

**St. Patrick's College,
Dublin City University**

**Supervisor: Dr. Michael O'Leary
Auxiliary Supervisor: Professor James Deegan**

**From Balkanisation to Boundary Crossing:
Using a Teacher Learning Community to Explore the Impact of Assessment
On Teaching and Learning in a Disadvantaged School**

Dissertation


by

**Zita Lysaght
B.Ed., M.Ed., M.Sc.**

**Submitted to the Department of Education, St. Patrick's College, DCU, in Partial
Fulfilment of the Requirements for the Award of the Degree of
Doctor of Education.**

April 2009

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of the degree of Doctor of Education, is entirely my own work and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed: 

ID No.: 54005595

Date: 16 April 2009

ABSTRACT

From Balkanisation to Boundary Crossing: Using a Teacher Learning Community to Explore the Impact of Assessment On Teaching and Learning in a Disadvantaged School

Irish education is currently witnessing a fundamental shift in focus from an emphasis on curriculum revision, professional development and policy implementation to one of review, with the concomitant effect that issues of evaluation and assessment are increasingly high on the agenda. This is evidenced by the publication in recent years of a series of evaluation reports which, in addition to identifying successes achieved, have highlighted a number of related iterative weaknesses in the system that continue to thwart progress. One is the apparent lack of assessment literacy which has been attributed mainly, but not exclusively, to teachers. Another is the growing body of international evidence suggesting that pedagogical practices in schools serving disadvantaged children are qualitatively different to those found in more advantaged settings. Finally, there is the evidence that low levels of literacy persist in Ireland, despite significant investment and innovation.

In responding to these challenges, this study examined the potential of a teacher learning community (TLC) as a vehicle of professional development, to bring about changes in teachers' understanding and use of Assessment *for* Learning (A/L), in order to improve the reading competency of a cohort of children attending a designated disadvantaged, junior school, in the Republic of Ireland. Employing a partially mixed, concurrent, equal status, quantitative/qualitative design, the study investigated three research hypotheses pertaining to (1) children's reading achievement, (2) their motivation to read/employment of A/L strategies when reading and (3) teachers' knowledge, skills and attitudes of/to A/L. Outcomes from the quantitative data with respect to the first two hypotheses indicated that there were no statistically significant differences in mean reading achievement between control and experimental groups following the intervention, although significant differences were found for reading strategies. The qualitative data relating to the third hypothesis revealed that important changes had occurred in teachers' attitudes and classroom practice over the duration of the project. The study concludes by drawing attention to the potential of a TLC, reconceived as a boundary zone, to challenge the traditional balkanisation of teachers' working lives.

**This thesis is dedicated to my mother, Theresa,
Who, together with my late Dad,
Nurtured in me the motivation and desire to learn.**

ACKNOWLEDGEMENTS

The time has finally come for me to write the acknowledgements and while I am tempted to say 'I did it all myself', I fancy having someone to party with, so I'll press on!

First and foremost, I extend deepest thanks to my family and friends who have loved and encouraged me over a very long period. Your steadfast support makes all the difference in the world.

To my supervisor, Dr. Michael O'Leary - who nurtured, shaped and honed my academic skills with such acumen and generosity over the past number of years – my sincere gratitude. In moving on, I will happily abide by your principle of 'paying forward' and if I can inspire a fraction of the confidence in others that you inspired in me, they will surely triumph.

To Professor James Deegan, my auxiliary supervisor and long-time friend, thank you for your insightful and timely contribution to this work and for going to bat for me on so many occasions over the years.

To the principal, teachers and staff of the school in which this research was undertaken – I owe an enormous debt. The successful completion of this project pivoted on your commitment to work with me conscientiously to the end. That you did so, with such unfailing professionalism, good humour and integrity, was truly remarkable. In this context, I also acknowledge, with thanks, the financial support to the project provided by the National Council for Curriculum and Assessment and the Research Committee of St. Patrick's College.

It would be inappropriate to finish without paying tribute to my colleagues in the Special Education Department and my co-journeyers in the inaugural class of the first taught doctoral programme in Ireland. To Ms. Maria Thornbury, who responded with such professionalism and kindness to all our concerns along the way, my heartfelt thanks. And finally, to Dr. Ciaran Sugrue, who convinced me that this was the doctoral programme for me - your stalwart tutelage and support over the period of this programme are very gratefully acknowledged.

TABLE OF CONTENTS

ABSTRACT

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

CHAPTER 1: A STATEMENT OF THE RESEARCH CONTEXT AND PROBLEM

Introduction.....	1
Organisation of the Chapter.....	2
Context of the Study.....	3
Research Evidence on the Efficacy of Assessment <i>for</i> Learning	4
Teachers' Assessment Literacy	6
Reading Standards in Disadvantaged Schools	8
Classroom Privacy and Challenges to Instructional Leadership	10
Teacher Professional Development	11
Policy, Implementation and Politics	13
Conclusion to Section	16
Study Scope, Focus and Interpretation of the Problem.....	16
Organisation of the Thesis.....	17

CHAPTER 2: LITERATURE REVIEW

Introduction.....	19
Organisation of the Chapter.....	20
Literature Review: Part 1 - Teacher Professional Development.....	20
Factors Influencing Student Achievement: The Role of the Teacher	22
The Current Status of Teacher Professional Development	23
Factors Mediating the Influence of Teacher Professional Development	24
Models of Teacher Change	27
A Four-Step Model of Teacher Change	28
The Interconnected Model	28

The Concerns-Based Adoption Model	30
The Efficacy-Based Change Model	31
Efficacy Theory	32
Attribution Theory	35
The Implications of Efficacy-Based Research	35
Conceptions of Teacher Knowledge: Implications for Teacher Learning.....	36
Emergent Theories of Learning.....	40
Towards a Social-Constructivist Framework of Situated Learning	40
Implications for Teacher CPD.....	43
Research on Professional Development: A Review from a Situative Perspective	44
Teachers and their Learning: Individual Focus	46
Processes and Activities of Professional Development: Group Focus	47
Teacher Learning and Research Methodology: Dual Focus	48
The Value of Multifocal Research Lens: Simultaneous focus	49
Best Evidence Synthesis Review	50
Assessing the Impact of Teacher CPD: The Need for Evaluation Tools	53
Summary of Literature Review: Part 1	56
Literature Review: Part 2 - Assessment <i>for</i> Learning.....	57
Towards a Definition of Terms	58
From Testing to Assessment	60
The Development of Interest in A/L	64
Formative Assessment Reviews	65
Inside the Black Box	68
The Six Sub-Sections of the Review	68
Significant Findings	70
Principles of A/L	71
The King's-Medway-Oxfordshire Formative Assessment Project	74
The Learning How to Learn Project	77
The Assessment is for Learning Programme	80
The Keeping Learning on Track Project	84

Delivering A/L at Scale	87
A/L in the Irish Context.....	90
Summary of Literature Review: Part 2.....	92
Conclusion: Towards a Conceptual Framework for the Study.....	93

CHAPTER 3: THESIS DESIGN AND METHODOLOGY

Introduction.....	95
Organisation of the Chapter.....	95
From Research Problem to Research Hypotheses.....	96
Research Paradigms.....	97
Paradigm Wars and the Incompatibility Thesis	97
Mixed Methods Research Typologies	101
Towards an Interpretive/Constructivist, Mixed Methods Framework	104
Research Study.....	106
Host School	106
Research Instruments and Data Collection: Hypotheses 1 and 2	108
Pre- and Post-Intervention Data: Control and Experimental Classes	108
Standardised Reading Tests	109
Elementary Reading Attitude Survey	109
Instrument Reliability	109
Scaled A/L Questionnaire	110
Instrument Reliability	111
Test Administration	111
Research Instruments and Data Collection: Hypothesis 3	111
Pre- and Post-Intervention Data: Teachers of Control and Experimental Classes	112
A/L Audit Instrument	112
Teachers' Learning Logs	113
Piloting of Instruments	113
Professional Development Programme and Timeline.....	114
The 'Content' of the A/L Professional Development Programme	114
The 'Process' of the A/L Professional Development Programme: The TLC	115
Video Footage of Teaching	116

Data Analysis.....	117
Quality Control.....	118
Ethical Considerations.....	120
Conclusion.....	122

CHAPTER 4: DATA ANALYSIS AND RESEARCH FINDINGS

Introduction.....	123
Organisation of Chapter.....	124
Characteristics of the Children's Sample.....	124
Research Variables.....	125
Hypotheses and Findings.....	128
Hypothesis 1: A nine-month, school-based intervention, employing Assessment <i>for</i> Learning principles and practice, would make a quantitative difference (i.e. effect size) to the reading achievement of a target group of children when compared to a similar cohort not involved in the intervention	128
The Paired-Samples T-Test	130
Hypothesis 2: There would be a discernible, positive impact on children's attitudes, motivation and approaches to reading	135
Hypothesis 3: Using the medium of a site-based, teacher learning community to provide a professional development programme on Assessment <i>for</i> Learning would have a positive impact on teachers' knowledge, skills and attitudes of/to AfL and, in turn, how they teach reading	141
Group Profile	141
Research Findings: Teachers' Data	142
Teachers' Assessment Audits: Pre- and Post-Intervention	143
Teachers' Learning Logs	149
Teachers' Group Response to the CPD Programme	155
Summary.....	161

CHAPTER 5: CONCLUSIONS AND RECOMMENDATIONS

Introduction.....	162
Research Hypotheses and Findings.....	163
Hypothesis 1	163
Hypothesis 2	164

Hypothesis 3	164
Limitations of the Study	165
Research and Policy Implications and Recommendations.....	167
Recommendation 1	167
Recommendation 2	170
Recommendation 3	172
Recommendation 4	174
Epilogue.....	176

