stratification of the embodied personality.

- (iii) A third kind of laminated system is that provided by a hierarchy of levels of scale. In particular, Bhaskar uses the idea of seven levels of scale to develop his critique of prevalent positions on methodological individualism and collectivism or holism and in his discussion of relationism (Bhaskar et al., 2010, p. 9).

He explains the constitution of these seven levels of scale as follows. The first level is sub-individual. This is the level at which we talk of motives and the level at which Freud talked of the unconscious. The second level is that of the individual. This is the level at which a novelist typically writes. The third level is what Bhaskar calls the micro level. This is the level of small-scale interaction as studied by, for example, ethno-methodologists. The meso level is the fourth level and is characteristic level of classical sociology. It is the level at which Marx talks about the capital-labour relationship. Weber and Durkheim also pitch most of their analyses at this meso level. The fifth level is the macro level, typically the level at which we talk about the properties of nation states or equivalent entities, such as the UK economy. The sixth or mega level is that of geo-historical trajectories and spans such as feudalism which cover whole swathes of space and time. The seventh and final level is global i.e. that of the planet as a whole.

- (iv) A fourth kind of laminated system is that provided by overlapping spatio-temporalities. An example of this is given by a case of intersecting times, where for example, a single event such as the opening of the UK's parliament by the Queen operates with components or aspects emanating from several different spatio-temporal epochs ranging from the very contemporary, (e.g. the Queen's speech being written by the Prime Minister's press officer), relatively modern (the institution of parliament), to the very old (e.g. the monarchy going back several hundred years). Or one could take another example, a case of spatio-temporal coincidence. Looking at a street in New Delhi, one can see camels and bullock carts, bicycles, cars, buses and, overhead, aeroplanes – all at the same time.

In this context we can 'ponder the extent to which, although we may live **for** the future, we live, quite literally, **in** the past' (Bhaskar, 2008a, p. 54).

In a school environment we do live quite literally in the past. Practices and paradigms which existed long ago survive today. The practice of external examinations (i.e. summative assessment) which has its beginnings in the meritocratic procedures of the Chinese for the selection of their top civil servants some two thousand years ago, remains embedded in schools in England even today.

### 5.3.4 Educational Triage

Gillborn and Youdell define ‘educational triage’ as ‘the rationing of educational opportunity’, and declare that this is operating in the secondary school system in England. They explain (Gillborn and Youdell, 2000, p. 199) that in medicine the principle of triage is applied for ‘sorting and prioritising those in greatest need’ and in a medical emergency ‘triage’ is used ultimately to direct effort towards saving those who could survive with medical help and allowing others less hopeful cases to die. They point out that in relation to education, such a triage operates ‘systematically’ so as to lead to the neglect of some pupils in order to devote resources to those who are deemed more likely to achieve the league table GCSE grades, A\* to C. They argue that this process, perhaps designed to maximise effectiveness of scarce resources, in practice affords privileges to some pupils based on features such as ethnicity and parental income, socio-economic status etc.

They argue that although the 11+ examination is not required for entry to a comprehensive school it remains operational in the guise of ability testing post-entry to the school. Thus, selection by *ability* remains the instrument used to operate the ‘educational triage’.

Stobart (2008) concurs with Gillborn and Youdell that the word ‘ability’ has replaced the word ‘intelligence’ in discourses in and about schools and that it carries the same meanings.

A note about Gillborn and Youdell's description of 'medical triage' seems appropriate here. In recent years it appears that this 'rationing' of medical treatment (i.e. medical triage) has now become a routine procedure in the National Health Service. In 2011 I was told that my painful big toe was not a priority for treatment at the local 'foot clinic' and was refused an appointment. On 28 February 2013, at the inquest of baby Peanberg King, the St. Pancras coroner, Dr Shirley Radcliffe said to Dr Shantikumar, the representative of the Harmoni medical service 'which has contracts covering eight million patients':

"If you are going to take on a task of **triaging** a seven-week-old baby, you have to do it responsibly" (Dunne, 2013). [My emphasis]

### 5.3.5 Assessment 'shapes' people

Stobart (2008) presents an up-to-date and direct critique of the notion of *ability* prevalent in much of school assessment, and argues convincingly that assessment in the form of tests and examinations is very powerful in that it shapes how individuals understand themselves. He states (page 6) that:

"Assessment shapes who and what we are and cannot be treated as a neutral measure of abilities or skills that are independent of society."

Stobart continues discussion on the 'myth' of ability (Stobart, 2014, p. 10) and argues forcefully that it is not ability but opportunities which count. He states:

"We don't talk about IQs in schools because of the doubtful past of IQ testing and its exponents. However, we are willing to talk about ability, aptitude and the gifted and talented as if these are very different concepts."

He gives the example of the Cognitive Ability Tests (CATs), a commercially produced test, taken by over two-thirds of 11-year-olds in England, on entry into secondary schools. He explains that CATs is a ‘repackaged intelligence test, with verbal, non-verbal and numerical sections’. It provides ‘ability scores’ and predicts how well a pupil should do at age 16 in GCSE examinations, and becomes the basis for tracking pupils’ progress. On one level it can be argued that there is nothing wrong in using CATs because the test result can give a good indication of developed cognitive ability as a result of how well the pupil has responded to education, but it can easily be ‘misinterpreted’ as a ‘measure of fixed ability – the underlying cause of their educational achievement, rather than a product of it’ (Stobart, 2014, p. 10).

Hart et al. (Hart et al., 2004) have shown in their book, *Learning without Limits*, how ‘ability labelling’ acts as an active force in schools and classrooms, helping to ‘create the very disparities of achievement that it purports to explain’.

### **5.3.6 Applying TMSA and the double articulation of complexity**

In considering the complex nature of the barriers to teachers moving away from summative assessment practices towards formative assessment, I use here the device of laminated systems as described in section 5.3.3. I begin with a laminated system organised in terms of seven levels of scale (the laminated system of the third type), and then at the meso level I employ a laminated system composed of different ontological levels. To a lesser extent, I will also use laminated systems of the second (four planar social being) and the fourth, the seven scalar laminated system which is



composed, as we have seen, of structures and mechanisms operating at, at least the following levels:

- global,
- mega,
- macro,
- meso,
- micro,
- individual
- sub-individual levels.

I discuss the global/international first.

### **5.3.7 Global/International factors/issues/influences**

James Callaghan's Ruskin College 'standards' speech in 1976 (chapter 1, section 1.1) was as we have seen, at least in part, a reaction to the 'black paper' writers. But he was also clearly motivated by the idea of economic competition, particularly with Germany. Was the then Prime Minister suggesting that the main motivation for learning should be enhancing economic competitiveness? Certainly, it seems obvious that there could be a better motivation for learning e.g. that of promoting the life chances and well-being of the pupils. The motive of economic competitiveness would suggest agonistic preparations emphasising rigorous imposed discipline, competition against others and winning. The motive of promoting well-being would

suggest internal aspirations and self-discipline, higher self-expectation and learning both for its own sake and as means for improving the quality of being and life, and in doing so, contributing to the progress and well-being of society. In other words, one might want to posit that a really good teacher or a person interested in education should be concerned with education and learning for their own sakes as well as a means to improving life and society.

Stephen Ball (Ball, 2008, p. 1) draws on ex-Prime Minister Tony Blair's pre-election speech in 1996 to state that education has become a 'crucial factor' in the context of 'informational capitalism':

"Education is now seen as a crucial factor in ensuring competitiveness in the context of 'informational capitalism'. In other words, education policy is increasingly thought about and made within the context of globalisation."

David Scott (Scott, 2011, pp. 14-15) describes how the Programme for International Student Assessment (PISA) influences national policies on curriculum and assessment. He states that:

"PISA is a performative device, in so far as the intention is not just to describe the skill/dispositions of children but to promote and thus contribute to national policy-making. Certain forms of performative knowledge become the norm. The instrument for measuring knowledge and skill levels of children becomes an instrument for determining what those knowledge levels and skills should be, and how they should be learnt".

Thus it is clear that there is a strong link between international competition (and the economic factors driving it) and globalisation on the one hand and on the other systems of assessment. That the competition is now truly global is borne out by the 2013 New Year's message of the British Prime Minister David Cameron, in which he emphasised the need to compete with China and India. When James Callaghan was Prime Minister the main competitor was Germany. Now, it is the emerging economies like Brazil, Russia, India and China (often referred to as the BRIC nations) and the competition is truly global.

This typically reinforces pressures to utilise summative assessment and assessment according to a school culture of grading and setting, responding to the needs of the international division of labour.

### **5.3.8 Mega level**

This is the spatio-temporal level at which we can observe practices and structures which existed more than a millennium ago continuing to be efficacious today. An example, as mentioned in section 2.2.1, is the use of competitive examinations for the selection of the Mandarins (the members of the elite Chinese Civil Service) in China more than 2000 years ago. In the relatively recent past, in the UK, the first competitive examination for the selection to the British Civil Service took place following the publication of the Northcote-Trevelyan report in 1853.

### 5.3.9 Macro level: National historical, social and political contexts

The summative tradition has a long history and is embedded in the English school system, and the processes of history have deposited the embedded structures and practices which we inherit. Thus it is important to recognise that we are dealing with the presence of the past (Bhaskar, 2008b). Replacing patronage by competitive examinations for entry to the British Civil Service was transformational to the consolidation of the British Empire and to ensuring the supremacy of the British economy: more generally, the institutions of external examinations were pivotal in the rise of meritocracy. Thus they are typically perceived as being ‘fair’. National policy in the last two decades has reinforced this perception by imposing national tests at ages 7, 11 and 14. However, this has further added to the perception that teachers’ own assessments cannot be entirely trusted and that their assessments lack rigour and reliability. Teachers themselves are all too aware that their assessments lack public acceptability.

The Norwood Report had recommended that teachers should undergo in-service training (Norwood, 1943) lasting for up to seven years, to enable them to assess the ‘whole’ achievement of a pupil. Despite this and numerous other reports before and after it, including Acland’s (Acland, 1911), warning about the limitations of external examinations, the tradition of examinations and tests has thrived and today pervades every aspect of school life. Moreover, for over a hundred years, this tradition has been instrumental in ‘shaping’ a learner’s identity by ‘labelling’ her/him as more or less ‘able’, on the basis of what can best be described as a pseudo-science based notion of *innate ability* emanating from the works of Galton and Burt.

### **5.3.10 Meso level: the school and the neighbourhood**

At this level I analyse the relevant mechanisms by using a laminated system of the first type, defined by different ontological levels. I therefore analyse both the school and the neighbourhood by looking at physical, biological, psychological, socio-economic, socio-cultural and curricular levels.

#### **5.3.10a The School**

In considering the school, these levels are contextualised as:

- (i) Physical – location, buildings, playing fields
- (ii) Biological – nutrition, hunger, puberty
- (iii) Psychological – expectations, encouragement, pastoral care, behaviour
- (iv) Socio-economic – resources (including teachers), libraries, classrooms, laboratories
- (v) Socio-cultural – language, expectation, attitudes to learning
- (vi) Curricular – Triage, unequal access, a compartmentalised curriculum

This analysis, however, first requires a consideration of the general concept of the learning environment of the school.

At the school level the senior managers are required to carry out self-evaluation of their schools as part of the ‘school improvement’ process. In this way they are locked into an ethos and ideology propagated at the national level. Given this, one question that arises is ‘Does the school provide a learning environment conducive to learning such that it facilitates the development of FA/AfL?’

In seeking to answer this question I first refer to Black et.al (Black et al., 2002, pp. 14 - 24). Reflecting on insights gained through the King's Medway and Oxford project (KMOFAP), they discuss notions of learning and the learning environment. They identify two crucial factors for learning to take place: active participation of pupils in 'creating their own understanding' and the involvement of pupils in peer – and self - evaluation. Teachers are expected to use AfL techniques such as 'good questioning' and 'good feedback', listening to pupils and acquiring up-to-date 'guild knowledge'. They state that:

“Our experiences in the project [KMOFAP] all point to the need to rethink a teacher's core aim – enhancing pupils' learning. To achieve this calls for a willingness to rethink the planning of lessons, together with a readiness to change the parts both teacher and pupils play in supporting the learning process”.

For the above activities and processes to happen they suggest that 'an appropriate learning environment' is split into two sections:

1. Principles and plans
2. Roles and expectations

Principles and plans require:

- careful forethought in planning to improve teaching actions
- care for the quality of responses that teachers make when engaging in dialogue with pupils or when marking their homework.

At a deeper level, they suggest that a learning environment has to be 'engineered' to involve pupils more actively in the lesson.

Under 'Roles and expectations' Black et al state that:

“It is one thing to plan new types of classroom activity; quite another to put them into practice in ways that are faithful to the aims that they were developed to serve. Here there are no recipes for all to follow in a uniform way.”

There are two aspects to this. One, in an open system where every situation is potentially a different one, we can't describe action adequately by mechanically applying algorithms. This means that we will have to use a type of machine that will be very intelligent and sensitive, that is, the human brain! The second point is that human creativity is what we actually need because there will be times when the situation is so radically different from anything experienced before that a completely novel response is required. That is why we require human creativity. Chomsky drew attention to the power of language as something we routinely draw upon to generate new sentences, that is, sentences we have never uttered before. So it is a remarkable thing. Even a person with a very limited vocabulary will generate new sentences in these situations. Teachers can use this capacity for language creatively in their classrooms.

Black et al continue:

“*Inside the Black Box* (Black, 1998) was clear in stating that the effective development of formative assessment would ‘only come about if each teacher finds his or her own ways of incorporating the lessons and ideas’ of formative assessment into her/his own classroom repertoire”.

They further state that the ‘learning environment envisaged’ would require a change in the ‘classroom contract’ between the teacher and the pupil.

For pupils it would mean that they would need to learn to be active not passive in lessons. For teachers it would require ‘courage’ to take on a new approach.

In addition, teachers will need to collaborate with colleagues and for this ‘**innovation**’ (i.e. FA) to succeed, support from school management will be essential.

In summary, Black et al make the following points in relation to the learning environment:

- active participation of pupils in ‘creating their own understanding’;
- involvement of pupils in peer– and self-evaluation;
- a change needed in the ‘classroom contract’ between the teacher and the pupil;
- a need to engage pupils in dialogue;
- a learning environment would need to be ‘engineered’ to engage pupils in active participation;
- a need for teachers to collaborate with colleagues.

In viewing FA as an ‘innovation’, Black et al appear to present a complete contrast to H.D. Black’s (1986) view that FA was an integral part of teaching not an innovation. H. D. Black maintained that FA is a part of good teaching and most



teachers engage in it at an informal level and in a less systematic manner. It would appear that in Black et al.'s view of 'innovation' the concepts and activities of FA are constructed externally in academic institutions and are offered to teachers for implementation. This is problematic for two reasons: firstly, it is one thing to say that the way teachers need to transform their practice (in terms of developing their best existing practices in the ways which benefit all their pupils) is to take on board ideas of FA and AfL; it is entirely another to mechanically apply a formula handed down from either the Ministry or the university academics. Secondly, describing FA as an 'innovation' underplays the fact that there has always been an element of FA in the process of teaching. By disassociating this concept from the normal practice of teaching there is a risk of undermining the role of the teacher in developing FA in their classroom by building on the element of FA in their current practice. However, for the widespread or systematic use of FA teachers will need to change their current assumptions about pupils' capacity to learn. They will also need to critique their current attitudes to professional learning. In other words, teachers will need to re-examine their adherence to the notion of *innate* ability and will need to look inwards towards their own professional needs in order to help 'unfold' what is 'enfolded' in their pupils.

Therefore, a **full** implementation of FA in practice would certainly be an innovation and viewed from this perspective there is no contradiction between the views expressed by HD Black and Black et al.

References to 'active participation' by pupils, the changes in 'classroom contract' between teachers and pupils and 'engaging in dialogues' with pupils, place Black et

al.'s arguments in the 'Student-Centred (social constructivist)' Learning Environments category in Jonassen and Land's scheme of categorising learning environments (Jonassen and Land, 2000). Constructivism's central position, as Brown (Brown, 2009) puts it, is the individual or social construction of knowledge. He uses the following quotation from Land and Hannafin (Land and Hannafin, 2000) to illustrate this point:

"... grounded constructivist learning environments, therefore, support individuals or groups attempting to negotiate multiple rather than singular points of view, reconcile competing and conflicting perspective and beliefs, and construct personally relevant meaning accordingly."

This is a judgemental relativism which Critical Realism rejects. The second objection according to Brown (*ibid*, 23) is that constructivists consider knowledge as the 'private thoughts of learners' but tend to ignore the public 'character of their concepts'. According to him "... learning environments are tiered, layered or laminar systems, where learning is an emergent property with multiple and tiered determinants."

Other authors, notably Michael Young (Young, 2008) and Leesa Wheelahan (Wheelahan, 2010) have also put forward strong arguments against social constructivist relativism and have emphasised the need for a knowledge-based curriculum for all. In the context of science, David Perks (Perks, 2006, p. 9) asks the question, 'What is science education for?' and discusses how Millar and Osborne's (Millar and Osborne, 1998, p. 12) distinction between science for the citizen and

science for the scientist is unhelpful. It has led to ‘scientific literacy’ courses intended for most pupils, who would probably be drawn from the middle and bottom sets in the school. He argues that ‘scientific literacy’ offers science without scientific knowledge and it is a ‘patronising approach towards young people’. He states that many educationalists have succumbed to the lure of ‘making science relevant’ because it is argued that school children cannot engage with concepts and ideas outside their immediate experience. Perks puts this argument as follows:

“School children ... cannot be expected to engage with concepts and ideas beyond their immediate frame of reference: educationalists, therefore, should relate the subject they teach directly to the language and ideas with which pupils are familiar in their everyday lives.”

He states that it is not only educationalists who have been misled in this way, others such as the Church of England and the Duke of Edinburgh’s Award scheme have also fallen for it. The former announced plans in 2005 to hold services everywhere ‘from skateboarding parks to pubs and cafes’ and then later ‘came up with slogans such as ‘NE14 Fun?’, ‘Wanna feel Gr8?’ and ‘Bored? U Wont B’ in an attempt to copy ‘teenage street speak’.

Perks argues rightly that it is important to give young people “... the fundamentals of a scientific understanding of the world that will stand them in good stead whether they pursue science further or not.”

I am now going to consider Brown’s (Brown, 2009) description of the ‘learning environment’ as layered or laminated where ‘learning is an emergent property with multiple and tiered determinants’. Brown refers to the World Health Organisation’s

conception of the human being (for health purposes) as ‘bio-psycho-social’ (WHO, 2001). He draws also on Roy Bhaskar and Berth Danermark’s (Bhaskar and Danermark, 2006) laminar conception of disability as a complex condition, requiring analysis in terms of a multiplicity of mechanisms situated at different ontological levels, for instance, physical, biological, psychological, psycho-social, socio-economic, socio-cultural and normative. Brown argues that learning environments are not only the physical construction of the classroom or the school building. He suggests that it is:

“... a semi-permanent, often episodic, complex ensemble of causal mechanisms that enable and constrain learning.” This is an example of the use of the first type of laminated system (see 5.3.3).

Brown therefore discusses five characteristics or properties of learning environments:

- Curricular
- Socio-cultural
- Psychological
- Biological
- Physical

I have added the sixth, the socio-economic level (which it must be said is in the spirit of his analysis).

Brown notes that the school is an open or at most semi-closed system so that there is much to be gained by rejecting the assumptions of the traditional (existing) view that schools and classrooms are closed systems. He stresses the learning environment is

partly but not wholly linguistic, and that the creation, reproduction and sharing of meanings are core activities of the learning environment. He also maintains that the learning environment is moral /political, in that it reflects decisions about what should be done.

Criticising both the dominant models/theories of learning, namely, objectivism and constructivism, Brown suggests that both have flawed philosophical bases. Brown argues helpfully that educationally, objectivism ‘captures’ the external (objective) meaning of the curriculum content to be learned, but fails to capture ‘the mental and social construction of meaning’. Constructivism, on the other hand, captures the work done socially and individually in constructing knowledge but fails to provide criteria for judging its ‘weight’ relative to external knowledge.

Objectivism captures elements of, in critical realist terms, the intransitive dimension while constructivism captures elements of the transitive dimension. The intransitive dimension, of course, is referring to the independently existing object of study whereas the transitive dimension is referring to the social process which studies it. According to critical realism there is no incompatibility between ontological realism, epistemic relativity (that is, accepting that what counts as knowledge is going to be variable over time and between contexts) and judgemental rationalism. This means that even if all our knowledge is relative and socially conditioned, we still have grounds for preferring some propositions, some ways of understanding the world, over others.

The main argument for this framework is that it is designed to enable us to assess whether schools as ‘autonomous’ organisations can, and if so how they can, create an environment conducive to learning such that it is enabled by formative assessment,

Barriers to ‘formative assessment’ and/or ‘assessment for learning’ may be identified as operating on all of these levels, that is on physical, biological, psychological, socioeconomic, socio-cultural and curricular levels.

### **5.3.10b      The Neighbourhood:**

The neighbourhood can also be studied at each of these six ontological levels, namely,

- Physical            -            Economically impoverished, poor housing
- Biological        -            Fast food outlets, hungry children
- Psychological -            depressing environment
- Socio-economic-            high unemployment
- Socio-cultural -            lower ambition
- Normative/curricular-            lower educational expectations

Schools and neighbourhoods are the sites of both **vicious** and **virtuous** circles.

Poverty in the neighbourhood, for example, may give rise to poor nutrition or a lack of own space to study at home. This contributes to poor test and examination results, and leads to pupils being classified as having low ability, manifested in being placed

in lower sets. This in turn leads pupils to have low expectations of themselves which leads to a lack of confidence and even worse results and then worse assessments of ability, as the pupils are grouped and graded into the lowest categories. They end up going back to their neighbourhood very much as they came from it, with their potential barely developed. This is a **vicious** circle. At the other extreme, you might have the children of wealthy parents who may each have their own study space, they go to expensive schools where they have a lot of individual attention and may be extra tuition, their self-confidence is built up and they end up becoming high earners, perhaps even a mini-Cameron or a Clegg clone – that's the **virtuous** circle. (Mr Cameron and Mr Clegg are currently the British Prime Minister and Deputy Prime Minister and both went to public schools).

This reference to **vicious** and **virtuous** circles is important because it describes how mechanisms at different levels interact and reinforce one another. Note that when one is describing the vicious and virtuous circle one may be describing not only elements from different ontological levels but also at different levels of scale (or planes of social being etc.), where, for example, lower expectations is psychological or sub-individual.

### **5.3.11 Micro level**

How does the isolation of the science classroom reflect the compartmentalised curriculum? To what extent does the physical barrier (the classroom) exclude other

related subjects and learning issues? How does the notion of *innate ability* operate at the classroom level? How well have teachers engaged with the notions of FA/AfL?

At the classroom level teachers may share some of the ethos identified at global, macro or meso levels, e.g. through their acceptance of the ideology of '*innate ability*'; and thus through their praxis come to reproduce those structures inimical to FA/AfL and indeed, learning. In doing so, they may contribute to Gillborn and Youdell's (2000) 'Educational Triage' in which a significant proportion of the pupil population has been trapped over many years. They may also be contributing to Stobart's 'shaping' of individuals through assessment.

### **5.3.12 (a) Individual level: Pupils as individuals**

It is at this level that the ideology of *innate ability* operates in its most vicious form. My fieldwork in the three case study schools suggests that prior attainment in English (and sometimes mathematics) is used to create bands and sets. The effects of this on pupils' access to resources and the curriculum are significant. Evidence has emerged that under the pressure of league tables of results and under some circumstances, teachers are under severe pressure to falsify summative assessment results for pupils, for the purposes of maintaining or enhancing the league table status of their schools. Schools are falsifying summative assessments, sometimes for a significant proportion of pupils. Thus, many children who are in the bottom sets suffer a double whammy - poor access to resources, curriculum and teaching on the



one hand and a falsified summative assessment at the end of compulsory schooling on the other.

The argument is not that pupils will not have different capacities to learn at any given time. Far from it: their learning capacity at any point in time will depend on many factors including their well being, their attitudes to the subject being taught, their liking of the teacher, and their motivation at that time. The argument here is, therefore, that teachers need to be aware of these factors and take them into account in their lesson preparation and their approach to pupils and their teaching. This is when sensitive teaching, positive feedback and intellectual or cognitive scaffolding are needed to build confidence, enhance learning, augment motivation and accelerate progress.

An important piece of research carried out by Dweck et al (Dweck, Trzesniewski and Blackwell, 2007) lends strong support to my view of pupils' 'learning capacity'. The research consisted of two studies; study one involved 373 students (198 females and 175 males) in four successive years of entry into 7<sup>th</sup> grade classes of 67 – 114 each, at a public secondary school in New York City. The second study involved 99 students (49 females and 50 males) drawn from a seventh grade class from a different public secondary school in New York City. The findings showed that the students who held an “expandable” or “growth” theory of intelligence (or ‘mindset’) had a positive view about effort, and chose “effort based strategies in response to failure, boosting mathematics achievement over the junior high school transition.”

Moreover, Dweck's research suggests that children with 'growth' mindset improve their attainment but crucially it shows that it is possible – relatively easily, to change pupils' mindset from 'fixed' to 'growth'.

It is perhaps because of the ingrained view of ability as a fixed genetically determined entity and perhaps because there is a habit linking this with the thread of thinking which initiated the tri-partite system of education - Grammar schools, Technical schools and Secondary Modern schools - that the GCSE did not live up to the rhetoric of 'positive achievement'.

This thread of thinking continues to dominate. This is evidenced by such statements as that in 1997 the White Paper Excellence in Schools:

“Unless a school can demonstrate that it is getting better than expected results through a different approach, we make the presumption that setting should be the norm in secondary schools.”

Similarly the 2005 White paper reinforced this policy, with Education Secretary of the day Ruth Kelly saying:

“We will further encourage setting and grouping pupils by ability” (Roberts, 2005).

Michael Wilshaw, Her Majesty's Chief Inspector of Schools, is quoted in the Evening Standard (Davis, 2012) as saying that:

“... high flying students are being held back by schools that put them into mixed ability classes that fail to 'stretch' them.”

Davis further states that:

“Schools that do not ‘stream’ will also be docked marks if the teachers fail to prove they can stretch all pupils.”

This obsession with ‘setting’ and grouping pupils by ‘ability’ in England runs counter to motivational theories developed in the US as long ago as the 1980s by researchers such as Dweck (Dweck, 1999) and Dweck and Leggett (Dweck and Leggett, 1988) and theories of intelligence by Dweck (Dweck, 2002) and Gardner (Gardner, 1983). Gardner argued a case for multiple intelligences whilst the other authors put forward and discussed an ‘expandable’ or ‘growth’ theory of intelligence.

#### **5.3.12 (b) Teachers as individuals**

Teachers as individuals are subjected to severe scrutiny by both the national monitoring procedures (e.g. Ofsted inspections, league tables of results) and the schools’ own monitoring systems. The absence of systematic provision for professional development leaves many teachers vulnerable in that they fail to keep up-to-date with their own ‘guild’ knowledge and are often unaware of improved pedagogical practices. In the literature, the concept of ‘good teaching’ is advocated by Wiliam (2012) because it is argued, and there is general agreement with this argument, that only ‘good teaching’ can bring about a transformation from the current dominance of summative assessment over formative assessment. The problem is how to define ‘good teaching’ and who defines what is ‘good’? There is a danger that such descriptions and emphases personalise the issues and place the burden on individual teachers and yet teachers are judged by the Ofsted’s view of what constitutes ‘good’ and the pressure of meeting the league table grades. It is

argued that the Ofsted notion of good teaching is compatible with achieving higher league table grades. However, teachers often do not accept this argument. Caught in this dilemma, teachers become vulnerable to the consequences of oppressive (or Power<sub>2</sub>) relationships (Bhaskar, 2008, p.402). This raises questions about the possibilities for the exercise of Power<sub>1</sub> or the transformative capacity of agents which will be discussed further in section 5.4.

### 5.3.13 Sub-individual level

This is the level at which pupils' motivation - feelings of well being (or helplessness), self-esteem, ambition and sense of belonging - operates. It is also the level at which expectations including self-expectations are operative. Self-expectations of pupils from poorer socio-economic groups are often low and generally reinforced by schools. It seems that those who are placed in the lower sets continue to achieve less and less and remain subjected to the **vicious circle** of low expectations and low achievement throughout their school education.

As mentioned in section 1.3 the Motivational Model of Achievement (Dweck, 1999; Dweck & Leggett, 1988; Dweck & Sorich, 1999) suggests that those students who hold the 'entity theory' of intelligence believe in a 'fixed' unchangeable intelligence, a 'thing' they have either a lot of or very little. This, according to the model, can orient them towards giving up making effort if they become convinced that they have only a little of this 'thing' called intelligence. The researchers suggest that such students develop 'helpless responses to failure'.

Truancy is arguably the most visible demonstration of pupils' 'helpless response to failure'. Many pupils can't take any more the humiliation 'bestowed' upon them by the school system, which still runs on the model of a potato processing factory where potatoes are graded for their sizes and quality and those the lowest quality are disposed of. The Guardian reported in June 2012 that DfES official figures show that around 3.7 million school days were missed the previous autumn as pupils truanted (went missing) from lessons without permission (Press Association, 2012). On a typical day in autumn 2011 around 55,600 pupils missed school, according to an analysis of government statistics. Around 48,000 children missed a month or more of lessons, making them "persistent absentees". Although this is a relatively small proportion of the population it is a large number (0.9% in 2011), and it may well represent the tip of the iceberg of compliance, dissatisfaction and failure.

The sub-individual level is also that at which teachers' motivation operates. In this context we could invoke the model of the four planar social being, the second type of laminated system, and consider how psychological depression at the plane of the stratification of the embodied personality could have a knock-on effect on social interactions and hence employability etc leading to stress and poor health of teachers.

I interviewed T<sub>6</sub> on 26 January 2012. He has been under pressure to improve and has been subjected to frequent lessons observations by the senior managers:

Q: I have known you for a while and you have been an enthusiastic teacher, thinking outside the box ... what is your situation now?

**A:** Yes, that's what I used to do. My classes get good results. I can't expect to get much better to be honest. Kids have got A\*s. I have done the camping. I have organised the discos, I have done the lunch time Chess Club, I have done all those things. I think, perhaps, **I am bored. Also, I am tired. I am perpetually tired.**

Monitoring of his lessons continued till he agreed to accept early retirement. He retired at the end of July 2012.

All the levels I have discussed above can contribute to what one experiences at the sub-individual level. All the levels contributing to facilitating or constraining learning can combine to motivate/de-motivate, enthuse/depress, skill up or de-skill individuals, and all these kinds of conditions can and do contribute to teachers' as well as to pupils' attitudes and responses to teaching, learning and assessment.

#### **5.4 Education and two concepts of power: Power<sub>1</sub> – Power<sub>2</sub>**

Bhaskar differentiates two concepts of power: power as transformative capacity (Power<sub>1</sub>) and power as oppressive (Power<sub>2</sub>). Power<sub>1</sub> is the transformative capacity intrinsic to the concept of action as such, whereas Power<sub>2</sub> is the capacity to get one's way against either the overt wishes and /or the real interests of others by virtue of structures of exploitation, domination, subjugation and control, i.e. what Bhaskar refers to as generalised master-slave-type relations.

“Around such relations hermeneutic and other more material (but still conceptualised) hegemonic/counter hegemonic struggles may be waged.” (Bhaskar, 2008a, p. 402).

The case of T<sub>6</sub> illustrates the point about Power<sub>2</sub>:

Q: What else are they expecting you to implement?

A: no, that's about it, apart from the endless drive to get better and better results.

Q: I see.

A: Which I think is just ludicrous. We are a comprehensive school. Where are our grade G kids? We deny that they exist. I think there is something seriously wrong with the exam criteria.

Q: Hum,

A: Or we are not a comprehensive school.

Q: Hum,

A: I have tried arguing that one with people and they don't listen. The management changed as well. They are very dictatorial. They don't listen. There is no forum. There is student's voice but there is no teachers' voice in the school. We are supposed to listen to what the kids say but the management don't listen to us. They don't care what we say. They dictate to us what is going on and we do it.

The factory- model of schooling instituted in the late 18<sup>th</sup> and early 19<sup>th</sup> centuries (see Earl, 2003) continues to hold sway today, while the power relations in the school system remain predominantly oppressive, and arguably increasingly so.

The structure of schooling, I would argue, displays Power<sub>2</sub> relations. The accountability structure, the monitoring procedures, targets, the league tables of results and Ofsted inspections can all be seen as and be experienced as manifestations of Power<sub>2</sub>, oppressive power exercised by the state and other institutions. I hope to show in my data that in addition to pupils and teachers, head teachers and other managers are also locked into oppressive structures. Moreover, in all these cases, Power<sub>1</sub>, the transformative power of agents, appears dormant. In

chapter six and in the remainder of this thesis I will be considering the question of what is required to re-activate Power<sub>1</sub>.

## 5.5 Summary of Chapter 5

In this chapter I have elaborated the main theoretical framework for this thesis. I have drawn on Critical Realist philosophy for my theoretical framework, in particular, the TMSA, the concept of ideology, the idea of open systems and the device of laminated systems. In particular, two of these, one defined by levels of scale and the other defined by ontological levels and the model of learning involving ‘unfolding the enfolded’; have been prominent.