APPENDICES

Appendix A. Brookhart's Review of Classroom Assessment From 1982-2002.....	177
Appendix B. The Ten Projects of the AifL Programme.....	179
Appendix C. 100+ Practical Classroom Techniques (Thompson & Wiliam, 2007).....	180
Appendix D. Formative Assessment in the UK: Review by Gardner (2008).....	181
Appendix E. List of Acronyms.....	184
Appendix F. Elementary Reading Attitude Survey.....	185
Appendix G. Scaled AfL Questionnaire.....	189
Appendix H. Scaled AfL Questionnaire Reliability Analyses.....	192
Appendix I. Teachers' AfL Audit Instrument.....	195
Appendix J. Teachers' Learning Log.....	206
Appendix K. Professional Development Programme.....	208
Appendix L. Supplementary Teaching Materials.....	209
Appendix M. Teachers' Learning Log Review.....	213
Appendix N. Teachers' Independent Review of the Professional Development Programme Provided.....	217
Appendix O. Teachers' Independent Review Findings.....	223
Appendix P. Teacher's Video Review Sheet – P.I.N.....	226

BIBLIOGRAPHY.....	228
--------------------------	------------

LIST OF TABLES

Table 1.	Factors Mediating the Influence and Application of
----------	----------------------------------------------------

	Professional Development.....	26
Table 2.	The CBAM - Dimensions and Key Stages (Hall & Ford, 1987).....	30
Table 3.	Conceptualisations of Professional Development, Change Models and Learning Theory.....	39
Table 4.	Research on Professional Development Programmes from a Situative Perspective.....	45
Table 5.	Teacher Reflection and Action: An Analytic Framework.....	55
Table 6.	Summative and Formative Assessment (Harlen, 2006a).....	59
Table 7.	The Incompatibility of New Views of Instruction and Traditional Views of Testing.....	62
Table 8.	A Representative Sample of Research on A/L Pre-Dating 1998.....	66
Table 9.	A Typology of Formative Assessment (Adapted from Wiliam, 2007).....	85
Table 10.	Aspects of Assessment <i>for</i> Learning (Thompson & Wiliam, 2007).....	86
Table 11.	Facilitating and Sustaining TLCs (Adapted from Thompson & Wiliam, 2007).....	90
Table 12.	Conceptual Framework for the Study.....	94
Table 13.	Research Paradigms and Theoretical Frameworks...	99
Table 14.	Towards an Interpretive, Mixed Methods, Research Framework.....	105
Table 15.	Quasi-Experimental Research Programme and Timeline.....	107
Table 16.	Descriptive Statistics: Control and Experimental Groups.....	125
Table 17.	Mary Immaculate College Reading Attainment Test Pre- and Post-Intervention Histograms for Standard Scores.....	126
Table 18.	Drumcondra Primary Reading Test Post-Intervention Histograms for Standard Scores.....	127
Table 19.	Descriptive Statistics for MICRA-T Pre- and Post-Intervention Standard Scores.....	129

Table 20.	Interpreting Effect Sizes (Adapted from Pallant, 2007)	130
Table 21.	Descriptive Statistics for the MICRA-T Pre- and Post-Intervention Standard Scores.....	131
Table 22.	Descriptive Statistics for the Drumcondra Primary Reading Test Post-Intervention Standard Scores.....	132
Table 23.	Descriptive Statistics for the MICRA-T Pre- and Post-Intervention Standard Scores (SEN Groups).....	133
Table 24.	Descriptive Statistics for the Elementary Reading Attitude Survey Post-Intervention Standard Scores...	136
Table 25.	Descriptive Statistics for the Scaled AfL Questionnaire Post-Intervention.....	138
Table 26.	Descriptive Statistics for the Scaled AfL Questionnaire Pre- and Post-Intervention.....	139
Table 27.	Teachers' Pre- and Post-Intervention AfL Audit Responses.....	144

LIST OF FIGURES

Figure 1.	Factors Influencing Student Achievement (Adapted From Hattie, 2003).....	22
Figure 2.	A Model of Teacher Change (Guskey, 2002).....	28
Figure 3.	The Interconnected Model of Professional Practice (Clarke & Hollingsworth, 2002).....	29
Figure 4.	The Efficacy-Based Change Model (Adapted from McKinney et al., 1999).....	32
Figure 5.	Factors Influencing Teachers' Success in Implementing Innovations (Stein & Wang, 1998).....	34
Figure 6.	The Black Boxes of Teacher and Student Learning (Timperley & Alton-Lee, 2008).....	51
Figure 7.	Assessment <i>for</i> Learning as a Cycle of Events (Harlen, 2006a).....	73
Figure 8.	KLOT: Content Areas and Learning Phases (Thompson & Goe, 2006).....	89
Figure 9.	Mixed Methods Research Typologies (Adapted from Leech & Onwuegbuzie, 2007).....	103
Figure 10.	From Balkanisation to Boundary Crossing - Video as A Boundary Object.....	171

CHAPTER 1

A STATEMENT OF THE RESEARCH CONTEXT AND PROBLEM

Introduction

This study focused on a challenge of significant, current, international interest: the potential of a teacher learning community - as a vehicle of professional development - to bring about changes in teachers' understanding and use of Assessment *for* Learning (AfL), in order to improve children's school achievement. Specifically, this research project was concerned with determining the effectiveness of AfL practices to enhance the reading competency of a cohort of children attending four Second Classes in a designated disadvantaged, junior school, in the Republic of Ireland. Schools in Ireland are designated disadvantaged according to criteria such as the number of pupils attending the school whose families are resident in local authority housing or non-permanent accommodation, hold medical cards, and are in receipt of unemployment benefit or assistance under schemes administered by the Department of Social Welfare.

As employed in this study, Assessment *for* Learning, which is used interchangeably with the term formative assessment, adopts the extended definition employed in the original short publication for teachers prepared by Black, Harrison, Lee, Marshall and Wiliam (2002), following the 1998 landmark review of evidence about the effects of improving formative assessment on students' performance (Cf. Black & Wiliam, 1998a). The rationale for adopting this definition is the same as that articulated by those who penned it - to avoid any confusion or misunderstanding of the term (Black, 2005). Hence, Assessment *for* Learning is to be understood as:

...Any assessment for which the first priority in its design and practice is to serve the purpose of promoting pupils' learning. It thus differs from assessment designed primarily to serve the purposes of accountability, or of ranking, or of certifying competence. An assessment activity can help learning if it provides information to be used as feedback, by teachers, and by their pupils in assessing themselves and each other, to modify the teaching and learning activities in which they are engaged. Such assessment becomes 'formative assessment' when the evidence is actually used to adapt the teaching work to meet learning needs. (Black et al., 2002, p. 1)

A teacher learning community (TLC) refers to a small, site-based, group of practitioners who meet regularly to share, critically review and reflect on their teaching practice and pedagogical knowledge and then use this learning to actively improve their practice for the benefit of children's learning. From the outset, this thesis acknowledges the challenge implicit in trying to define with acuity related terms such as TLC (McLaughlin & Talbert, 2006), professional learning community (Stoll & Louis, 2007) and communities of learners (Cochran-Smith, Feiman-Nemser, McIntyre, & Demers, 2008). However, it should be noted that discussion on TLCs in this work draws – albeit not always explicitly, given limitations of space – on the interrelated and complementary ideas that underpin these overlapping concepts including, *inter alia*, the sharing and co-creation of knowledge, understood both in the epistemological sense (as moving from the tacit to the explicit) and in the ontological sense (as moving from the individual through to the organisational and, potentially, inter-organisational levels), as outlined recently by Harris (2008).

It is not insignificant, in this context, that both formative assessment and teacher learning communities share a common understanding of learning as being essentially social in nature, with knowledge being constructed - by adults and children alike - through their active engagement with, and support of, one another. These ideas, which are associated with social constructivist theories of learning and situated cognition, reflect emerging paradigms of learning and knowing which have significant import for this work; as a consequence they are explored at some length in Chapter 2.

Organisation of the Chapter

This introductory chapter outlines the problem that this research study sought to address, with particular reference to the context in which it was conceived and the scope of the work undertaken. The contextual discussion has a dual focus; it is intended to introduce important factors influencing the decision to conduct the study while, at the same time, distinguishing those that were central to the research questions addressed from those that were both relevant and important, but deemed nonetheless, to be beyond the scope of this work.

Context of the Study

Irish education is currently witnessing a fundamental shift in focus from an emphasis on curriculum revision, professional development and policy implementation to one of review, with the concomitant effect that issues of evaluation and assessment are increasingly high on the agenda. This is evidenced by the publication in recent years of a series of evaluation reports seeking to determine the extent to which the *Revised Primary Curriculum* (Department of Education and Science [DES], 1999) has been implemented, and the effectiveness of the attendant programme of professional development to marshal the recommended changes. These reviews, in addition to identifying successes achieved, have served to highlight a number of related - but nonetheless distinct - iterative weaknesses in the system, that have frustrated progress to-date.

The first is the perennial problem of an apparent lack of assessment literacy (DES, 2005a, 2005b, 2005c; NCCA, 2005b, 2005c; O'Leary, 2008), which has been attributed mainly, but not exclusively, to teachers (Hall, 2000). The second is the growing body of international evidence suggesting that pedagogical practices in schools serving disadvantaged children are qualitatively different to those found in more advantaged settings (Oakes & Lipton, 1999; Shiel, Forde, & Morgan, 1996), coupled with evidence that low levels of literacy persist in Ireland, despite significant government investment (Eivers, Shiel, & Shortt, 2005; Shiel, 2007). Linked to this is a dearth of research evidence and, hence, a paucity of information, on what *actually* happens both within Irish classrooms (Conway, 2002) and at the level of instructional school leadership (Elmore, 2004); in turn, these lacunae raise questions about the basis on which recent nation-wide programmes of professional development, purported to respond to the professional needs of school personnel, were designed. Underpinning each of these are the power plays within the process of negotiation of Irish educational policy which, as reported, have been complicit in silencing a discourse around issues of quality and standards, with the result that the system is deemed poorly positioned to monitor the process of policy implementation (Sugrue, 2004c).