Section 5.2 of the chapter presents a critique of the ideology of *innate ability* and using the theoretical framework indicates how the ideology of fixed innate ability, in its many guises, continues to dominate state school education in England and blight the education and life chances of a significant number of pupils.

Section 5.3 sets out the structural barriers to FA and their role in the reproduction of what I call the factory-model of schooling, then considers the intensification of the ‘standards agenda’. This has resulted in a regime characterised by what I have called ‘double triage’ and an acute crisis in morale and motivation of school science teachers. The following quotation given by Gillborn and Youdell (Gillborn and Youdell, 2000, p. 197) captures the teachers’ dilemma succinctly:

“How do the league tables impact on us? It puts permanently more, more, more, more pressure on us, you know. Flog us to death, flog us to death, flog us to death. You do that anyway. But if at the end of the day you’re going to be told, ‘Hang on, we haven’t got X percent of A-to-Cs, so you must be bad teachers’ – ah ah, no.”



“... they want us to do well in this so their league tables are all right. They don't seem to care about what we want to do for ourselves. Just seem to care about us on the league table.”

Finally, I contextualise these issues by considering the power relations in schools and set the scene for Chapter 6.

## **Chapter 6 A Critical Realist Theory - Enabling Assessment:**

### **Building on FA and AfL:**

#### **6.1 Introduction**

In chapter five I considered the barriers to the implementation of formative assessment/assessment for learning. Here in chapter six I will be examining the extent to which FA/AfL are themselves insufficient and need to be set in the context of a broader model. I formulate this: a theory of ‘enabling assessment’ in which FA/AfL is supplemented by the important procedure of diagnostic assessment, which considers what has *not* been learnt by the learners and the teacher’s practical activity of scaffolding. In chapter seven I will be bringing this theory to bear on my data, examining the specific inadequacies of science teaching.

The idea of ‘Enabling Assessment’ is a new concept which builds on the ideas of formative assessment and diagnostic assessment and is designed to help address some of the overlaps, gaps and confusions found in the literature about FA and AfL. In doing so, it is hoped that the new theory will provide an approach to assessment with empowering potential for both pupils and teachers, such that teachers can become, to employ a Bhaskarian concept, ‘transformed transformative agents’ (Bhaskar, 2008a, p. 120), that is, agents apt for bringing about a transformation from summative to formative assessment. The discussions which follow are intended as a contribution towards developing a theory of enabling assessment informed by Critical Realism and including Bhaskar’s notion of ‘unfolding the enfolded’.

### 6.1.1 Shortcomings of the concepts of F A and AfL

For the most part, Formative Assessment and Assessment for Learning appear to be used interchangeably, and in recent years the term formative assessment has largely been replaced by the term assessment for learning (AfL) in discussions of assessment. As mentioned in section 3.2, the origin of formative assessment is associated with the child-centred education advocated by the mid-nineteenth century pioneers Charlotte Mason and Helen Parkhurst. The notion arose from their efforts to develop child-centred education and required a change in the nature of assessment including among other things that all written work was to be corrected (Black, 1986). This was to help the teacher to find out what more the children needed to learn. The term ‘formative’, as mentioned in chapter 3, was first used by Scriven (Scriven, 1967) in connection with curriculum development, when he discussed the differences between summative and formative evaluations.

This notion of marking pupils’ work to gauge ‘gaps’ in pupils’ learning has all but disappeared from recent articulation of formative assessment. This is so much so that Professor Wiliam’s Inaugural Lecture at the Institute of Education, London (Wiliam, 2009) with the title, ‘ Assessment for Learning: why, what and how?’ makes very few references to formative assessment and almost no references to the marking of pupils’ work. In Figure 1, page 12, Wiliam (2009) presents a table showing ‘Aspects of formative assessment’ but there is no mention of marking pupils’ work or diagnostic assessment to identify ‘gaps’ in learning, or of the need for teachers’ professional learning. A similar lack is clearly visible in Figure 4, page 32 of his lecture which is described as a ‘Logical model of Keeping Learning on Track’

(KLT). There is no mention of teachers' own professional learning despite the fact that the table in Figure 4 mentions six 'Teacher outcomes'.

Thus FA appears to be marginalised and AfL is couched mainly in terms of 'sharing objectives', 'sharing criteria for success', 'classroom discussion', 'self-assessment and peer-assessment'. The evidence from the case study schools used for this thesis show that self-and peer-assessment remain at superficial and knowledge-free levels, whilst discussions remain at the 'question and answer' level. This involves very few pupils and encourages 'guess what is in my head?' amongst pupils, an approach which does not develop into 'dialogic' discussion. This is illustrated in chapter 7 under the rubric of the 'inadequacies of science teaching'.

Professors Dylan Wiliam and Paul Black do acknowledge that their efforts have been mostly in vain and have criticised the way AfL is currently applied in schools (William, 2012). Despite the fact of sales of 'tens of thousands' of copies of the booklet *Inside the Black Box* which Professors Black and Wiliam co-authored and which promoted the value of formative assessment, it seems that very little FA has been happening in practice. The TES (13 July 2012) states:

"But the problem, says Professor Wiliam, is that they have not understood it properly."

Quoting Professor Wiliam the article goes on:

"There are very few schools where all the principles of AfL, as I understand them, are being implemented effectively."

Wiliam blames the Labour government for telling schools that "... it was all about monitoring pupils' progress, it wasn't about pupils becoming owners of their own learning."

This sentiment is shared by Professor Black from King's College who said in 2010 that the AfL techniques were not being used in many schools. He too blamed the government for "... emphasising measurement of pupils' progress."

Going back to the TES article (2012), Professor Wiliam also partly blames himself for underplaying a "really crucial aspect" of AfL: "designing your teaching on the assumption that pupils aren't going to get it all the time".

Perhaps this is the most significant admission by Professor Wiliam, his acknowledgement of a lack of presence of the absences – the learning which has not happened, pupils "aren't going to get it all the time".

He states:

"The big mistake that Paul Black and I made was calling this stuff 'assessment' ... because when you use the word assessment, people think about tests and exams. For me, AfL is all about better teaching."

Perhaps the big mistake was the absencing of one of the key components of FA which the pioneers Charlotte Mason and Helen Parkhurst had identified, which was to 'correct' pupils' work so as to discover what has not been learnt i.e. what is absent or lacking or missing. This reveals the ideological pre-disposition to focus on the present, positive and actual, which Bhaskar has critiqued as ontological monovalence (Bhaskar, 1993, 2008, chapter 2).

AfL emphasises classroom discourse, peer assessment, self assessment without fully exploring the mechanisms for identifying 'gaps' in learning and thereby failing to

notice fully the teacher's needs in this regard. In much of the literature there is an assumption that the key elements of FA are already in place so that AfL can build on them to encourage classroom discussion, self-assessment, peer-assessment etc. Gaps in pupils' learning, in AfL, are to be identified mainly through classroom discussion but to be effective this would require building pupils' knowledge base in many cases and also addressing teachers' own professional learning needs.

The articulation of AfL underplays two 'crucial elements' of FA, i.e. marking and feedback and also the role of diagnostic assessment which Wiliam acknowledges as partially his fault for underplaying the need to pay heed to the 'absence' of learning. Generally both require supplementation with diagnostic assessment, which focuses on what has not been learnt, and identifying the impediments to learning it.

My new theory of assessment is an attempt to include what is either absent or less emphasised in the articulation of FA and AfL and to make explicit the need for teachers' professional learning and for a change in teachers' 'mindset'.

## **6.2 Enabling Assessment**

'Enabling Assessment' (EA) combines FA and diagnostic assessment (DA) with the notion of '*scaffolding*' put forward by Wood et al (Wood, Bruner and Ross, 1976), drawing on Vygotsky's ideas about supporting learners.

Key elements of formative assessment/assessment for learning are:

- Effective feedback to confirm pupils' positive achievement. This would involve good and constructive marking and evaluation of pupils' work in addition to gathering information of pupils' learning through observations of their work, for example in a laboratory and/or working on problems in lessons and through listening to pupils.
- Identifying the next steps for pupils and discussing these with them on the basis of a good knowledge of pupils' strengths and weaknesses.

Successful formative assessment is dependent on:

- pupils' self-esteem
- a belief in their own ability to succeed

These in turn are heavily influenced by:

- teachers' expectations of pupils
- teachers' beliefs in the pupils' ability to succeed
- a supportive classroom culture
- a supportive school ethos

The purpose of EA is to enable both pupils and teachers to take the next steps in learning. For pupils it should confirm what they already know, can do and understand and enable them to take the next steps by identifying what they do not yet know, which is where diagnostic assessment comes in. For teachers it should inform and thus enable them to identify their own professional learning needs in order to support i.e. scaffold pupils' learning more effectively. It should also challenge their 'mindset' by placing an enabling responsibility on them.

By emphasising the enabling purposes of teaching, EA is intended to empower teachers to respond to pupils' learning needs by applying and continually updating their own professional knowledge and skills and challenging and reflecting on their assumptions and beliefs about children's ability. This would enable teachers to avoid falling into the trap of taking received ability (or, more generally the outcomes of summative assessment) as an indication of potential, which often leads to bypassing rather than tackling pupils' learning needs. In other words, EA is intended to help teachers overcome and transcend Sir Cyril Burt's (Burt, 1912) legacy of achievement linked to *innate ability* as a spurious, out of date, debilitating and de-skilling construct.

In my articulation of EA, AfL is subsumed in formative + diagnostic assessment. The concept of diagnostic assessment is borrowed from TGAT but modified to mean identifying pupils' needs in order to take the next steps and also teachers' needs in order to successfully scaffold the pupils' next steps. Thus the notion of Enabling Assessment can be expressed as follows:

$$\begin{aligned}
 \text{6.2.1 'Enabling Assessment'} &= \text{Formative Assessment} + \text{Diagnostic} \\
 &\quad \text{Assessment} + \text{Scaffolding} \\
 (\text{EA} &= \text{FA} + \text{DA} + \text{SC})
 \end{aligned}$$

The idea of formative assessment in the above formula is based on the 'learner-centred' work of the pioneers of FA and is similar to that given by TGAT, i. e. 'so that positive achievements of a pupil may be recognised and discussed and



appropriate next steps may be planned'. Here, FA subsumes AfL and contains an element of SA. As mentioned previously the notion of 'scaffolding' is borrowed from Wood et al's (Wood, Bruner and Ross, 1976) metaphor for Vygotsky's description of the activities which the adult can do to support the child's learning.

The relationship between FA, DA and SC is a cyclic one. FA indicates the next steps for the learner. DA identifies the 'gap' between the learner's current knowledge, understanding and skills and the next steps. It also identifies the teachers' learning needs for scaffolding the learner. EA altogether thus involves the following steps:

- To confirm the learning that has taken place in terms of knowledge, understanding and skills, requiring robust and adequate summative assessment (which the concept of EA is not ruling out);
- To indicate the next steps;
- Diagnosis of the 'gap' between the learner's current learning and the desired learning and the identification of the learner's needs for taking the next steps;
- Diagnosis of the teacher's needs for scaffolding the learner in his/her endeavour in taking the next steps;
- To provide support and feedback to the learner, continually confirming the learning gained and showing the way to what is yet to be learned.

### 6.2.2 EA – a cyclical and iterative process

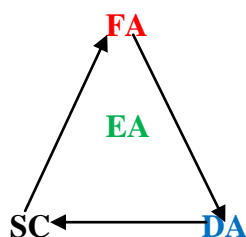


Fig. 6.2.2

The use of a triangle emphasises the action points at the vertices.

### 6.3 ‘Unfolding the Enfolded’

Bhaskar’s notion of ‘unfolding the enfolded’ (Bhaskar, 2011a) may be counter-posed to the notion of fixed ‘*innate ability*’ in order to develop a more appropriate theory of assessment which will help to overcome the effects of ability grouping on pupils and teachers, so that current expectations of pupils’ potential for learning can be transformed. Bhaskar refers to Plato’s theory that education involves a process of ‘anamnesis’, that is, ‘bringing out something’ which is already there. What is to be learnt is already ‘implicit’; it is ‘enfolded’ in the learner. The purpose of teaching is to unfold this enfolded potential. Learning happens when this potential is ‘actualised’, when as we say ‘the penny drops’ or the learner says, ‘Ah, now I can see it’. This is sometimes called the ‘eureka’ or ‘aha’ moment.

Bhaskar (Bhaskar, 2012, p. 116) describes the following five cycles of creativity:

**Calling** – initial inspiration, corresponds to being – and **1M** ontology

This corresponds to the level of Platonic anamnesis. This knowledge is unmanifest, enfolded, so necessitates a process of ‘education’, including external stimulus for the bringing out or unfolding of what is already implicit or known, in an unmanifest, enfolded form, in the learner.

**Creation** – It is implicit in being and explicit in becoming and corresponds roughly to **2E** of negativity and emergence. It is actually the process of learning, that is, internalising it, in which even though it may be a well-attested piece of knowledge, it is for the learner something new, a creation of the past which the learner must re-create in the present to not only to make his own, but to see what it is that is the ‘knowledge’.

At this point the knowledge is ‘heteronomous’, something externally given, albeit internally reflected, which comes to the student’s conscious mind as discovery, the acquisition of something new. At 2E we have the level of becoming. This is the moment of discovery or ‘eureka’, or we are taught it. Either way, this is the moment of learning or re-creation; as well as creation.

**Formation** – connotes elaboration or binding. It is presupposed by evolution and corresponds roughly to **3L** of totality. This moment in the process of creation is the assimilation of the new knowledge so that the learner can apply it on his own account. It is the moment in which he/she binds it back into their own being, shaping it, learning its implications and application, what to do with it and how to use it, until the point when it is thoroughly assimilated and becomes part of their innermost being. At 3L we bind it, we make it our own, so that it becomes part of us. This is also the moment of formation or elaboration, self-formation of the subject and elaboration of the discovery. This is the sphere of our discipline and service.

**Making** – this is transformative praxis (poesis) i.e. objectification which corresponds roughly to **4D** of agency. The learner has now made his discovery a part of his own being, and can apply it in all sorts of novel ways, and in such a way that the result now stands apart from this making. At 4D in turn we use it or apply it – this is the making or objectification of what we know.

**Reflexivity** – this relates to the theme of unity of theory and practice; having emerged in human form from matter, it reflexively (reflectingly) looks back on the creator (god), in traditional theocosmologies. This is the **fifth moment** – the objective result can be more or less adequate to the agent's intentionality and returns to him as the fulfilment of his will, or the consequences of his action in the world.

These five cycles of creativity characterise all the processes of learning and equally of discovery; they are characteristics of science, art, craft - indeed the whole tapestry of human agency.

The process of 'unfolding the enfolded', in my view, contrasts sharply with any notion of fixed *innate ability*. It lends itself to being utilised by teachers to 'unfold the enfolded' within all learners by engaging them in dialogue, helping them to diagnose or discover what they have learned and what to learn next and then to 'scaffold' their learning. Burt's concept of *innate ability*, on the other hand, tends to close the door on teachers and school managers investing time and energy to help those learners labelled as pupils with 'low ability'. *Unfolding the enfolded* requires

sensitive and systematic teaching and empathetic scaffolding to develop pupils' learning capacity and potential.

Another seriously debilitating point about the notion of fixed *innate ability* is that it has, over many years enabled policy makers to direct their efforts in modelling schools on a factory's production line which 'processes' and 'packages' pupils. It insufficiently focuses on pupils' learning and their all-round development. It is a wasteful process. For example, most pupils do not reach even the 2L level (the Aha moment) of learning in science at the end of this process.

Whilst idea of fixed *innate* intelligence affects teachers, school managers and policy makers in limiting their expectations of children's potential for learning, its effect on learners themselves can be devastating because it influences their 'mindset' as Dweck et al. (2007) have pointed out. Their research about intelligence and motivation has shown that student's self-theories about intelligence have a profound influence on their motivation to learn. Dweck has put forward an "expandable" or "growth" theory of intelligence, opposed to Burt's theory of 'fixed' intelligence. She considers intelligence to be 'malleable' and expandable with effort and application. She explains that those students who hold a 'fixed' theory of intelligence are mainly concerned with 'how smart' they are. Such students prefer tasks which they can already do well and avoid tasks in which they may make mistakes and not look 'smart'. In contrast she argues, students who believe in an "expandable" or "growth" theory of intelligence want to challenge themselves to 'increase their abilities, even if they fail at first'.

The purpose of EA is to enable both pupils and teachers to take the next steps in learning together – not only pupils’ learning, but teachers’ professional learning too. If both the teacher and the pupil believe in the “expandable” or “growth” theory of learning then taking the ‘next steps, and ‘scaffolding’ the learner can be greatly enhanced.

The idea of EA is developed further using Bhaskar’s notion of ‘unfolding the enfolded’ and of the ‘transformational model of social activity’. Gordon Brown’s concept of ‘learning environment’ as enunciated in his paper on ‘The Ontological Turn in Education’ (Brown 2009) and discussed in section 5.3.10, is also drawn on to augment Black’s (Black, 1993b) idea of ‘a classroom culture of success’.

Here it is worth noting that leaving out diagnostic assessment (taking FA on its own) is just taking the positive without the negative, begging the question, ‘What are the things which are blocking learning?’ That is why DA is called for because it allows us to look at the absences. DA is needed because it identifies the absences – the learning that has not happened. It must be said that according to Dialectical Critical Realism this reflects a characteristic failing in Western philosophy and social thought, which is always to assume positivity and disregard the negative.

The notion of diagnostic assessment in my theory is different from that given by TGAT in one important respect. It is the same as the TGAT’s (1987) notion of

identifying learning difficulties but it also considers the scaffolding needed for the learner to take the next steps. Moreover, it is different in that here the information gained is to inform scaffolding by the current teacher, not that of informing the next teacher. Here, it is to inform the current teacher and more generally all subsequent teachers, for the purpose of scaffolding. In this way we are moving towards an organic and more continuous conception of the process of learning. The purpose of diagnostic assessment is to identify learning that has not yet been unfolded.

My use of scaffolding includes providing feedback as a part of sensitive and systematic teaching and assessment. The purpose is to build knowledge, confidence and self-esteem through feedback sensitively provided, with a view to developing and shaping self-evaluating learners able to manage their own learning. I interpret the process of scaffolding as an *iterative process* in which both the learner and the teacher learn through dialogue and feedback. The dialogue enables the learner to progress towards her/his learning goals and it can also enable the teacher to learn how to respond better to the learner's needs and support pupils' next steps.

AfL rightly emphasises the importance of dialogue in the classroom, thus addressing dialogic social interactions both in respect of relationship between teachers and pupils and between pupils and their peers. Where I think there is a need for a much greater emphasis is in the marking and feedback aspects of assessment and the need for the teachers to discover what pupils do not know, are unsure about, and do not understand, so that they are better able to support pupils' next steps.

The purpose of ‘Enabling Assessment’ is to enable the learner to see what to do next and how, and to enable the teacher to know what to teach next and how. It will also inform the teachers’ own professional learning needs. For teachers’ learning too, the Platonic notion of ‘anamnesis’ applies. Helping children and young people to learn is the main intuitive driving force for a teacher but often it can remain ‘enfolded’, especially when the teacher is subjected to ‘productivity’ (Ball, 2008, p. 16) pressures.

Exploring the notion of ‘Enabling Assessment’ further, one can relate it to Bhaskar’s model of normative change, DEA. Here, ‘D’ refers to description which corresponds to formative assessment; ‘E’ refers to explanation and corresponds to diagnostic assessment, explaining the sources of the difficulties, blocks etc. that the learner may have. This then gives rise to action and corresponds to scaffolding which is the ‘A’ in the DEA model (Bhaskar, 2011b). A related model, the Describe, Explain and Transform (DET) model can also be drawn on when considering the transformation of the barriers to formative assessment, thus moving it towards Enabling Assessment. The process of transformation will require teachers to be agents of change. In this process both the DEA and the DET models are useful; diagnosing learning needs, explaining the issues involved and then taking appropriate action, which is integrated into teaching. Thus, EA has the potential to overcome the overlaps and confusion between FA and AfL and address their shortcomings.



#### 6.4 EA, FA and AfL - a comparison

Here it is possible to consider the transformative potential of EA compared to FA and AfL in their standard formulations: in FA the confirmation of the positive achievements of a pupil is explicit but in AfL it is hinted at or is implicit. In EA it is explicit. The diagnosis of pupils' learning difficulties is similarly explicit in FA but implicit in AfL. In EA it is explicit. The provision of scaffolding to close the 'gap' between current learning and the desired goal is implicit in both FA and AfL. In EA it is explicit. Both in FA and AfL the need to teach pupils how to engage in self- and peer-evaluation is implicit and not emphasised. It is made explicit in EA. Creating an interactive classroom is made explicit in AfL but remains implicit in FA. Similarly, promoting independence and autonomy among pupils is explicit in AfL and implicit in FA. In EA they are explicit. The need for a culture of success in the classroom is identified in both FA and AfL, but the process of creating this culture is absent in both. In EA both are explicit: the process as well as the need.

In relation to teachers' learning (teachers' professional learning) there is an absence in both FA and AfL of an emphasis on 'teachers' subject knowledge' and 'teachers' pedagogical subject knowledge', an omission highlighted by the National Strategy's Science Pilot's CPD module on Assessment for Teaching (AfT) which emphasised both of these. The module was designed to develop teachers' subject knowledge of physics, chemistry, biology and earth sciences and their pedagogical subject knowledge i.e. the knowledge of the age-related content of the National Curriculum and its associated progression and assessment (see Assessment for Teaching in section 3.6.2). EA addresses this 'absence' by asking teachers to focus on and to learn from pupils' learning needs and identify and address their own learning needs.

This will require a collegiate culture of professional learning for teachers and access to appropriate CPD. The efforts of school managers and national policy makers will then be rightly directed to facilitating such a programme of CPD.

The idea of self-reflective, confident learners is only realised when there is a good theory of the self that is learning it. This category of shaping self-reflective confident learners pre-supposes a robust theory of **the learning self**. Dweck & Leggett (1988) and Dweck et al (2007) have produced convincing research evidence to show that those learners who believe that ‘ability’ can be developed through their effort apply themselves to challenging tasks that build up their skills of overcoming their difficulties. As mentioned earlier she has put forward an “expandable theory of intelligence” according to which intelligence is not ‘fixed’, it is ‘malleable’ (Dweck, 2007). She explains that self-reflective learners develop a ‘mindset’ which adheres to the ‘malleable’ notion of intelligence. Such learners apply themselves to new learning situations with more determination and confidence and thus succeed in achieving their ‘goals’.

## **6.5 What needs to happen for the practice of EA to occur?**

Four things need to happen for change to occur and progress to take place.

According to Critical Realism, for genuine structural transformation to occur one would need the following (Bhaskar, 2008, p. 262 – 5):

- an explanatory critique;

- concrete utopianism – a vision of a better alternative which is what I am proposing – my theory of ‘enabling assessment’;
- a theory of transition;
- a good organic relationship with the practice of agents involved – a good relationship between theory and practice.

My data analysis is intended to show where we are now and where we need to get to. This will be at the level of the explanatory critique. The concept of Enabling Assessment is itself an exercise in concrete utopianism, the third and fourth steps are a theory of transition and a good organic relationship with the practice of agents involved. Ideally, the practice of agents should be oriented to transition and the theory of transition should be rooted or grounded in the practice of agents (a situation pre-supposed by Gramsci’s well known concept of an organic intellectual).

## **6.6 Towards an explanatory critique**

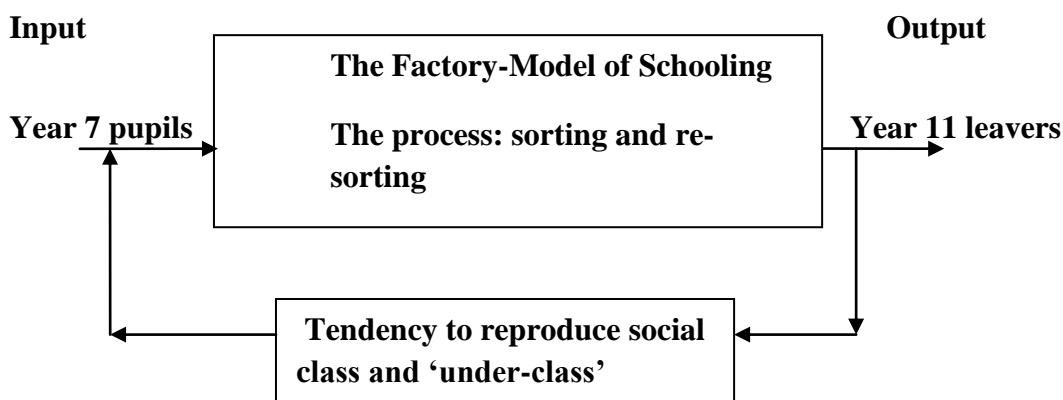
### **6.6.1 The ideology of ‘innate ability’ and its debilitating consequences**

The ideology provides a generative mechanism which, through tests and examinations (i.e. summative assessments), tends to contribute to the reproduction not only of the class structure, but also of the excluded and the disengaged, a group which is sometimes referred to as the ‘under class’. There are actually four components of this mechanism. First is the notion of innate ability, and second the notion that innate ability is measured or revealed in performance, i.e. in summative

assessments. The third component is that this performance forms the basis for setting or grouping pupils into different classes and fourthly that this setting procedure then forms the basis for the distribution of educational resources. This leads to ‘educational triage’ and the further impoverishment of the already poorly resourced.

These generative mechanisms facilitate and underpin the ‘factory model of schooling’ depicted in Fig 6.2.

### 6.6.2 The Factory-Model of Schooling



**Fig. 6.2** Factory-Model with a social feedback loop

The Year 7 intake pupils are aged 11 at the beginning of their secondary school. Many come in with insecurity, poverty, poor attainment and low self-esteem. Many Year 11 school-leavers leave aged 16 at the end of their compulsory schooling with these insecurities (absences relating to lack of education and well-being) intact or even magnified. Thus there is a tendency for the continuing regeneration of a *vicious cycle*.

TMSA, as explained in section 5.3.6, states that the human agency through its primary day-to-day activities reproduces or transforms the structures it inherits. The factory model of schooling is underpinned by the ideology of *innate ability*. Using this ideology, the factory model tends to continually reproduce the *vicious cycle* of **‘poverty - low expectations - low attainment – poverty’** for a large minority of pupils in the state comprehensive schools in England.

The pedagogy which informs the factory-model of education is also limited to what is best described as a ‘pragmatic’ model, focused on producing the proportion of GCSE grades which count towards the **league tables** of results. The remaining nearly 40 to 50 percent of pupils are simply professionally ‘managed’ or contained. This is illustrated with the aid of data in chapter 7.

### **6.6.3 Agency: teachers’ complicity in the ideology of *innate ability***

Teachers’ praxis is central to the ‘process’ represented in Figure 6.2. I argue that teachers have been complicit in the process to the extent that they have bought into the ideology of *innate ability* and routinely engage in practices which have the effect of tending to reproduce the structures which have perpetuated inequality in educational provision. This is not to deny the contributions of those teachers who toil hard to overcome all the pressures imposed on them to provide as good an education as they can for their pupils. The argument here is that the long standing structures (i.e. summative assessment) underpinned by long held ideologies (i.e. the notion of *innate ability*) create conditions under which social structures are reproduced, as a consequence of which unequal access to education tends to continue.

#### **6.6.4 Agency: Schools' Senior Managers**

Most people would agree that a school's managers must play a crucial role in the smooth running of the school, involving the pastoral care of both pupils and teachers. They help provide a conducive environment for learning and are charged with the responsibility for improving 'standards' and providing 'value for money'. Ofsted inspections and league tables of results have been imposed on them. However, what has not been, and cannot be, imposed on head teachers is their professional obligation to pedagogy and learning. As pedagogical leaders they are free, in theory at least, to provide a curriculum, within the national framework, that meets the needs of their pupils. They are also free, in theory, to alter the pedagogical processes in their schools, again, to meet the learning needs of their pupils. In other words, it can be argued that they do not have to 'ape' Ofsted inspections or slavishly follow the requirements of the league tables of results. They can, if they choose, deploy the resources at their disposal in ways which would support teachers and pupils, all pupils.

However, there are severe limitations placed on the capacity of head teachers in relation to innovation or experimentation with pedagogy. The pressure to climb up the league tables of results is immense and most headteachers are focused on achieving just that, whether 'by hook or by crook'. Besides, it is part of the economic conventional wisdom that triage is a sensible thing to do. Vertigan draws attention to this when discussing 'managerialism' in his thesis 'Ethics and Business of Schooling' (Vertigan, 2007, p. 150).

### 6.6.5 Agency: 'Double Triage'

A major finding of this study is that teachers too are being subjected to a process of 'triage'. Thus we have, in effect, 'double educational triage'. This concept of 'double triage' has emerged from both the theoretical and the empirical aspects of this research. Teachers themselves have, in this way, been subjected to severe scrutiny both at national and school levels. Many are thus victims of 'double educational triage' and having been a party to 'educational triage', find themselves in a quandary as to how to respond to their own victimisation.

There is little or no systematic CPD provision to skill up experienced teachers who in many cases have served the school for more than 20 years but have not updated their skills as desirable. Low expectations of pupils have played a debilitating role as far as teachers are concerned. Very often teachers have not seen the need (nor have the school managers drawn their attention to it) to skill up because they have been able to explain away their pupils' failure by citing 'ability' as the main causal determinant of their performance.

For a radical change to happen head teachers would need to re-think the factory model of schooling and construct a more learner-centred school environment which would include more inclusive and more learner-friendly educational processes and pedagogies. The pre-requisite for this, in my view, is that they will need to confront their mindset about the notion of *innate ability*. Moreover, the factory-model is antiquated and even modern business enterprises have outgrown this model.

## **6.7 A theory of transition: what needs to change?**

In section 6.6 above I have outlined the key features of the current state of schooling and its' facilitating generative mechanisms. In this section I wish to discuss theories for change. In this context Dylan Wiliam's notion of 'good' teaching and 'good' teachers, Lambert and Lines ideas about teachers' 'beliefs and perceptions' and Spendloves's notion of teachers' 'mindset' need to be considered.

### **6.7.1 The notion of a 'good' teacher**

Dylan Wiliam (2012) argues that 'good' teaching and, by implication, 'good' teachers will be required in order to bring about a transformation to formative assessment. He discusses ways of achieving this. He recommends sustained professional development of existing teachers in order to achieve this. He acknowledges, however, that this transformation could be achieved also by getting rid of 'bad' teachers and employing new 'good' teachers. He cautions against the latter approach on the basis that it is harsh as well as unreliable because there is no guarantee that today's good teacher will remain a good teacher tomorrow. Besides, the 'hiring and firing' model of school improvement, in his view, is not desirable.

### **6.7.2 The notion of teachers' 'mindset'**

Lambert and Lines (2000) have argued that 'teachers' beliefs and perceptions' play a vital role in action. Spendlove (2008) points out that it is the teachers' 'mindset' which needs to change for transformative action to take place. However, for action to occur there must be beliefs and wants. Wants depend on attitudes and what informs



attitudes includes values, so inner values have to change. One could say that the 'mindset' is something like beliefs including perceptions and attitudes informed by values. In the context of education one would need to stop looking at a particular child, for example, as a slum-dweller or of low ability, but as one who requires good care and good teaching. So the emphasis is to be placed on good teaching, good care, a welcoming school environment and a nurturing and enabling learning environment. Mindset also includes the views of pupils' ability which a teacher holds - in terms of Dweck's idea, a 'fixed' entity or of an 'expandable' one.

Clearly for such a transformation in the mindset of teachers they need to acquire an interest in formative assessment, assessment for learning and particularly in what I propose, 'enabling assessment' (EA).

Teachers can be, as Paulo Freire thought they should be (Freire, 1972) true (albeit partial) agents of transformation in society. Conditions for the real transformation of teachers' mindsets must be such that they have effectively both the disposition and the space (both physical and mental) to introduce these changes. They have to see why the change is necessary and also how it will help them, so that they come to have an interest in the change. On top of this they must have the space and security to do it. It is intended that EA should be an essential component of teaching, and, as such, could play an enabling role in transforming teachers and teaching. For this to happen, teachers need to engage in the debate and indeed confront the issue of pupils' failure and to consider why, despite their best intentions, so many pupils fail. In other words, teachers will have to become aware of and in some sense confront,

the barriers to formative assessment we have been analysing in this thesis. Teachers would need to consider why:

“...just one-in-six teenagers gained decent GCSE in at least five traditional academic disciplines – English, maths, science, a language and either history or geography. But among pupils eligible for free school meals, the proportion dropped to only one in 25.” Telegraph (Paton, 2012).

## **6.8 Compensation or Transformation?**

Finally, with regard to developing a theory of transition, Bernstein’s argument that ‘education can’t compensate for society’ and Wiliam’s argument (Wiliam, 2011) that ‘good’ education can compensate society, a possible resolution is through a critical realist synthesis.

### **6.8.1 A possible resolution: A Critical Realist synthesis?**

Good education can perhaps ‘compensate’ for society. However, the argument of this thesis is that what is required is not so much compensation as transformation. Teachers can help in partially transforming the educational structure. This can then help to transform society and in such a way that it will reinforce their own efforts in education. Thus, two social systems, a relatively small educational system and the larger social system in which it is embedded, going all the way up to global levels, are involved. The need is to move them all in the right direction.

### 6.8.2 Education and Power 1 and Power 2

Teachers are currently subjected to severe procedures of monitoring and accountability in which many suffer a huge amount of anxiety and stress but are unable either individually or collectively to mount an effective defence against the ‘double educational triage’, a process used for their subjugation. The recent changes (2012) to teachers’ pay and conditions introduced by the Secretary of State for Education, Michael Gove, have made it easier for teachers to be sacked. Thus Power<sub>2</sub> in the guise of the repressive, exploitative relations between the state and the school managers and between the school managers and the teachers has increased, whilst Power<sub>1</sub>, the effective capacity of teachers to use their transformative power, has continued to be diminished.

For Power<sub>1</sub> to increase, teachers will need to re-examine their adherence to the existing paradigm (or ideology) which is supported and underpinned by summative assessment and which enables the educational triage to take place with its devastating consequences for almost half of the pupil population, a point made powerfully by Marshall (Marshall, 2013, p. 3) quoted in section 5.2, repeated here for the convenience of the reader:

“... we do have a good idea of what happens to the 40% who do not achieve five good GCSEs. If you are part of this group then there is greater than one in four chance that two years later you will be NEET, that is, not in any kind of employment, education or training. If you are one of those (4%) who had gained no GCSEs at all, there was a greater than one in two chance that you would be NEET.”

Teachers have been party (consciously or otherwise) to the educational triage and they will need to confront this issue. They will need to reject the ideology which culminates in this ‘educational triage’ and re-confirm their commitment to education for **all** in a way which rejects all forms of the ‘tri-partite’ model of schooling – a model based on Burt’s 1925 paper on the education of ‘Backward Children’ discussed in 5.2. Teachers will need to see that, surely, this model is out of date and cannot be fit for the purpose of schooling in the 21<sup>st</sup> century?

In other words, teachers will need to become ‘transformed transformative agents’ of change. The national and school structures will need to change in order to promote real equality of opportunity and a real level playing field by adopting and sustaining enabling structures. Changes in education have the potential to reinforce changes in society.

## 6.9 Summary of Chapter 6

In chapter 6 I have developed the concept of Enabling Assessment. This is a theory of assessment which builds on the ideas of formative assessment and assessment for learning by including a new definition of diagnostic assessment, the concept of scaffolding and is informed by Bhaskar’s concept of ‘unfolding the enfolded’.

$$\text{Enabling Assessment} = \text{Formative Assessment} + \text{Diagnostic Assessment} + \text{Scaffolding}$$

$$\text{EA} = \text{FA} + \text{DA} + \text{SC}$$

Here, FA includes the key ideas of confirming learning and considering the next steps in discussion with pupils. DA applies to all learners and to all teachers; in my definition it has no association with ‘remedial’ work or ‘intervention’ strategies. Its purpose is to diagnose the next steps in learning and diagnose the next steps in the teachers’ learning, so that the teacher is able to scaffold the learners’ next steps more effectively. The relationships, between FA and SC (scaffolding i.e. sensitive teaching) are iterative, with a dynamic feedback between the two processes happening concurrently.

EA can become a tool for transformation if the ideology of *innate ability* which underpins the factory-model of schooling can be challenged and overcome; in this teachers will need to play a key role.

## **Chapter 7 Data and Analysis:**

### **The current state of science at Key Stage 3 (ages 11 -14)**

#### **7.1 Introduction**

This chapter discusses the current provision of science in the case study schools. The focus of analysis is Year 8, though samples of pupils' exercise books from years 7 and 9 were also scrutinised. The analysis considers the grouping of pupils, marking and feedback, the current state of assessment in science, the quality of teaching, the science curriculum, staffing in science departments and the nature and the quality of CPD for science teachers. Section 7.2 contains brief descriptions of the case study schools; section 7.3 describes pupils' grouping - banding and setting - in the schools; section 7.4 discusses assessment in science in practice while section 7.5 considers the nature and the quality of feedback to pupils. In section 7.6, I discuss the quality of teaching and in section 7.7 I explore the curriculum – C. P. Snow's two cultures and the issue of coverage of the National Curriculum. Consideration of staffing, teachers' workload and the current state of provision for CPD is covered in 7.8. Section 7.9 presents a summary of the key points of the chapter.

#### **7.2 The case study schools**

The case study schools are described in detail in section 4.9, so here I recall some of their main features. Curie High School (CHS), is the principal case for this study, has some features which make it an especially educative case. It is a large comprehensive school housed in a building that is over 70 years old. Historically the standards had been low and consistently below national averages.

The school's neighbourhood is relatively poor and its ethnic mix has changed considerably in the last decade. The Ofsted reports (2005, 2008), state that almost all students come from the local neighbourhood which has high levels of deprivation. The proportion of pupils entitled to free school meals, an index of poverty, is more than twice the national average, laying the ground for the kind of *vicious circle* analysed in chapter 5. Teachers at Curie High School have been subjected to severe scrutiny through mini-inspections of their lessons every term by the school's senior managers. As a consequence, teachers' morale has been low and some have lost their jobs through this process.

Downton Comprehensive School (DCS), in contrast, is a relatively new school opened in 2002. Its special features make it an 'ideal' case. It is housed in an impressive new, purpose built building which has a welcoming and expansive 'façade' with open access to the school reception. There is a very spacious front foyer and a cafe open to pupils, parents and the public. The science department was well-resourced and with only 359 pupils and at that time, 'hand-picked' staff, the school provided an ideal setting for studying marking and feedback. By 2008 the school had 1316 pupils in total, with 156 (11.9%) in the sixth form, (Ofsted report 2008).

Westfield Comprehensive School (WCS) is an appropriate case for comparison with Curie High and Downton Community schools. It was on two sites about a mile apart at the time of the Ofsted inspection in 2002 when it was graded as 'satisfactory'. Twenty nine per cent of the pupils were eligible for free school meals, which was

above the national average, indicating the degree of economic deprivation of the neighbourhood. Twenty three per cent of the pupils had English as an additional language. The Ofsted report had noted that “provision for pupils with English as an additional language was unsatisfactory” and was ‘barely satisfactory’ for those pupils with special educational needs.

In 2006 the school moved to a single-site and the 2007 Ofsted report noted that nearly half the pupils were from minority ethnic backgrounds whilst English was not the first language of one third of these pupils. These proportions were well above average and growing. This was also true of the proportion of students entitled to free school meals. The school had historically low standards and low ‘expectations’. The school ethos is less academic and it remained in the Ofsted category ‘satisfactory’ throughout the period of my fieldwork. The school, unlike the other two, has a strong focus on pupils’ ‘behaviour’.

As mentioned in Section 2.4, when my main study began in September 2005 the National Curriculum had been in place for 16 years, the National Tests had been in place for 12 years and league tables for results had been published for 9 years. Ofsted inspections had been taking place for 10 years and the Key Stage 3 National Strategy had been going for 3 years. In short, the Government’s ‘standards agenda’ was fully in place. My fieldwork began with the hope of discovering greater and improved use of formative assessment focusing particularly on marking, feedback, teachers’ involvement with professional learning, and the impact on pupils’ learning of science and their enjoyment of it. Instead, the data showed that, far from improving the situation, the Standards Agenda had further entrenched summative assessment and the barriers to formative assessment.



### 7.3 Pupil grouping: banding and setting

In this section pupil groupings are considered in some detail because all three schools used banding and within the bands subject departments are allowed to set. This process has a big impact on the school experiences of both the pupils and the teachers. As discussed in chapter five, the ideology of innate ability is still very strong in practice in these schools. Consequently there remains considerable inequity in provision, with the effect that the full benefit of a comprehensive school education is denied to more than half of the pupil population. One of my key findings is that the use of prior attainment in English is the main arbiter of banding and setting. This device alone deprives many pupils of access to good teachers and good resources in science, even though their prior attainment in science might have been good.

#### 7.3.1 Pupil grouping at CHS

Indeed, Curie High School (CHS) very closely follows Burt's notion of the treble-track. Year7 pupils in the school are put into ability bands using prior attainment levels obtained from the Key Stage 2 (age 11) national tests and Cognitive Ability Tests (CATs). Three bands plus a 'behaviour' group are created. The groups are as follows:

**Fast Track (A):** 4 Forms – (5 sets in science), 120 pupils

**Upper Band (B):** 4 Forms – (5 sets in science), 120 pupils

**Middle Band (C):** 2 Forms – (3 sets in science; Mu, Ml, and AS), 60 pupils (Mu is middle upper, Ml is middle lower and AS is academic support)

My information on banding and setting at CHS was obtained from the Key Stage 3 science co-ordinator and was confirmed by scrutiny of the timetable and an interview with the Deputy headteacher. In science, the four Fast Track forms are grouped in 5 sets to give class sizes of 24 in each set. There is a similar arrangement for the upper band. The Middle Band is effectively the bottom band and is divided into three sets. Thus in Year 8 science 320 pupils are put into 14 *sets* at CHS.

Table 7.2.1 - Attainment on entry to CHS for 2009/10

NC Levels	Number of pupils	Percentage (%)
Level 5 + Level 4	146 + 116 = 262	82
Level 3	25	8
Level 2 + B + N + F	1+2+3+1 = 7	2
Pupils with no Key Stage 2 results	25	8

Total = 319

B = below level 2      N and F = below level 1.

An analysis of the prior attainment (as measured by Key Stage 2 tests) of the incoming pupils to CHS for the year 2009/10 illustrates the unnecessary rigidity and illogicality of creating up to 14 ability-based groups (Table 7.2.1).

The groups could have been organised differently: for example,

(1) **L5 + L4** = 262 pupils in 12 parallel groups

(2) **L3 + L2 + B + N + F** = 32 pupils in 2 groups

(3) 25 pupils with no KS2 results could be given an equivalent KS 2 test and assigned appropriately to one of these groups.

This would produce 14 groups (same as the current) – with class sizes of approximately 24 pupils in each class and would add no extra cost to the school.

In creating 12 parallel groups teachers might worry that the spread of ‘ability’ will be too wide. However, if they consider the knowledge, understanding and skills gap those would be narrower than the National Curriculum sub-levels suggest. In fact, the gap is likely to be made of knowledge of some physics topics, the concept of particles and in Sc1 ‘presenting and interpreting’ evidence. A teaching programme carefully constructed to address these areas will meet the needs of most pupils. Teachers will need support in teaching such a programme. The benefits, however, are likely to outweigh the negative impact of setting as practiced at CHS.

Indeed, the Deputy Headteacher (DHT) at CHS defends the 14 ability sets by saying that the grouping has been done on educational grounds. I quote at length below from my interview with the DHT because her views are expressed strongly and appear, it transpired, to be representative of the other case study schools:

Q: So let me ask you this question, why do you focus so much on ability?

A: Because you have to give children appropriate tasks. You have to set them appropriate places to maximise their achievement. If you look in science at the moment, I mean, in the whole school we have bands. We have a Fast track, an

Upper and a Middle band. And the Bands are very focused in methodology, in pedagogy, and as they get into Key Stage 4 in terms of their options.

Here, the DHT asserts that the rationale for the banding in the school is linked to ‘methodology’, ‘pedagogy’ and to options (pupil subject choices) at Key Stage 4 (the GCSE). However, the evidence shows favoured provision for Fast Track pupils in relation to the allocation of teachers, access to resources and curriculum and access to other opportunities available in the school. For example, in science,

- Fast Track pupils are allocated subject specialist teachers. The Upper Band and the Middle Band are not.
- Fast Track do triple Award Science, others do Core science or BTEC science
- For revision the Fast Track and the top set in the Upper Band are given ‘Revision Guides’, while the others are not.
- When a guest speaker (such as a poet, writer, artist or scientist) comes to the school only the Fast Track pupils are invited to listen to him/her.
- Fast Track pupils are encouraged and expected to achieve good results, the others much less so.

My interview with the DHT continues,

Q: OK,

A: It’s [banding] crucial to our success.

Q: So, you have got banding on the basis of what?

A: Ability

Q: Ability?

A: And, strictly on, initially on ability. However, if you are able but we can’t access that ability, may be your behaviour is a handicap, you will find that you are not in

the Fast Track because our focus at that stage has to be improving your behaviour. You may be in the behaviour modification group. And what we don't do at our school is that we do not sacrifice 32 children for 1. We are not prepared to do that. We will work with the one and every child does matter but we will work with the one without disadvantaging 32 and we are very clear on that.

The school is ruthless in protecting the privileged position of the Fast Track pupils.

The following interview with two Year 8 pupils shows how the promotion and demotion from Fast Track works at CHS.

Demi (not real name) has recently been promoted to Fast Track from Upper Band set 1 and Clive (not real name) has recently been demoted from Fast Track to Upper Band. (Demi is middle class black, Clive is working class white):

Q: Tell me how well you have done since our last conversation.

Demi: I have done well. I have got two Level 7s and one high 6 for my last assessment.  
[*Demi is very confident*]

Q: You also said that you had moved up, did you say that in our last conversation?

Demi: Yes, I am moving up again to Fast Track. [He had previously moved up from a lower set in the Upper Band to set 1]

Q: Tell me how did that happen? Did you work extra hard?

Demi: Yes, I worked extra hard and they re-set us due to the exam that we took.

Q: Ok, what made you make that progress? Tell me more about how hard you worked and what you did.

Demi: Homework, extra revision, we [have] got our own books that we can look at – look at things we need [in] chemistry to physics and biology.

Q: You have got your own books?

Demi: Yes we have, *given to us by the teacher*

Q: Were you given books by your teacher and were those Revision Guides?

Demi: Yes, Key Stage 3 Revision Guides, CCG something like that [It is actually CGP (Gannon, 1998)]

Q: That's right, ok, did that help you?

Demi: Yes,

Q: Did you get help from home, your family?

Demi: Yes,

Q: What kind of help did you get from your family?

Demi: My dad sat with me and at exam time we did revision. He asked me questions, what does potassium make? What's that symbol?

Q: [Addressing Clive] Now, you introduce yourself then tell me how well you have done in Year 8.

A: I am Clive Denton

Q: How has your Year 8 gone? [*Clive is quiet and subdued but well spoken*]

Clive: It has gone alright but the last few tests I have done rubbish in. I did not revise for them.

[Demi; because he was not in]

Q: The test did not go well for you?

Clive: No. I done alright with the stuff I could do but there was some stuff that I did not know how to do.

Q: Ok, is it because you did not have time to revise?

Clive: No, I had, I had time to revise I just I thought we were doing a physics test. I did not know we were doing chemistry and biology.

Q: Ok, yes, so, what does it mean for you? Did you not get the level you wanted to get?

A: I am probably not going to get the levels. I don't really know.

Q: Has Miss not told you what you have got?

Clive: No, I don't think she has marked them yet.

[Demi: The only reason he did not get his levels is because he was in the hospital for about a month.]

Q: Are you well now Clive?

Clive: I am better now.

Q: So, what do you think you will be doing in Year 9?

Clive: I will do a computer course

Q: In science what will you do, do you know?

Clive: No, I don't. I am not doing the same stuff as him because he has gone up to Fast Track. I am not, so, -

Q: Are you staying in this group?

A: Yes.

Q: And you don't know what kind of science you will be doing next year or do you know?

[Demi: Miss said that the Upper Band will only do, they are not going to do like the Fast Track because Fast Track, they get a teacher for biology, science, (Clive: yes), I mean biology, chemistry and physics, one teacher for each where in Upper Band, (Clive: one teacher), you only get one teacher. You only get one teacher for all three subjects.]

Q: Is that good or bad?

Clive: It is **crap** [quietly].

Demi: Bad because one teacher could be better in biology than physics and chemistry.

Q: Yes ...?

Clive: I don't see the point. So, I just don't do it, the same for Upper Band as well. [Clive seems to be *disheartened*!]

Q: Yes, what do you think is the reason then?

Demi: They are basically set in order. There is AS which is for the bad behaved children and there is another band for children that are kind of below [not clear] and there is Upper Band which we are in at the moment and then there is Fast Track which I am going into.

Clive: They act as if everyone who is not in Fast Track is *stupid, really*.

Q: Who acts like that?

Clive: Most of the teachers. They don't really say it and I don't really know that they mean to do that but most of them just act like it.

Q: What kinds of things do they say or do which make you say this?

Clive: One teacher told me to shut up unless I should get to Fast Track because I was *stupid*. The teacher called me stupid. I did not even know the teacher.

Demi: Bearing in mind that he [Clive] got kicked out of the Fast Track because of his behaviour.

Clive: Yes, I was in Fast Track. I just got moved down.

Q: But how can anybody say anything like that to a pupil?

Demi: He might be annoying the teacher

Clive: I am not annoying

Q: You can be annoying and people can be annoying right?

Clive: Yes, everyone can be annoying

Q: Tell me more about why you feel that they think that you are, or some of them think that you are, stupid if you are not in the Fast Track.

Clive: What, just me or ... ?

Q: No, pupils who are not in Fast Track

Clive: Because they are always going on about Fast Track and ...

Q: But you were in the Fast track weren't you?

Clive: Yes, I could have stayed in but I just had a bad day,



Demi: Misbehaved,

Clive: Yes, in one lesson.

Q: What do you mean when you say, 'they are always going on about the Fast Track? What do they say?

Demi: Basically, if you stay in Upper Band there is a boundary that in some subjects you can get the highest in and the teaching will be focused on that level. But if you are in Fast Track, they will be teaching you higher and you have to catch up faster like I am expected to get an **A** because ...

Q: Right, good, I am also interested in you because you were in the Fast Track

Clive: Yes,

Q: And then you got moved to Upper Band because of bad behaviour ... ?

Clive: Yes, Fast Track was easy. For me I don't find much difference in Fast Track or Upper Band because ...

Q: For you it isn't difficult

Clive: No.

Demi: Because he is smart, but he just does not behave well.

Clive: Not all the time. Like in science I don't misbehave too much.

My interview with the DHT continues:

Q: How do you measure that ability?

A: We have an assessment system that is pretty rigorous, I would say. Every half term teachers have to do a tracking assessment and they have to put in a progress report towards the target. Erm, every half term that is reviewed and we would look at children's performance in all their subjects to get a general picture and actually have a look if there are, is a subject or a child in a particular area that is a concern. Now that happens five times a year. And that's how we will keep those bands. So, there is movement between the bands. We are not afraid to move you and children certainly know that.

The DHT asserts that the assessment system at her school is 'rigorous', while at the same time conceding that the quality of teaching is barely satisfactory. Teachers who are notable to produce good teaching somehow organise assessment which the DHT accepts as rigorous. However, the evidence shows that teachers know the content of the end of module tests and they make sure that those topics are covered. There is evidence also that teachers mark up the tests to ensure that the targets are met.

Working under a severe accountability regime, it would be a brave teacher who would not ensure that the targets were met. Strong evidence emerged from the school that teachers were reporting 'adjusted' marks for the 'tracking assessments' in Key Stage 3 (Years 7, 8 and 9) and also in Year 10. By the time pupils reach Year 11 panic sets in and those pupils (up to 30% at CHS) who are unlikely to get grades A to C in the forthcoming GCSEs are switched to BTEC in science so that they can be awarded two minimum grades Cs by hook or by crook.

My interview with the DHT continues:

Q: What's the starting point then?

A: They come in Year 7. Initially they had the Key Stage 2 SATs. This year there wasn't that many [Compulsory KS2 Science Tests had been withdrawn] so we had to do our own internal tests and we used those as our initial yardsticks. Our Fast Track, if you are in our Fast Track we would anticipate very clearly from your Key Stage 2 results that you are very strongly within an A\* to C category projection for five years down the road.

Q: Right,

A: And if you are in the Upper Band you will be a D, possibly, yes, you will be a D or below. If you are in our Middle Band that means you have come in and we have real worries about your literacy. Your literacy is very poor or you are on P scales or you are a Level 2. You will need a lot of intervention and support. So, we put you in the Middle Band because we operate a literacy focus group.

Q: What is a literacy focus group?

A: So, you will have a *limited* number of teachers. You have one teacher for English, history, geography, and civics and languages with the real focus on pulling up your English.

In practice very little focused teaching of literacy happens in the ‘literacy focus group’. One teacher teaches five subjects. Every pupil in this group is given a ‘Literacy Booklet’ to work through on her/his own, ostensibly supervised by the Form Tutor who has to perform numerous duties during the form time and hardly ever gets to look at the Literacy Booklet.

Being in the Literacy group excludes pupils from:

- interacting with other pupils who have good English and so being denied the opportunity to learn from them;
- practical subjects such as science, physical education and technology where you would have opportunities to learn English;
- attending normal English lessons where you would be taught by teachers trained to teach English.

Science in particular is generally recognised for its emphasis on practical laboratory work and, through this, the opportunity to learn technical and scientific words. Also, pupils work in pairs during laboratory work which gives them time and an incentive to share ideas and discuss results. This also contributes to learning and improving English.

My interview with the DHT continues:

Q: So, the system allows that

A: Yes, absolutely.

Q: Going back to the starting point what subjects do you assess to create the bands?

A: English and maths generally. If there was a pecking order between English and maths **English usually wins.**

This seems to be illogical and irrational. One can argue that it should not in fact be revealed ability, i.e. performance, but potential that counts anyway; but if you wished to select, for economic reasons, the most accomplished in science for extra attention, then the most accomplished i.e. the best performing in science, not the most accomplished, i.e. the best performing in English, should have been selected.

*The use of English for selection immediately disadvantages those pupils who were born abroad or whose parents came from abroad and also the native poorer white pupils. By default, this is a de-selection on the basis of ethnicity and class which cannot be morally justified.*

Interview with the DHT continues:

Q: Right,

A: Because of the impact it has on other subjects. But we would put someone who is very good in maths with poor English, although that's quite rare in the Fast Track. But we will look very closely at it.

Q: Right, therefore, you got Fast Track which is your top,

A: Absolutely

Q: Then, you have got

A: **Upper**

Q: Upper Band, then you have got,

A: The **Middle Band**

Q: Have you got any others?

A: Yes, we have got **AS** (Academic Support group). That's our behaviour modification group. That's part of the Middle Band.

Q: part of the Middle Band

A: It's **10 Mu, 10 MI** and **10 AS or 8Mu etc.**. There will be three, one of which would be the behaviour modification group.

[Here, the weakest pupils are further divided into three groups; AS, MI, and Mu.]

Q: Right! So one would not be wrong in saying that the leadership is doing its best to create an environment where teachers can actually perform well?

A: Yes, yes,

Q: Because you are dealing with behaviour in such a way that a teacher who wants to do well can focus on learning.

A: Yes, absolutely, we think teachers need to teach and we try to make the teacher's job as easy as we possibly can. If you have a mixed ability class (but) unless you have super teacher who is a mixture of **Mother Teresa and Genghis Khan** it is very difficult for you to be effective. So, we try to make it possible for work to be tailored, tailored to meet the needs of a class where the ability levels aren't from one extreme to the next.

The above assertions by the DHT are not borne out by the data on pupils' prior attainment profiles summarised in Table 7.2.1. Most pupils are on average or above average attainment levels and according to their targets, as forecast by the Autumn Package, most are expected to achieve Level 5 or above at age 14 and a minimum of grade C in their GCSE examinations. The DHT claims that there is movement to and from Fast Track. Whilst it is true the evidence shows that only a few pupils move up to Fast Track, and movement down from Fast Track is used as a punishment for minor disruptive behaviour, as evidenced by the above interview with Clive and the DHT statement below.

Here the DHT's contempt for 'mixed ability' comes to light especially when one looks at the pupils' prior attainment profile (Table 7.2.1) which shows that 82% of pupils have levels 4 and 5. The difference in the prior attainment is of only one level and not 'from one extreme to the next' which the DHT asserts. Secondly, by the DHT's own admission, the teaching in the department is still *barely satisfactory* indicating that the creation of numerous 'ability' bands and sets had not made the teaching any better. One reason could be that the adherence to the 'setting regime' encourages low expectations of pupils in the middle and bottom sets.

Continuing the interview with the DHT:

Q: Then you have the assessment throughout.

A: That's right. And we will change bands. And if your behaviour is stopping others from learning you will be removed.

The DHT had asserted that the banding was based on ability and was focused on methodology and on pedagogy. The evidence above indicates that banding and setting only serve to 'triage' access to education such that a significant proportion of pupils are disadvantaged and denied access to the National Curriculum.

### **7.3.2 Pupil grouping at DCS**

At Downton Community School (DCS) Year 7 pupils are grouped into two parallel populations, with 4 mixed ability science classes in each population. The school allows setting in mathematics, science and French. In Years 8 and 9 the two bands continue, but in science pupils are then set into 4 ability groups using prior

attainment data from Year 6 (Key Stage 2 levels) and end of Year 7 examinations. Similar groupings operate in Year 9 science. At this school, however, the targets are less ambitious than those at CHS. Here, the Key Stage 3 (age 14) targets are below the Autumn Package minimum targets which are based on two levels of progress between Key Stages 2 and 3. The school uses sub-levels (e.g. Level 4 has sub-levels 4a, 4b and 4c, 4a being nearer to 5c) and the targets are generally two sub-levels below those indicated in the Autumn Package. These are expectations which fall below the expected rate of progress given pupils' prior attainment i.e. they are low expectations. This is surprising because it appears to be contrary to the school's 'vision' of 'success for all' (Staff Handbook 2008, Appendix A3).

The pupils' attainment profile in the table below indicates that pupils could be grouped differently and this might be more beneficial to learning.

**Table 7.2.2 Attainment on entry to DCS for 2010/2011**

NC Levels	Number of pupils	Percentage (%)
Level 5 + Level 4	83 + 104 = 187	76.6
Level 3	34	14
Level 2 + B + N + F	1+1+1+ 0 = 3	1.4
no Key Stage 2 results	20	8

Total = 244

A much more inclusive grouping, for example would be 8 parallel groups comprising Level 5 and Level 4 pupils, then two parallel groups comprising Level 3, Level 2, B and N. Pupils with no Key Stage 2 levels (20 pupils) could be given a Key Stage 2 level test and allocated across the ten groups. This will incur no extra cost to the school and would bring all the benefits outlined in section 7.2.1.

### **7.3.3 Pupil grouping at WCS**

At Westfield Comprehensive School (WCS) setting is done at the end of the first term in Year 7, on the basis of prior attainment (Key Stage 2 Test or TA levels) and teacher assessment of pupils in the first term at the school. Pupils are grouped into two bands and then within each band they are put into six sets in science. Similar setting arrangements are used in other subjects.

Although there are only two bands dividing pupils into two populations, there is setting throughout the school and in every subject. The new Deputy Headteacher acknowledges that the school is more ‘stratified’ than other schools where he has taught. He is a linguist and can see some merit in setting. In his words:

“What I will say is that this school is one of the more stratified schools that I have taught in, in terms of setting.”

And he continues,

“... and I think that you can, I think what we are not seeing is the fruits of, or one impact of, setting which can definitely impact on the quantities of A and A\* that you see at the top end in top sets and we are not seeing the fruits of that at this school”.



**Table 7.2.3:** Attainment on entry to for WCS (2010/2011)

NC Levels	Number of pupils	Percentage (%)
Level 5 + Level 4	124 + 120 = 244	82
Level 3	24	8
Level 2 + B + N + F	3+2+3+ 0 = 8	2.6
With no Key Stage 2 results	21	14.1

Total = 297

Given the prior attainment of the pupils, banding and setting at WCS are both rigid and unnecessary. It would make sense to have 10 broadly parallel groups comprising L4 and L5, with the end of Year 9 targets set between Level 6 and 7.

Similarly, two parallel groups comprising L3 and L2 could give two smaller groups with targets between L4 and L5. The remaining 21 pupils could be allocated across the 12 groups on the basis of a school set Key Stage 2 test. This would add no further cost to the school but would provide appropriate pedagogical challenge to teachers and pupils to meet or exceed those targets. Again, as outlined in the previous two sections, great benefits could be gained in terms of pupil self-belief and motivation and a curriculum more equitable distribution of resources.

The DHT at WCS acknowledges the link between ‘setting’ and low expectations, when he says:

“You know, because I have taught in four different schools with different approaches to setting and different approaches to streaming and things like that. I think all the research is indicating that it does contribute to low expectations.”

Low expectations are clearly indicated when the targets set by the school for pupils age 14 are examined. In contrast to CHS where targets are set according to the Autumn Package, at WCS the targets are below the Autumn Package expectations. 65 pupils out of 298 have no targets, 88 have targets at least 2 sub-levels below and 48 pupils have targets one sub-level below those given in the Autumn Package.

#### **7.4 Assessment in practice in science: Summative and Formative**

This section describes the nature and the quality of summative and formative assessment in science. Summative assessment remained dominant even after the national tests at age 14 were withdrawn at the end of 2008. One of my findings is that the quality of summative assessment has deteriorated because there is a tendency to ‘adjust’ upwards the results of school tests and exams to keep attainment in line with the targets. Evidence was found of such ‘adjustment’ of levels when reporting Key Stage 3 results at the case study schools and in relation to BTEC science coursework.

The balance between FA and SA is similar in all the three case study schools; SA is dominant and FA/AfL is superficial. For example, what is intended to be learnt in the lesson (i.e. the learning objectives) is rarely made explicit in terms of knowledge, understanding and skills to be taught and learned. There was no consistency in the use of learning objectives. Teachers used aims, LO, or objectives for learning objectives. In one school (DCS), one teacher used WILF (What I am Looking For)

for learning objectives, when in fact it represents success criteria (or learning outcomes) (Clarke, 2001). Just one teacher used WALT (We Are Learning To) correctly. At CHS, teachers use what they refer to as AfL assessment sheets with 20 multiple choice questions. The marking of these questions, however, is summative, with marks out of 20 (Appendix A4). The National Strategy's National Director and the Regional Director had both suggested in their interviews that the introduction of Assessing Pupils' Progress (APP) would address the AfL issues effectively. However, the evidence of the use of APP obtained from the schools shows very little formative use of it.

For instance, as we have seen, feedback is a key element of FA/AfL, and an example of feedback provided by the deputy head of science at WCS (T<sub>31</sub>), who is also the Local Authority's lead teacher on APP, is telling. It exemplifies the poor quality of feedback in terms of supporting learning and indicates the lack of teachers' own knowledge in this regard. The teacher wrote the following comments on the work done on 7 April 2010 by P<sub>22</sub> (Appendix A5):

*"To get L6 on your conclusion – learn to compare results properly interpreting data consistent with your evidence (results) and adding detail scientific explanations".*

The work on which this comment was made was on "Neutralisation of indigestion tablet". The following results had been obtained:

**Results Table:** (Taken from a pupil's book, pupil P<sub>22</sub>)

Type of tablet	Volume of acid used (ml)	Time taken to change colour with methyl orange
Antacid	35	35 secs
Remegel	11	3.42 min
Alka	50	2.00 min
Rennie	4	30 secs

What in fact was the teacher expecting the pupil to do in response to the feedback?

Given the above results, where 'type of tablet' is a discrete variable and both 'volume of acid' and the 'time taken' are continuous variables, what learning was the teacher asking the pupil to do? What was it which the pupil could/should have compared?

If the task had been to compare the mass of the tablets with either the amount (ml) of acid used or time taken then there would have been an opportunity to interpret the data suggesting a relationship between the variables, for example, between mass of tablets and volume of acid or mass of tablets and time taken for neutralisation. None of this is clear. However, it must be said that this teacher was one of two teachers in my sample of teachers (20) whose pupils did some practical activities in their lessons. Evidence of this was found in pupils' books and in my lesson observations.

At Curie High School pupils in Years 7 and 8 do end-of-module tests every six to seven weeks. Marks obtained in these tests are entered in the school's electronic database which is used to report progress of pupils in the Fast Track (the top Band) and the Upper Band (the second top – in fact, middle Band) in relation to their end of year targets. The report card contains numbers 1 to 4, where 1 represents better than expected progress and (the pupil is expected to exceed her/his target), 2 represents the expected level of progress (the pupils is expected to meet her/his target), while 3 represents less than expected progress (the pupil is likely to narrowly miss her/his target). Number 4 represents poor progress and indicates that the pupil is likely to miss their targets by a wide margin. These numbers are used to report progress to pupils and parents. For the school's own records the numbers are given colour codes (1 = white, 2 and 3 = green and 4 = blue) in the schools' record of progress. It is claimed that this system provides a quick and easy method for identifying pupils who are 'falling behind'. However, the colour codes as well as the number codes are easily manipulated and are open to abuse. Targets for the so-called Middle Band (in practice the bottom band) or in the 'identifiable behaviour support' group are not taken seriously.

Since 2009 when the National Tests at age 14 were withdrawn no guided revision lessons have been provided to pupils, as used to be provided for over eight weeks prior to the national tests. Thus, pupils miss out on the extra attention and teaching they used to receive.

As mentioned above there is no guided revision. Pupils are merely told to revise. There are also end of Year examinations in Years 7 and 8. In Year 9 pupils did two trial (mock) examinations and the national tests (tests in May) until 2008, and since then they also have done an end of year examination. At DCS the Head of Science has

been a senior marker for the national tests so all tests and examinations in the department mimic the national tests. He is keen to see that these internal tests and examinations follow the 'rigour' of the national tests, but, he has found it difficult to maintain this 'rigour' given the pressures of the targets. This pressure has meant that reported levels are 'amended' upwards from the test results to be in line with the set targets.