Viewed individually, each of these issues presents formidable challenges of its own; however, taken collectively, they give a sense of a complex national context, the import of which must be considered, both in designing and evaluating any study on assessment because, as Black and Wiliam (2005) assert:

...In each country assessment practices have impacts on teaching and learning that may be strongly amplified or attenuated by the national context. Indeed, the overall impact of particular assessment practices and initiatives is determined at least as much by culture and politics as it is by educational evidence and values. (p. 260)

Hence, in the next section, these issues are introduced briefly as a way of contextualising the study described in this Irish thesis. This is prefaced by a snapshot of the international research evidencing the potential of AfL to bring about significant and sustainable improvements in children's learning, particularly those at most risk of educational under-achievement - the socio-economically disadvantaged and those with special educational needs. Reference is made, in this context, to recent Irish publications on assessment that leverage this research to guide teachers' assessment practices in the classroom.

Research Evidence on the Efficacy of Assessment for Learning

The argument in support of AfL as a warranted strategy to effect significant change in children's educational achievement, especially lower-achieving students, has been well documented. Five seminal reviews (Black & Wiliam, 1998; Crooks, 1988; Kluger & DeNisi, 1996; Natriello, 1987; Nyquist, 2003), which collectively synthesized in excess of 4,000 research studies over a forty-year period, reached one overarching conclusion: "...when implemented well, formative assessment can effectively double the speed of student learning" (Wiliam, 2007, p. 36). More specifically, the original Black and Wiliam research review of 1998 attested to typical effect sizes of between 0.4 and 0.7 in classes where the use of formative assessment was improved, in comparison to classes where it was not used as effectively. In Irish terms, an effect size of 0.4 would mean that "...an average pupil taught by a teacher using high quality formative assessment would record the same score on the DPRT (Drumcondra Primary Reading Test – *my insertion*) or MICRA-T (Mary Immaculate

College Reading Attainment Test – *my insertion*) reading tests as another pupil in the top 35% whose teacher was not using high quality formative assessment” (O’Leary, 2006, p. 11).

The significance of these gains, notwithstanding, the phrase - “when implemented well” - demands critical attention because, as the research also testifies, formative assessment is something of a *Trojan Horse* (Black & Wiliam, 2006); as a concept that is “...more complex than it might appear at first sight” (York, 2003, p. 478), it demands changes in classroom practice that are seductive in their apparent simplicity but which, research has proven again and again, are deceptively difficult to achieve. This is because, if formative assessment is to be implemented well, in accordance with the *spirit* of AfL, then changes of a radical and fundamental sort will ensue including, critically, changes in the roles of both teachers and pupils, and the manner in which they relate and interact with each other, as they strive to bridge the gap between pupils’ existing knowledge and what they are aiming to achieve. To bring about changes of this nature, teachers require more than formative assessment tools and techniques; they require regular and routine opportunities to engage with colleagues in professional development, the process and content elements of which are becoming increasingly well understood.

This is the point at which teacher learning communities enter the debate as vehicles of professional development with the potential to satisfy, most effectively and efficiently, teachers’ ongoing needs for site-based, classroom-focused, reflective enquiry with colleagues. As this juncture, there are strong arguments being made that they reportedly offer “...the most effective, practical method for changing day-to-day classroom practice” (Wiliam, 2008, p.38), with the critical additional attraction of being implementable at scale.

The publication, in late 2007, of the long-awaited guidelines on assessment in the primary school in Ireland (NCCA, 2007), coupled with the promise of professional development opportunities for teachers, reflect a growing appreciation both of the potential of AfL to raise standards and the need to support teachers – not just to adopt new strategies and techniques – but to embrace the challenges and opportunities that follow. Significantly, at the launch of the guidelines, reference was

made to the *Revised Primary Curriculum* (DES, 1999) handbooks as offering the “why” of formative assessment, while the guidelines purportedly provide the “how” (Irish Times Newspaper, 2007). It must be said that a publication of this kind has been long sought (Eivers et al., 2004; O’Leary, 2006; Sugrue, 2004c) and, in light of recent audits of teachers’ assessment literacy appears timely, if not overdue. While it is deemed premature to undertake a review of the impact of the guidelines at this juncture - because they have been circulated to schools only very recently and hence, have not been road-tested to any significant degree - some kind of textual deconstruction is warranted. At this point in the thesis, this takes the form of what has been described by Codd (1998) as an explicit recognition of the context in which such policy documents develop:

Policy document...do not have a single authoritative meaning. They are not blueprints for political action, expressing a set of unequivocal intentions. They are ideological texts that have been constructed within a historical and political context. The task of deconstruction begins with the explicit recognition of this context. (p. 244)

It is to a review of the peculiarly Irish situation that the focus now turns, with a commitment to revisit the guidelines briefly, and the supporting programme of professional development, when examining the import of this study in the final chapter.

Teachers’ Assessment Literacy

In 2005, the Department of Education and Science (DES) inspectorate and the National Council for Curriculum and Assessment in Ireland reported on independent evaluations of the implementation of the *Revised Primary Curriculum* of 1999, focusing on English, Mathematics and Visual Arts (DES, 2005a; NCCA, 2005b; 2005c). In the same year, the DES published a review of the standards of literacy and numeracy in disadvantaged schools (DES, 2005c) on foot of a report they had commissioned a couple of years previously entitled *Reading Literacy in Disadvantaged Primary Schools* (Eivers et al., 2004). A key finding to emerge from these studies was that assessment policy and practice in Irish primary schools require significant attention and improvement.

The executive summary, of what has come to be known as the LANDS report (*Literacy and Numeracy in Disadvantaged Schools: Challenges for Teachers and Learners*) (DES, 2005c), began by acknowledging the particularly challenging contexts in which schools designated as disadvantaged work, characterised by higher rates of absenteeism, lack of parental involvement and worryingly low rates of academic achievement. Although the report stressed the critical need for "...best quality planning and optimal teaching practices..." in such schools, the findings indicated that a significantly higher number of children had standardised test scores in the bottom quintile in literacy and numeracy, compared with children in more affluent areas, with in excess of 50% of the children having "...very low scores in reading, while almost two-thirds of children were extremely weak in mathematics" (DES, 2005c, p. 9; p. 57). This raised questions in relation to teaching and learning, assessment, learning support and instructional school leadership.

In a chapter devoted to the topic of assessment in literacy and numeracy, the report distinguished between whole school and classroom level approaches and acknowledged some aspects of good practice, particularly with regard to the assessment of children with special educational needs and learning difficulties. However, the overall conclusion reached by the team of inspectors was that, although a variety of assessment tools were being employed, assessment policy and practice required "...significant attention and improvement..." (DES, 2005c, p. 62) because the assessment data were not being used effectively to inform teaching and learning.

While one might justifiably argue that these findings were based on a very small sample of just eighteen schools and, hence, may not be either representative or generalisable, they correspond very closely to those of a subsequent study conducted by the DES in which the practices of teachers in eighty-six primary schools were assessed in relation to their implementation of the revised curricula in English, Visual Arts and Mathematics (DES, 2005a). The overlap in findings between these reports led to similar recommendations being reached including the need for school management bodies and principals to actively encourage and promote high expectations of children, backed up by co-ordinated school and class-based planning, to ensure that formative assessment would be used to guide and inform improvements in teaching and learning. In acknowledgement of the challenges this would present

for schools, the reports suggested that support should be forthcoming from the DES, the NCCA and other school support services to enable schools nationwide "... to engage further with assessment issues, clarify what should be assessed, and specify the assessment tools that can be used" (DES, 2005b, p. 54).

It is noteworthy that the issue of professional development for teachers received particular attention in light of teachers' remarks that they were unaware of any courses that would support them in developing their skills in the teaching of literacy or numeracy. The conclusion was reached that "...there is a need, therefore, to plan school-based professional development programmes that would support coherent whole-school responses to the particular needs and contexts of each school..." (DES, 2005c, p. 62).

Coupled with the reports by the NCCA (2005b; 2005c) and Eivers et al., (2004) which again drew similar conclusions, the overall picture is one of a teaching force that requires significant, sustained, professional support if it is to come to grips with best practice in assessment planning and practice. Furthermore, as reported, the lack of assessment literacy is a pervasive problem; it affects schools across the socio-economic divide and not just disadvantaged settings, although, clearly, there is greater urgency to address this problem in schools where a disproportionate cohort of the attending children consistently exhibit very serious academic difficulties, the full extent of which has just recently emerged.

Reading Standards in Disadvantaged Schools

The twin problems of educational disadvantage and low standards in reading have been the focus of many Irish studies over recent years (Archer & O'Flaherty, 1991; Shiel, Cosgrove, Sofroniou, & Kelly, 2001; Weir, 2001). However, a notable difficulty with these studies was that they failed, collectively, to provide a description of the reading achievements of pupils in a broad range of disadvantaged schools either because they focused on extreme cases of disadvantage or too narrow a sample of schools. Hence, the commissioning of a study by the DES in June 2002, to obtain data from a representative sample of designated disadvantaged schools of the reading achievement of children at three class levels: First, Third and Sixth (Eivers et al.,

2004). Given the unique focus of this study, the findings reported below are derived exclusively from it.

Almost 6,500 pupils across 93 designated disadvantaged primary schools participated, their reading achievements being assessed using a standardised reading test. In the context of this study, the following findings are particularly noteworthy:

- At each of the three grade levels, pupils in the survey were out-performed by pupils in a representative national sample who had participated in the standardisation of the test in the previous year, the difference being approximately two-thirds of a standard deviation at each grade level;
- The percentages of pupils who achieved scores at, or below, the 10th percentile were considerably higher than at time of standardisation, with 27% of First and Sixth Class pupils, and 30% of pupils in Third Class, achieving scores at, or below, the 10th percentile (compared to 10% at each class level in the standardisation sample);
- Assuming the 10th percentile to be the cut off point for identifying pupils in need of learning support, between a quarter and three-tenths of pupils surveyed would qualify for additional help;
- Fewer than 5% of pupils at any of the three grades achieved scores at, or above, the 90th percentile.

This is a very significant and lengthy report, the contents of which deserve much closer scrutiny than is being afforded here, not least because it also gave consideration to the status of assessment literacy in the schools surveyed. A number of important observations were made. First, the practice of curriculum-based assessment, including documenting pupil learning outcomes, seemed to be indicative of practices in learning support and special education contexts, rather than in ordinary classrooms. In response, the authors noted that it is "...perhaps regrettable in this context that the implementation of the 1999 Primary Schools Curriculum has not been supported to date by a stronger emphasis on, and in-career development in, the use of classroom-based assessments to enhance pupil learning" (Eivers et al., 2004, p. 16).

It was also remarked that although summative assessments featured in ordinary classroom practice, other forms of assessment were less common, with almost 75% of teachers indicating that they had never used curriculum profiles. Perhaps most worryingly, the majority of class teachers surveyed indicated that they did not view increasing time spent on pupil assessment as a priority, with the authors drawing the conclusion that "...this may be because the value of formative assessment has not been sufficiently highlighted in Ireland" (Eivers et al., 2004, p. 168).

Given the findings, it was estimated that it would take a minimum of ten years to effect a substantial change in reading standards amongst this population of children, the target being "... to halve (to 14-15%) the proportion of pupils at or below the 10th percentile in designated schools..." within 10 years (Eivers et al., 2004, p. 168).

Classroom Privacy and Challenges to Instructional Leadership

One of the major obstacles to be overcome in trying to improve children's learning - in Ireland as in many other countries - is that what goes on inside the classroom is largely unknown; hence the "black box" analogy (Black & Wiliam, 1998). While this is always a controversial issue, it becomes a crucial one when faced with educational under-achievement of the scale and severity of that reported in the previous section. Conway (2002) articulates the problem well:

The moment-to-moment transactions between students, teachers and curriculum have received insufficient attention in various efforts to address educational disadvantage. In sum, conflict and debate about pedagogical and curricular concerns (i.e. what is taught, how and why it is taught and its impact) remains marginal in both educational research and debates on educational disadvantage. (p. 63)

Much has been written about classroom privacy. In this context, some researchers have provided evidence that teachers in disadvantaged schools demonstrate low expectations of children - for example, in the nature and range of academic tasks set - as a consequence of which children experience less challenging learning environments than their more advantaged peers, culminating in reduced

educational achievement (Oakes & Lipton, 1999; Shiel et al., 1996). Other researchers argue that school leaders tend to protect teachers from being subjected to external scrutiny in the event that there is any perceived weakness in practice at the classroom level. Richard Elmore (2000), who has written extensively on this theme, makes two critical arguments, the first in relation to what he terms the “technical core” of education:

...Detailed decisions about what should be taught at any given time, how it should be taught, what students should be expected to learn at any given time...and perhaps most importantly, how their learning should be evaluated - resides in individual classrooms, not in the organizations that surround them. (Elmore, 2000, pp. 5-6)

Second, he asserts that if the “technical core” is perceived to be weak, the primary function of the administrative layer of the system - including principals and boards of management - becomes one of intentionally buffering this weak internal core from outside inspection; as a consequence, educational leadership becomes “... not the management of instruction but the management of the structures and processes around instruction” (Elmore, 2000, p. 6). Recent comments in the Irish context seem to suggest that this practice might not be uncommon here:

...Whenever even veiled criticism of teachers and teaching is advanced, it is accompanied by ritual condemnation of structural inadequacies such as school architecture, lack of equipment and resources, as well as inadequate provision for professional development. (Sugrue, 2005c, p. 187)

As a consequence, the ‘black box’ of teaching and learning remains unexamined, school leaders’ efforts are directed away from the core function of instructional leadership and, *ipso facto*, children’s achievement. Moreover, teachers’ professional needs and concerns remain backgrounded - unheard and unresolved.

Teacher Professional Development

It is worth remembering that the studies reported so far came on foot of the introduction of the *Revised Primary Curriculum* (DES, 1999), which advocated strongly the use of formative assessment, coupled with a prolonged and extensive programme of professional development to support curriculum implementation. Yet,

the assessment challenges, as outlined, do not appear to have changed substantially from those reported previously by Hall and Kavanagh (2002) who found that "...much of the assessment conducted..." appeared to be "... intuitive and impressionistic rather than systematic and detailed..." resided "... in the teachers' heads..." (p. 266), and, therefore, was not grounded in evidence that could be easily shared with others, including parents and children.

This raises concerns about policy messages - their clarity and impact - which are signalled in the next section. However, perhaps more importantly, it raises questions about the nature and quality of the programme of professional development that was offered in an effort to bring about changes in teachers' knowledge, skills, attitudes and practices.

In 2005, following a review of the range of measures and programmes of support to tackle disadvantage over the previous two decades, the DES introduced the DEIS programme, *Delivering Equality of Opportunity in Schools*, a new action plan to "...build on the success of existing measures for tackling educational disadvantage, while identifying and effectively addressing the issues that reduced their overall effectiveness in the past" (DES, 2005d, p. 2). Laying out their programme of work, they made a strong case for replacing "...the traditional off-site "training course" model (of professional development – *my insertion*)..." with a model that would be based "...in school and classroom settings... where opportunities (would be – *my insertion*) provided for teachers to be actively involved in the analysis of teaching and learning..." (DES, 2005d, p. 61). This represented a significant departure in DES policy given that so much of the professional support offered by the Primary Curriculum Support Programme (PCSP) - the function of which was "...to mediate the Primary School Curriculum for teachers towards enabling them to implement it in their schools..." (PCSP, 2008) was that of the rejected model. Indeed, a recently commissioned review of the national programme of professional development offered to Irish primary teachers, in support of their implementation of the Revised Curriculum (DES, 1999) by the PCSP, questioned the logic of replicating "...the same seminar hundreds or even thousands of times with small groups of participants..." in the context that while "...the financial, material and personnel costs of this approach

are obvious; the observable benefits are less clear” (Murchan, Johnson, Loxley, & Fitzgerald, 2006, p. 24).

It is significant in this context that an over-arching conclusion drawn from a recent all-Ireland conference entitled *Teacher Education and Schools: Together Towards Improvement* (SCoTENS, 2006) was that “...most of the problems... with regard to professional development activities in OECD countries derive from their fragmented nature, their tendency to be unrelated to classroom practice, and the lack of intensity and follow up” (p. 7). In attempting to address these weaknesses, they urged that future policy developments in Ireland should attend to four issues:

- The need to ensure that incentives are provided for all teachers to encourage their participation in professional development, to include a combination of entitlement-based professional support, incentives-associated promotional opportunities at the individual level, and school development activities for all staff;
- The development of coherence and consistency between individual teachers’ needs and school development priorities;
- The provision of opportunities for schools to develop as professional learning organisations characterised by routine reflection and peer-review;
- The requirement for more coherent frameworks to document and certify teachers’ engagement in professional development coupled with the possibility of broadening the range of activities available to include “...peer review and action research, mutual school visits, and the development of teacher and school networks – activities that in order to be successful require sufficient time and resources” (SCoTENS, 2006, p. 8).

Policy, Implementation and Politics

McLaughlin (2006) makes the interesting point that “...how a policy problem is framed - what a policy concern is assumed to be a “problem of” - arguably is the most important decision taken as a policy is developed...” (p. 3). Her observation was particularly important in the context of this study because the focus, design and

scope of this research reflected a particular interpretation of the challenges facing Irish education currently, and Irish assessment and professional development policies, in particular. This interpretation was influenced both by the nature and urgency of the issues discussed thus far, as well as by the broader contextual issues that underpin them. While it is clearly beyond the scope of this chapter to discuss these broader issues in any depth, flagging them affords a greater understanding of the importance and timeliness of the research study undertaken. Hence, four final issues are raised.

First, the argument has been made that the 1991 OECD report, *Review of National Policies for Education: Ireland*, had a catalytic effect on Irish educational discourse, spawning an era of unprecedented debate, analysis and policy development (Clancy, 2005; Coolahan, 2003). However, it has also been argued that any groundswell of collective reflection on the shaping of educational debate, the process of policy decision-making or the task of policy-analysis itself has been conspicuously absent during this period (Hall, 2002; Looney, 2001; O'Sullivan, 1992; Walsh, 1997). Hence the comments that "...despite unprecedented change... there has been relatively little discussion about primary schooling, its goals and pedagogies" (Sugrue, 2004c, p. 168), as a consequence of which, "...the broad sweep of curriculum policy (in Ireland – *my insertion*) remains under-analysed and... under-theorised" (Looney, 2001, p. 2).

Second, against a backdrop of performativity (Lyotard, 1984), "high-stakes, standardized, test-based reform" (Thompson, 2001, p. 361) and the "re-regulation" of education (Ball, 2003) - manifest particularly in the US and the UK - one might anticipate that the assessment-related challenges facing Ireland currently would be similar in nature. However, research specific to Irish assessment, though limited, does not necessarily support this view. What has been suggested is that assessment policy at primary level is characterised by epistemological confusion in policy text (Hall, 2000, with reference primarily to *A Programme for Reform: curriculum and assessment policy towards the new century*, NCCA, 1993). This is coupled with a tendency to be "...stronger on progressive sentiment and much more vague on what this would look like in practice" (Sugrue, 2004c, p. 197, with reference to the *Revised Primary Curriculum*, DES, 1999). In the context of the former, the conclusion was reached that:

...Policy reports since the 1990s have assumed that there is an automatic and simple link between diagnostic assessment and the capacity to promote learning. They have assumed that since teachers are already engaged in teacher assessment for formative purposes, through such informal means as observations, teacher-set tasks, classroom interaction and the like, they are doing it well, and therefore, further guidance is unnecessary. (Hall & Kavanagh, 2002, p. 262)

Third, in sharp contrast to other jurisdictions, teachers in Ireland exercise considerable influence on educational policy debates. Indeed, the fact that it took almost fourteen years to move from a position where "... an extensive use of attainment tests..." was deemed "inappropriate ...and... especially prejudicial to the needs of disadvantaged pupils" (Primary Curriculum Review Body, 1990, pp. 80-81) to the announcement by the DES of "...a controversial and unexpected proposal to introduce standardised testing in literacy and numeracy for all in compulsory education" (Looney, 2006, p. 351), suggests that, traditionally, there has been a clear preference to *induce* rather than *mandate* reform (McDonnell & Elmore, 1987; Stone, 2002). In the context of the partnership approach to Irish decision-making, this is not altogether surprising. Clancy's comments (2005) are illuminating in this context, in particular his observations that:

...Irish teacher unions (and other special interest groups such as managerial bodies) enjoy a virtual veto on the formulation of educational policy.... One of the implications of involvement in statutory bodies such as the NCCA and investigative bodies such as the Primary Education Review Body is that it serves to prevent certain issues getting on the policy agenda and circumscribes the range of solutions which are considered. (p. 92)

This commentary links to the fourth, and final, point to be raised in this section, the fact that - since the early 1990s and arguable to the present day - "...the evidence suggests that teachers have consolidated their "domination" of the field of curriculum development and increasingly also, their position in the allied field of professional learning" (Sugrue, 2004c, p. 195). This observation is supported by the fact that, as members of the curriculum working parties of the NCCA, teachers co-authored the *Revised Primary Curriculum* (DES, 1999) and, subsequently, were largely responsible for the delivery of a programme of nation-wide professional development to guide their teaching colleagues in implementing the curriculum as intended.