Summative and Formative Assessment at WCS follow similar patterns to those in the other two schools. Pupils in Years 7, 8 and 9 are given end of unit tests approximately every 4 weeks and results of these tests are entered into the school's electronic data base. Years 7 and 8 also do end of year examinations in June. Year 9 did a 'trial' (mock) examination in December and KS3 national tests in May when the National Tests were in place. Since the withdrawal of the national tests year 9 pupils have also done an end of year examination. The deputy head of science is also the Key Stage 3 science co-ordinator. Like the Head of science at DCS, she is also very keen to maintain the integrity of the tests and examinations in the department and reports the results of Key Stage 3 (years 7, 8 and 9) exams 'accurately and honestly' to the head of department. However, between her reporting and the DHT office receiving the results they are 'adjusted' to meet or nearly meet the targets. The following conversation with the Key Stage 3 science co-ordinator at WCS reveals some tensions:

Q: Do you think that the Key Stage 3 Teacher Assessment gives an accurate percentage of pupils achieving Level 5+ as compared to the national tests?

A: I tried as much as possible to mark last year's, the current Year 10, when they were in Year 9. I gave them 2010 National Test sample papers and their teachers did not have to mark their classes. So, I gave them out randomly to teachers to mark.

- Q: So, a common marking scheme but teachers did not mark their own classes?
- A: Yes, that was one strategy that I used. After that we all moderated the papers, each other's marking.
- Q: Right,
- A: So we could give actual levels to the children. But then I was told that most schools were doing, well giving higher marks but I am not going to manipulate the results. You know this is what they got, this is what they got. In a nutshell you know, you don't really know where to ... it is a very interesting thing but that's how I did it and I *feel comfortable* with that. Nobody can query that.
- Q: Because you actually carried it out robustly with good marking which was moderated?
- A: Yes, if you mark Key Stage 3 papers you might get papers from Birmingham or from Scotland, you don't mark your own student's ones do you?
- Q: So, what percentage of pupils got Level 5 or over (L5+)?
- A: 57%, I reported 57% Level 5+.

*By the time the results reached the DHT office, via the head of department, the percentage of those gaining Level 5+ had been increased ('adjusted') to 70% L5+. The target was 75% (Appendix A6).*

In the league table of results schools were ranked on the percentage of pupils achieving Level 5 or above. The school was not achieving its Autumn Package targets while the national tests were in place despite three months of intensive revision and booster classes before the tests. With the withdrawal of the national tests at the end of 2008, schools were able to report attainment at age 14 by 'adjusting' the actual results (attainment) obtained from internal KS3 type tests (without the intensive revision associated with the national tests) and **making it up** to at or near the targets set.

## 7.5 The nature and quality of feedback to pupils

This section focuses on the quality of marking of pupils' books to discover the nature of written feedback through comments in the books and/or reports to pupils and parents. In addition, lessons were observed for episodes of discussion and dialogue.

Information to (parents and) pupils is generally provided in all the three schools in terms of the national curriculum levels. However, it makes no reference to knowledge, understanding and skills, which is a key weakness of the system. Pupils are told that they are at level 4C or level 5B, or level 6C etc., and often it is left at that. Often pupils are asked to work harder towards the next level, but how is the pupil to know *how to progress* to the higher levels, if the teachers only give these levels and do not spell out what specific knowledge, understanding and skills are required to progress to the next level? This point about feedback is also made by Ofsted's own evaluation of the impact of assessment for learning:

"In response to initiatives from the Secondary National Strategy, there has been some improvement in teachers' use of assessment for learning in science ... however ... the emphasis has been on summative assessment and preparations for tests and examinations. Students generally knew what their targets were in science, but lacked sufficient guidance to be certain about what they needed to do to meet them." (Ofsted, 2008b)

My conversation with three pupils in July 2011 (recorder reference 086) typically reveals the nature of feedback:

Q: Do you know how well you are doing?

A: Yea, I am on target



Q: What's your target in science?

A: 6a

Q: Are you on target to achieve it?

A: Yes,

Q: How do you know?

A: Because I am at progress Level 2

Q: Progress Level 2? Explain to me what that means.

A: 4 means you are not going to reach your target, 3 means you might fall short of your target, erm, 2 means you are going to achieve it with little bit help and 1 means you have already achieved it.

Q: Ok, how often do you get told what your levels are?

A: After the test, every half term.

Q: So, every half term you get a piece of paper which gives a number, 1 to 4

A: Yes, and the target

Q: Alright, how do you find out how well you are doing in the lessons, in the homework, in the topics, do you get your books marked and the comments given to you?

A: (one pupil): Not very often.

A: (Another pupil): Not very often but we get a chance to speak to a teacher once in the year.

Q: You get a chance to speak to a teacher once every year?

A: Yes, at parent's evening basically

Q: Parent's evening, but during normal lessons, through homework, week by week do you get some feedback?

A: (one pupil): No;

A: (another pupil): Not very much;

A: (first pupil): Get shouted at

Q: Don't you get some feedback, say if you are doing magnets and you are having any problems, then any comments?

A: **No.**

Q: Do you get your science corrected in your books?

A: **No,**

A: (another pupil): **No**

A: (another pupil): **Not really.**

There is a surprising degree of consistency with respect to this between the schools. Pupils are told the national curriculum levels at which they are currently working and what their end-of-the year targets are. For pupils in the lower sets, neither the levels they are at nor their targets are generally mentioned.

### **7.5.1 Marking**

The nature of written feedback was studied through the scrutiny of pupils' books and also through interviews with teachers and pupils. A pro-forma used for analysing marking is shown in Appendix A1. Marking was found to be cursory and written feedback trivial, non-specific and often erroneous. Two findings emerged: pupils in the top two sets get their books marked, although the quality of marking is poor. Errors in science are seldom corrected and aspects of 'Scientific enquiry' (Sc1) such as tables and graphs are rarely marked. The books of the pupils in the lower sets are rarely seen by teachers.

Generally, marking was given a low priority in the day-to-day work of the teacher. There was evidence of intermittent or haphazard marking and often evidence of no marking. There appeared to be ‘events’ in the marking calendar and by this I mean either ‘book-checks’ by senior managers or an upcoming Ofsted inspection. At the approach of such events there appeared to be a flurry of marking activity, then a long period of little marking. During these ‘quiet’ periods, some marking of work of pupils in the top sets was seen. Even in the best case example, i. e. of DCS in the Exploratory Study, when the school was new, there were fewer than 300 pupils in total, with ‘hand-picked’ teachers and with the headteacher and deputy headteacher themselves teaching science, the marking provided little constructive feedback or corrections of scientific errors. On the whole, subject knowledge weaknesses of teachers were evident in marking. The topics covered too were indicative of the teachers’ own areas of strength and the ‘gaps’ perhaps showed their areas of weakness. The following two tables showing an analysis of teachers’ marking illustrate these points.

**Table: 7.5.1a: Analysing marking:**

Pupil P<sub>36</sub>, Book 1 contains work from 16<sup>th</sup> Sep.' 10 – 7<sup>th</sup> April' 11

**Book ID:** (P<sub>36</sub>) - 8Sc6

**Teacher Code:** T<sub>32</sub>

<b>Themes</b>	<b>Present Yes or no</b>	<b>Details</b>
<b>Learning Objectives (LO)</b>	Y	Example: To describe how enzymes break down our food – 21/10/10.
<b>Learning outcomes</b>	N	
<b>Success Criteria</b>	N	
<b>Science corrected</b>	N	There were various opportunities for correcting science but none taken
<b>Literacy corrected</b>	Y	Spelling: pinut corrected to peanut on 30 <sup>th</sup> September 2010.
<b>Tasks given to complete</b>	Y	“Make sure you check over your spellings and include key words in your descriptions.”
<b>Pupils’ response to tasks set</b>	None	2/12/10 – In copying the LO, pupil copies oxygen and respiration.  11/11/10 – Title: Digestion Assessment. Pupil copies the LO, ‘assesment’.
<b>Comments only</b>	Y	<b>S</b> - well done for completing the experiment.  (S stands for star)
<b>Marks or grade only</b>	Y	For end of unit tests on ‘Food’ (3/3/11). pupil gets 11 marks out of 15. T assigns a Level <b>5c</b>
<b>Comments and marks/grade</b>	Y	
<b>Ticks only</b>	Y	Ticks for tests on Enzymes
<b>Ticks and comments/grade/marks</b>	N	
<b>Teacher follows up to check completion of task/s</b>	N	

**Coverage by T<sub>32</sub>:**

16 <sup>th</sup> September 2010	–	Food Digestion
18 <sup>th</sup> Nov. 2010	-	Respiration
13 Jan 2011	-	Microbes and disease
10 <sup>th</sup> March	-	Ecological relationship

The following table was stuck on the inside of the front cover of this pupil's book:

<b>My KS2 Level was:</b>	<b>My Y8 Level is:</b>	<b>My KS3 Target is:</b>
<b>4b</b>	<b>5c</b>	<b>6c</b>

In addition, this information was entered on the last page of the book:

‘Test Achievements: End of Unit (Food) Test (3/11/11), 11/15, Level 5c’.

**Table 7.5.1b: Analysing marking:**

**Pupil P<sub>36</sub>, Book 2:** contains work from 20<sup>th</sup> Jan'2011 – 21<sup>st</sup> June 2011

**Book ID: P<sub>36</sub> - 8 Sc6**

**Teacher Code: T<sub>34</sub>**

Themes present	Yes or no	Details
Learning Objectives (LO)	Y	LO: Example: investigating what material is the best conductor – 20/01/11
Learning outcomes	N	
Success Criteria	N	
Science corrected	N	A strange graph of 'independent variable against dependent variable' – not marked.  [Normally you have a named independent and a named dependent variable]
Literacy corrected	N	T spells excellent as exelent (20/01/11) and again on 01/02/11. Pupil writes three sentences numbered 1, 2 and 3 with 'particals', then 'particales', then, repeats 'particales'. Each sentence is ticked in red and at the end the comment is 'Excellent'.
Tasks given to complete	N	
Pupils' response to tasks	N	
Comments only	Y	
Marks or grade only	N	
Comments and marks/grade	N	
Ticks only	Y	
Ticks and comments/grade/marks	Y	
Teachers' follow up to check completion of task/s	N	

In Book 2 graphs are not marked, worksheets are not marked (e.g. worksheet 8Kc/1 on mirror image), work on refraction is poor and is also not marked, e.g. ‘Why do we get a rainbow?’ (5/04/11)

**Coverage by T<sub>34</sub>: (Information from Book 1 of pupil, P<sub>36</sub>)**

20/1/2011 - Heat Transfer

17/3/2011 - Light

26/5/11 - Rocks

Overall, this pupil (P<sub>36</sub>) has remained at Level 4, his primary school level, 18 months after leaving his primary school.

In both books there are gaps in the coverage of physics topics and graphical work and there is some low level (primary school Year 4/5) practical laboratory work.

At CHS some evidence of more marking is found in the books of pupils in the top sets, but almost no marking was found in the books of the middle and bottom sets. Even when the books were marked, graphical work went unmarked. Possible teachers’ weaknesses in subject knowledge were evident in the instances of errors in science which were not corrected, and in teachers’ ‘corrections’ which were often palpably wrong. The coverage of topics indicated a lack of balance for the lower ‘ability’ sets – physics topics were scarcely covered and Scientific enquiry (Sc1) was rarely taught.

At DCS similar patterns in marking to those observed at CHS emerged. This is despite an apparently systematic programme of ‘book-checks’ by the head of department and the school’s senior managers (an example in Appendix A7). Here,

two more issues emerged; one was the link with the changes in Ofsted's practice with regard to 'inspecting' marking, and the second one was the advent of 'content free' science emanating from Science for the 21<sup>st</sup> Century with the emphasis on 'how science works'. As an example, in practice it meant some classes spent half a term looking at the extinction of the Woolly Mammoth.

At WCS the marking also showed a similar pattern. Priority was given to the marking of books of pupils in the top sets. Like the other two schools low expectations were clearly visible in the comments, ticks, and incorrect science noticed in the marking. It emerged that the Key Stage 3 co-ordinator who is also deputy head of science department had very weak knowledge of Scientific enquiry (Sc1).

Thus, while marking for the top sets is given higher priority, it is not good and marking for the lower sets is poor. The lack of expectations, the low level of work covered, the overlooking or even praising of poor work, all evidence of neglect or 'triage'.

### **7.5.2 Feedback through discussion in the classroom**

Discussions in classrooms often remained at the question-answer level. Mostly it was the teacher who asked questions, only some pupils answered and then the teacher would move on to the next question. The best example was one in which the teacher was discussing neutralisation, acids, alkali and PH value with a Year 8 class. Pupils had mini-white boards and some were engaging with the discussion:



P: Neutral

T: Neutral, well done! Excellent! So, what colour, sorry! How will you describe the yellow colour then?

P: Oh,

T: What's yellow colour then? If you say acid is red, neutral is green, what's the yellow colour then? 5, 4, 3, 2, hold on I just check, excellent, excellent, excellent, not very strong acid, weak acid is also the correct answer. What's purple then? 5 seconds.

P: Slightly ... (inaudible)

T: 5, 4, 3, 2, 1 hold it up.

No, it's an alkali but not weak. Alkali or strong alkali is the good one. Wrong spelling – well bless you! A, l, k, l i, strong alkali, alkali, well done. Ok, what colour do you think is number 1 on the PH scale? Ye, you see .. 5, 4, 3, 2, 1 up they go. Oh, no, not black, it is red, red, red, red. Go for a different one. What is represented by 11? 5, 4, 3, 2, 1, up they go. No, no, yes ok I give you the colour blue excellent even better to say alkali. What sort of alkali? Medium strength alkali I like it well done! I didn't say what colour is it. I said what represented number 11. The answer says medium strength alkali which is even better.

T: Alright! Knowing what you know now from the last question I expect you to get the next one right. What's represented by maybe a PH4?

Five seconds, 4, 3, 2, 1 up they go, oh, that's better. That's what I want to see. Excellent, excellent, come on you should be learning from what we said before.

T: Ok, alright then one more. Let me make it the hardest. What would happen if I mix PH 3 solution with another one which is PH 11?

P: Oh, I know.

T: If I mix PH 3 with perhaps a PH 11/12

P: What will it be or what colour?

T: No, not what colour. You can say colour but you can't say tell me what it is. You can tell me as a PH, you can tell me as a colour or you can tell me as a name. Ok? This one is most difficult. 5, 4, 3, 2, 1.

- P: (Inaudible)
- T: No, no, no, you are so close 1 PH off. Excellent, yes, yes, no, no, yes, no, yes, neutral yes,
- P: Miss ...
- T: No, it is neutral. It is green. Sorry Tracey, universal indicator usually goes up to 14. So, I was trying to pick one that was same strength acid and same strength alkali. Listen, without using your boards can anyone tell me what would happen if I mix a weak acid with a very strong alkali? Will I get neutral or not?
- P: No.
- P: You will get a weak alkali.
- T: Yes, because you will need them to balance each other out you see otherwise you don't get the neutral.
- P: If the other one ... (*inaudible*),
- P: Yea,
- T: Most probably, depends if they are exactly the same or not. If you have a strong alkali and a weak acid you end up slightly alkalised because the weak acid can't neutralise everything.
- P: What about a strong and strong one?
- T: Strong and strong one? Can't neutralise, got to be exactly the same strength.
- P: Like a strong acid and a weak alkali make a weak acid.
- P: Nno.
- T: If I get a strong alkali like this one here
- P: And yes, strong acid
- T: Strong acid?
- P: No, strong alkali and one below strong acid will make it weak acid won't it?
- T: Yes, there still be a tiny bit of acid left therefore making it weak.
- P: Why won't it be strong?

T: Because it is not strong enough. Yes, it has to be precise equal strength to neutralise it, a bit like mixing 3 gram to 2.5 gram. It has to be exactly the same to balance each other out. Good questions guys. May I have the things back?

*There was dialogue between T and a few pupils. Most did not really get involved.*

*The results of the end of Year 8 examinations showed that only seven out of 28 made some progress compared to their Primary school levels.*

### **7.5.3 Feedback through reports to pupils and parents**

At CHS and DCS progress cards were issued every term but contained number codes, 1, 2, 3 and 4 indicating above target, on target, just below target and well below target respectively. In relation to teachers talking, teachers did not have much time to talk to pupils about their work. In reply to the question “How often does your teacher talk to you about your work?” one pupil said, ‘At the parents’ evening’ (see section 7.4).

Feedback through written reports was limited to the National Curriculum Levels. Progress was reported with a number as mentioned above. This has a ‘masking’ effect. It masks the gaps in learning in relation to knowledge, understanding and skills. How were pupils and parents to know how to improve if progress was presented in terms of numbers and not expressed in terms of the knowledge, understanding and skills needed to be learnt?

## **7.6 The quality of teaching in science**

Observing the quality of teaching was not one of the aims of my fieldwork, and these findings are by-products of the fieldwork. My scrutiny of marking and feedback to

pupils contribute to a general picture that the quality of teaching in science is at best only satisfactory. Two further findings emerged: teaching to the test for the ‘able’ pupils – the two top sets - and somehow or other ‘managing’ the rest. The Ofsted reports on the case study schools confirmed that the overall quality of teaching in the three science departments was no better than satisfactory (Ofsted judges the quality of teaching as ‘outstanding’, ‘good’, ‘satisfactory’ or ‘inadequate’). Interviews with teachers, heads of departments and deputy head teachers confirmed the Ofsted judgements.

Though, as I have said, examination of the quality of teaching had not been an aim of my study, attention to this proved unavoidable when my fieldwork began. Ofsted reports had raised the issue of quality of teaching, stating year on year that the quality of teaching in science nationally was barely satisfactory. The Ofsted’s annual report for 2008 went further and stated in section 78:

“There is rather more unsatisfactory teaching in science than the other core subjects.”  
(Ofsted, 2008a)

The senior managers in each of the case study schools were very mindful of the need to improve teaching and openly discussed this with me when I interviewed them. I corroborated their views with the views of the heads of department. The similarity in the quality of teaching across the schools was surprising in terms of the sources of evidence I was using, namely, book scrutiny, lesson observations and interviews with managers, teachers and pupils.

### 7.6.1 The quality of teaching at CHS

In 2005 the quality of teaching in the science department was deemed by the Ofsted inspectors to be inadequate. In 2008 a further inspection judged it to be generally satisfactory. The school's senior managers agreed with the findings. The DHT explained in response to the question, "How good is the quality of teaching in science?"

"Overall quality is satisfactory but I would say too much inadequate teaching." She continues, "Satisfactory overall because the results are overall satisfactory – but too much last minute catch-up, intervention etc. by another teacher – so the results become ok."

The DHT also added here that switching pupils from GCSE to BTEC is also a factor in producing results which are deemed satisfactory. BTEC is a coursework based qualification, with the marking done by the teachers at the school, thus enabling the school to 'adjust' the results in such a way as to assign a grade C to most pupils, C being the minimum league table grade (see Table 7.6.1 below). The following excerpt from my interview with the DHT (T<sub>3</sub>) (28.01.11) illustrates this point:

Q: In terms of the standards the inspectors thought, that compared to the national, they were satisfactory. Obviously a mixed bag, some better than others. Where did science fit in?

A: In terms of triple science physics, chemistry and biology incredibly well, way above the national average. This is superb. Then you are creaming off your level 6s and above. In terms of the Single Award science the results are very disappointing. And they were only salvaged by the fact that we converted them to BTEC as the course was underway. So, children then walked away with two good science GCSEs at grades A to C as BTEC, as opposed to a Single Award science.

Q: When you consider that the GCSE is supposed to be for 80 – 90 % of the pupils, why do you enter around 40% for the BTEC exam?

A: In order to salvage it. But we had to, because, otherwise, those kids would not have come on to an A to C grade in the Single Award Science. So, there are real issues there.

Q: With the BTEC they managed to get a grade C.

A: *Two.*

Q: Two? Hmm - that is obviously a currency worth having for the kids.

My conversation with one of the science teachers (T<sub>22</sub>) revealed how the switching worked. I had asked the teacher about the impact of the CPD (on AfL ) that teachers were provided with twice a term by the school (section 7.7.2). The teacher responded with another assessment issue.

Q: So, what is the other issue?

A: The other issue is about results and perceived results – (a long pause!) I have spoken before about the idea of manipulation of students' work or the assessment of students to give them high grades: at a recent CPD forum we were talking about or we were being talked to about the way we record our assessments, about whether we allocate a 1 for outstanding performance, a 2 for satisfactory performance of a student. And we were told to bear in mind – what we were putting down for each child in terms of how that would reflect – the way it was said seemed to imply that we should not be at all pessimistic in what we predict to be the outcomes for our children.

I have worries about that because if target-setting is not accurate and appropriate at the outset when a child is first set their targets at the beginning of a Key Stage and the teachers are being optimistic about the results of those and they are teaching in terms of the grades the children will actually achieve in the end of Key Stage 3 then when they reach Key Stage 4 the targets which the children had been set are unnaturally high and the achievement towards those targets has been assessed in a

way that is unnaturally high and therefore by the end of Key Stage 4 when external assessment occurs any deviation from those high targets is then seen as **a failure of the teacher concerned**. [bold – my emphasis]

Q: So, are you saying that the assessments are being done in such a way as to give the child a higher level or higher grade than the one they will achieve through an accurate assessment?

A: Yes, basically. And then again in Key Stage 4 children have not achieved their target grades specifically at the border grades in the past, but now it is because of the actual changes in the way the school assesses. It has moved across to the whole grade range, any child that is not achieving their target grades by the time they reach Key Stage 4 ... **pressure** is brought upon the teachers involved with marking coursework to ensure that the grades which are allocated [given] are appropriate for the level [target] the child is predicted to achieve and that happens in the subjects which are either **100% or majority** coursework [this] can have a significant impact on the outcome for that child. [My emphasis]

Q: Are you saying that the teachers are being asked to amend or give extra marks to a piece of work to meet the targets?

A: Not generally asked to just simply amend marks, but give extra assistance to a child to make sure their mark is better. I mean that could be done in a **very appropriate way by spending more time teaching and giving the child more opportunities to** absorb the material through extra lessons, extra time outside and inside school, homework etc. and then to assess again. But, generally it seems to be much more efficient simply to give assistance during the **[coursework]** assessment. [My emphasis]

Q: So, that might put a question mark on the reliability of the summative (Coursework) assessment.

A: Yes, a particular example – in one of my Year 10 classes there is a boy who is, erm, his targets were not particularly high, grade Cs were his targets and I don't think, I can't recall finding a student whose targets were less than grade C at GCSE which in itself is a bit of an anomaly. But this child is *functionally illiterate* and every lesson I have with this child there is a Teaching Assistant - there are a couple of children in the class who are also in a poor state as far as their literacy is concerned - and so the

Teaching Assistant (TA) will sit with the child and assist them – scribing for them or assisting them in their writing.

When I have read the work that the TA has actually submitted, or the child has submitted with the help of the TA, the quality of the writing is that of an adult. And it's clear to me that the child is not responsible for what has been written on the paper. The clarity of thought is far too deep. When I am marking a piece of work of a child who is illiterate I expect to see lots of indications of that lack of experience with literacy which I don't see in that type of work. Even if the kid is literate at that age their experience, their breadth of knowledge is quite limited by comparison to an adult and their ability to write coherent sentences and paragraphs should still show signs of that immaturity, and since I don't see it there in the work where a Teaching Assistant has been involved, I do suspect that perhaps the child hasn't achieved the standard in their science work that the piece of work in front of me suggests they have achieved.

Q: So, in the case of this boy you are talking about you think his literacy is poor, what about his science?

A: Erm, his access to the, to the teaching material is limited by his literacy. He has had a consistent impairment to his progress as far as access to material is concerned. His science ability certainly does not match up to C grade GCSE or pass level of BTEC.

Q: Has he got a grade C?

A: Having passed the work on to those who supervise the course it has been recommended that the work is of a suitable standard and suitably reliable and therefore ... he should get a grade C equivalent pass.

Q: Do you think the boy ...

A: In this case, no, in this particular case **definitely not**. [**bold** – my emphasis]

Q: What kind of grade would probably be most appropriate for him in terms of his knowledge?

A: There is no grade in BTEC below grade C. There is no grade below a pass.

Q: I see. Therefore, you either get a grade C or you fail.

A: Exactly right.



Q: In your judgement would he fail?

A: *Absolutely*. [My emphasis]

My interviews in July 2013 with two other science teachers reveal an even more worrying aspect of this practice. The conversation with two teachers involved in getting Year 11s to complete BTEC coursework which should have been completed in April, reveals a desperate situation in the science department at CHS with the stress on teachers and direct evidence of how BTEC results are ‘adjusted’ to boost GCSE results in science.

T<sub>19</sub> is the teacher in charge of BTEC science and T<sub>15</sub> is the Head of Science, and they are helping pupils to complete their Coursework.

I entered T<sub>19</sub>’s classroom. The classroom was full of Year 11 pupils working on their coursework, and the conversation went as follows:

T<sub>15</sub> comes in to ask T<sub>19</sub> something about a pupil’s coursework. They are both working to get the pupils to complete their coursework (which should have been done well before this point in time).

T<sub>19</sub>: Have you looked at P’s folder?

T<sub>15</sub>: No, I have not checked it yet. Is he having problems with graphs?

T<sub>19</sub>: Yes

T<sub>19</sub> begins to help a pupil with how to plot a graph showing simple things like drawing the axes, labelling axes, and choosing the right scale – (very basic Year 8/9 stuff).

T<sub>19</sub>: (addressing me): Last year these pupils had a succession of supply teachers. So they had not done their coursework, had not been taught much. Some of these pupils’ work was not marked last year.

Q: How many are doing BTEC in Year 11?

T<sub>19</sub>: About 120

Q: You said these pupils had supply teacher after supply teacher. Why was that?

T<sub>19</sub>: This boy had 8 science teachers [pointing to a pupil].

Q: 8 teachers in Year 10?

T<sub>19</sub>: Yes, in years 9 and 10.

Q: Extraordinary!

My conversation with the pupil confirmed this.

T<sub>19</sub> continues,

T<sub>19</sub>: In November they changed the time-table when I and Mr B began to teach them.

Q: So that was in Year 11?

T<sub>19</sub>: Yes.

Q: At that time were they doing GCSEs?

T<sub>19</sub>: Yes. It was March, or was it April? ... when they were moved to BTEC

Q: Which March, this March?

T<sub>19</sub>: Yes, some had proper teachers, some had supply teachers...

[T<sub>19</sub> begins to help a pupil do a graph – the pupil does not

know how to choose scales for a line graph.]

T<sub>19</sub>: Now the exam board has asked for more samples. Nobody had bothered to get these kids to complete their coursework – just gave them a C. Now we are having to help them complete their coursework. I mean **copy** or **we tell** them what and how to do it, like doing this simple graph.

[I then walked up to T<sub>15</sub>'s room.]

Q: Why is the BTEC in such a mess?

T<sub>15</sub>: We are now just getting them to finish their coursework somehow – there are three pieces of course work still to do. Mr. E did not do any marking. He had some of these kids last year. He did not bother to mark their work. He did not have time to do that. He was concentrating on the triple Award and AS level kids. SMT [Senior Management Team] encourage this.

[Knock at the door] T<sub>19</sub> comes in and shows T<sub>15</sub> a page from a pupils' coursework.

T<sub>19</sub>: Have you got the original copy of this?

[There is a lot of tension between the teachers. Both are trying to get the BTEC

coursework completed, both are finding it difficult - there is a last minute urgency.]

T<sub>15</sub>: I thought I gave the original back to you.

T<sub>19</sub>: No, you didn't

T<sub>15</sub>: I have got a photocopy

T<sub>19</sub>: I want the original first page for this.

T<sub>15</sub>: I have got a photocopy of that.

T<sub>19</sub>: Where is the original one of that?

T<sub>15</sub>: I gave that back to you in there

T<sub>19</sub>: No, you didn't.

T<sub>15</sub>: That one [shows a page of pupils' work]

T<sub>19</sub>: No, you didn't. This is different from that one.

T<sub>15</sub>: There has been quite a commotion (? Not audible) this morning.

T<sub>19</sub>: I know [angrily!]

T<sub>15</sub>: I gave it to you back in there (meaning in T<sub>19</sub>'s room)

T<sub>19</sub>: No, you have not given that to me (raised voice!)

[Conversation becomes very tense]

T<sub>15</sub>: I have not got it here (raised voice). That's it. What am I going to do with it – what a waste! [shuts the door on T<sub>19</sub>]. ... ridiculous! Tell you what I am going to do is go to the Union and get them on workload on this because this workload has come from senior management. We have been **directed** to do this, getting kids in there to do the stuff.

Q: You mean the coursework?

T<sub>15</sub>: This is the BTEC

Q: That should have been done last year?

T<sub>15</sub>: (A bit exasperated). They started doing this in March this year.

Q: This March?

T<sub>15</sub>: Yes, Year 11 swapped on to do BTEC in March this year, six weeks before they leave the school.

Q: But the course (BTEC) is a two year course

T<sub>15</sub>: A two year course, yes.

Q: How can you do it in ...

T<sub>15</sub>: Deputy headteacher clicks her fingers and said we have to. It was obvious they were not going to finish their ... course. Now these kids have left school, they were not coming in. She said they had to come in, otherwise all our results were going to go through the floor. So, we have got to falsify the ... results. It is going to come down to falsifying the entries and I am not bloody doing it.

Q: So who will do it then?

T<sub>15</sub>: I don't know. With this group I have been told that they have got to get their coursework, you can't get the kids in. You have to make phone calls, that what's missing, look at this [T shows a piece of coursework with the top sheet filled in but no work inside]

Q: It is all missing!

T<sub>15</sub>: These kids have not been seen. Look, she is not coming in. This one is not coming in, this one has disappeared, M, we have not seen him at all, but he has been in school, let us get him in here. I have got at least 9 kids working in there (in T<sub>15</sub>'s room – another colleague supervising them). I have got a couple who have finished. I have got five or six done out of 16. And that is put on you now. That is absolutely ridiculous.

Panic management, panic management!

Q: But what does it mean for the kids?

T<sub>15</sub>: There is nothing they are going to get out of this. They are getting nothing out of this, copy that, copy that, copy that then we are marking it and if it gets moderated

and we are told it is not right we will have to call the kids in again effectively. I will just slip it through, sod it.

Q: For how long will you be calling them back? Do you have to get them to a minimum grade or something?

T<sub>15</sub>: They have got to do a lot of units – that is two years work, it is BTEC, a BTEC is like a GCSE. So, it should take two years.

Q: This is almost unbelievable.

T<sub>15</sub>: It is panic management. If we had started in November of last year we would have had a fighting chance. We had told them that they were not going to get their grades in the GSCE exams.

Q: So, what grades are you expecting them to get now?

T<sub>15</sub>: Grade C

Q: Grade C?

T<sub>15</sub>: Two grade Cs

Q: In your professional judgement, what's going on here?

T<sub>15</sub>: I am playing the game. I am just playing the game. ... All this chopping and changing - They have built everything on it – where are your GCSE A to C grades coming from? From the BTEC which is being abused.

Q: So, if you remove the BTEC element what kind of GCSE results would you be looking at?

T<sub>15</sub>: They'll go through the floor. That's why it has got to be, but you can't just swap over the course six weeks before the kids leave, can you?

T<sub>15</sub> continues,

T<sub>15</sub>: They were told that but we were told no. ... [Deputy headteacher] panicked and at the end of March says I want them changed to BTEC. Absolutely disgusting!

Q: And the amount of stress it is producing on you lot?

T<sub>15</sub>: Yes, it is too much. The kids are coming in now. If it does not get done, it does not get done the kids are going to suffer because they won't have any results.

Q: So how many Year 11s this year will be doing BTEC?

T<sub>15</sub>: 16 in this group.

Q: And in total in science?

[T<sub>15</sub> checks on the computer and prints out a list]

T<sub>15</sub>: Well, count, 116?

Q: But, the red lines will have to be removed won't they?

T<sub>2</sub>: Yes, take them out then, call it 110

Q: 110 kids?

T<sub>15</sub>: Yes, 110 kids now you work that out

Q: Nearly what?

T<sub>15</sub>: 30%, isn't it?

Q: Yes,

T<sub>15</sub>: 30% of your grades, if they don't come in, if they [exam board] put a block on that then 30% of your grades will go down the swanee.

Taken together with the table below (Table 7.6.1) in which 100% A to C passes in BTEC science is shown, the interview *indicates that the practice of 'adjusting' the marking to produce 100% A to C grade passes is a 'normal' practice at the school.*

**Table 7.6.1 Table of results for 2010 for Vocational Courses****CHS Level 2 Vocational Courses - Results for 2010**

<b>Subjects/Exam Board</b>	<b>Entries</b>	<b>4 A* - C</b>	<b>3 A* - C</b>	<b>2 A - C</b>	<b>1 A - C</b>
<b>ART / OCR</b>	16	10		6	
<b>Business Studies /OCR</b>	105	8		90	
<b>ICT /OCR</b>	250	153	6		90
<b>PE / Edexcel</b>	15			11	
<b>Science BTEC/Edexcel</b>	<b>102</b>			<b>102</b>	

It would appear that moderation by the exam boards must be minimal given the 100% A – C passes in BTEC science. The above exchanges and the table of results show how around 30% of the Year 11 pupils are neglected and then used in the most cynical way to boost the school's GCSE results, and enhance its league table position.

The following interview with the GCSE science co-ordinator T<sub>21</sub> at the school (CHS) shows the benefit to the school of the BTEC 'fiddle':

Q: Right, ok, so, the overall Key Stage 4 performance in terms of exams, where is the department at? Is it at the average, is it at below average?

A: We got 61% A\* to C.

Q: Is that counting the BTEC?

- A: Yes, I think so.
- Q: And without the BTEC?
- A: It is about 46%
- Q: Will that be below...
- A: Below average, yes.

Thus BTEC contributed 15% extra GCSE passes to the department's total of 61%. Other subjects such as ICT, as shown in the above table are also obtaining close to 100% GCSE passes. How much 'fiddle' is involved in this? ICT was not my focus, so I did not pursue this.

The above was an accidental discovery towards the end of my fieldwork. Nevertheless it shows the extent of these distortions which have crept into the assessment processes at the school level. Coupled with the findings that the schools are marking up Key Stage 3 (age 14) levels of attainment, the accuracy and reliability of schools' internal summative assessments must be in doubt.

It is a sobering thought that if such practices are being repeated elsewhere, then the headline figure of 68.1 % national A\* to C grade passes at GCSE (METRO Friday, August 23, 2013) will perhaps need to be significantly revised downwards to take into account the BTEC in science 'fiddle-factor'.



### **7.6.2 The quality of teaching at DCS**

At DCS, despite the fact that it has a modern building with spacious laboratories the quality of teaching in science has remained only at a satisfactory level. The school professes to be pro-pupil, monitoring by senior managers is considered efficient and the school prides itself on being data-rich (maintaining an electronic record of the results of unit tests and school exams) yet the quality of teaching in science nevertheless has remained weak.

The quality of teaching in the science department at DCS was deemed to be satisfactory by the Ofsted inspectors in 2005, the deputy headteacher (DHT) and the head of department also made the same judgement. The head of department concurred. Here, too, the grades are ‘adjusted’ upwards by awarding two grades Cs to pupils doing BTEC (80 pupils out of 240).

During the 2008 (Inspection date: 6/11/2008) Ofsted inspection no lessons were observed, or pupils’ books scrutinised (features of Section 5 inspection framework). Thus inspectors did not give a subject specific judgement on science. Schools’ own monitoring gave the science department only a satisfactory category.

My findings indicate that the main reasons for weaknesses in the quality of teaching are:

- weak marking and feedback
- dominance of tests and exams

- low expectations of pupils – most pupils only do the Single Award science at GCSE
- teachers' lack of knowledge especially of Sc1 and Sc4, and
- teachers' lack of knowledge of progression in science as outlined in the National Curriculum, (often referred to as pedagogical knowledge).

### **7.6.3 The quality of teaching at WCS**

The quality of teaching in science at WCS has been historically at the borderline between satisfactory and inadequate. The 2007 Ofsted report noted that the overall quality of teaching in the school was only satisfactory:

“Teaching and learning are only satisfactory overall reflecting inconsistent practice”.

The report also noted that,

“... two issues identified in the last inspection report remain areas for further action; specifically provision for students with special educational needs or for whom English is an additional language and, secondly, the provision for personal, social and health education (PSHE) and tutor time. This is symptomatic of the lack of opportunities for students to develop spiritual awareness or a corporate identity”.

In my interview (21/07/11) with the deputy head teacher he confirmed that the quality of teaching in the science department has been between inadequate or barely satisfactory, stating:

“That is a good question.... in January I would have said it was clearly inadequate with at least 50% of the department judged as inadequate in the review that I commissioned with the

Local Authority. Since then we have been doing a lot of work via the project board which we set up to address the issues within science.”

In relation to science my findings indicate that teachers’ subject knowledge and pedagogical subject knowledge are the main issues. These are reflected in poor marking and inadequate feedback. Given these issues the school’s focus on pupils’ behaviour is an example of what, to appropriate the Foucauldian concept of a gaze (Rabinow, 1984), we may call a *misplaced gaze*.

### **7.7 Restricted access to the science National Curriculum - a curriculum divide: C.P. Snow’s ‘two cultures’**

Unequal access to the curriculum operates on various levels. Primarily, any weakness in English leads to exclusion from much of the ‘entitlement’ curriculum. Pupils in the lower sets do not have access to the full range of the National Curriculum. This is because the available teacher resources are geared towards the top sets by the mechanism of triage. Moreover the increasing emphasis on ‘How Science Works’ as distinct from the substance of ‘Scientific enquiry’ means that discussion of science in the news and talking about how science develops has practically replaced laboratory investigations and ‘hands on’ practical activities for most pupils. Consequently, most pupils lack laboratory skills, and even the higher sets find it difficult to complete GCSE coursework, which requires laboratory and investigative skills, without ‘excessive’ help from teachers. A lack of physics specialist teachers contributes to the lack of coverage of more than a quarter of the National Curriculum (Sc1 and Sc4). Lower sets do not generally have access to

specialist physics or chemistry teachers and frequently have supply teachers who are paid on a daily rate and leave every few weeks.

The publication of *Beyond 2000* (Millar and Osborne, 1998) which argued for a distinction to be made between ‘science for the scientist’ and ‘science for citizen’ has contributed to the unequal access to the National Curriculum in science observed during my fieldwork. After the initial opposition to the imposition of the National Curriculum in 1989, science teachers had begun to see the advantages of an ‘entitlement curriculum’ for all pupils. I was Head of Science at Wanstead High School, London, at that time. There was one significant advantage of the National Curriculum which was that for the first time in the history of science education in England all children between ages 5 to 16 were expected to have access to the main disciplines in science: biology, chemistry, earth science and physics. This would offer better opportunities for girls to do physical sciences and for boys to do life sciences. It would offer all children access to a fuller curriculum in science – an ‘entitlement curriculum’. I along with others saw this potentially as a wonderful opportunity, but the funding implication for in-service training of teachers was massive. This raised a doubt as to whether there would be sufficient funding to support the significant programme of in-service training needed to re-train the thousands of science teachers whose subject specialism only included one out of the three main disciplines; biology, chemistry and physics. In particular, would there be the political will to provide the funding needed to undertake the training programme to retrain biologists in chemistry and physics, in the light of the fact that there was an acute shortage of physics teachers, a serious shortage of chemistry teachers. There

was another serious question, would secondary science teachers be willing to undertake such a training programme?

In January 1991 I became a senior lecturer in education (science) at Anglia Higher Education College Brentwood, Essex. Here I provided in-service training courses in science for serving primary school teachers. Government funding was available for such training and the College offered a 20-day science course spread over two terms. (This was later reduced to 10-day and then to 5- day courses). No such courses were offered to secondary school science teachers. It was thought that there was a need for such courses for primary school teachers as they had to teach 10 subjects and most did not have a suitable qualification in science. Although in my experience such a need existed in secondary science too, it was not generally acknowledged. Secondary science teachers themselves were also reluctant to acknowledge such a need. Thus, there was no programme to provide in-service training to serving secondary school science teachers.

From January 1997 I moved on to be Science Inspector with a Local Authority (LA), where I had responsibility for promoting science across all the schools. With two colleagues, one primary and one secondary, we provided in-service training to both primary and secondary school teachers. As a LA we also offered 10 day and 5 day courses to primary school teachers and demand for such courses remained strong even when the funding was reduced. Many primary school teachers were keen to improve their knowledge of science and enjoyed the opportunity to do so.

In contrast to the enthusiasm shown by the primary school teachers, the secondary school science teachers did not show much enthusiasm for in-service training or Continued Professional Development (CPD). The lack of interest in CPD was not because the need was not there, but because the culture of CPD in secondary schools was minimalist. Secondary school science teachers considered themselves specialists. However, their specialism was generally limited to one discipline and it was often biology. Teachers coped with the lack of chemistry and physics by relying on textbooks and other resources without necessarily understanding the key concepts and skills involved in these subjects.

### **7.7.1 The emergence of ‘How Science Works’**

One of the curricular changes which has happened in the last six years is the introduction of ‘how science works’. This comes directly from the ‘science for the citizen’ category and seems to have replaced ‘Scientific enquiry’ which was the main theme under which pupils learnt how a scientist works.

How science works emphasises ‘science in the news’ and the use of secondary data to interpret the results of experiments pupils have neither done themselves nor seen demonstrated. So, pupils talk about ‘science in the news’ which is often at a superficial level because the knowledge and understanding of the topic being talked about have not been taught. Pupils become vaguely familiar with the subject matter but lack a good knowledge of it.

A clear example of this approach was witnessed at DCS, the ‘outstanding’ school in my sample, where two science teachers, T<sub>26</sub> and T<sub>27</sub>, taught their Year 8 classes for

two terms without any experiments or investigations. T<sub>26</sub> was in her third year of teaching and T<sub>27</sub> was the deputy head of science. In addition, there was a complete lack of rigour in what was taught or learned during this time. Take graphs for example. One attempt by P<sub>1</sub> to draw a line graph using secondary data contained many errors and showed a lack of the requisite basic skills. It was not marked.

Samples of work from the books of four pupils' (P<sub>1</sub>, P<sub>2</sub>, P<sub>3</sub> and P<sub>4</sub>) showed physics work is either incomplete or incorrect. No science is corrected. P<sub>2</sub> and P<sub>4</sub> have regressed from their primary school Key Stage 2 (Year 6) attainment level, going down from level 3a in Year 6 to level 3b in Year 8 (P<sub>2</sub>) and down from level 4a to level 4b (P<sub>4</sub>). The expected average progress is 4 sub levels, so they should be at 5c and 6c respectively. P<sub>1</sub> has stayed at the same level, 5c, which he gained in Year 6. He should have moved up to at least 6b. Only one out of 4, P<sub>3</sub>, made one level progress moving from 5c to 6c, but did not reach level 6b which would be average progress.

### **7.7.2 The National Curriculum – an ‘entitlement’?**

The scrutiny of pupils' writing books shows that many pupils in the case study schools have only a limited access to the science National Curriculum (NC) Programme of Study (PoS). Over a period of more than five years pupils' books were scrutinised to discover how well the PoS was covered. The focus was particularly on two Attainment Targets; Sc1, ‘Scientific enquiry’ and Sc4, ‘Physical processes’ (DfEE/QCA, 1999a), as my exploratory field work had indicated poor coverage of the National Curriculum Programme of Study in these areas. This

evidence arose from the scrutiny of pupils' books from four teachers at DCS in the exploratory phase. The teachers were: T<sub>1</sub> (Headteacher), T<sub>2</sub> (Deputy Head teacher), T<sub>3</sub> (Head of Department) and T<sub>4</sub> (Key Stage 3 science co-ordinator). Given that it was a new school with new, well-resourced laboratories, only 300 pupils in total and the four most senior teachers teaching, I expected the quality of marking, the quality of coverage and the quality of teaching all to be good, so it would provide a good baseline for my fieldwork.

The King's College survey of science teachers' CPD needs (2000) had highlighted subject knowledge of science teachers as the main issue raised by head teachers. It had also reported that newly qualified teachers' had identified Sc1 (Scientific enquiry) as a priority area for their further development through CPD. Bearing these in mind and noting that teachers T<sub>2</sub>, T<sub>3</sub> and T<sub>4</sub> were all biology graduates and T<sub>1</sub> was a Geography graduate, a question arose about whether the national curriculum was being covered.

The following provisional findings emerged from this initial book scrutiny:

- There was very little coverage of Sc4 (physics)
- Teachers were unsure of their own knowledge of topics in physics and indeed their comments contained errors
- The work in Sc1 and Sc4 remained at a low level – L4 i.e. the average for the end of primary phase.

When my main study began in 2005 I continued a periodic sampling of books from new teachers to check if my initial findings were borne out. Small samples of books



in 2006/7 and 2008/9 and a focused and relatively large sampling of books in 2009/10 confirmed the above findings. The coverage of Sc1 and Sc4 continued to be very weak and remained at Level 4. The main reasons appeared to be teachers' inadequate subject knowledge of both Sc1 and Sc4 and of the progression in the National Curriculum in science, which Black et. al. (2003 and 2004) had described as 'pedagogical subject knowledge'. By the description 'pedagogical subject knowledge' I mean the curriculum, the National Curriculum Programme of Study and pupils' prior knowledge, understanding and skills acquired in science in their primary schools, together with knowing what should be taught at what age in the secondary school.

## **7.8 Science Teachers' workload and the nature and the quality of CPD**

Teachers' workload is high with most teaching 21 out of 25 lessons per week and seeing upwards of 250 pupils during the week. The marking of tests and school exams takes up a significant amount of teachers' time. CPD is haphazard and generally focused on the preparation for Ofsted inspections which in turn are focused on 'tips and tricks' intended for the Ofsted visit, with little or no improvement in overall pedagogy. In addition, there is a huge imbalance of subject specialism amongst teachers in all the case study schools. At DCS there has been no physics teacher during the whole period of my field work (Table 8.3.2).

### 7.8.1 Teachers' workload

I carried out an in-depth interview with a science teacher (T<sub>24</sub>) at DCS on teachers' workload. I examined the time-table of the participating teachers at all three schools. The teachers had very full time-tabled commitments.

In addition, the end of module tests and the end of Year exams alone run into over a thousand tests and exam scripts in a year altogether for the six case study teachers. The marking of these is mandatory and marks have to be entered on the schools' database by the deadline provided for each test and exam.

Marking of homework and carrying out of formative assessment are also considered a teacher's professional duties but are not mandatory. Given the burden of 'mandatory marking', the marking of homework and marking for feedback are often neglected. This is summed up by T<sub>24</sub> from DCS as follows:

"Official assessment as required by Year 8 Scheme of Work - 8 question exam – entire lesson then an hour to mark. Then there are biology, chemistry and physics units in each module. At the end of each unit there is an hour multiple choice exam which gives them a level. You have to get 6 out of 10 right to get that level. It does not really show whether they have understood it or not, followed by 16 questions SAT's style end of module paper. This is repeated for the next module three times and a module test. This is the same for Years 7 and 8. It does seem excessive."

He continues,

“Marking of books and h/w: I try once every half term. Taking it one at a time [marking each piece as they are done] does not work. It can leave you a bit lacking in energy when you have done 8.15 am to 6 pm every weekday”.

This teacher taught 39 lessons out of 50 over a two week period.

### 7.8.2 CPD

At CHS there is a ‘Learning and Teaching forum’ which takes place for two hours twice a term (Appendix A8). All teachers are expected to attend this and it is the most organised CPD programme compared to those in either DCS or WCS. The new DHT at CHS set this up and is determined for it to succeed. The ‘Learning and Teaching’ forum focuses on an interpretation of ‘Assessment for Learning’ which the DHT claims as based on Dylan Wiliam’s work (Wiliam, 2009). However, it is prescriptive and it has been used as a tool for coercive monitoring and effectively ‘bullying’ teachers. See interview with T<sub>21</sub> below:

“At the beginning of the last term I was given a couple of inspections in very quick sequence because in the first of those inspection was deemed that my lesson had been inadequate. I wasn’t, the lesson methodology wasn’t actually following that which is deemed correct for the school at this time. The way that delivered the lesson I think was correct for about five years ago. I have not actually updated my, the way I deliver lessons to match with the CPD that we have been receiving recently. It is harping back to that idea that if you receive enough CPD which is not actually taken on board and seen as serious by the school then eventually it all becomes much of a one- ness – this is another CPD session, very interesting, and in fact it was very interesting. The ideas seem to be very good to me but how to actually implement them in the class? There was no impetus to actually going through the effort of doing that. So, when I was inspected at the beginning of this year none of the ideas which were in the CPD were actually present in my lessons.

Following lesson, another observation saw that actually this was a pattern and that I was really not implementing the practice that has been demanded of us. So, for the last ten weeks I have been on a programme of observations and feedback from senior management to try to modify my approach to teaching.

Q: Hmm,

A: Feedback was not given as much as negative comments but advice on what changes need to happen. So, that has been very successful. Unfortunately, still very, very stressful, so, for the last ten weeks I have been coming to the school feeling sick, not wanting to come into school, going home and being very tired, having had a very difficult time, so, it has been a very difficult period for me.”

Following this sequence of monitoring the teacher was persuaded to leave.

At DCS there is a more informal arrangement of CPD. There are meetings where Heads of Department discuss aspects of pedagogy. The teachers’ Handbook provides detailed guidance on curriculum, pedagogy and assessment. On pages 41 and 42, for example, it describes the purposes of assessment, quoting from the TGAT report, and outlining five purposes:

- formative
- diagnostic
- ipsative
- summative
- evaluative

As expressed by T<sub>24</sub> earlier, (section 7.7.1) there is very little formative or diagnostic or ipsative assessment taking place. There is no time for them to happen.

At WCS there are whole staff meetings at the beginning of the academic year in September and also at the start of the spring and summer terms. Some whole-school issues are discussed. Prominent among these are 'behaviour management' and preparation for Ofsted inspections. The school hires external consultants to provide 'support' to teachers often for 'quick fix' purposes, as the interview with T<sub>31</sub> indicates (3<sup>rd</sup> February 2011):

Q: So, after the inspection do you feel something positive has come out of it?

A: No. We were given a de-briefing. There was no positive comments made whatsoever. It was all negative. You know, it was said that we can fix these things and we are going to try and fix these things and the stuff that we can fix – it is not going to be difficult. However, there was nothing mentioned the fact that we have now, before there was no Scheme of Work for Key Stage 4, you know, and now we have one in place that everybody is following. Teachers are on similar sort of lessons. The kids who are doing Applied Science, is now focused this year. It was not very focused last year. Different teachers were doing different experiments for different units of Course Work. This year everybody is doing the same experiments. So, we have all had the opportunity to discuss it. We had a CPD where we discussed it. We looked, you know, thrashed out all the issues, practical issues, logistics and working with the course work and it is much more focused this year in terms of Key Stage 4. Strategies have been put in place to, you know, for improvement and that was not recognised. It was just that – it was recognised what the issues were, which we already knew and we had started to make changes. It was not recognised that it takes a while and that was not going to happen overnight.

Q: Is that because, it can't be a quick fix?

A: yes,

Q: And the 'quick fix' is required for what purpose then?

A: For the results

Q: For the results? And the results are required for?

A: For the League Tables

Q: League Tables

A: Not for the pupils.

The CPD at CHS is organised but is very narrowly focussed on tips and tricks intended to satisfy Ofsted requirements and is being used to 'push teachers out'. CPD at DCS and WCS is unsystematic, unorganised, unfocused and infrequent. At WCS there are attempts to provide CPD with the aim to 'fix' it and the consultants claiming that the 'stuff that we can fix – it is not going to be difficult'. This is to achieve 'quick fixes' and short term gains linked to the perceived requirements of Ofsted.

## **7.9 Summary of chapter 7**

This is the main data chapter. It has not always been possible to cut out the interview data and replace them with a summary because many of the participants express their deep feelings through their words. As the thesis set out to discover the teachers' voice and pupils' voice, it feels right to present some of the telling words as spoken by the participants for the attention of the reader. This makes the chapter a relatively long one.

In this chapter I have considered the available data from the case study schools to analyse the nature of assessment both formative and summative, the quality of

teaching, pupils' access to the National Curriculum and teachers' engagement and opportunities for CPD,

### **7.9.1 Poor quality of marking and feedback**

The scrutiny of pupils' books shows that the quality of marking and the quality of written feedback remain poor. Teachers' workload is heavy and pressure to meet deadlines for summative assessment is huge. Given these constraints, teachers say that they are unable to mark well and provide good feedback. This point is succinctly made by T<sub>24</sub> (7.7.1) in relation to marking and feedback.

### **7.9.2 Poor quality of teaching**

The following conversation with T<sub>21</sub> summarises the issues regarding the quality of teaching:

Q: Have you been observing lessons as part of the school programme of lesson observations?

A: Yes, I have.

Q: What have you found in terms of the quality of teaching in your department?

A: Lots of satisfactory teaching, a couple of unsatisfactory teaching, mostly because of a lack of interaction with the students, and one or two good lessons.

Q: Have you done an analysis of where the unsatisfactory lessons were and with which sets?

A: Generally speaking the ones I looked at which were unsatisfactory were generally in Key Stage 3.

Q: Ok,

A: Or BTEC

Q: Hmm.

A: I had seen three unsatisfactory lessons observing with the Deputy headteacher, (we went for joint observations), and they were all either with Years 7, 8 or Year 9.

Q: I see. And the good ones you saw where did you find them?

A: They tended to be specialists teachers teaching their specialist subject.

### **7.9.3 Poor coverage of the National Curriculum**

Over more than five years of fieldwork I was able to scrutinise more than 40 pupils' books from the classes of 25 teachers which clearly indicate that Sc1 and Sc4 are inadequately covered. Teachers cover their specialist discipline reasonably well, but even there they often lack pedagogical subject knowledge of the Programme of Study. Thus much of the work covered does not connect with or build on what pupils have done in their primary school. Big gaps in pupils' knowledge, understanding and skills, especially in Sc1 and Sc4, remain, which could explain the lack of A/A\* passes at GCSE and the relatively high dropout rate in A level Physics at CHS. In addition, the introduction of 'How Science Works' has further weakened the coverage of 'Scientific enquiry' (Sc1) so much so that teachers feel they have to tell them what to do in the coursework as was the case at DCS (T<sub>24</sub>'s evidence) and the 'fiddling' of BTEC results at CHS.

### **7.9.4 A lack of CPD culture**

There is no real culture of CPD at the case study schools rather there is resistance to CPD. The growth of the accountability culture has eroded whatever possibility there was for teachers' self-reflection and desire to seek and attend appropriate CPD.



Teachers' professionalism has been further devalued and teachers are often fearful of asking for help (evidence given by T<sub>22</sub>).

#### **7.9.5 Key Issues**

A number of important issues immerge:

- Summative assessment continues to dominate
- Summative assessment is distorted under the pressures of accountability and the League tables of results
- FA and AfL barely feature in the science teaching of the case study schools
- Arrangement for teachers CPD is generally inadequate
- Appropriate CPD provision would be required to improve teachers' skills in assessment, not only in FA and EA but also in SA.

## **Chapter 8 Data and Analysis continued:**

### **The misplaced gaze: the absences and their consequences**

#### **8.1 Introduction**

This chapter considers the inadequacies described in chapter 7 and discusses their consequences. It describes how a ‘misplaced gaze’ has guided and determined the course of school science in England for over a century and a half and continues to do so. I consider first the historical misplaced gaze (8.2) then the current misplaced gaze (8.3) because it is not that the historical misplaced gaze disappears; the current misplaced gaze is overlaid on the historical one. I then consider in section 8.4 the ‘absences’ resulting from this misplaced gaze. In section 8.5 I discuss the causally efficacious consequences of the ‘absences’ on pupils and teachers and consider their possible impact on the wider society. Section 8.5 gives a brief summary of chapter 8.

In chapter 7 I have analysed the nature of assessment, the quality of teaching, the provision of CPD and the access to the National Curriculum in the case study schools. In this chapter I continue the data analysis with the aid of a model of schooling which fits the features of the current provision, i.e. the factory-model. My findings suggest that the factory-model of schooling has been intensified during the last two decades with the standards agenda tightening its grip on every aspect of school education. One of the key instruments of the ‘standards agenda’ is the monitoring of schools and teachers through regular Ofsted inspections. Another key plank of this agenda is the so-called ‘School Self-Evaluation’. I have described it as

‘so called’ because my findings suggest that it is not self-evaluation but more akin to self-congratulation with schools playing the ‘Ofsted game’ using fair means or foul. What the Ofsted inspectors and the school managers fail to see, i.e. the ‘absences’ from their gaze is the subject of consideration in this chapter. Here the nature of the monitoring by school managers and Ofsted inspectors is discussed with regard to what is monitored and, more importantly, what is absent from the monitoring process. The output is monitored. The bare minimum of training is provided to teachers in the operation of the factory-model. Teachers’ productivity is measured.

Given the above context, I consider the interplay between theory and data in terms of four emergent headings: (1) the ideology of ‘innate ability’; (2) the standards agenda; (3) the inadequacy in science teaching and the generation of a false dichotomy between ‘science for scientists’ and ‘science for citizens’ and (4) the complicity of teachers in a disabling status quo. I present an immanent critique (Bhaskar 1986) of the current ideology of assessment and its consequences, focusing on the unequal access to science education and the role of summative assessment in sustaining inequality of opportunity and continuing to underpin the ‘factory-model’ of schooling outlined in chapter 5. Together, these factors have led to a systematic neglect of the science education of at least half of the pupils in the case study schools. The national ‘standards agenda’ with its primary focus on the output in terms of GCSE grades, league tables of results, Ofsted inspections and its consequent obsession with the role of the leadership, as opposed to the importance of the contributions teachers make and need to make, is repressive, ill-focused and a prime example of misplaced gaze. At the same time the silence of teacher

organisations (trade unions or the subject associations) in the face of such monumental failure by the state comprehensive schools is deafening.

## **8.2 Misplaced gaze:**

### **Historical focus on ability leading to ‘triage’ & ‘double triage’**

Historically, the ‘tripartite’ system of schooling has focused on revealed performance, which has been assumed to accurately measure ability or potential, which in turn has been presumed to be more or less innate. This is an example of misplaced gaze on a colossal scale when one considers how effective it has been in denying a clear majority of pupils, opportunities of a good science education at school over a period of more than six decades in England. The notion that humans can be divided into three categories; ‘academic’, ‘technical’ and ‘backward’, on the basis of a spurious process of measuring, justified by ideas and concepts originating from Francis Galton’s eugenics, is preposterous but nevertheless has been powerful. The mode of schooling based on such an out of date idea in the form of the ‘tripartite’ system thrives even today in state comprehensive schools. And now the intensified accountability culture unleashed by the ‘standards agenda’ is encouraging the practice of ‘triage’ of pupils and moreover, as one of my key findings shows, has brought into play a ‘double triage’ where teachers are subjected to this unjust, inefficient and unnecessary process.

So, the ideology of innate ability has resulted in a system of schooling which promotes ‘triage’ and ‘double triage’. Furthermore, it fails to identify either pupils’ real learning needs or teachers’ professional learning needs. Pupils are subjected to ‘triage’ with respect to access to resources, curriculum and attention. Only the pupils

in the top sets receive adequate resources, have access to a satisfactory curriculum and receive teachers' attention regarding their learning and progress. Others are contained within the system, receive unsatisfactory teaching, have access to an inadequate curriculum and receive very little attention and guidance from teachers. As a very thoughtful Year 8 pupil (P<sub>29</sub>) at CHS observed:

"If you are not in the Fast Track the teachers think that you are stupid. I know it because I was in the Fast Track and now I am not."

Another pupil (P<sub>28</sub>) said, "... they only care about the Fast Track."

### **8.3 The misplaced gaze: the Standards Agenda:**

A key part of the 'standards agenda' is the monitoring of schools by Ofsted. Schools are inspected between every three to five years. Schools in the 'satisfactory' category are inspected every three years. Schools in the 'inadequate' category are visited every six months and inspected every year. Schools in the 'good' and the 'outstanding' categories are inspected every five years. However, if a school's annual league table position deteriorates it can be inspected much sooner within the three or five year's window. During the period of my field work the Ofsted inspection framework changed from being Section 10 inspections to Section 5. This changed the rules governing inspections in ways which required fewer inspectors. The inspection of individual subjects was dropped so specialist subject inspectors were no longer required. Fewer lessons were observed and fewer books were scrutinised with the effect that in science, for example, the quality of teaching, the coverage of the curriculum and the nature and quality of marking ceased to be scrutinised by science specialist inspectors.

Noting that the science Programme of Study (PoS) should be the minimum entitlement curriculum for all pupils and that monitoring of teachers and schools are fully in place, it is hard to explain how this lack of access to the minimum entitlement curriculum has been allowed to continue unchallenged for so many years and continues to be so. So, if not this, then what is being monitored?

### **8.3.1 Monitoring by the schools' senior managers**

The case study schools have put in place systematic monitoring procedures which are, it is claimed, focused on improving the quality of teaching and learning and aimed at satisfying Ofsted requirements. Thus when Ofsted Inspectors (under the Section 10 inspection framework) used to scrutinise pupils' books for assessing the balance and quality of coverage and the quality of comments written in the books, the school managers would monitor the 'quality of marking' through regular and, sometimes frequent, 'book-checks'. Some attention was also paid to the subject content because there used to be subject inspections by subject specialist inspectors. Section 10 inspections were replaced by Section 5 inspections in 2006 (see time-line section 2.4), with the new inspections relying much more on the schools' own self-evaluations. This meant the senior managers' view of how good their school was provided the starting point and basis for the inspection. Schools now employed various means of presenting their best image to the inspectors who had no subject specialist inspectors and largely saw what the school managers selected for them to see (see the description on how schools managed inspections below).

### 8.3.2 Monitoring at DCS

DCS has experienced a higher than average turnover of science teachers. There is a huge imbalance of subject specialism in the science department. However, the school managers have failed to recognise or have overlooked the imbalance of subject specialism in their science department and thus failed to address the CPD issues it should have identified.

**Table 8.3.2: Staffing at DCS (Science) 2010/11**

Subject	Percentage (%) of teachers
Biology	81
Chemistry	13
Physics	6
Other	0

The subject specialism of science teachers is consistently heavily biased towards biology. Such a significant imbalance of subject knowledge remained apparently unnoticed and certainly unaddressed throughout the duration of my fieldwork. The head of science had obtained a certificate in physics from the Open University on his own initiative, was motivated by his intention to emigrate to New Zealand which he subsequently did.

### 8.3.3 Monitoring at WCS

The science department at WCS has also experienced an above average turnover of staff and, as in the case of DCS, the imbalance in the subject specialism has been overlooked.

**Table 8.3.3: Staffing at WCS (science) 2010/11**

Subject	Percentage (%) of teachers
Biology	46
Chemistry	15
Physics	15
Environmental Science	7.6
Psychology	7.6
Biology and Psychology	7.6

### 8.3.4 Monitoring at CHS

At CHS the balance of subject specialism was much better as the table below shows. However, due to a lack of pedagogical knowledge of the teachers and a shared approach to the policy of ‘triage’ a better subject balance was not utilised for all pupils.



**Table 8.3.4: Staffing at CHS (Science) 2010/11**

<b>Subject specialism</b>	<b>Percentage (%) of teachers</b>
Biologist	28.5
Chemists	36
Physicists	28.5
Other	7

These proportions compare well with the following Royal Society's 2007 data:

**Table 8.3.5: Royal Society's data (2007)**

<b>Subjects</b>	<b>Percentage (%) of teachers</b>
Biology	47.6
Chemistry	27.5
Physics	20.1
Other sciences	4.8

The staffing data at CHS shows that a better balance of subject specialism alone may not be sufficient to promote AfL or EA. A weak pedagogical knowledge of

progression in the National Curriculum and historically low expectations of pupils are also crucial ingredients.

### **8.3.5 Monitoring by Ofsted**

As mentioned earlier during the course of my fieldwork there were two main changes to the Ofsted Inspection framework. The Section 10 inspections were replaced by Section 5 inspections (2006). The Section 5 Ofsted Framework was modified in September 2013, after my fieldwork was completed. The main features of section 10 inspections were that there was a large team of inspectors which included subject specialists. Inspections lasted for four days and a summary report was presented to departments and the school managers on the fifth day, usually a Friday. Inspectors observed lessons and carried out book scrutiny to assess the quality of marking and written feedback and also the quality of coverage of the National Curriculum.

The Section 5 Inspection Framework was a much reduced inspection both in scope and with regard to the number of inspectors and the amount of time spent in the school. The team usually consisted of 3 to 4 inspectors. Subjects were no longer inspected and the whole inspection lasted only for 2 days. The absence of subject specialist inspectors meant that the coverage of the National Curriculum Programmes of Study was not scrutinised nor were pupils' books checked for the quality of marking.

#### **8.4 What is absent in the process of monitoring?**

Despite the fact that there is an elaborate monitoring structure in place as described on previous pages, what is absent from the monitoring process is any analysis of teachers' subject specialism, their work load, their CPD needs, whether or not the Programme of Study is being fully covered and whether or not all pupils are receiving the entitlement curriculum. What is not seen is what pupils are receiving as a normal 'diet' of science teaching. There is no scrutiny of whether teachers themselves have a secure knowledge of science. What is also absent is any analysis of the effect on pupils of the coercive procedures of ability grouping, and the effect the severe and relentless accountability processes are having on teachers.

##### **8.4.1 Excessive importance given to the leadership – resulting in a top heavy structure in which teachers become invisible.**

During the last two decades successive governments' have placed more and more emphasis on the role and the importance of the headteacher and the leadership of the school. Ofsted's annual reports for the past 10 years have accentuated the importance of these, while downplaying the professionalism of the teacher. Issues relating to learning and teaching have tended to be seen in the legitimate domain of 'managerialism', while the efficiency of the factory-model of schooling has been assumed to be secured by a regime of extreme competition. Thus the gaze of the policy makers and their appointed monitors remains away from the needs of both the pupils and the teachers.

### **8.5 Impact on pupils and teachers of the absences arising from the misplaced gaze**

The progressive tightening of the ‘standards agenda’ has created a culture of fear in schools. In such a coercive environment teachers are reluctant to disclose their CPD needs. They are forced to ‘falsify’ the results of the summative assessments they carry out, thus masking the learning needs of pupils which the summative assessments could reveal if carried out honestly and rigorously. School assessments have become even more intensive instruments of ‘triage’ in the last decade than previously and teachers, having been party to them, find it difficult to defend themselves when confronted with the process of triage themselves. Under pressure from the league tables of results school managers have raised their ‘expectations’ without dismantling the structure of ability grouping which, in fact, consolidates and continually re-generates low expectations. These are inherent in the ideology of innate ability, the ideology on which ability grouping in the schools is based, and almost inevitably contribute to pupils under-achieving.