Conclusion to Section

This potted review of some of the key principles of AfL and teacher learning communities, in tandem with the broad sweep commentary on the current Irish context, was intended to situate the study for the reader. The key issues raised in the previous sections may be summarised as follows: (i) research on AfL has provided quantitative data to suggest that it improves children's achievements and is particularly beneficial for weak children; (ii) in Ireland, as in other jurisdictions, teachers engage in assessment practices routinely, however, they do so with varying degrees of skill and success, with assessment *of*, rather than *for*, learning tending to be emphasised; (iii) many children in Ireland are continuing to experience serious literacy and numeracy difficulties, particularly children attending disadvantaged schools, despite years of intensive investment; (iv) although the elements of AfL associated with improving children's learning are known, in Ireland currently, teachers are neither sufficiently assessment literate, nor is there adequate knowledge of existing classroom life, to exploit this pedagogy with success; (v) in general, approaches to professional development in Ireland seem to have been at odds with what is recommended internationally, being largely removed from the school context, irregular and discontinuous; hence, the need to test-case an alternative regime which is context-based and scalable; (vi) Irish assessment policy lacks a robust epistemological foundation that manifests in a critical under-estimation of the challenges to be met in implementing formative assessment effectively; this is exacerbated by – if not in part the result of – a system of policy development and implementation roll-out in which teachers exercise enormous influence, with mixed results.

Study Scope, Focus and Interpretation of the Problem

As signalled earlier, the previous section was also intended to distinguish between issues of primary and secondary interest to this thesis, thereby defining the scope of this work, while also setting some boundaries. So, to begin with some boundaries: this study was not about policy formation, interpretation or implementation; it did not seek to address, *per se*, reading pedagogies, the nature and quality of education afforded children from socio-economically disadvantaged areas,

or any attenuating links to low literacy standards in schools designated disadvantaged, although it did fully acknowledge the powerful mitigating influence of these issues. Rather, this study foregrounded two issues - AfL and teacher professional development - and sought to investigate if a one-year, intervention project, based in a designated disadvantaged, junior school, could effect changes in teachers' assessment literacy to the extent that it would have a positive impact on children's reading achievement. A more elaborate articulation of the problem and linked research questions follows the literature review.

This statement of the problem reflects the researcher's interpretation of the challenges presenting as being - first and foremost - an absence of research knowledge in the Irish context. For although AfL has been road-tested, with varying degrees of success in other countries, to-date, there has been no intervention in Ireland of the scope and scale of that undertaken by this research. Nor has there been any comparable study that tested the potential of a teacher learning community to support teachers' incremental learning in AfL, to the benefit of children's learning, especially in disadvantaged schools. Given the bodies of research evidence attesting to the nature and degree of underachievement amongst children in disadvantaged schools on the one hand, and the potential of the effective use of AfL to bring about significant improvements in children's learning, particularly the traditional 'low-achievers', on the other, one might even argue that a study of this kind carries a moral imperative.

Organisation of the Thesis

This introduction constitutes the first of five chapters. Chapter 2 focuses on literature and previous research in the areas of Assessment *for* Learning, professional development for teachers and emerging paradigms of learning and cognition. The chapter engages Maxwell's conceptualisation (2005b) of a literature review as an integral element of the research design and conceptual framework of a thesis, rather than as the basis or starting block from which a thesis progresses. Hence, rather than attempting an exhaustive review of the literature, the chapter attempts to draw selectively and judiciously from key research to highlight the distinction between the

process and content elements of AfL, thereby informing the research questions raised in this chapter.

This is followed, in Chapter 3, by a detailed exposition of the methodology employed in undertaking this study. With reference to current debates about the applicability of qualitative and quantitative research to research in the social sciences, the chapter begins by mapping three research paradigms and their associated theoretical frameworks, by way of introduction to the conceptual framework of this study. The chapter progresses to discuss the mixed methods study design, including the methodological instruments employed at each stage. It concludes by outlining the ethical issues raised by the study and how these were addressed.

Chapter 4 presents an analysis of the findings from the quantitative and qualitative data and uses this information to test, in turn, each of the hypotheses of the study with reference, as appropriate, to previous findings of comparative research.

The final chapter offers a critical review of the study, its findings and its implications, in the light of previous research and literature, and policy and implementation concerns specific to the Irish system at this time. The thesis concludes by acknowledging the potential strengths and weaknesses of the work undertaken and the opportunities for future research that may emerge as a consequence. In particular, the traditional barriers to learning such as the professional isolation of “egg-crate” schools (Lortie, 1975) and the associated “balkanisation” which can ensue (Hargreaves, 1992) are foregrounded using boundary crossing as a metaphor for understanding the potential of a TLC and video as an artefact to mediate teachers’ experience of implementing AfL in their classrooms and beyond.

CHAPTER 2

LITERATURE REVIEW

Introduction

In the context of the interpretation of the research problem - and in light of the research questions generated - this chapter focuses on literature and previous research initiatives relating to *Assessment for Learning* and continuing professional development (CPD) for teachers, with reference also to emerging paradigms of learning and cognition. Recent exchanges between Boote and Beile (2005; 2006) and Maxwell (2006) highlighted a significant divergence of opinion within the educational research community regarding the proposed form and function of literature reviews. As noted:

This division is between faculty who expect a thorough review of the research literature in the area of a dissertation (the traditional view), and those who want a selective review of the literature that relates directly to what the student plans to do, showing these works' implications for the proposed study.... (Maxwell, 2006, p. 30)

Rejecting, what he interpreted as, Boote and Beile's (2005) advocacy of a traditional, foundationalist approach, and a literature review conceived as the basis or starting block for a thesis, Maxwell (2006) proposed that what a doctoral study requires is a conceptual framework that develops from traditional lines of inquiry, but serves a very different purpose:

Examining, assessing, and connecting published research is an important source for this conceptual framework, but the goal is an integrated set of theoretical concepts and empirical findings, a model of the phenomena they are studying that informs and supports the research, rather than a review of a body of literature. (p. 30)

In keeping with Maxwell's (2006) interpretation, this chapter does not purport to offer an exhaustive review of literature; rather it attempts to draw, selectively and judiciously, from key literature and research, in order to highlight some of the more significant debates that have characterised international discussions on the core

themes of A/L, CPD, learning theory and cognition, in recent years. The chapter concludes by presenting a conceptual framework that acts as a bridge between the research problem identified, the literature reviewed and the research study undertaken.

Organisation of the Chapter

The chapter is in two parts; part one considers the issue of teacher professional development in the context of the changing paradigms of learning. Part two introduces A/L and teacher learning communities as approaches that, amongst others, are increasingly promoted as offering the greatest potential to meet the content and process needs of educational reform, leading to increased student achievement. The decision to begin this review with reference to professional development was taken because it offers an opportunity to signpost some of the bigger themes that shaped and contextualised this thesis including, critically, the relatively recent realisation within the research community that it is *within*, rather than *between*, school factors that significantly determine student achievement, i.e. it is the teacher that counts (Darling-Hammond, 2000; Hattie, 2005). The decision to place the discussion of A/L and teacher learning communities in the second part of this chapter can be interpreted as an acknowledgement that, while these are currently 'hot topics' in the research literature, they are not the only approaches offering potential for school reform and increased student achievement, a point readily acknowledged by some of the most vociferous advocates of these approaches (e.g. Black & Wiliam, 2006; Wiliam, 2008).

Literature Review: Part 1 - Teacher Professional Development

Teacher development has been defined as "...the professional growth a teacher achieves as a result of gaining increased experience and examining his or her teaching systematically" (Glatthorn, 1995, p. 4). While it is acknowledged that professional development programmes vary in terms of process and content (Reeves, McCall, & MacGilchrist, 2001; Wilson & Berne, 1999), it is argued that most share the common purpose of effecting changes in the "...professional practices, beliefs, and understandings of school persons toward an articulated end" (Griffin, 1983, p. 2), the most obvious end being the improvement of student learning (Guskey, 2002, p. 381). There is a problem, however: some commentators suggest that teacher

professional development may have failed to deliver (Borko, 2004; Elmore, 2004; Fullan, 2006; Wiliam & Thompson, 2007). Reflecting on his review of the field, Fullan (2006) observed recently:

Even reform efforts that had millions of dollars and political will behind them, along with focusing on many of the right strategies (standards, assessment aligned with standards, curriculum revision, plenty of professional development for teachers and principals and even professional learning communities) have failed to make much of an impact on the classroom.... What is going on here? We finally get jurisdiction to take the reform literature seriously and we still get halting reform efforts. (p. 11)

These remarks belie expectations and beliefs commonly expressed in the literature about the centrality of teacher CPD for school improvement; the argument, as presented, may be framed as follows. The success of school improvement and reform initiatives hinge, to a large degree, on the qualifications and effectiveness of teachers (Darling-Hammond, 2000; Tharp & Gallimore, 1988) because "...the single most important variable in the amount of progress that a student makes at school is the quality of the teacher..." (Hattie, 2005, p. 14). The best way to increase teacher effectiveness in the classroom is through regular, high quality, professional development (Hanushek, Kain, & Rivkin, 1998; Killion, 1999). However, despite a range of professional development models and approaches (Guskey, 2000; 2002), traditionally, the professional development available to teachers has been "woefully inadequate" (Borko, 2004, p. 3) and "...flew in the face of what the research says about what makes for effective professional development" (Wiliam, 2007, p. 4). Contemporary developments in learning theory demand a reconceptualisation of the role of the teacher as an active learner, and the professional development designer as facilitator and enabler of a socially-interactive environment, roles that are new and challenging (Putnam & Borko, 2000; Shepard, 2000). Hence, as Borko (2004) argues, "...we have a full research agenda ahead of us to gather information to guide professional development policy and practice" (p. 3).

In the sections that follow, each of the elements of this argument is unpacked in turn. Beginning with an outline of the shift in focus from school improvement and school effectiveness to teacher improvement and effectiveness, the argument leads to an examination of the findings of a number of reviews of relevant research literature.

In this way, the opportunities and challenges that lie ahead, with regard to the design and implementation of an efficient and effective system of CPD for teachers, are highlighted.

Factors Influencing Student Achievement: The Role of the Teacher

Conceptualisations of educational change, school improvement and the role of the teacher have been co-evolving, propelled, at least in part, by advances in technology that have empowered researchers to examine hierarchical, or nested, data structures, including those relating to school effectiveness and improvement, with great efficiency and reliability. In turn, the analysis of disaggregated and longitudinal school databases have yielded crucial information – not just in relation to *between-school* effects - but, more critically, in relation to the *value-added* element of schooling, attributable to individual teacher's efforts, within schools, as shown in Figure 1.

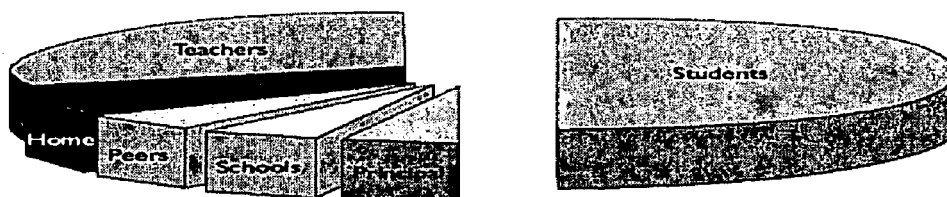


Figure 1.

Factors Influencing Student Achievement (Adapted from Hattie, 2005)

By decomposing the major sources of variance on student achievement – such as school leadership, school climate and culture, teachers, home and student' attributes - researchers have uncovered a number of notable patterns. Critical among them is the finding that variability in student achievement is "...far greater between classes within a school than between schools" (Hattie, 2005, p. 15), thereby replacing the school with the teacher as the fulcrum of educational achievement, and shifting the focus from school improvement to teacher improvement, and, ipso facto, creating a demand for effective teacher professional development.

The last fifteen to twenty years have witnessed significant change in the way professional development of teachers is conceptualized, with one of the most significant changes being a shift from a technical “menu” approach, including singular-session workshops, activities and techniques that focus on generic skill development, to an espousal of the principles of lifelong learning (Sugrue, Morgan, Devine, & Raftery, 2001). Moreover, this is reflected in a change in terminology from talk of ‘in-service’ education to continuing professional development; Hogan and Smith (2006) present the case well:

Continuing Professional Development (CPD) for teachers, as for other occupations, is increasingly placed in the context of lifelong learning by the international literature on social policy (e.g. OECD) and by the literature of educational research. On the one hand, or at one end of the spectrum, CPD could be viewed as a series of incremental ‘upskilling’ activities, tailored to the emergent requirements of educational systems, and carried out at intervals over the duration of a career in teaching. By contrast, it could mean something more visionary, holding more fruitful promise: an unprecedented enrichment of the daily environments of teaching and learning in schools and colleges, sustained through networks in which teachers are actively involved. (p. 79)

In an attempt to capture more clearly the key elements that might constitute this revisioned model of CPD, Putnam and Borko (1997) undertook an in-depth review of what was being advocated by writers in the field of teacher professional development; four common themes emerged:

- An emphasis on the social-constructivist nature of teachers’ learning;
- A commitment to a situative perspective aimed at exploiting the various school contexts in which teachers routinely engage to encourage colleagues to review and debate their practices in order to increasingly develop their understanding and knowledge of, and about, teaching;
- An acknowledgement of the need to demonstrate respect for teachers as professionals towards their increased empowerment;

- A recognition of the importance of modelling the design and delivery of professional development so that it reflects what the teachers are asked to create in their own classrooms, with their own students.

These “mantras” or “truisms”, as Putnam and Borko (1997) refer to them, fit with many of the theories of professional development being espoused in the literature, such as Ball and Cohen's (1999) “practicebased” theory of professional development, for example. According to this theory, professional learning for teachers should emphasize long-term, active engagement, connections between teachers' work and their own students' learning, and opportunities to practice and apply what students learn in a real-world context. The emphasis is on a continuous cycle of exploration of new issues and problems, creating cognitive dissonance between existing practices and beliefs and emerging theories of learning, engaging in collaborative discussions, constructing new understanding and improving professional practice. The importance of developing and sustaining conversations and activities, which encourage a critical and investigative stance towards teaching and learning, is a recurring theme in the literature (Ball & Cohen, 1999; Darling-Hammond & McLaughlin, 1995; Little, 1993; Loucks-Horsley & Stiegelbauer, 1991), as is the belief in the potential of learning environments for teachers that are more collaborative and ‘centred in practice’ (Little, 1993; Cochran-Smith & Lytle, 1999a; 1999b), both of which reflect an evolving understanding of the socio-constructivist nature of learning.

Factors Mediating the Influence of Teacher Professional Development

The current body of research on the relationship between professional development, teachers' instructional practices and student achievement is limited (Darling-Hammond, 2000). As observed by Loucks-Horsley and Matsumoto (1999), to-date, the majority of research on professional development has focused primarily on examining changes in instructional practices, teachers' knowledge and teachers' beliefs - variables that are very important - but which may be only indirectly linked to student achievement. Furthermore, the fact that this literature is “...largely self-referential to the perspectives of the adults and not to the benefits for children...”

suggests to some that "...in the interests of both children and their teachers, there needs to be far more attention to the nature and the outcomes for both teachers and students of professional development, and the links between these (William, 2006, p. 360).

As a consequence, notwithstanding the difficulties and expense involved in undertaking research on the specific relationship between professional development and student achievement, repeated calls are emerging for research linking CPD with tangible improvements in student learning. Such calls find voice in the 'process-product' association of increased teacher effectiveness and student achievement (Adey, 1995). Timperley and Alton-Lee (2008) articulate this view very forcibly when they state:

The outcomes-linked evidence about effective professional learning poses challenges to researchers and teacher education scholars in relation to the approach to be taken both to knowledge building in this field and to providing professional development. It is timely to reflect on the moral purpose of education and teacher knowledge. If we accept that our schooling systems are ultimately in the business of educating students, not teachers..., then we must use (student – *my insertion*) outcomes as the criteria for effectiveness of our various improvement efforts. (p. 360)

The linearity of this argument, as expressed, is not uncontroversial however. In a salutary note, other researchers refer to the 'dilution' effect of professional development, the idea that what a teacher actually does - his or her practice - is but one of a host of interrelated factors that influence children's learning. Smith and Gillespie (2007) make the case well:

Teachers do not exist in a void; they are individuals with different backgrounds and ambitions who work in varied school and system contexts. In the same way as student achievement is affected by factors other than the instruction they receive (including socioeconomic status, race, and class size), teacher change is also affected by individual and school factors that influence how they improve instruction. Although the teacher is always the link between professional development and student achievement... the actual impact of the professional development is diluted by all the other factors that support or hinder teachers from making change.... (pp. 225-226)

Assuming the position that the process of professional development does not always result in the product of student achievement, Smith and Gillespie (2007) introduced three frameworks for evaluating the factors that mediate the influence of professional development on teacher change and, potentially, student achievement; for ease of analysis and comparison, the key elements of each framework are presented in Table 1. This model is intended to simplify those of Guskey and Sparks (1996) and Ottoson (1997) and the extent to which this is achieved is not at issue here. What is important, however, is the extent to which their framework spotlights teacher-specific factors that serve to mediate change.

Table 1.

Factors Mediating the Influence & Application of Professional Development

Guskey & Sparks (1996)	<p><u><i>Factors that mediate the influence of CPD:</i></u></p> <ul style="list-style-type: none"> ➤ Content characteristics ➤ Process variables ➤ Context characteristics
Ottoson (1997)	<p><u><i>Factors that affect the application of what is learned through CPD:</i></u></p> <ul style="list-style-type: none"> ➤ Educational Factors ➤ Innovation ➤ Pre-dispositional factors ➤ Enabling factors ➤ Reinforcing factors
Smith & Gillespie (2007)	<p><u><i>Factors that mediate the influence of CPD:</i></u></p> <ul style="list-style-type: none"> ➤ Individual teacher factors: <ul style="list-style-type: none"> ➤ Teacher motivation for professional development ➤ Teacher concerns ➤ Teacher self-efficacy ➤ Teacher cognitive styles or “ways of knowing” ➤ Teacher reflectiveness ➤ Teacher formal education and years of experience ➤ School, programme and system factors: <ul style="list-style-type: none"> ➤ Leadership ➤ Coherence between professional development topic and school reform ➤ Collegiality within the school

Detailed analysis of the factors identified within the three frameworks is not offered at this point because they are unpacked in the context of a review in the next

section of four models of teacher change that demonstrate how these mediating factors have variously influenced CPD models, their conceptualisations and design.

Some twenty years ago, Stein and Wang (1988) observed that:

Despite the fact that conventional wisdom and research data have pointed to teachers and teacher training as key ingredients of school improvement, data on teacher characteristics associated with implementing innovative programs is sparse. Particularly lacking is information on the motivational factors undergirding change efforts, including how and why innovative programs are adopted and maintained by teachers. (p. 171)

As a review of some of the current models of teacher change in the next section suggests, in the intervening period since 1988, a number of researchers have given serious consideration to the issues raised here by Stein and Wang.