Many teachers have continued to believe in the notion of innate ability and the ‘deficit model’ of pupils, despite messages of hope and high expectations emanating from publications as old as *Fifteen Thousand hours* (Rutter et al., 1979) and numerous publications since, notably Carol Dweck’s *Mindset* (Dweck, 2007). Many science teachers remain trapped in the ideology of innate ability, in the mindset of Sir Cyril Burt’s notion of IQ and intelligence. The creation of the ‘tripartite’ system of schooling in England and Wales cemented such a mindset and the factory-model of schooling continually regenerates it. It is therefore a cruel deception to expect

teachers to meet output targets without helping them overcome the effects of the innate ability mindset with the structural and ideological changes needed to negate such an ideology.

The impact on pupils of the shortcomings described in this chapter is also devastating. All pupils are adversely affected by the processes operating at the case study schools. The most able have access to better resources and opportunities but are subjected to 'teaching to the test'. The head of biology at DCS (T<sub>28</sub>) described it as 'spoon feeding' which does not adequately prepare them for Advanced Level study in science. This, he said, was the reason for the relatively high 'drop out' rate in biology 'A' level at the school. A similar high 'drop out' rate was reported at CHS in physics.

Pupils in the middle and bottom bands receive inferior resources including non-specialist and supply teachers and restricted access to the curriculum. They are also denied opportunities which are readily available to those in the top band, for example, visits to museums, to prestigious universities and invitations to hear outside speakers or authors. In the case of the bottom band, access to books such as revision guides for Key Stage 3 Science is also restricted.

In addition, the effect on the pupils in the bottom sets is desperately disabling. For example, the disaffection of some pupils who are sent to the local college at the end of Year 10 on life skills courses is so deep that when they come to the school for two

days a week they create ‘mayhem’ and “urinate in the fountains” (T<sub>5</sub>). Many of the bottom set pupils who remain at school do a BTEC programme, being switched between GCSE and BTEC and enduring a constant change of teachers (often supply teachers). These pupils receive very poor teaching in their BTEC courses but the results are often ‘fiddled’ so they are given a GCSE grade C which is a pass grade (see 7.5.1 for evidence given by T<sub>22</sub>, and T<sub>15</sub> and T<sub>19</sub>).

The model of schooling represented by the case study schools appears to present barriers to learning at every step. By dividing pupils in the top, middle and bottom bands, akin to the tri-partite system and often on the basis of pupils’ prior attainment in English, the schools ration access to the curriculum and resources including access to specialist teachers and text books. Within the bottom band the model produces, in one case study school, a group even below the lowest set, the ‘underclass’ or ‘les misérables’ (the wretched ones) equivalent to a group below the three main socio-economic groups in society. This process of classification is carried out with ruthless efficiency in all three of the case study schools.

### **8.5.1 Impact on pupils at Curie High School (CHS)**

At CHS the boundaries between the bands are maintained rigorously and the status of the ‘Fast Track’ is defended vigorously against any entry to it by ‘undesirables’. All the case study schools operate this system but it is most explicitly demonstrated at CHS, a school in the Ofsted category ‘good’ and trying desperately (see reference to the BTEC science ‘fiddle’ above) to gain Ofsted’s ‘outstanding’ status. My conversation with a science teacher (T<sub>5</sub>), who is also a member of the Governing

Body of CHS, as a teacher representative, revealed how pupils in the 'Fast Track' enjoy access to the best resources whilst the rest of the pupils have to settle for much inferior provision.

Teacher T<sub>5</sub> teaches Set 1(top set) in the Upper Band. She says that some pupils from her set will be promoted to Fast Track but most will remain in the Upper Band. My interview goes as follows:

Q: I understand you have two classes in Year 8 or is it only one class?

A: Only one. It is the best of the Upper Band, so, it is actually a very interesting group.

Q: When you say it is the best of the Upper Band what is it, is it Set 1?

A: Yes, it is set 1 of the Upper Band. [There are five sets in the Upper Band]

Q: Ok.

A: So the children who are most likely to go into Fast Track when they re-group just before they go into Year 9. Don't forget, English can still stop a child from going into the top Band.

Q: Hmm, hmm,

A: So a child who has not got particularly neat handwriting and isn't a particularly good writer they don't get into Fast Track. So the Upper Band has the children with the biggest spread of ability. You have kids who are outstandingly good who could get an A or B at GCSE in science but they are in the Upper Band because the English department felt that their literacy was not good enough for Fast Track.

Thus EAL and EFL pupils and the poor white pupils whose literacy is likely to be weaker are automatically barred from the Fast Track, irrespective of their attainment in science.

In addition, there is the behaviour or a 'good pupil' criterion as another teacher (T<sub>6</sub>) discovered. She described her battle with senior managers on behalf of a pupil in her tutor group (Form group) who she felt should be in the Fast Track but he was not.

T<sub>6</sub>: Yes, he has got an end of Key Stage 3 target of Level 7 which I think is accurate. In the Key Stage 3 work I have done with him he is getting high Level 6s and 7s. He is in the bottom set, Upper band. The reasoning for that when I spoke to his Head of Year is behaviour based. I actually tried to get him moved to Fast Track and I was told that wouldn't happen because his behaviour wasn't good enough. So he stays in the Upper Band and he does the same subjects as the Upper Band kids. He will end up I don't know, doing BTEC IT, probably Single Award GCSE science having done BTEC this year. I don't know what else he will end up doing. But there won't be much opportunity for him to stay on in the 6<sup>th</sup> Form here because our 6<sup>th</sup> Form students come predominantly from the Fast Track.

So, although in theory there is movement between the bands, in practice this rarely happens because the system is based on completely inappropriate grounds i.e. attainment in English not in science. In addition, behaviour is used to exclude pupils even with a mild behaviour issue which might well be borne out of boredom at poor teaching/lack of challenge.

Continuing interview with T<sub>5</sub>:

The teacher (T<sub>5</sub>) says, reflecting on her comments:

"I think that is so true it is painful."



Now, teacher (T<sub>5</sub>) opens up. Her feelings of anger and frustration which she had been concealing so far bubble out. She describes how pupils in the Fast Track are the 'elite' and the privileged:

"But the Fast Track gets treated differently. The Fast Track are the privileged ones. If there is a famous author coming in to talk about how they get inspiration to write their books. 'Ok, Fast Track you can do that. Sorry, Upper Band, Middle Band, go to lessons'."

T<sub>5</sub> describes how the Middle Band, the Upper band and even the lower sets in the Fast Track, get inferior provision:

A: The English department is very keen on getting authors in because we are very keen on the Library. We have invested thousands in the Library.

Q: Yes,

A: Pupils in the Middle Band go to the Library when it is raining so they don't have to go outside. Those in the Upper Band go to the Library because they have got some homework they have been set and they have not got resources at home. The Fast Track pupils go to the Library because they feel relaxed because they can sit round a book or Encyclopaedia. There is a bit where you can sit with magazines and they talk quietly. So, it is a social thing. It is no big deal to go in the Library.

Q: But when the authors come, are only the Fast Track invited?

A: Nobody else. Well, look at the Student Council and the senior students. The only people who can be senior students in Years 9, 10, 11, 12 and 13 are Fast Track. The school has said the only people who can be senior students to represent the school at events, to go to meetings in the borough, to go to meetings outside of the borough have to be Fast Track. There are some real good thinkers, some real good talkers in the Upper Band. No, you can't be one. You are not in the Fast Track. Fast Track, go

on let's get you going to university. Oh, let's go on to London. Oh, let's go to Cambridge, let's go to Chelmsford. Guess who goes on those? (T<sub>5</sub> thumps the table).

Q: How many pupils are there in the Fast Track in any Year group? You have around 300 in each Year group in total.

A: Yes,

Q: How many are in the Fast Track currently, do you think?

A: 150 pupils, something like that?

Q: 150 will be 50%.

A: Something like that. But, then with the famous authors only the top two sets in English are invited.

Q: Even in the Fast Track?

A: Oh, yes. Oh yes.

The teacher continues and states that the status of the Fast Track is defended vigorously. For example, there is zero tolerance of lack of attendance. She states,

“If there are kids who are poor attendees for a variety of reasons, they are thrown out of the Fast Track. Poor attendance, you are out.”

She continues,

“So, it's a bit scary really.”

According to T<sub>5</sub>'s testimony, 150 out of a cohort size of 310 are in the top band, the Fast Track. There are five sets in the Fast Track graded from set 1 to set 5. The remaining 160 pupils are in the Upper Band and Middle Band.

Thus, only around 20% of the pupils, the top two sets of the Fast Track, are the proper elite who would be in a grammar school in the tripartite system of schools. The remaining (around 80% of pupils) would be considered as Secondary Modern or Technical school pupils.

The above description of how banding and setting are organised, shows that CHS is a comprehensive school only in name. It operates a tripartite model of schooling within its boundary walls. Four fifths of its pupils are not valued enough to be invited to listen to an outside author or speaker.

In my interview with the two boys, Demi and Clive, Demi explain the contrasts between the Fast Track and the Upper band:

Demi: Like Fast Track, they get a teacher for biology, science, (Clive: yes), I mean biology, chemistry and physics, one teacher for each where in Upper Band, (Clive:one teacher), you only get one teacher. You only get one teacher for all three subjects.

Clive: It is crap [**quietly**]

Demi: It is bad because one teacher could be better in biology than physics and chemistry.

Demi explains the setting:

“They are basically set in order. There is AS which is for the bad behaved children and there is another band for children that kind of below [not clear] and there is Upper Band which we are in at the moment and then there is Fast Track which I am going into.”

Demi's description of setting is confirmed by the teacher (T<sub>5</sub>):

"We give them an exam and set them in October. In addition, there are those who follow the Literacy focus limited curriculum with only one teacher for up to three years. These pupils go into subsets of the Middle Band and are variously labelled as MU, AS and ML. Some of these pupils are promoted to the Upper Band in Year 10 but are kept in labelled groups such as UU (10Upper Upper) or 10 MU (10 Middle Upper)".

Clive states that:

"Most teachers act as if everyone who is not in the Fast Track is stupid, really."

The 'Upper Band' at CHS is actually the middle band and contains some of the most angry pupils. A few from set 1 of the Upper Band are moved up to the lowest set in the Fast Track but most remain in the Upper Band. The teacher says that many of these pupils feel let down by the school.

My conversation with the teacher again:

Q: Will these pupils go on to do GCSE in Year 10?

A: Well, yes, the Upper Band do BTEC in Year 9, then they will do GCSE in Year 10, and anybody who just does not get a level, they could finish off doing BTEC in Year 11.

The teacher then describes how her pupils in Year 9 Upper Band feel and behave:

"The Year 9 Upper Band, who are technically heaps better than the Middle Band are a miserable group. Half of them got promoted to Fast Track and they are now doing GCSE. The other half of the Year 9 Upper Band has had to carry on with BTEC knowing that they

are going to start GCSE in September. They are **bored**, they are **aggressive**, they are **unresponsive** and they feel they have been **ignored**.”

Q: Hmm, and what about the Fast Track?

A: Everybody is going to do GCSE something, but only Fast Track can do Triple Award science, you can't do Triple in the Upper Band.

The Middle Band at CHS is essentially a 'sink' group which includes all the children who spend up to three years in the Literacy Support group where they have a restricted curriculum taught by one teacher. When they come out of this 'support group' they go into Years 8, 9 or 10 according to their age and are given the label MU (Middle Upper). There is also a group within this Band which is below MU and it is labelled ML (Middle Lower). This Band includes the Academic Support group (AS) who are often referred to as the 'behaviour' support group. I asked the teacher (T<sub>5</sub>) about the Middle Band children. My conversation proceeded as follows:

Q: What happens to the Middle Band pupils?

A: You know, some of the Year 9 Middle Band have gone up to an Upper Band group but it has been designated UU. They are kept together. They are not scattered across the Forms in the Upper Band. They go from 9 MU (9 Middle Upper), to 10 UU (10 Upper- Upper). 10 UU is the best of the two Middle Band classes. Those left in the Middle Band have to do BTEC.

Q: BTEC in Year 10 and what about Year 11?

A: Oh yes. Because in Year 10 they will just do, they will do what's left of their BTEC and in Year 11 they will be sent to the college for skilled based learning. And they will be back in the school two days a week. For them their sort of schooling will finish at the end of year 10.

Q: What do they do for the rest of the time?

A: What in year 11?

Q: Those kids who go to college.

A: Well, the kids who go to college they come in to school 2 days a week – school pays the college. My understanding is that the school gives a proportion of what they get for each child to the college to cover the cost of them being taught at college. It keeps the college classrooms busy. And we get rid of them because they feel they have nothing left to learn (at school) that would be useful to them.

Q: Is this at the end of Year 9?

A: No, Year 10, they are not terribly keen at the end of year 9 but at the end of year 10 they are utterly switched off and when they are on the two days a week that Year 11 Middle Band are not at college and come in to the school they create unbelievable amount of mayhem. It's them that **urinate** in the water fountains. It is them that set off the fire alarm. It is them who ride round the site on bicycles. Oh, woeful!

The pupils who go to the college become invisible most of the time. The schools do not track their progress. They do not appear on the results sheets. At *CHS about 19% are not on the results sheet*. In my interview with the two boys (Demi and Clive), Clive had said that 'they don't care about anybody other than the Fast Track'. I had asked him, what makes you say that? He said that they were always going on about the Fast Track. So, I asked the teacher (T<sub>5</sub>) about Clive's comment. This is what the teacher said:

A: Fast Track have better teachers, better trips, better facilities, made to feel special, separate assemblies. Children aren't stupid. The one thing I have learnt in nearly 40 years of teaching if you want to know what's going on in the school, ask the children.

Q: So, the ethos, from listening to all this makes me uncomfortable.

A: *It makes me uncomfortable as well because I thought we were here for all children and **every** child mattered.*

It appears that once you are in the Middle Band at CHS and in the bottom bands in DCS and WCS there is no chance to move up. The route is closed. Many kids become bitter and angry and begin to hate the school.

One pupil wrote in his exercise book':

***"I hate iS aload Op RuBish"*** – P<sub>32</sub>'s book (Appendix A9)

### **8.5.2 Impact on pupils at Downton Community School (DCS)**

A similar but less elaborate process operates at the outstanding school, DCS.

Expectations of the top sets are much higher than the other sets. The marking of their books is better. In contrast, the pupils in the lower sets are not expected to achieve very much. Even when they are put in a 'mixed ability' group the teaching and expectations remain poor. For example, the attainment of pupils in year 8 taught by one 'good' and one 'outstanding' teacher showed that at the end of Year 8 they (the pupils) had made little or no progress from their attainment at the end of their primary school. For example, of the six pupils from these teachers' group who took part in the research only one improved by 1 level (P<sub>3</sub>) compared to his primary school level. My interview with a teacher (T<sub>30</sub>) who taught one of the two top sets in the same department revealed how pupils in the top set were treated as privileged pupils and the other pupils were largely ignored:

Q: Thank you very much for agreeing to take part in this research. Just talk to you about your Year 8 classes last year and particularly if you had the top set. What set was it?

A: Set 1 – 8R1

Q: How well did they do at the end of the year?

A: How well?

Q: Yes,

A: They did, most of them did very well. We had in the end of Year tests quite a lot of them had Level 7c, then mainly Level 6s. We had one 7b as well.

I asked the teacher about the resources used to prepare pupils for the national tests:

Q: What kind of resources did you use for the revision programme? I am talking about textbooks and things like revision guides.

A: For revision, I used board work quite a lot, I used BBC Bite-size and got them to answer questions using computers. I used Test-Base for practising questions and I used the CGP revision guides but the questions, you can get a book of questions. I used them.

Q: Did pupils have access to the CGP books?

A: No. At parents' evening I recommended that they got a book or use BBC Bite-size but I guess if parents can't afford to go out and buy a book then it is hard.

My conversation continued:

Q: What was the language level of the pupils like?

A: Em, very high. There were no EAL [English as Additional Language] pupils

Q: Hmm, and were they able to use scientific language?

A: Yes, yes, yes.



Q: And what were the expectations like in the classroom? Did they expect to do well?

A: Yes, they were very competitive about grades and they did want to get, you know, they were top set and they wanted to do well. *Sometimes, the Head of Year would come in and say that you are the top set. You need to do well.* People are competing for places in the top set. So, you want to maintain your place in the top set. You are the top of, you know, top of Year 8 and we want you to stay that way. *So, there was always a lot of motivation I think within the school.* [My italics]

I asked the teacher if she taught any other year 8 class:

Q: Did you teach any other sets in Year 8?

A: I taught the bottom set as well.

Q: How many kids did you have in the bottom set?

A: About 20.

Q: And what was their motivation like?

A: Em, it was very, very poor.

Q: What kind of things made you think that their motivation was poor?

A: I think there were two pupils in that class – this is 8 J5, the bottom set. Two pupils in the class who have the reading ability, when they were tested in Year 7 they came out the top pupils in the school. And it was through not being willing to write and not wanting to learn that they were in the bottom set. So there were kids in there with much lower ability than they were and they were sort of mixed in with kids who could not speak English very well, could not read.

Q: And what support was available for you for that class?

A: I had a TA to support the weaker pupils.

Q: What kind of responses did you get from the parents of these kids?

A: From the bottom set?

Q: Yes,

A: [Pause!] Em, how do you mean like?

- Q: I mean were there parents who expected their children to do well or did they come and talk to you, or anything?
- A: They came and spoke to me, very sort of passive! Would sort of sit there and take on the information that you are saying to them. Em, say, we are pleased that their grades were improving. They had good grades in the class.
- Q: What levels did they achieve at the end of Year 8?
- A: Sort of levels 4 to 5 [Level 4 is the average end of primary school level]
- Q: What was the behaviour like?
- A: Em, there were a lot of ... very, very weak pupils but their behaviour was fine. The pupils who were much higher ability and should not have been in that set, their behaviour was very bad.
- Q: What kind of resources did you use for them, the weaker, the bottom set? Did you use anything different?
- A: More sort of cut **and stick activities, more word search**, more literacy based things, highlight key words as you do reading activity, em, **gap fills** rather than questions.
- Q: Hmm, and what kind of books did you use with them?
- A: *We did not have any books.*
- Q: You didn't have any books? Science books I mean.
- A: There were none because you know we were using the Wiki science scheme.
- Q: I am asking about books because there are a lot of science revision guides in the market. As you have mentioned, the CGP books - would those probably have been useful for the weaker pupils? What were the main issues which kept them in the bottom set?
- A: Literacy, they could not read the questions. And they didn't care.
- Q: *And did the science department have a policy for improving their literacy through science?*
- A: No.

Q: No?

A: It was that I had to find the information. So, I had to develop my own literacy tasks.

Q: That seems to leave a lot for the individual teacher doesn't it?

A: Yes, yes

Q: What kind of time-table did you have? How many periods of contact time did you have?

A: 80% teaching

Q: Not unfair but quite high

A: I had, because I had double top set on a Tuesday and double bottom set on the Wednesday and the majority of my time, you know, I have Key Stage 4 and A-levels, was spent preparing.

It is clear from the above conversation that there was very little effort made by the school or the department to help improve the literacy levels of the pupils in the bottom set. There are good inexpensive science books available in the market which could have been used to improve science as well as literacy. What does it say about the attitudes of the school and the department towards the learning needs of these pupils?

### **8.5.3 Impact on pupils at Westfield Comprehensive School (WCS)**

At WCS there are two bands created on the basis of pupils' prior attainment in English and in science there are 6 sets in each band. Pupils are set at the end of their first term in the school on the basis of tests administered by the science department and marked by the science teachers. The DHT and the Head of Department agree that the quality of teaching in

the department is barely satisfactory. The following comments on 21 July 2011 by the DHT of the school illustrate this point:

Q: What's the overall quality of teaching in science, in your view?

A: Well, (a sigh!) that's a very good question really because it is quite fluid and back in January I would have said it was clearly inadequate with at least 50% of the department judged as an inadequate in the review that was commissioned with the Local Authority.

The generally poor quality of marking in the department reflects this judgement. The quality of marking at WCS is better for the top sets. In the middle sets (sets 3 and 4) the quality varies between poor and satisfactory and in the bottom sets (sets 5 and 6) the quality of marking is invariably poor.

There is clear evidence of triage in that the marking of the top set is better than for the lower sets. A comparison, for example of the books of P<sub>21</sub> and P<sub>22</sub> of the top set, (teacher T<sub>31</sub>), with those of P<sub>36</sub> and P<sub>44</sub> of set six, the bottom set, (teacher T<sub>32</sub>) reveals that the quality of marking of 8J1 (top set) is barely adequate but at least there is some marking. In contrast, there is little or no marking of the books of 8J6 (bottom set) between October 2010 and July 2011.

Over the year, before each test the pupils in the top set are given extra tuition at lunchtimes and after school. Other pupils do not receive such help. Teachers say that 'they (the pupils) are not interested'. The teachers, however, do not acknowledge their own role and the role of the school in contributing to this 'lack of interest'.

My conversation with the deputy head of science at WCS (T<sub>31</sub>) illustrates this point:

Q: Schools used to organise systematic revision programmes for Year 9 as preparation for the National Tests, do you still do an intensive revision programme?

A: Yes, it was for the whole of May before they did the National Tests.

Q: Do you still do them or do booster classes after school?

A: Individual teachers do it. Kids will come in, *obviously the top set students* would come in at lunch time, you know break times, and after school. Some come in early in the morning before lessons start.

Q: *And what about the other sets?*

A: The other sets, the middle and bottom end of the cohort just did teacher revision during lessons. That's all they did. *They don't want to do anything.*

Q: *Who? The teachers didn't want to do anything?*

A: *No, the students.* They don't want to *help* themselves. So, you are physically like forcing them or you know or you keep on reminding them of it. It is revision. You have to revise for the exams.

Q: Yes,

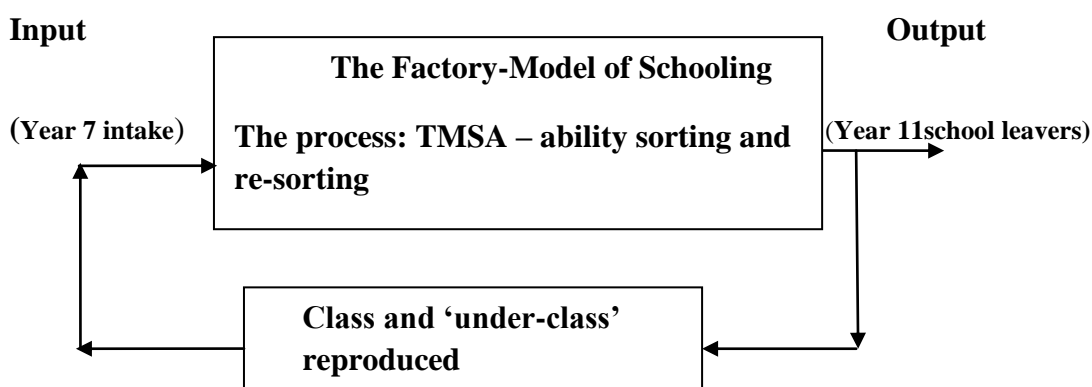
A: And because they already know in the news that SATs is cancelled. You find that at the top end they understood that they still need to do that form of assessment and they have to be secure in public examinations. So the message was quite clear to the top end but coming to the, you know, the *lower* sets they don't *understand why* we are doing it because they keep on saying oh! Miss, why do we have to do SATs when the government says, you know, blah, blah, blah? Therefore, *even telling them to revise was a problem.* [My italics]

What seems to be absent from the responses of this teacher is any thought that the teachers and the school might have had an effect on the lack of interest shown by pupils in the lower sets. She happened to be in charge of Key Stage 3 and was also 2<sup>nd</sup> in charge of the Science Department.

## 8.6 The Factory-Model of schooling: continued regeneration of the vicious cycle

As described in chapter 5 the factory-model of schooling has been intensified in the last two decades with adverse effects on all pupils attending the case study schools. Those deemed ‘able’ are taken through a ‘mill’ of ‘teaching to the test’ where learning by remembering, often without the experience of experimental science or scientific investigative skills, is emphasised and encouraged. Those labelled ‘less able’ are denied access even to what should be the minimum entitlement science curriculum and resources.

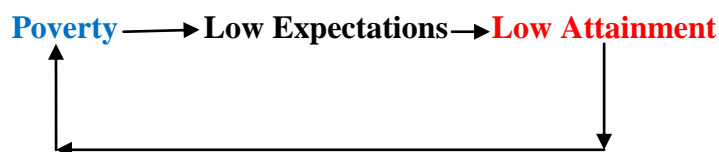
### 8.6.1 The Factory-Model of schooling



**Fig. 1**

The factory-model reinforces the vicious cycle of poverty, low expectations, low attainment and back to poverty and it continually reproduces it for a significant minority of pupils.

### 8.6.2 The Vicious Cycle:



**Fig. 2**

## 8.7 Summary of Chapter 8

This chapter describes the nature of the misplaced gaze both historical and current, heavily influenced by a historical misconception created by false theories of intelligence and IQ. These have been augmented in relation to education in England by the discredited claims of Sir Cyril Burt that ‘intelligence’ was innate and IQ is fixed and can be measured.

Burt’s very influential paper on ‘backward children’ published in 1927 became the backbone of the ‘tripartite’ system of schooling in England and Wales following the 1944 Education Act. My findings suggest that despite decades of campaigning to get rid of the 11+ and for the creation of a fully comprehensive system of schooling, the influence of summative assessment and the ‘tripartite’ system remain dominant within the walls of comprehensive schools in England, through rigid systems of streaming and setting.

Modelled on a factory production line and triaging the allocation of resources, these schools continue to generate the vicious cycle of poverty, low expectations, low attainment and poverty, aided by the misplaced gaze of government and school leadership and the absence of a focus on the learning needs of children and teachers.

In Chapter 9 I explore ways to break this vicious circle and consider transformational possibilities.



## **Chapter 9 Conclusion: Findings and a possible way forward:**

### **A new vision and transformational possibilities**

*"We cannot solve our problems with the same thinking we used when we created them."*

- Albert Einstein

#### **9.1 Introduction**

This chapter presents the findings of my research and suggests the possibility of creating a radical alternative to the current state of school science. As described in chapters 7 and 8, the evidence from the three case study schools paints a bleak picture of the role assessment plays in the state of science education in comprehensive schools. If what has been observed in this small sample of schools were to be repeated in even a minority of schools nationally then we should have serious concerns about the quality of the experience many of our pupils have in their science learning, and about our failure to nurture and develop their potential. Findings show science teaching for most pupils is weak, there is very little formative assessment/assessment for learning and a huge emphasis on the outcomes of summative assessment leading, in the worst case scenario, to a situation in which pupils are given qualifications which have little value except in relation to the school's position in the league tables, and reinforcing low expectations and the cycle of poverty/deprivation.

Given the above, I argue for a new vision, a transformed gaze, and a new approach to our pupils in which Burt's 'treble-track' will be fully rejected as we contemplate a future where all pupils will receive at least the minimum entitlement curriculum and will have equitable access to all the resources of the school. Teachers would also enjoy professional freedom on the basis that their professionalism is underpinned by professional development that meets their learning needs and by commitment to all their pupils.

Thus, in section 9.2 I outline the findings of the study. In 9.3 I consider a possible way forward from the sorry state that science education is arguably in. In section 9.4 I outline a new vision. I re-iterate the features of Enabling Assessment and the EA cycle and discuss how Bhaskar's five stages of creativity can provide a strong theory of learning to support the new vision. In 9.5 I posit a 'learner-centred model' as opposed to the 'factory-model' and section 9.6 proposes a new vision for teachers, a new concept of professionalism. Together these would be capable of facilitating the growth of a truly comprehensive system of schooling, by negating Sir Cyril Burt's 'treble-track' provision which underpinned the 'tripartite' model and which is still being used in the case study schools. Section 9.7 highlights how the system is still failing half our children, and section 9.8 suggests what is needed for us to move towards a genuinely comprehensive school system. 9.9 is a summary of the chapter and in section 9.10 I suggest some ideas for further research.

## 9.2 The findings:

Teachers are familiar with the term Assessment for Learning (AfL) in the case study schools but unaware of Formative Assessment (FA). AfL is largely interpreted to mean self- and peer-assessments and class discussion. Discussions, however, remain at a question-answer level and seldom progress into dialogue.

Marking is routine with ticks and smiley faces and appears to be simply paying lip-service to the whole notion of written feedback. Errors in science are often not corrected and when corrected the corrections are frequently wrong. Comments are trivial, often misleading and fail to show or indicate how pupils could improve scientific knowledge, understanding and skills. There is a complete absence of criteria for marking. The books of the top bands showed some evidence of marking, the books of the bottom bands, however, showed little or no marking. There is little or no feedback on what has been learnt and what is next to be learnt, or on how to improve. ‘Gaps’ in learning are not identified. Pupils are told the National Curriculum levels they are at, but there is little or no conversation with pupils about the scientific knowledge, understanding and skills they have gained or need to gain next or where they are making errors.

When asked, ‘What more can your teacher do to help you learn better’?

A pupil in set 4 (of 5) of the Fast Track at CHS said:

“Tell us whether our answers are correct”

Such words were spoken also by pupils at DCS and WCS.

There is evidence of 'objectives-led' lessons which the National Strategies had emphasised but the 'objectives' evidenced in pupils' books had little connection to building on pupils' previous learning or the need to address errors and misconceptions noticed during marking and/or class discussion. This was precisely because there was little marking worthy of the name and there was little attempt to observe and capture misconceptions or 'gaps' in pupils' learning either through dialogue, observation or through marking.

Summative assessment is dominant in the case study schools but the outcomes/results are unreliable. There is clear evidence that results of these assessments are 'marked up' to meet the targets and/or pupils are taught to the test as the teachers know what is in the end of unit tests or in the end of year exams. All three case study schools reported the Key Stage 3 (age 14) assessment results to be around 15 – 20 % higher than those that pupils actually achieved.

Assessment arrangements at the school level are complex and are geared to the collection of pupil performance data every term. This takes up a considerable amount of teachers' time as the marking of tests and exams and recording of the results on the schools' data-base are mandatory duties.

The overall quality of teaching in science was barely satisfactory and was generally inadequate to poor in the lower 'ability' sets.

In many cases science teachers' subject knowledge and their knowledge of the National Curriculum and its progression (pedagogical knowledge) are weak. There is an acute shortage of physics teachers and serious shortage of chemistry teachers. The subject specialism of science staff is in the ratio of 5:2:1:1 of biologists: chemists: physicists: others.

The coverage of at least two Attainment Targets of the National Curriculum in science, Sc1 (Scientific enquiry) and Sc4 (Physical sciences) is poor. Many pupils have a limited access to the science National Curriculum. The bottom bands mainly do biology topics and even those often poorly.

There is a lack of a CPD culture amongst science teachers amounting almost to a denial of the need for CPD, and a 'quick fix' approach to CPD at the school management level in response to Ofsted reports.

The 'tripartite' model of schooling underpinned by the ideology of innate ability is still dominant in the case study schools, and is facilitated through the mechanism of triage. Coupled with the standards agenda, this has resulted in an intensification of the factory-model and has led to increasingly coercive monitoring procedures, low morale of teachers and double triage.

The government's monitoring and accountability programmes are strong and are clear examples of **misplaced gaze**. Teachers' agency is weak – science teachers' continued adherence to the notion of *ability* and thereby acceptance of the 'factory-

model' of schooling has weakened their professionalism, and collectively, they have failed to address the vicious cycle of '*poverty – low expectations – low attainment – poverty*' outlined in chapter 8. This weakness allows repressive national and school structures to dictate and dominate every aspect of school life to the detriment of both pupils and teachers.

### **9.3 A possible way forward**

Bearing in mind Albert Einstein's widely quoted advice which is stated at the beginning of this chapter, it is important to pause and reflect on his words. In order to find solutions to the crisis in science education as depicted in this thesis, our thinking must change. We must think afresh. We must develop a new vision, a new gaze.

In this thesis I argue that the tripartite system of schooling which was based on an out-of-date and damaging ideology still provides the basis for educating children in some or perhaps even many of our comprehensive schools. In arguing for a change, a transformation, it is crucial that we move away from this debilitating ideology and move towards the notion of 'unfolding the enfolded', a Platonic idea which Roy Bhaskar has rearticulated.

#### **9.3.1 'Unfolding the Enfolded'**

Unfolding the enfolded should be the main job of teachers. This should also provide the basis for their professionalism. The seeds of what we learn and the ability to

learn are already inside us as children, the job of the adults in our lives is to help bring out what is already inside us and nurture its growth.

### **9.3.2 Towards high expectations to eradicate the ‘vicious’ cycle**

‘Unfolding the enfolded’ will require the end of the misplaced gaze and its replacement by a new vision and approach. There will be a need to look at every child as a learner where the challenge for the teacher is to ‘unfold’ their learning. The challenge for the rest of us - school managers, policy makers, parents and academic writers and researchers - would be to support teachers in their endeavour to bring about learning through the five stages of creativity and use of the *EA cycle*. The collective aim would be the eradication of the vicious cycle.

## **9.4 A new vision - Enabling Assessment and the EA cycle**

### **9.4.1 EA and the EA cycle**

Here, I re-iterate the formulation of EA and of the EA cycle as well as describe the four stages of creativity. As described earlier in the thesis there are three key components of Enabling Assessment. These are: Formative Assessment (FA), Diagnostic Assessment (DA) and Scaffolding (SC). As stated in 6.2.1 in this definition FA subsumes AfL and contains an element of SA.

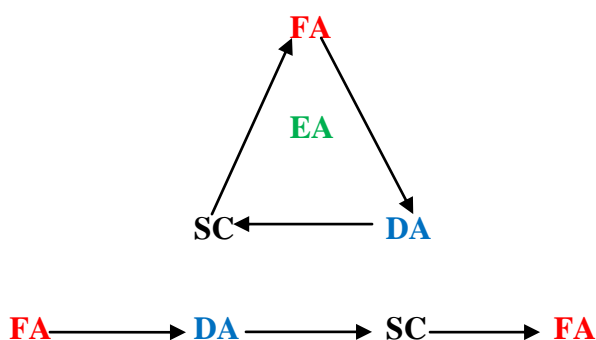
‘Enabling Assessment’ (EA) combines FA and DA with the notion of ‘*scaffolding*’ put forward by Bruner and colleagues, drawing on Vygotsky’s ideas about

supporting learners, and can be further strengthen by being informed by Bhaskar’s model of ‘Unfolding the Enfolded’.

**Enabling Assessment = Formative Assessment + Diagnostic Assessment + Scaffolding**

$$\text{EA} = \text{FA} + \text{DA} + \text{SC}$$

**Fig: 9.4.1: The EA cycle:**



The EA model of assessment is underpinned by the notion of creativity and learning put forward by Bhaskar (Bhaskar, 2012). Creativity is the making of something new. As he puts it:

“Creativity is the production of something new; but since whatever is, is already implicit or potential in what was there before, ... what is created is always a new synthesis of the new and the old, of the manifest and unmanifest, of the ephemeral and the eternal.” (Bhaskar, 2012, p. 105)

Also on pages 105/106,

“But what is created also transcends, surpasses, negates (i.e. absents and transforms) what is already there ... what is negated is not what was implicit, involved, enfolded or ingredient as



a potential in what pre-existed it. It emerges, consisting in the suprainposition of a transcendent cause upon an immanent ground, taking shape, form.”

In the new vision, the new approach, I envisage the teacher as the ‘creator’ of new learning and as such they would work through the following five circles of creation:

#### **9.4.2 The five circles of creation (Bhaskar, 2012, p. 112)**

- **The initial impulse –**

traditionally called the ‘word’ or ‘calling’, or the lightening flash or emancipation, the creator’s will to impose on the world.

- **Creation proper itself –**

the articulation and emergence, de novo, of something which has never existed before.

- **The circle of formation –**

this is the shaping or binding, the activity or labour of love, desire or worth in which the imaginative creative impulse of thought is now shaped into an object about to be released in the world. This is the circle of formation.

- **The circle of making –**

its physical objectification, the manifestation of the initial impulse or intentionality of the creator.

- **The circle of reflection, or reflexivity –**

which corresponds to a moment of self consciousness or return (namely of the result or consequences of his objectification) to the creator/maker. (Bhaskar, 2012, p. 112)

### **9.4.3 The five moments of the dialectic of creativity**

Put differently, the stages in the cycle in the previous section correspond to five moments of creativity or learning which are as follows:

#### **The first moment**

This corresponds to the level of Platonic anamnesis. But this knowledge is unmanifest, enfolded, so necessitates a process of ‘education’, including external stimulus for the bringing out or unfolding of what is already implicit or known, in an unmanifest, enfolded form in the learner.

Here, the teachers’ role would be to provide the ‘external stimulus through good, teaching.

#### **The second moment**

This is actually the process of learning, that is internalising it, in which, even though it may be a well-attested piece of knowledge, it is for the learner something new, a creation of the past which the learner must re-create in the present not only to make it his own, but to see what it is that is the ‘knowledge’. At this point the knowledge is ‘heteronomous’, something externally given, albeit internally reflected, which comes to the student’s conscious mind as discovery, the acquisition of something new.

This is the stage where teachers' expert knowledge facilitates the process of 'internalising' through diagnostic assessment and sensitive scaffolding.

### **The third moment**

This moment in the process of creation is the assimilation of the new knowledge so that the learners can apply it on their own account. It is the moment in which they bind it back into their own being, shaping it, learning its implications and application, what to do with it and how to use it, until the point when it is thoroughly assimilated and becomes part of their innermost being.

This is where the role of feedback and scaffolding becomes crucial. Teachers' expert knowledge can come into play here in scaffolding the 'assimilation of new knowledge'.

### **The fourth moment**

The learners have now made their discovery a part of their own being, and can apply it in all sorts of novel ways, and in such a way that the result now stands apart from this making.

Here, the teachers' role as an expert observer and empathetic evaluator and a guide becomes crucial. As the learners begin to apply their knowledge in various ways teachers would be able to provide guidance and encourage pupils in the process of expanding their knowledge and understanding.

### **The fifth moment**

The objective result can be more or less adequate to the agent's intentionality and returns to him as the fulfilment of his will, or the consequences of his action in the world.

This offers teachers opportunities to rejoice in the success of their pupils and also to address what has not yet been fulfilled.

## **9.5 Science in society: a '*learner-centred model*' to replace the 'factory-model'**

This new vision would require a new gaze, a new way of viewing the learner, full of hope and high expectations. It would mean a complete rejection of the factory-model and its replacement by a '*learner-centred model*' of education. The aim of schooling, for example in science education, would then be to equip the learner with the scientific knowledge, understanding and skills to open the door to the wonderful and exciting world of science, enabling them to develop these creatively to their full potential and then use them with confidence in whatever field of learning or work they choose to pursue.

### **9.5.1 The '*learner-centred model*'**

In this model the focus of the school will be exclusively on the learners, their learning history and their learning needs. The gaze will move away from the level they have obtained to the learning they have accomplished and the 'gaps' remaining. This is so that the best possible provision can be made. The focus will be on the

quality of care and teaching they receive. The searchlight will move to the inadequacies in provision. Teachers will be valued and supported in improving subject knowledge, pedagogical subject knowledge and their skills in scaffolding pupils' learning. The ethos of the school will be welcoming to learners and their parents. In this model, every parent would matter, every child would matter and indeed every teacher would matter.

There will have to be full access to the entitlement curriculum with good diagnostic assessment and good scaffolding. Teachers will need to skill themselves up for this purpose. And in such a model there would be zero tolerance of the 'deficit model' of the child or the parent as an explanation of failure – the misplaced gaze would be replaced by a searching scrutiny of provision. Neither 'triage' of pupils nor of teachers would apply.

## **9.6 A new vision for teachers**

This will require a new gaze upon teachers. Just as every child matters and every parent matters, in the new vision every teacher will also matter. Double triage and hiring and firing of teachers are unacceptable ways of treating them and there is no convincing evidence that this approach produces good quality learning. Teachers need to be viewed with respect and high regard and should be treated as such, with opportunities for good professional development, a reasonable workload taking into account time for preparation, marking, feedback and scaffolding and with a reasonable work environment free of excessive pressure and fear of inspection. In return, teachers will need to adopt a new professionalism which will be derived from

their commitment to their pupils and a strong commitment to learning, their pupils and their own.

### **9.6.1 A new professionalism**

As mentioned in the previous section, the new professionalism would combine (a) a commitment to ‘unfolding the enfolded’ (b) a commitment to ‘scaffolding’ learners in their pursuit of learning and (c) a commitment to professional learning. Teachers will derive their professionalism from their concern for their pupils – *all their pupils*.

## **9.7 The crisis in education: risking ‘Half Our Future’**

“Our pupils constitute, approximately, half the pupils of our secondary schools; they will eventually become half the citizens of this country, half the workers, half the mothers and fathers and half the consumers. Disraeli once said that on the education of the people of this country its future depended and it is in this sense that we have entitled our report

“Half Our Future” (Newsom, 1963, p. xiii)

Under that heading, ‘Education for All’ the Newsom report asked, ‘who are the boys and girls with whom this report is concerned?’ and answered,

“... ‘our children’ as we came to call them, cannot be defined as any exact percentage of population; and full description of what we have learnt about them from our survey must come later. But if we ask where they are to be found at present time, they are the boys and girls who form the majority of pupils in the secondary modern schools, or who are in the middle and lower forms of comprehensive schools” (Newsom, 1963, p.3).

*Today, in 2014, half a century after the above words were written ‘our children’ still come from ‘the middle and lower forms of comprehensive schools’.*

Then and now these are the children who are being failed by the system education.

### **9.7.1 A continuing danger to ‘half our future’**

The above quotations from the Newsom report painfully illustrate the state of schooling in England today as exemplified in this thesis by the state of science education. ‘Half our future’ is still in a perilous state. It seems clear that the factory-model of schooling based on the ideology of fixed innate ability and the ‘tripartite’ system has failed our children and has continued to underpin the vicious cycle generation after generation. The 11+ is less visible but its principles continue to occupy a prominent place in the minds of teachers, school managers and policy makers. Indeed, the government’s principal monitoring agency actively promotes the more visible *ability* grouping. For example, notice the emphasis in the following Ofsted report:

“The school has worked hard over the last year to improve the performance of more able students by grouping and teaching students in a way that is highly tailored to their needs.”  
(CHS Ofsted report, May 2013, p.4).

Here grouping means the top band (or Fast Track) pupils who get the best teachers and the best resources whilst the rest of the pupils are subjected to ‘triage’ and receive inferior resources, non-specialist teachers and sometimes a string of supply

teachers. To add insult to injury these pupils in the lower sets are given unearned ‘certificates’ in the form of BTEC GCSE passes which mean next to nothing in terms of their knowledge, understanding and skills.

In a much wider context, notice the threat Michael Wilshaw, the HMCI, has issued in the Evening Standard (Davis, 2012):

“Sir Michael Wilshire also said high-flying students are being held back by schools that put them into mixed ability classes that fail to “stretch” them”.

He added,

“Schools that do not “stream” will also be docked marks if teachers fail to prove that they can stretch all pupils.”

The message is clear: if you ‘stream’ it is assumed that the high-flying pupils are being stretched but if you don’t stream the assumption is that these pupils are being held back.

## **9.8 Every child matters, every teacher matters**

The Labour government (2002 – 2010) made huge strides in attempting to make all children visible and accounted for in the school system by policy initiatives such as ‘Every Child Matters’ (DES, 2003) - from which emanated the ‘Healthy Schools’ initiative, the provision of breakfast clubs in schools, Children’s Centres and the Sure Start programme. Indeed the last government set out to ‘transform’ education for all children by making an unprecedented amount of resources available for education. The National Curriculum was clearly presented as an ‘entitlement curriculum’, with very high ambition to provide every child with access to the main areas of knowledge in science.



However, its ambitions achieved only partial success due at least to two main factors. The first factor was the Government's continued adherence to an out-of-date ideology of educational provision on the basis of 'innate' ability'. This, as we have seen, removes incentives from schools and teachers to endeavour to *unfold the enfolded* in pupils and to genuinely focus effort and resources on meeting their learning needs. It furthermore removes the onus from the teacher to engage in activities and initiatives which would assist them in *unfolding the enfolded* for every pupil, especially in the secondary sector where the notion of innate ability remains dominant.

Moreover, the government's monitoring structures are punitive and ineffective and are a consequence of misplaced gaze. An alternative would be to develop a partnership with teachers in a project (praxis) of change from an outmoded ideology to one which enables teachers to decisively and purposefully grasp the nettle of 'untapped potential' in all our children. There is a need to invest in teachers and in teaching – teaching all our children.

### **9.8.1 A fight back for a truly comprehensive school system**

The new vision, the new gaze and the new approach (a learner-centred model) I have described in this chapter offer a unique opportunity to re-engage in campaign for a truly Comprehensive Schools and within that framework, a good science education for **all**.

In the words of Bloom and colleagues (Bloom, 1976, p. 17):

“... human nature is not the barrier to educational and cultural development that philosophers, politicians, social scientists and educators have frequently alleged. The characteristics of the students and of the instruction ... are alterable, and if this is so, changes in the school environment can relatively quickly (in a single decade) make great changes in the learning of students.”

A creative future in education and particularly in science education is not only possible but is practically achievable with a collective effort and the new vision and a new gaze.

## **9.9 Summary of Chapter 9: from crisis to a creative future?**

This chapter has attempted to summarise the findings of my study drawing attention to the root causes of the crisis in science education it has uncovered. The main focus of the chapter has been on the ‘way forward’ which emphasises the need for a transforming gaze, a new vision and a new approach. These are to be informed by the dialectic of creativity as expounded by Bhaskar and the Platonic notion of anamnesis as articulated by him.

The chapter re-iterates the EA cycle and proposes a ‘learner-centred model’ to replace the factory-model of schooling and discusses how the factory-model and the underpinning ideology of *innate ability*, can be overturned by the Bhaskarian notion of the practically infinite potential of every child and learning as the process of ‘unfolding the enfolded’. In the final sections, the chapter contemplates a new approach for education which will provide an enabling structure in which both

teachers and pupils can flourish. The chapter concludes with a hopeful vision of science education and school education in general by proclaiming that in the new vision of education every child will truly matter, every parent will matter and every teacher will matter.

The new theory of Enabling Assessment as proposed includes the merits of both formative assessment and assessment for learning and (also finds a place for summative assessment), and adds the essential elements of diagnostic assessment, as involving the identification of what is absent from learning, and of the teacher scaffolding the pupil. The realisation of the new vision will of course depend on the transformation of practices and of social structures, practices and structures (described section 5.4.3) which have hitherto inhibited the implementation of formative assessment/assessment for learning. Even when the new vision has persuaded teachers and others concerned with education that an alternative way is possible, this will have to be worked out in concrete specificity and in detail. Thus, for example, teachers will need work out how they are going to improve their CPD, etc. This can be considered, in the critical realist way, as a ‘concrete utopian exercise’ (Bhaskar, 1993, 2008, p. 395).

### **9.10 Ideas for further research**

In this section I include a few other ideas for further research. Although my research has produced strong evidence for my findings, it relates to practices at a small sample of schools. However, it does draw from national structures, ideologies and

frameworks. It also draws on national statistics on the plight of children who truant and the national GCSE results which show that only 25% of children from the lower socio-economic groups and the traditional working class backgrounds manage to achieve 5 good GCSEs. Roughly 15 – 20 % of pupils who are sent to the local for Life-Skills training become invisible. The schools do not monitor their progress. They do not appear on the schools' table of GCSE results. For example, the Y11 cohort size at CHS in 2010 was 320. However, GCSE tables show results for only 264 pupils. What happened to the missing 56 pupils? They constitute around 17% of the cohort. Is their progress tracked?

A suitable research question could be, 'How is the progress of (15 to 20%) of the cohort who attend the College monitored and to what extent the College experience contributes to their success and to their life-chances?'

Other areas of investigation could be focused on science teachers regarding their view of the CPD they would value coupled with the question, 'To what extent the weaknesses in knowledge, understanding and skills of the pupils reflect those of their science teachers?'.

My experiential evidence and a small amount of empirical evidence suggest that the areas of weaknesses the pupils have in science are reflections of the subject knowledge weaknesses of the teachers. Most primary school teachers are non-scientists. Most secondary school teachers are biologists. The areas of specific weaknesses in pupils' knowledge, understanding and skills generally fall within the National Curriculum Attainment Targets of Scientific enquiries (Sc1) and Physical processes or Physics (Sc4) and to a smaller extent Materials and their properties,

chemistry (Sc3). Most science teachers in the state schools are biologists or environmental scientists. It is likely that their subject knowledge in the above areas will be weak.

Pupils' weaknesses are pronounced in topics such as the kinetic theory, chemical changes, many physics topics and in Scientific enquiry (Sc1) in presenting evidence (tables, charts, and graphs), considering evidence (diagrams, lines of best fit, identifying and describing patterns and relationships in data). Most primary school teachers are likely also to have such weaknesses. Most secondary school teachers too, seeing that a clear majority of them do not have physics and chemistry as parts of their subject specialism, are likely also to have these weaknesses.

Thus in the words of Bloom:

"... variation in learning and the level of learning of students are determined by the students' learning history and the quality of instruction they receive." (Bloom, 1976, p. 16)

Thus pupils' learning history will be inadequate throughout the period of compulsory education.

It is a case of profound misplaced gaze that these obvious issues are not addressed. These are known absences. King's College's research drew attention to these areas of weakness in secondary school science teachers' subject knowledge in 2000. And still these have not been addressed.

I have not been able to pursue this research further but feel that it is an extremely important area for investigation.

For researching in this area a suitable RQ could be:

Do the pupils' lack of knowledge, understanding and skills in science reflect those of their teachers?'

Other areas for future research could be the absence of practical work in science. Schools appear to focus on Controlled Assessment (CA) at the expense of normal classroom practical activities. Pupils are often given the answers to the CA. Often teachers set up the experiments, do the measurements and give the pupils all the answers. Research into this field likely to be very fruitful for the future of science education in England

School self-evaluation is another critical area for investigation. Are schools being honest in their self-evaluation? Can they be honest in an environment of severe scrutiny?

Finally, research into the stress level of teachers is crucial for the future recruitment and retention of teachers, for the future of the profession and most importantly, teachers' health.

## Bibliography

- Acland, A. (1911), Board of Education Consultative Committee Report on Examinations in Secondary Schools. London: HMSO.
- Allen, M. and Ainley, P. (2007), *Education make you fick, innit?* London: The Tufnell Press.
- ARG. (1999), *Assessment for Learning: Beyond the Black Box*. Cambridge: University of Cambridge School of Education
- ARG. (2002), *Assessment for Learning: Ten Principles*. [Online]. Available at: [www.assessment-reform-group.org.uk](http://www.assessment-reform-group.org.uk).
- ARG. (2009), *Assessment in schools: Fit for purpose?* London: Institute of Education, University of London.
- Association, P. (2012, 14 June 2012), 'Truants miss 3.7m school days in one term, official figures show'. *The Guardian*.
- Atkinson, P., Coffey, A. and Delmont, S. (2003), *Key themes in qualitative research: Continuities and changes*. New York: Rowan and Littlefield.
- Ball, S. (2008), *The Education Debate*. Bristol: The Policy Press.
- Beloe, R. (1960), *The Beloe Report - Secondary School Examinations other than the GCE* (211). London: Ministry of Education
- BERA. (2004), *Revised Guidelines for Educational Research*.
- BERA. (2011), *Ethical Guidelines for Educational Research*.
- Bhaskar, R. (1975/2008), *A Realist Theory Of Science*. London: Verso.
- Bhaskar, R. (1979/98), *The Possibility of Naturalism*. (3rd ed.). London: Routledge.
- Bhaskar, R. (1993, 2008), *Dialectic: The Pulse of Freedom*. London: Routledge.
- Bhaskar, R. (2008a), *Dialectic: The Pulse of Freedom*. London: Routledge.
- Bhaskar, R. (2008b), *A Realist Theory Of Science*. London: Verso.
- Bhaskar, R. (2011a), *From Science to Emancipation*. London: Routledge.
- Bhaskar, R. (2011b), *Philosophy and the idea of Freedom*. Oxford: Routledge.
- Bhaskar, R. (2012), *The Philosophy of Metareality Creativity love and freedom*. Abingdon, Oxfordshire: Routledge.
- Bhaskar, R. and Danermark, B. (2006), 'Metatheory, interdisciplinarity and disability research: a critical realist perspective'. *Scandinavian Journal of Disability Research*, 8 (4), 278-97.

- Bhaskar, R., Frank, C., Hoyer, C., Naess, P. and Parker, J. (eds) (2010), *Interdisciplinarity and climate change*. Oxford: Routledge.
- Biggs, J. (1998), 'Assessment and Classroom Learning': the role of summative assessment '. *Assessment in education*, 5 (1), 103 - 110.
- Black, H. (1986), 'Assessment for learning'. In D. J. Nuttall (ed.), *Assessing Educational Achievement* (pp. 7-18). London: Falmer.
- Black, H. D. and Broadfoot, P. M. (1982), *Keeping Track of Teaching*. London: Routledge and Kegan Paul.
- Black, H. D. and Dockrell, W. B. (1984), *Criterion-Referenced Assessment in the Classroom*. Edinburgh: SCRE.
- Black, P. (2003), The Nature and Value of Formative Assessment for Learning, *AERA*. Chicago.
- Black, P. (2013), 'Formative and Summative Aspects of Assessment : Theoretical and Research Foundations in the Context of Pedagogy'. In J. H. McMillan (ed.), *SAGE Handbook of Research on Classroom Assessment* (pp. 167-179). Thousand Oaks: SAGE Publications, Inc.
- Black, P., Harrison, C., Hodgen, J., Marshall, B. a. and Serret, N. (2010), 'Validity in teachers' summative assessment'. *Assessment in Education Principle, Policy & Practice*, 17 (2), 215-232.
- Black, P., Harrison, C., Lee, C., Marshall, B. and Wiliam, D. (2002), *Working inside the blackbox: Assessment for learning in the classroom*. London: King's College.
- Black, P., Harrison, C., Lee, C., Marshall, B. and Wiliam, W. (2003), *Assessment for learning: putting it into practice*. London: Open University Press.
- Black, P. and Wiliam, D. (1998a), 'Assessment and Classroom Learning'. *Assessment in Education*, 5 (1), 7 - 76.
- Black, P. and Wiliam, W. (1998), *Inside the blackbox*. London: King's College, London.
- Bloom, B. J. (1976), *Human Characteristics and School Learning*. New York: McGraw-Hill.
- Bloom, B. S., Hastings, J. T. and Madaus, G. F. (1971), *Handbook on formative and summative evaluation of student learning*. New York: McGraw-Hill Book Company.



- Broadfoot, P. (1986), 'Alternatives to Public Examinations'. In D. L. Nuttall (ed.), *Assessing Educational Achievement* (pp. 54 - 77). London: The Falmer Press.
- Broadfoot, P. (1996), *Education, Assessment and Society: a Sociological Analysis*. Buckingham: Open University Press.
- Brown, G. (2009), 'The Ontological Turn In Education: The Place of the Learning Environment'. *Journal of Critical Realism*, 8 (1).
- Bryman. (1994), *Analysing Qualitative Data*. London: Routledge.
- Burgess, R. G. (ed.) (1984), *The Research Process in Educational Settings: Ten Case Studies*. London: The Falmer Press.
- Butler, R. (1988), 'Enhancing and undermining intrinsic motivation; the effects of task-involving and ego-involving evaluation on interest and performance'. *British Journal of Educational Psychology*, 58, 1-14.
- Chadwick, E. (1864), 'Statistics of educational results'. *The Museum*, 3, 479-84.
- Charmaz, K. (2006), *Constructing Grounded Theory*. London: SAGE Publications.
- Clarke, S. (1998), *Targetting assessment in the Primary School*. London: Hodder and Stoughton.
- Clarke, S. (2001), *Unlocking Formative Assessment*. London: Hodder and Stoughton.
- Clarke, S. (2005), *Formative Assessment in the Secondary Classroom*. London: Hodder Murray.
- Collins. (1998), *Collins English Dictionary*: Harper Collins Publishers.
- Corrigan, D., Dillon, J. and Gunstone, R. (eds) (2011), *The Professional Knowledge Base of Science Teaching*. London: Springer.
- Crooks, T. j. (1988), 'The impact of classroom evaluation practices on students'. *Review of Educational Research*, 58, 438-481.
- Daugherty, R. (1995), *National Curriculum Assessment: a Review of Policy 1987 - 1994*. London: Falmer.
- David, P. (1981, 27 November), 'Multiple-choice under fire'. *The Times Educational Supplement*.
- Davis, A. (2012, 20.9.2012), 'Taking Exams Early Hits Grades'. *The Evening Standard*, pp. 8.
- Daws, N. and Singh, B. (1996), 'Why is formative assessment so impoverished?' *School Science Review*, 76 (23), 16 -32.
- DES. (1965), Circular 10/65: HMSO, London.

- DES. (1970), Circular 10/70: HMSO.
- DES. (1988a), Better Schools - a Government White Paper. In Education (ed): HMSO.
- DES. (1988b), HMI Report. In E. a. Science (ed): HMSO.
- DES. (1989), *Science in the National Curriculum*. London: Department of Education and Science.
- DES. (2003), Every Child Matters. In Education (ed): HMSO.
- DES and WO. (1985), *General Certificate of Secondary Education: A General Introduction* London: HMSO
- DES/WO. (1988), Science for ages 5 to 16. In D. o. E. a. S. a. t. W. Office (ed): HMSO.
- DES/WO. (1991), Science in the National Curriculum. In D. o. E. a. S. a. W. Office (ed): HMSO.
- Dewey, J. (1916), *Democracy and Education: an introduction to the philosophy of education*: Macmillan.
- DfEE. (2001), *Key Stage 3 National Pilot (science)*. London: Standards and Effectiveness Unit.
- DfEE/QCA. (1999a), *National Curriculum for England - Science*. London: jointly by DfEE & QCA.
- DFEE/QCA. (1999b), *The National Curriculum: Handbook for secondary teachers in England key stage 3 and 4* (First ed., pp. 200): HMSO.
- DfES. (2001), *The Autumn Package* (pp. 32): DfES Publications.
- DfES. (2004), *Assessment for Learning, whole school training materials: Guidance for senior leaders, Key Stage 3 National Strategy Assessment for Learning whole school training materials*. London: HMSO.
- Dhillon, R. (2000), *Professional Development Needs of Science Teachers*. London: King's College.
- Dunne, J. (2013, 28 February 2013), 'Baby death coroner says doctors are ticking boxes, not assessing patients'. *London Evening Standard*, pp. 5.
- Dweck, C. (1999), *Self-theories: Their role in motivation, personality, and development*. Philadelphia: Psychology Press.
- Dweck, C. (2002), 'The development of ability conceptions'. In A. Wingfield and J. Eccles (eds), *The development of achievement motivation*. New York: Academic Press.

- Dweck, C. (2007), *Mindset: The New Psychology of Success*. New York.
- Dweck, C. and Leggett, E. L. (1988), 'A social-cognitive approach to motivation and personality '. *Psychological Review*, 95, 256-273.
- Dweck, C., Trzesniewski, K. H. and Blackwell, L. S. (2007), 'Implicit Theories of Intelligence Predict Achievement Across an Adolescent Transition: A Longitudinal Study and an Intervention '. *Child Development*, 78 (1), 246 - 263.
- Earl, L. M. (2003), *Assessment As Learning*. California: CORWIN PRESS, INC.
- Earl, L. M. (2013), *Assessment as Learning: Using Classroom Assessment to Maximise Student Learning*. (2nd ed.). California: Corwin.
- Fairbrother, B., Black, P. J. and Gill, P. (eds) (1995), *Teachers Assessing Pupils: Lessons from Science Classrooms*. London: The Association for Science Education.
- Ferriman, B., Lock, R. and Soares, A. (1994), 'Influences of the National Curriculum on assessment in Key Stage 3'. *School Science Review*, 76 (274), 116-120.
- Freire, P. (1972), *Pedagogy Of The Oppressed* (M. B. Ramos, Trans.). (1st ed.). London: Seed and Ward Ltd.
- Gadsby, C. (2012), *Perfect Assessment for Learning*. Carmarthen, Wales: Independent Thinking Press.
- Galton, F. (1865), 'Hereditary talent and character'. *Macmillan's Magazine*, 12, 157-66; 318-27.
- Galton, F. (1869), *Hereditary Genius*. (reprint of 1892 edition, first published in 1869 ed.). London: Friedman.
- Gannon, P. (1998), *Key Stage Three Science: The Revision Guide*. Newcastle upon Tyne: Coordination Group Publications Ltd.
- Gardner, H. (1983), *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.
- Giddens, A. (1984/2003), *The Constitution of Society*. Cambridge: Polity Press.
- Gillard, D. (2011), *Education in England: a brief history*.
- Gillborn, D. and Youdell, D. (2000), *Rationing Education: policy, practice, reform and equity*. Buckingham. Philadelphia: Open University Press.
- Gillham, B. (2000), *Case Study Research Methods*. London: Continuum.
- Gipps, C. (1990), *Assessment: A Teachers' Guide to the Issues*. London: Hodder and Stoughton.

- Gipps, C. V. (1994), *Beyond Testing: Towards a Theory of Educational Assessment*. London: The Falmer Press.
- Hammersley, M. and Gomm, R. (2008), 'Assessing the Radical Critique of Interviews ', *Questioning Qualitative Inquiry* (pp. 89 - 100). London: SAGE.
- Harlen, W. (2004), A systematic review of the evidence of impact on students, teachers and the curriculum of the process of using assessment by teachers for summative purposes. In Research Evidence in Education Library. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.
- Harlen, W. (2006), 'The Role of Assessment in Developing Motivation for Learning'. In J. Gardner (ed.), *Assessment and Learning* (1 ed., pp. 61 - 80). London: SAGE Publications Ltd.
- Harrison, C. (2001), *KMOFAP Project*. Paper presented at the BERA, University of Leeds.
- Harrison, C. (2005), 'Teachers Developing Assessment for Learning'. *Teacher Development*, 9 (2), 255-263.
- Hart, S., Dixon, A., Drummond, M. J. and McIntyre, D. (2004), *Learning without Limits*. Maidenhead, Berkshire: Open University Press.
- Henry, J. (2002, 12 July 2002), 'Pupils lose a year in Tests'. *Times Educational Supplement*, pp. 4.
- Hitler, A. (1925), *Mine Kamph*. Berlin: Eher Verlag.
- James, M. (ed.) (2007), *Improving Learning: How to Learn* (First ed.). London: Routledge.
- James, M. and Pedder, D. (2006), 'Professional Learning as a Condition for Assessment for Learning'. In J. Gardner (ed.), *Assessment and Learning* (pp. 27-43). London: SAGE.
- Johnston, P., Guice, S., Baker, K., Malone, J. and Michelson, N. (1995), 'Assessment of teaching and learning in literature-based classrooms. ' *Teaching and Teacher Education*, 11 (4), 359-371.
- Jonassen, D. H. and Land, S. M. (eds) (2000), *Theoretical Foundations of Learning Environments* London and Mahwah, NJ: Lawrence Erlbaum.
- Jones, A. and Cowie, B. (2011), 'Moving Beyond Deconstruction: Teacher Knowledge-as-Action'. In J. D. Deborah Corrigan, Richard Gunstone (ed.),

- The Professional Knowledge Base of Science Teaching* (pp. 51-63). London: Springer.
- Jones, O. (2011), *CHAVS*. London: Verso.
- Joseph, S. K. (1984), 'Catastrophe or Watershed'. *Oxford Review of Education*, 10 (2), 137-146.
- Kamin, L. J. (1974), *The Science and Politics of IQ*: Eribaum.
- Kingdon, M. and Stobart, G. (1988), *Science GCSE - A Guide for Teachers*. Milton Keynes: The Open University Press.
- Kingdon, M. and Stobart, G. (1988 ), *GCSE EXAMINED*. London: The Falmer Press.
- Land, S. M. and Hannafin, M. J. (2000), 'Student-centred learning environment'. In D. H. J. a. S. M. Land. (ed.), *Theoretical Foundations of Learning Environments* (pp. 1 - 24). London: Lawrence Earlbaum.
- Lloyd-Jones, R., Bray, E., Jonson, G. and Currie, R. (eds) (1986), *Assessment: From Principle s to Action*. London: Macmillan.
- Lofland, J. and Lofland, L. H. (1995), *Analyzing social settings*. (3rd ed.). Belmont, CA: Wadsworth.
- Marsden, W. E. (1987), *Unequal educational provision in England and Wales: the nineteenth - century roots*. London: Woburn Press.
- Marshall, B. and Drummond, M. J. (2006), 'How teachers engage with assessment for learning: Lessons from the classroom'. *Research Papers in Education*, 21 (2), 133-149.
- Marshall, P. (ed.) (2013), *The Tail* (1 ed.). London: Profile Books.
- Millar, R. and Osborne, J. (1998), *Beyond 2000*. London: King's College, London.
- Moss, C. M. (2013), 'Research on Classroom Summative assessment'. In J. H. McMillan (ed.), *SAGE Handbook of Research on Classroom Assessment* (pp. 235-256). Thousand Oaks: SAGE Publications, Inc.
- NAHT. (2001), Circular No. BR 56/01, LA 42/01: NAHT.
- Newsom, J. (1963), *Half Our Future* (S.O. Code No. 27-361). London: Ministry Of Education.
- Norwood. (1943), *Curriculum and Examinations in Secondary Schools* (210). London: Ministry of education.
- Nuttall, D. L. (ed.) (1986), *Assessing Educational Achievement* (1st Edition ed.). London: The Falmer Press.