Models of Teacher Change

Guskey (2002) argued that, although the failure of programmes of professional development may be attributed to a range of factors, two are crucial: the motivating factors that encourage and sustain teachers' involvement and the process by which change in teachers typically occurs:

The crucial point is that it is not the professional development per se, but the experience of successful implementation that changes teachers' attitudes and beliefs. They believe it works because they have seen it work, and that experience shapes their attitudes and beliefs. Thus, according to the model, the key element in significant change in teachers' attitudes and beliefs is clear evidence of improvement in the learning outcomes of their students. (pp. 383-384)

In line with this thinking, a number of change models have been developed in recent years, four of which are outlined here. The selection of models for review was dictated by the fact that, collectively, they offer a range of insights into various interpretations of teacher change. Additionally, they serve to introduce issues in relation to efficacy, attribution and motivation, factors that have significant import for students' learning, as will be discussed in the second part of this chapter.

A Four-Step Model of Teacher Change

Guskey (2002) developed a four-step model of teacher change that proposed an alternative change sequence to that typically employed. Based on the belief that change in student learning acts as a catalyst and motivator for change in teachers' beliefs and attitudes, rather than being a product of such change, Guskey (2002) devised the following model of teacher change.

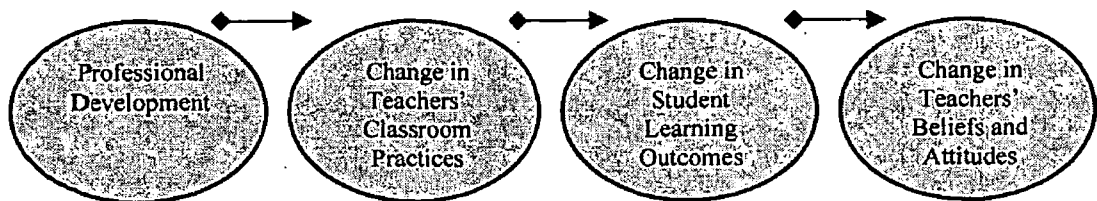


Figure 2.

A Model of Teacher Change (Guskey, 2002)

Guskey's (2002) model draws attention to the importance to teachers of evidence that the changes they are making, on foot of professional development, really benefit students' learning. Hence, he argues, practices that yield results are maintained and those that do not are abandoned. However, Guskey (2002) also acknowledges that change is not easy and he highlights three important principles when trying to implement change. First, there needs to be an explicit acknowledgement that change is demanding and takes place over time; successful implementation, then, is not an event; it is a process. Second, routine feedback on student achievement supports and motivates teachers' on-going engagement. Third, ongoing pressure must be coupled with on-going support; both are vital.

The Interconnected Model

While elements of this model have been influential - notably the emphasis placed on the need for teachers to adapt classroom practices in light of professional development - the linearity of Guskey's interpretation of change has prompted some adaptations by other researchers, including Clarke and Hollingsworth (2002). The

Interconnected Model of Professional Practice builds on the previous model by retaining the four elements of change, albeit using different terminology; however, it replaces the linearity of Guskey's (2002) model with an emphasis on change occurring between four domains: the external domain, that lies outside the teacher's personal world and the practice, consequences and personal domains, respectively, as depicted in Figure 3.

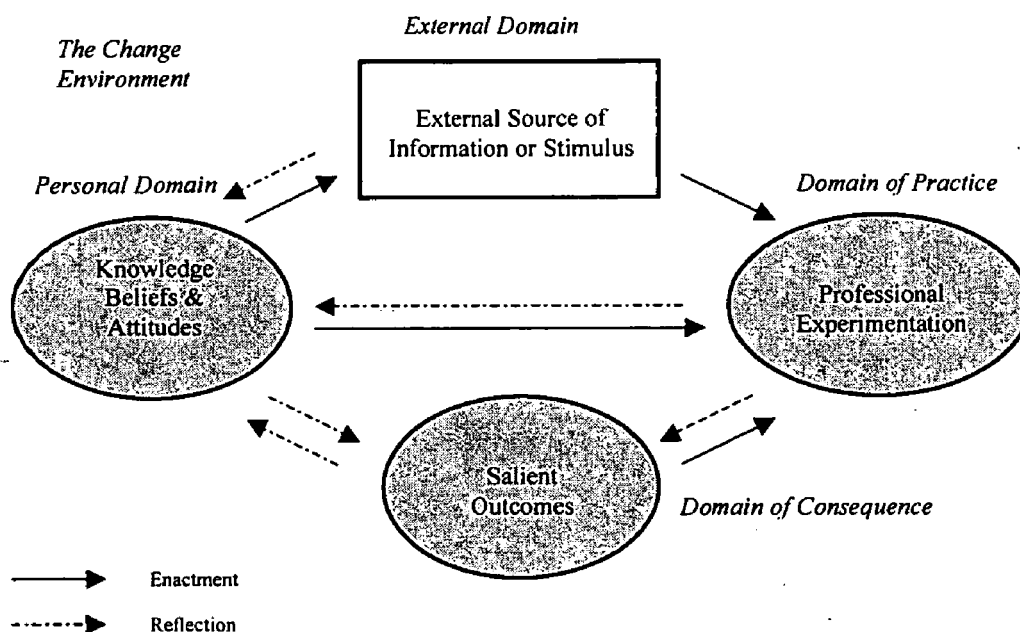


Figure 3.

The Interconnected Model of Professional Practice (Clarke & Hollingsworth, 2002)

Critically, Clarke and Hollingsworth's (2002) model emphasises that the sequence of change is non-linear; it can occur in any one of the four domains as determined by the nature of change that takes place. Hence, changes in teachers' knowledge, beliefs and attitudes occur within the personal domain, whereas, classroom implementation change is located in the domain of practice. According to this model, inter-domain change is mediated through the processes of teacher "reflection" and "enactment", with the term "enactment" being used to "...distinguish the translation of a belief or a pedagogical model into action from simply "acting", on the grounds that acting occurs in the domain of practice, and each action represents the enactment of something a teacher knows, believes or has experienced" (Clarke & Hollingsworth, 2002, p. 951).

Significantly, a change model developed some fifteen years previously, emphasised what Clarke and Hollingsworth (2002) later called the personal domain, and, specifically, the role played by teachers' concerns, or more crucially the support offered in response to the concerns identified, in determining the success of an innovation.

The Concerns-Based Adoption Model

The Concerns-Based Adoption Model (CBAM) was designed originally by Hall and Ford (1987) as an instrument to evaluate the role of school leadership in the change process and as a guide to teachers' concerns about new innovations. Arguing that change naturally causes a person to question and express concerns, Hall and Ford (1987) proposed a model that represents the kinds of concerns that are most often expressed during the course of an intervention. Defining teachers' concerns as "... the feelings, thoughts and reactions individuals have about a new program or innovation that touches their lives", they argued that:

...Being concerned about change is universal, even though the nature of the concerns varies from person to person. Concerns exert a powerful influence on the implementation of a change, and they determine the kinds of assistance that teachers find useful. (Hall & Ford, 1987, p. 30)

The CBAM categorises teachers' concerns in seven stages under three dimensions that are distinct, but not necessarily mutually exclusive, i.e. individuals may have to negotiate more than one stage within a dimension at any given time.

Table 2.

The CBAM - Dimensions and Key Stages (Hall & Ford, 1987)

Dimension 1: <i>Self Concerns</i>	Dimension 2: <i>Task Concerns</i>	Dimension 3: <i>Impact Concerns</i>
- Stages -		
<ul style="list-style-type: none"> ➤ Awareness ➤ Informational ➤ Personal 	<ul style="list-style-type: none"> ➤ Management 	<ul style="list-style-type: none"> ➤ Consequence ➤ Collaboration ➤ Refocusing

According to the framework, as teachers progress through intervention initiatives, their concerns change from initial preoccupation with themselves (Awareness, Information, Personal), through task related issues (Management), to issues relating to the impact of the innovation on children (Consequence, Collaboration, Refocusing). This is reflected in changes in the nature of the questions they raise. In turn, while variations in the rate and nature of individuals' progression are to be expected, research shows that progress is directly attributable to the "...pattern and intensity of change... that is directly affected by the nature of the change and the kind and, especially, the amount of support provided" (Hall & Ford, 1987, p. 32). Traditionally, the CBAM has enjoyed success within the schools change movement as an instrument that could be used variously as a diagnostic and/or evaluative instrument for facilitating change. More recently however, the CBAM has been incorporated into the Efficacy-Based Change Model (EBCM). As conceived by McKinney, Sexton, & Meyerson (1999), this provides a more inclusive framework of the change process, indicating the implementation of change over time, the internal processes of the participants and the common and idiosyncratic aspects of change; hence, it is considered in detail here.

The Efficacy-Based Change Model

According to the EBCM, the genesis of people's concerns lies in their self-beliefs or self-efficacy that, in turn, is derived from the attribution process. Taken in concert, the components of concerns, efficacy and attribution contribute to the initiation, implementation and refinement stages of an innovation, as shown in Figure 4 overleaf.