- Ofsted. (1996), Annual Report of Her Majesty's Chief Inspector of Schools: standards and quality in education 1994/95 In DES (ed): HMSO.
- Ofsted. (2004), *Key Stage 3 Strategy: evaluation of the third year* (HMI 2090). London: Office for Standards in Education.
- Ofsted. (2008a), *The Annual Report of Her Majesty's Chief Inspector of Education, Children's Services and Skills 2007/08*. London: Department for Education, Children's Services and Skills.
- Ofsted. (2008b), Assessment for Learning: the impact of National Strategy support. In DCSF (ed).
- Ofsted. (2009), Ofsted Annual Report. In DCSF (ed): Ofsted.
- Paton, G. (2012, 26 Jan 2012), 'GCSE league tables: warning over 'shocking waste of talent'. *Telegraph*.
- Perks, D. (2006), 'What is science education for?' In T. Gilland (ed.), *What is Science Education For?* (pp. 119). London: Academy of Ideas Ltd.
- Rabinow, P. (ed.) (1984), *The Foucault Reader*. London: Penguin.
- Ramprasad, A. (1983), 'On the definition of feedback'. *Behavioural Science*, 28, 4 - 13.
- Rieg, S. A. (2007), 'Classroom assessment strategies: What do students at-risk and teachers percieve as effective and useful?' *Journal of Instructional Psychology*, 34 (4), 214-225.
- Roberts, C. (2005, 9th December 2005), 'Setting'. *The Times Educational Supplement*.
- Robson, C. (2002), *Real World Research*. Oxford: Blackwell.
- RS. (2007), *A 'state of the nation' report: The UK's science and mathematics teaching workforce*. London: The Royal Society.
- Ruiz-Primo, M. A. and Li, M. (2013), 'Examining Formative Feedback in the Classroom context'. In J. H. McMillan (ed.), *The SAGE Handbook of Research on Classroom Assessment* (pp. Print pages 215 - 233). Thousand Oaks: SAGE Publication Inc.
- Rutter, M., Maughan, B., Mortimore, P., Ousten, J. and Smith, A. (1979), *Fifteen Thousand Hours: Secondary schools and their effects on children*. Cambridge, Massachusetts: Harvard University Press.
- Sadler, R. (1989), 'Formative assessment and the design of instructional systems'. *Instructional Science*, (18), 119 - 44.

- Saksena, S. (2000), *Manusmriti* (pp. 522). New Delhi: Manoj Publications.
- Satterley, D. (1991), *Assessment in Schools*. Oxford: Blackwell.
- Schön, D. (1991), *The Reflective Turn*. New York: Teachers College Press.
- Scott, D. (2011), PISA under Examination: Changing Knowledge, Changing Tests and Changing Schools., *Critical Realism Seminar*. Institute of Education, University of London.
- Scriven, M. (1967), 'The methodology of evaluation '. In R. W. Tyler (ed.), *Perspectives of Curriculum Evaluation* (pp. 39-83). Chicago: Rand McNally.
- SEC. (1986), 'Science GCSE - A Guide For Teachers'. Milton Keynes: The Open University Press.
- Simon, B. (1988), *Bending the Rules: The Baker 'Reform' of Education*. London: Lawrence & Wishart.
- Smith, C. and Elger, T. (2014), 'Critical Realism and Interviewing Subject'. In P. K. e. a. Edwards (ed.), *Studying Organizations Using Critical Realism - A Practical Guide* (pp. 109 - 131). Oxford: Oxford University Press.
- SNS. (2007), Getting Back OnTrack. In DfES (ed): Secondary National Strategy.
- Spendlove, D. (2009), *Putting Assessment for Learning Into Practice*. London: Continuum.
- Stewart, W. (2012, 13 July 2012), 'Think you've implemented Assessment for Learning? Think again - most schools are doing it wrong, says its creator'. *TES magazine*, 24.
- Stiggins, R. (2001), *Student-involved classroom assessment*. (3 ed.). Upper Saddle River, NJ: Prentice Hall.
- Stobart, G. (2008), *Testing Times: the uses and abuses of assessment*. London: Routledge.
- Stobart, G. (2014), *The Expert Learner: Challenging the myth of ability*. Maidenhead, Berkshire: The Open University Press.
- Strauss, A. L. and Corbin, J. M. (1990), *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. London: Sage.
- Swain, J. (1995), 'Survey on National Curriculum Statutory Assessment in Science'. In B. Fairbrother (ed.), *Teachers Assessing Pupils* (pp. 125 - 128). London: Association for Science Education.

- Torrance, H. and Pryor, J. (1998), *Investigating Formative Assessment: Teaching, Learning and Assessment in the Classroom*. Buckingham: Open University Press.
- Vertigan, S. (2007), *Ethics and the Business of Schooling*. University of London, London.
- Wenger, E. (1998), *Communities of Practice*. Cambridge: Cambridge University Press.
- Wheelahan, L. (2010), *Why knowledge matters in the curriculum: a social realist argument* London: Routledge.
- White, J. (2006), *Intelligence, Destiny and Education: The ideological roots of intelligence testing*. London: Routledge.
- WHO. (2001), World Health Organisation, International Classification of Functioning, Disability and Health (ICF): WHO.
- William, D. (2009), *Assessment for Learning: why, what and how?* London: Institute of Education.
- William, D. (2011), *Embedded formative assessment*. Bloomington: Solution Tree Press.
- William, S. (2012, 13 July 2012), 'Think you have implemented Assessment for Learning? Think again - most schools are doing it wrong, says its creator'. *Times Educational Supplement (TES)*, pp. 24.
- Willis, P. (1977), *Learning to labour*. Farnborough: Saxon House, Teakfield Limited.
- Wood, D. (1998), *How children think and learn* (2nd edition ed.). Oxford: Blackwell.
- Wood, D. J., Bruner, J. S. and Ross, G. (1976), 'The role of tutoring in problem solving'. *Journal of child psychology and Psychiatry* 17 (2), 89-100.
- Young, M. F. D. (2008), *Bringing knowledge back in: from social constructivism to social realism in sociology of education*. London: Routledge.
- Zweig, F. (1948), *Labour, Life and Poverty*. London: Gollancz.



## **Appendices**

### **A0 i: Email requesting a science teacher to participate in the research**

**Dear science teacher**

Thank you for agreeing to take part in my research programme. I trust you will agree that we have discussed the nature of the research and I have explained the activities which I will ask you to take part in. I have also discussed with you issues of ethics and confidentiality relating to my fieldwork in your department and in relation to working with you. I have set out the full details in my letter to the headteacher and head of the department, a copy of which is available if you wish to see it.

I value your participation in this research and would like to re-iterate that your involvement is entirely voluntary, enabling you to withdraw from it at any stage without giving reasons and without prejudice to my professional relationship with you, and in complete confidence.

Sincerely

Birendra Singh

April 2008

## **A0 ii: Letter to pupil participants**

**Birendra Singh (SIP, Barking & Dagenham LA)**

**Research towards PhD/MPhil Degree**

**Institute of Education**

**University of London**

**November  
2009**

**Pupil's name:**

**Dear .....**

Thank you for agreeing to take part in this research. It is part of my MPhil/PhD study at the Institute of Education, University of London. This will involve you in group discussions with me about your work in science for the next two/three terms. The information gathered will be used only for research purposes. There will be no trick questions, and your name will not appear in any communication or publication arising from this research.

I do hope that you will enjoy taking part in the research and contributing to the discussions. However, your participation in this research is entirely voluntary and you can withdraw from it at any stage without giving reasons and without upsetting either me or your teacher.

**I agree to take part in this research.**

Signed by pupil.....

Signed by class teacher .....

**Dear Parent/Carer**

I have obtained the permission of the Headteacher and the Head of Department, and the class teacher's agreement for the above educational research. Your permission is required for your son/daughter to participate in this research. I hope you will be able to give your consent.

**I/we consent to.....participating in the above research.**

Signed.....

**BS/Nov'09**

## **A0 iii: Letter to the headteacher**

### **Request to conduct research in the Science Department**

**April 2008**

Alongside my role as General Inspector (science) with Barking and Dagenham Local Authority I have undertaken research for a PhD with the Institute of Education, London. I have completed all the requisite courses in research methods and have conducted some research in your science department with your oral permission and the agreement of the Head of Science and a number of other teachers who have agreed to take part. This letter is to request written confirmation of the oral permission.

The working title for my research is 'The nature and content of assessment in Key Stage 3 Science'. My supervisor is Professor Michael Barnett.

I am asking you as the Headteacher of the school to agree to allow me to conduct research for the sole purpose of my PhD study. I will seek similar agreement with your Head of Science and any other teachers who participate.

#### ***The research will involve the following:***

- An approach to teachers to invite them to participate in the research
- Requesting access to policy documents on assessment
- Requesting access to examples of written feedback
- Requesting opportunities to observe episodes of question/answer sessions and dialogue
- Requesting information on test and examination procedures
- Requesting opportunities to attend departmental meetings relating to assessment issues
- With teachers' agreement talking to pupils

#### ***Potential benefits to the science department of participation in this research***

My work with your science department in my capacity as the science inspector has given me a good understanding of the work of the department. I have developed good professional relationships with the head of department and his/her colleagues in the department. My research is intended to provide a deeper understanding of one key aspect of the overall pedagogy of the department; at the end of the research process the findings might offer helpful indicators for improvement. The process of the research in itself is likely to provide participants with opportunities to think about and reflect on their current practice. The interviews and discussions may lead teachers to a better understanding of their own pedagogy and practice. Through the research process I am likely to gain deeper insight into practice which in turn will enable me to offer better and more bespoke advice and guidance to the department in my role as an adviser.

#### ***The nature and purpose of my research***

The research is about understanding the nature and the purposes of assessment as practised in the science department in Key Stage 3. There is a national framework for assessment within the requirements of the National Curriculum. There are national tests, the results of which are used for comparison through league tables and also for setting targets. Thus the assessment activities of the department take place in a broader national context. I am not starting with the premise that the department or individual teachers 'ought' to be engaged in a particular way of assessing pupils' work - I am, rather, interested in learning how they assess and why.

#### ***My research methodology***

The research will be a qualitative study of assessment practice in Years 7, 8 and 9. The procedures will consist of discussions and interviews with teachers, observations of teaching, examination of policy documents, the study of procedures and studying the nature of marking and feedback

Observations of teaching are intended to capture episodes of dialogue. They will not lead to judgements on the quality of teaching. The interviews and discussions with teachers are intended to elicit teachers own perceptions and their reasons for the assessment activities. Discussions with pupils will focus on their learning in science. The observation of episodes of dialogue and the interviews with teachers will be recorded with the prior consent of teachers. Copies of transcripts will be given to participants for their scrutiny and comments.

***Considerations in relation to ethics and confidentiality***

- I have already discussed the nature and the purpose of my fieldwork in the department with the head of science, the Key Stage 3 co-ordinator and the other teachers who have agreed to participate in the research I have explained my role as a researcher as opposed to my role as the science inspector. I have taken care to explain to the teachers that their participation in the research process is voluntary and they can withdraw from it at any stage of the process. The withdrawal will remain confidential and will in no way prejudice my professional relationship with them. I will also write to the members of the department to confirm the above points.
- If I observe any meetings where issues relating to assessment are being discussed I will seek the prior permission of those attending the meeting.

There is a potential conflict of interest between my role as a researcher and my role as the science inspector. I have been alert to this and it is important that I continue to be so throughout the period of my fieldwork in the department. As I have stated earlier none of my activities as a researcher will have any judgemental aspect. I have, as far as possible and with meticulous care, separated my two roles and will continue to do so. However, there might arise situations which would require careful consideration in distinguishing what is legitimate data and what is not.

My current thoughts in relation to anticipating possible conflicts relate to

- my 'Extended Visits' to the department which are monitoring visits and the outcomes of which are formally reported to the department and the head teacher
- any review of teaching or any other aspect of the work of the department I am asked to carry out
- any competency procedures which I might be asked to monitor
- any incidences of possible breaches of safety or child protection procedures I happen to observe.

In the case of breaches of safety rules I will need to discuss the issues with the teacher/s concerned. In so far as breaches of child protection procedures are concerned the confidentiality rule must be deemed not to apply and I will bring the matter to the attention of the appropriate member/s of the school.

Any information in relation to my research topic (assessment), gathered from my visits in my role as the science inspector will need to be considered on its merit so as not to contaminate the data and its use will be subject to permission from the teacher/s concerned.

Those who agree to participate in the research will have access to the data generated through interaction with them. The data will be confidential to me and my supervisor/s. No other person will have access to the data. The data obtained from interviews and observations or from scrutiny of

pupils' work will not identify the individuals who helped to generate them. Neither will the name of the school be identified in the thesis arising out of the research. The same will apply to any publications arising from the thesis.

I anticipate and hope that the research will be of some value to the department and perhaps also to the school.

Birendra Singh

April 2008.

## **A0 iv: Request to Head of Department**

**March 2008**

**Dear Head of Department**

Thank you for your agreement to my request to conduct research in your department. We have discussed the topic of my research and you know that alongside my role as General Inspector (science) with Barking and Dagenham Local Authority I have undertaken research for a PhD with the Institute of Education, London. The research work with your department is part of my PhD studies.

As you know I had your headteacher's permission to conduct research in the department subject to your agreement which you readily granted. I have also discussed the progress of the research with the headteacher. He has re-affirmed his permission, subject to your agreement, which again you have readily given..

This letter is a formal request to confirm this agreement in writing. As you know the research topic is Assessment at Key Stage 3 and we have discussed the nature of the fieldwork in your department. For the purposes of clarity and for my records I re-iterate them as follows:

The research will involve an approach to teachers and pupils to invite them to participate in the research and request for the following:

- access to policy documents on assessment
- access to examples of written work
- opportunities to observe episodes of question/answer sessions and dialogue
- information on test and examination procedures
- opportunities to attend departmental meetings relating to assessment issues

### **Potential benefits to the department of participating in this research**

My work with your department in my capacity as the science inspector has given me a good understanding of your policy and practice. I think that I have developed good professional relationships with you and your colleagues. My research is intended to provide a deeper understanding of one key aspect of the overall pedagogy of the department; at the end of the research process the findings might offer helpful indicators for improvement. The process of the research in itself is likely to provide participants with opportunities to think about and reflect on their current practice. The interviews and discussions may lead teachers to a better understanding of their own pedagogy and practice. Through the research process I am likely to gain deeper insight into practice

which in turn will enable me to offer better and more personalised advice and guidance to the department in my role as an adviser.

### ***My research methodology***

The research will be a qualitative study of assessment practice in Years 7, 8 and 9. The procedures will consist of discussions and interviews with teachers, observations of teaching, examination of policy documents, the study of procedures and studying the nature of marking and feedback

Observations of teaching are intended to capture episodes of dialogue. They will not lead to judgements on the quality of teaching. The interviews and discussions with teachers are intended to elicit teachers' own perceptions and reasons for the assessment activities. The observation of episodes of dialogue and the interviews with teachers will be recorded with the prior consent of teachers. Copies of transcripts will be given to participants for their scrutiny and comments.

### ***Considerations in relation to ethics and confidentiality***

- I have already discussed the nature and the purpose of my fieldwork in the department with you, your Key Stage 3 co-ordinator and the teachers who have agreed to participate in the research. I have explained to you my role as a researcher as opposed to my role as the science inspector. I have taken care to explain to the teachers that their participation in the research process is voluntary and they can withdraw from it at any stage of the process. The withdrawal will remain confidential and will in no way prejudice my professional relationship with them.
- If I observe any meetings where issues relating to assessment are being discussed I will endeavour to seek the prior permission of all those attending the meeting.

There is a potential conflict of interest between my role as a researcher and my role as the science inspector. I have been alert to this and it is important that I continue to be so throughout the period of my fieldwork in the department. As I have stated earlier none of my activities as a researcher will have any judgemental aspect. I have, as far as possible and with meticulous care, separated my two roles and will continue to do so. However, there might arise situations which would require careful consideration in distinguishing what is legitimate data and what is not.

Any information in relation to my research topic (assessment), gathered from my visits in my role as the science inspector will need to be considered on its merit so as not to contaminate the data and its use for research will be subject to permission from the teacher/s concerned.

Those who agree to participate in the research will have access to the data generated through interaction with them. The data will be confidential to me and my supervisor/s. No other person will have access to the data. The data obtained from interviews and observations or from scrutiny of pupils' work will not identify the individuals who helped to generate them. Neither will the name of the school be identified in the thesis arising out of the research. The same will apply to any publications arising from the thesis.

I anticipate and hope that the research will be of some value to the department and perhaps also to the school.

**Birendra Singh**

**March' 08**

## **A0 v: Initial interview questions**

### **Deputy Headteacher**

Your school was praised for your assessment procedures by the Ofsted Inspectors

1. What aspects of your assessment policy did they like most?
2. What's your view of the role of assessment in your school?
3. You have adopted AfL as a whole-school policy, what does this mean in practice?
4. How does this affect a teacher of Year 8?
5. You are in charge of the whole-school assessment system. What does it involve?
6. How does school assessment link to the National Curriculum?

### **Head of Science:**

1. The school has adopted AfL as a whole school policy - how does this apply to your department?
2. What is your assessment policy in Year 8?
3. Is the policy different for Year 9?
4. What is your view of the role or function of assessment in years 8 and 9?
5. How does assessment in the department link to the National Curriculum?

### **Class teacher:**

1. What assessments do you carry out for your Year 8 class?
2. What assessments do you carry out with your Year 9 class?
3. What is your view of the role and function of assessment in your classes?
4. The school has adopted AfL as a whole- school policy, how has this affected your assessment of year 8 pupils?
5. How much time do you spend on assessment?

## A1: Marking Pro-forma A

Feedback	Yes	No	Description
Criteria used			
Objectives used			
Grades:			
<ul style="list-style-type: none"> <li>Numerical grade</li> </ul>			
<ul style="list-style-type: none"> <li>letter grade</li> </ul>			
Comments:			
<ul style="list-style-type: none"> <li>encouraging</li> <li>on science</li> <li>on presentation</li> <li>reprimand</li> <li>reward</li> </ul>			
Corrections:			
<ul style="list-style-type: none"> <li>Science corrected</li> <li>Other corrections</li> </ul>			
Tasks set:			
<ul style="list-style-type: none"> <li>Response by P</li> <li>Response checked by T</li> </ul>			



## A1: Marking Pro-forma B

Grid for analysing marking

Work covered:

Book ID:

Teacher Code:

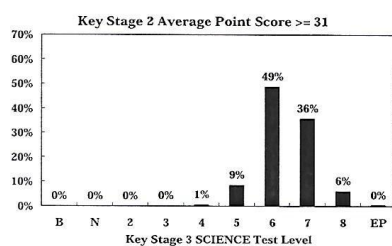
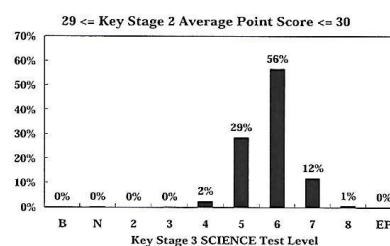
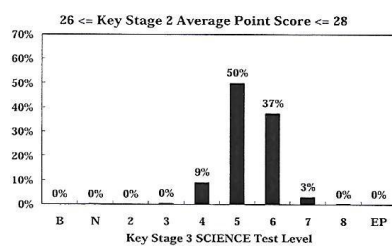
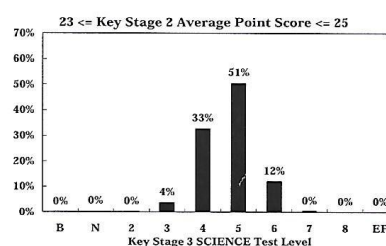
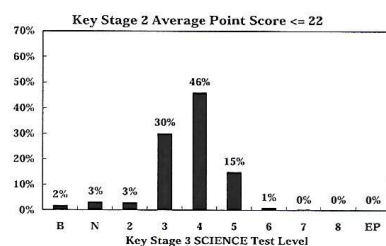
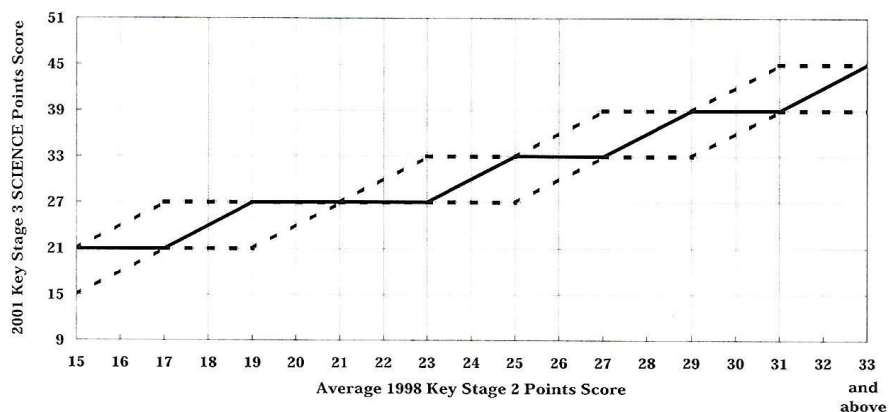
Themes	Yes or no	Details
Learning Objectives		
Learning outcomes		
Criteria		
Science corrected		
Literacy corrected		
Tasks given to complete		
Pupils' response to tasks		
Comments only		
Marks or grade only		
Comments and marks/grade		
Ticks only		
Ticks and comments/grade/marks		
Teachers' follow up to check completion of task/s		

## A2ai: 2001 KS3 National Value Added information – Science

Source: Autumn Package [www.standards.dfes.gov.uk](http://www.standards.dfes.gov.uk) Ref: DfES 0718/2001

### 2001 KEY STAGE 3 NATIONAL VALUE ADDED INFORMATION

Graph 2.3: 2001 Key Stage 3 Science Test Level Value Added Line and Progress Charts



## A2aii: 2001 KS3 National Benchmark information

Source: Autumn Package [www.standards.dfes.gov.uk](http://www.standards.dfes.gov.uk) Ref: DfES 0718/2001

### 2001 KEY STAGE 3 NATIONAL BENCHMARK INFORMATION

Table 3.4: Non-selective schools with pupils known to be eligible for FSM of:

***more than 9% and up to 13%***

#### Percentage of pupils achieving Level 5 and above

	95%	UQ	60%	Median	40%	LQ	5%
English	87	77	71	68	65	59	46
Mathematics	83	76	72	71	69	66	57
Science	83	76	74	71	69	65	55

#### Percentage of pupils achieving Level 6 and above

	95%	UQ	60%	Median	40%	LQ	5%
English	57	41	33	29	27	22	12
Mathematics	60	50	46	44	41	38	28
Science	51	41	37	34	32	28	17

#### Average KS3 points score achieved<sup>1</sup>

	95%	UQ	60%	Median	40%	LQ	5%
English	36.9	34.7	33.6	33.2	32.6	31.8	29.6
Mathematics	37.6	35.8	35.2	34.8	34.4	33.7	31.9
Science	35.9	34.6	33.9	33.6	33.2	32.6	31.0

Table 3.5: Non-selective schools with pupils known to be eligible for FSM of:

***more than 13% and up to 21%***

#### Percentage of pupils achieving Level 5 and above

	95%	UQ	60%	Median	40%	LQ	5%
English	84	70	65	61	57	51	35
Mathematics	78	70	66	63	61	57	46
Science	79	70	65	63	60	56	44

#### Percentage of pupils achieving Level 6 and above

	95%	UQ	60%	Median	40%	LQ	5%
English	50	33	28	24	21	16	8
Mathematics	54	43	39	36	34	29	19
Science	45	34	29	27	24	20	12

#### Average KS3 points score achieved<sup>1</sup>

	95%	UQ	60%	Median	40%	LQ	5%
English	35.9	33.5	32.5	32.0	31.5	30.6	28.2
Mathematics	36.4	34.6	33.8	33.4	33.0	32.1	30.2
Science	35.0	33.4	32.7	32.4	31.9	31.2	29.5

<sup>1</sup> Use 'Ready Reckoner' provided on page 31, to calculate your school's average KS3 point score.  
[www.standards.dfes.gov.uk/performance](http://www.standards.dfes.gov.uk/performance)

## A3: Extract from DCS Staff Handbook, pages 40-41

### 2.4 Assessment and recording

The policy and practice is founded in the major principles:

- assessment, reporting and recording is central to the teaching and learning process. It provides a means of identifying strengths, targeting weakness, evaluating progress and gathering data for a range of audiences
- assessment assists teachers in the evaluation of teaching and learning
- assessment should inform long and short term planning. Effective planning depends upon well-managed assessment strategies, coherent monitoring practice and regular evaluation of outcomes
- assessment is a process that continues in a cycle of teaching and learning. It is built into a planned programme of work. It should involve dialogue and communication between teacher, students and parents. Students should be made aware of assessment criteria
- assessment should reflect the concept of multiple intelligences. Teachers should ensure they include a wide variety of assessment activities
- the process is clear and manageable
- assessment supports equal opportunities

#### Aims

- to inform and dynamise the teaching and learning process
- to generate a wide range of evidence
- to celebrate a wide range of achievements
- to inform a range of audiences
- to raise standards of achievements

#### Purposes

"The assessment process itself should not determine what is taught and learned.... It should be an integral part of the educational process, continually providing both feedback and 'feed-forward'. It therefore needs to be incorporated systematically into teaching strategies and practices at all levels. Since the results of assessment can serve a number of purposes, these purposes have to be kept in mind when arrangements for assessments are designed."  
(Task Group on Assessment and Testing, 1989)

The purposes of acquiring information for assessment should be:

- formative - so that the positive achievements of students may be recognised and discussed and to aid decisions on how learning can be advanced so that information can be used as feedback to modify teaching and learning strategies
- diagnostic - so that learning difficulties may be scrutinised and classified in order that the cause of the difficulty may be identified and appropriate teaching and learning strategies are adopted
- ipsative - so that individual progress can be built upon and target for improvement against 'personal best' can be set
- summative - so that the overall attainment of the student can be recorded in a systemic way
- evaluative - so that aspects of the work of the School can be judged and reported upon

#### Process

The overall purpose of assessment is to promote student learning, leading to higher achievement by providing motivation to learn. The three main elements for the processes are:

##### 1 - For learning

Assessment should support individual students in their learning through the process of review, evaluation and target setting and forward planning.

The process should seek to:

- recognise, acknowledge and celebrate individual progress and achievement
- identify learning difficulties and barriers to learning and inform decisions about the next and future steps to learning
- help students to take responsibility for and to participate in their learning
- provide students with clear and positive feedback on their progress
- assist students in setting targets for their personal and academic achievement

## A4: CHS multiple choice AFL Assessment sheet

## AFL ASSESSMENT ANSWER SHEET

Name \_\_\_\_\_ Title variation Date \_\_\_\_\_

- |  |  |  |   |
|--|--|--|---|
| 1) <u>B</u> <input checked="" type="checkbox"/>  | 6) <u>  </u> <input checked="" type="checkbox"/> | 11) <u>D</u> <input checked="" type="checkbox"/>               | 16) <u>  </u> <input checked="" type="checkbox"/> |
| 2) <u>A</u> <input checked="" type="checkbox"/>  | 7) <u>  </u> <input checked="" type="checkbox"/> | 12) <del>  </del> <u>D</u> <input checked="" type="checkbox"/> | 17) <u>  </u>                                     |
| 3) <u>Bd</u> <input checked="" type="checkbox"/> | 8) <u>C</u> <input checked="" type="checkbox"/>  | 13) <u>B</u> <input checked="" type="checkbox"/>               | 18) <u>  </u>                                     |
| 4) <u>A</u> <input checked="" type="checkbox"/>  | 9) <u>D</u> <input checked="" type="checkbox"/>  | 14) <u>  </u> <input checked="" type="checkbox"/>              | 19) <u>37%</u>                                    |
| 5) <u>  </u> <input checked="" type="checkbox"/> | 10) <u>C</u> <input checked="" type="checkbox"/> | 15) <u>  </u> <input checked="" type="checkbox"/>              | 20) <u>  </u>                                     |

Total = 7/16

### A5: An example of a teacher's marking

Neutralisation of indigestion tablets Friday 7<sup>th</sup> April 2010

Put units on the heading of the table only

Type of tablet	Volume of acid used (mL)	Time taken to change colour <del>it</del> with methyl orange
Antacid	35	37 <del>secs</del>
Remeigel	11	3:42 <del>min</del> <i>convert this to seconds</i>
Alka	50	2:00 min
Rennie	4	0:30 secs

conclusion: From my <sup>result</sup> ~~test~~ I found out that Rennie is the ~~fastest~~ best tablet because it took the quickest time to neutralise the acid while Remeigel was the worst.

AF5  
L5

(T) To get L6 on your conclusion

- Learn to compare results properly interpreting data consistent with evidence (result) and adding detail scientific explanations.



## A6: Meeting with DHT (WCS)

Meeting with DHT (WCS) 1/07/11

Review of WCS Science Dept.

Jan 2011

Indicators  
2010

	School	LA	Nat
5 A*-C in EN & MA	43	51.6	53.4
KS4 250 indicators	24	50.8	61.6
3 L Program	39	45	53
⇒ KS3 L5+	70	73	80
KS3 L6+	31.6	36	48

But  
T31  
reported  
57%

T32	3	Lesson observation Grades
T2	2	
T6	4	
Tc	4	

2 = good

3 = Satisfactory

4 = Inadequate

L/Walk  
26/06/11

Focus Cover

- comp in lessons to enthuse students
- Lo led lessons
- Effective use of Success Criteria
- Appropriate pace and teaching expectations.

low expectations

Lack of marking

Use of criteria X weak

KS3 Co-ordinator had reported L5+ = 57 %

on 21/10/10

See T31  
Interview

**A7: An example of book audit by Head of Science**

	Issues	Comments
<b>Progress over time</b>	Not much progress No clear progression - Very fragmented	Without checking there is no way to see how students are progressing
<b>Relevance of tasks/challenge</b>	Very low level No clear progression	
<b>Opportunities for extended writing</b>	None	Back of book not being used
<b>Standards</b>	Not checked: not identifying key areas. (HL, SK. Hardly anything in books!	Standards not being checked. There is no detailed marking therefore <u>no</u> follow up.
<b>Effectiveness of marking</b>	SK not marked!! (13/1) HL hardly marked!! (15/19) RK hardly marked!! (16/1)	Best a tick for <u>one</u> HW - no mark or comment
<b>Presentation</b>	RK - OK - diagrams are few! Not on a single page	Poor DCS policy needs to be followed.



## A8: Learning and teaching CPD forum - CHS

# Workshop 11

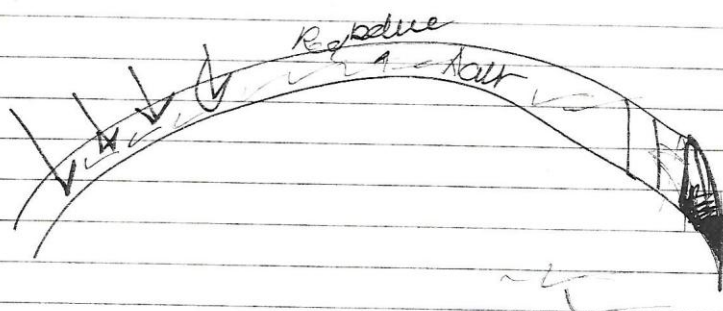
## Agenda

Item	Description	Time
1	<b>Introduction</b> <b>By the end of the session we will:</b> <ul style="list-style-type: none"> <li>To know a range of different group work strategies to use as a starter or plenary activity</li> <li>To be able to teach students to explain effectively using group work strategies</li> </ul>	
2	<b>Starter activity</b> SMi <ul style="list-style-type: none"> <li>Paired test led from the front</li> </ul>	10 mins
3	<b>Using group work strategies to support literacy at word level</b> SMi (supported by group leader) <ul style="list-style-type: none"> <li>Activity one: Heart attack!</li> <li>Activity two: Group work Evaluation exercise</li> </ul> <b>Plenary</b> <b>Sketch and Spell</b> using whiteboards	35 mins
4	<b>Using group work strategies to support literacy at text level</b> SMU (supported by group leader) <ul style="list-style-type: none"> <li>Group activity on 'Miss X looking for love' Feedback will be taken.</li> <li>Individual activity on 'Miss X looking for love'</li> <li>Paired work. A consensus must be reached</li> <li>Feedback to group</li> <li>Decision making and recording</li> </ul> <b>Plenary</b> This will be led from the front.	35 mins
5	<b>Feedback</b> (led by group leader supported by challenger) Each member of the group will be asked to feedback on: <b>AfL techniques you observed and</b> <ul style="list-style-type: none"> <li><b>AfL techniques you use.</b></li> </ul> Questions to think about as you actively listen to feedback: How did it help students' learning? What was formative about it? How do you know it is working? (You may want to refer to <b>My personal action plan</b> from the previous session.)	10 mins
6	<b>Personal action planning</b> (led by group leader supported by challenger) Complete <b>My personal action plan</b> , choosing which techniques you are going to use, including the chosen <b>group work strategy</b> , and <b>what you will do less of</b> <b>Select a peer from your table or speak to Sophie Dryden to request an SLC to observe you. On sheet C record the time and date of the peer observation.</b> You must select: <ul style="list-style-type: none"> <li>a class;</li> <li>a time to be observed using agreed group work strategy and</li> <li>a peer to observe you.</li> </ul> The observation is to be a maximum of 15 minutes with 10 minutes planned to feedback the observed evidence for the teacher's reflection. Share your plan with the group and ensure your observation schedule has been agreed by the	10 mins

A9: Sample page from Year 8 pupil's book (P<sub>32</sub>)

there are 3 primary colors red blue  
~~and that~~ green. Com 1500 use red to violet

Thate | S abroad OP  
 RuBush



### A10: Sample pupil progress card

		Form: 8FT				
Subject	Trg	PR1	PR2	PR3	PR4	PR5
Art	7c			2	2	2
Dance	7c			2	2	2
Drama	7c	1	2			
Design and Technology	7b	4	4	4	4	4
English	7c	3	3	3	3	3
French	7c	3	2	2	3	3
Geography	7c	2	2	2	2	2
History	7c	2	2	2	2	2
Information Technology	7b	2	2	2	2	2
Mathematics	7b	2	2	2	3	2
Music	7b	2	3	4	3	3
Physical Education	7c	2	2	2	2	2
Science	7c		2	2	3	2