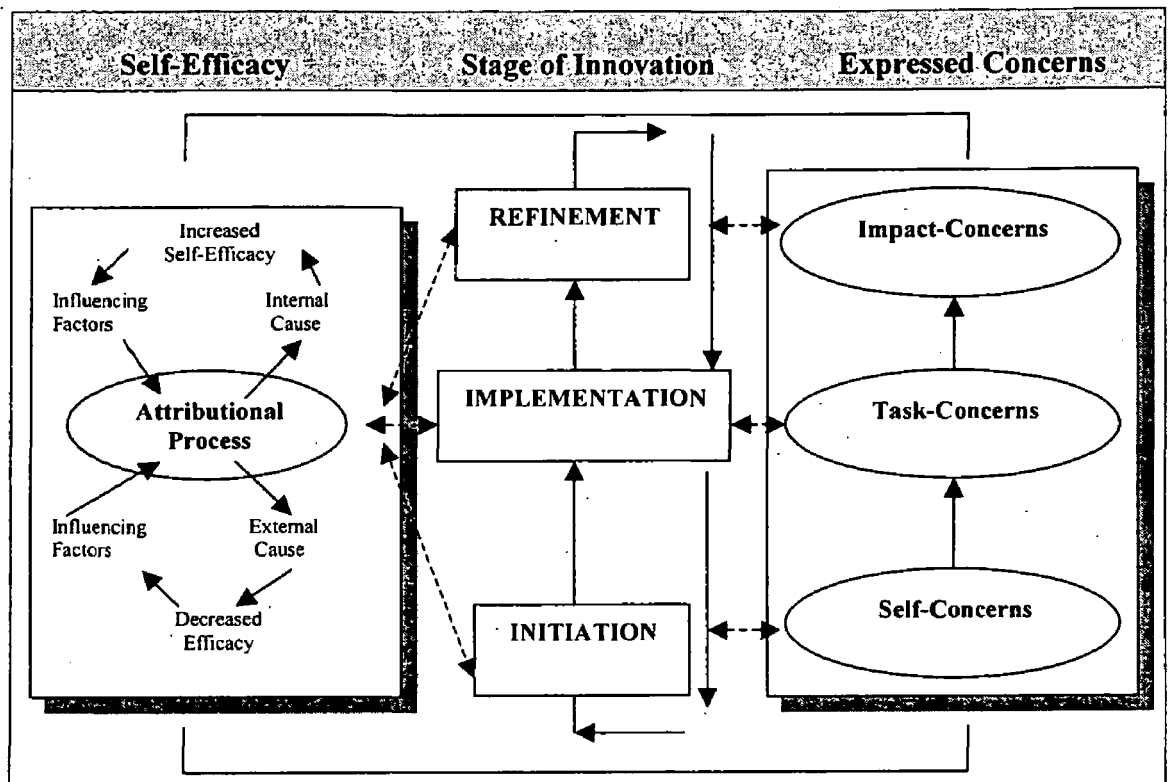


Figure 4.

The Efficacy-Based Change Model (Adapted from McKinney et al., 1999)

To appreciate the EBCM, some further comment on the interplay of forces such as efficacy and attributions is required.

Efficacy Theory

Educational researchers have long grappled with the construct known as teacher efficacy. According to McKinney et al. (1999) efficacy has been portrayed as "...the central mediating self-belief that influences the initiation and maintenance of any new task behaviour, e.g. an educational innovation" (p. 474). Research identifies two branches of efficacy: personal or self-efficacy and teacher efficacy; the former refers, in a general way, to "...how people feel, think, and motivate themselves to behave" (Bandura, 1993b, p. 118). In contrast, teacher' efficacy refers specifically to a teacher's perception of his/her ability to (a) teach well and (b) to do so to the benefit

of his/her students. A critical distinction has been made in the literature between these two facets of teacher efficacy:

Teachers hold two independent beliefs, the belief that they can teach, and the belief that student' outcomes are due to their teaching. The independence of these beliefs suggests that they may differentially influence teachers' instructional decisions. (Soodak & Podell, 1996, p. 409)

Hence, teachers with high-efficacy in their professional ability tend to assume greater responsibility for students' learning, believe they can effect improvement and are open to innovations and new ideas that will support them in this task. Hence, the research findings that (a) teacher efficacy correlates positively with receptiveness to educational change (Guskey, 1998), and (b) teacher efficacy is higher in schools in which teachers participate in making decisions regarding instruction and curricular co-ordination (Rosenholtz, 1989). Research links efficacy to two other self-beliefs: outcome expectation (the likely consequences of a certain behaviour), and outcome value, both of which have been found to correlate with a teacher's perceived efficacy to implement and embrace a suggested change.

In one of the very few empirical studies to examine this correlation, Stein and Wang (1998) devised a conceptual model for the study of factors related to teacher success in implementing innovative programmes and practices, as shown in Figure 5 overleaf. In many respects, the ideas presented here reflect the ideas introduced previously with reference to Smith and Gillespie (2007); it emphasises the interplay of factors, such as the nature of the proposed programme/innovation, its perceived value to students, and teachers' concerns and attitudes regarding their ability to implement the practice as required.

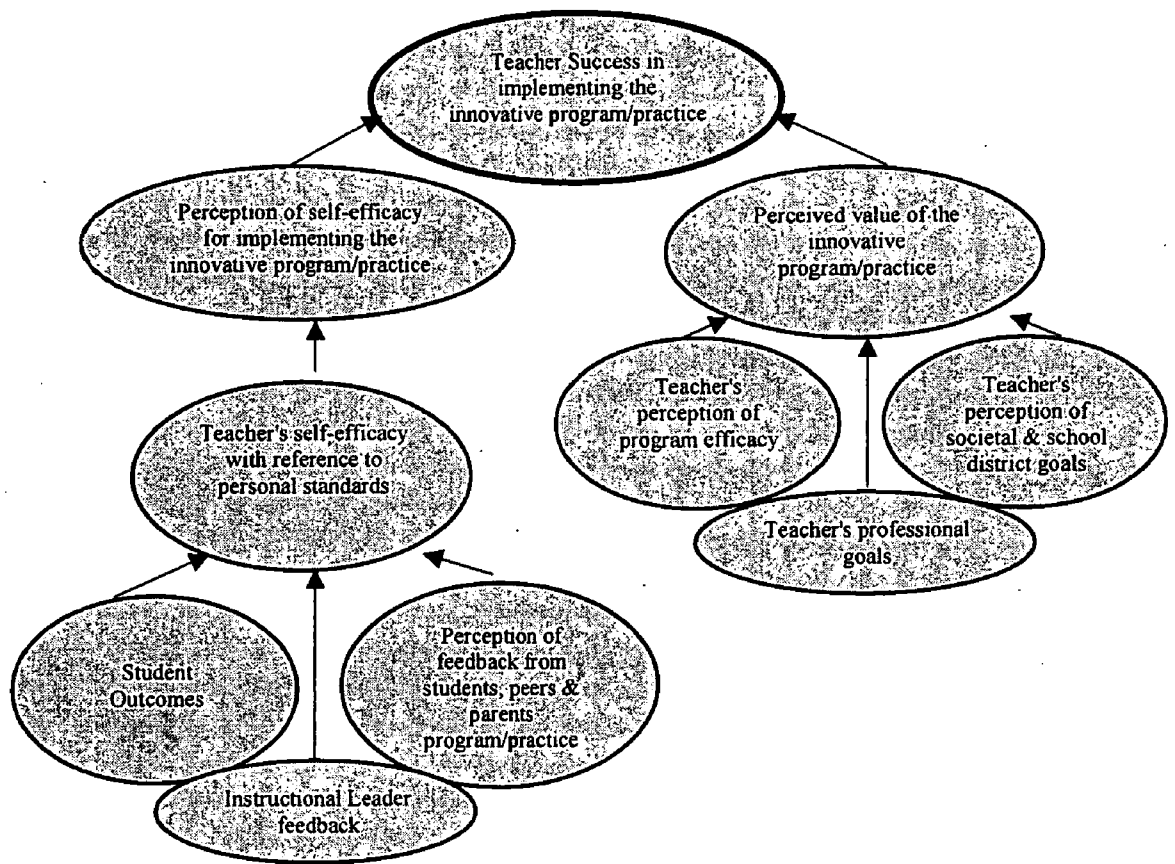


Figure 5.

Factors Influencing Teachers' Success in Implementing Innovations (Stein & Wang, 1998)

The research by Stein and Wang (1998) verified empirically a number of key relationships:

- The values and goals of an innovation need to be congruent with those of the teacher if substantive change is to occur;
- Regardless of a teacher's individual perception of his/her self-efficacy for implementing change, change is contingent on an implicit belief that the programme can, and will, work;
- Where consonance exists between a teacher's professional goals and the perceived goals and/or practices of an intervention, the teacher is more likely to accept and endorse the new programme.

Attribution Theory

Research links much of what is known today about attribution to work conducted by Rotter (1966) and Weiner (1985). In the context of research on motivation, Weiner (1985) determined that the decisions a person makes today are often influenced by the outcomes of similar decisions taken in the past. He used the term 'causal attribution' to describe the causes one attributes to past successes or failures. Accordingly, both adults and children are said to categorise causal attributions in relation to:

- Locus - meaning internal (ability/effort), or external (luck/task difficulty), causes;
- Stability of causes over time. Attribution theory associates a person's expectation of success with comparable prior experience. Past success engenders confidence that one will enjoy success in the future and vice versa;
- Controllability - perceived ability to control the cause. Weiner (1985) suggested that individuals who perceive events to be within their control are more motivated to persist in their involvement than those who feel that they can exert no power or influence over a situation.

The Implications of Efficacy-Based Research

In 1999, McKinney et al. reported how they incorporated research related to teacher' efficacy, attribution and concerns within a comprehensive model of change. In keeping with their predictions, their research confirmed that as participants negotiate an innovation they:

- Move through stages of initiation, implementation and refinement;
- Express different concerns in relation to the efficacy process that are influenced by causal attributions;
- Enjoy greater efficacy when they have fewer concerns. Equally, those with lower efficacy beliefs had more concerns characteristic of those in an early stage of change.

Similar findings were reported in comparable research that employed the same change model and supported teachers' change by acknowledging and responding to their expressed needs and concerns (van den Berg et al., 2000).

As argued at the beginning of this section, and as indicated by each of these models in turn, teacher CPD is mediated or diluted by a variety of factors ranging from internal perceptions of self-efficacy, to the context and conditions in which the innovation takes place, to the perceived value placed both on the *raw* intervention and the speed with which tangible effects on student achievement are observed. Complementing the emphasis on factors that mediate teacher change, there has been a growing appreciation in recent years that in order to design appropriate programmes of CPD for teachers, much closer attention must also be paid to contrasting - and sometimes conflicting - perceptions about both *what* teachers need to learn in order to *teach better* and *how* this might be achieved.

Conceptions of Teacher Knowledge: Implications for Teacher Learning

From Cochran-Smith and Lytle's work (1999a), it is clear that a changing or emerging view of what counts as knowledge for teaching influences both the way teacher learning opportunities are conceived and how CPD is organised:

...Within various change efforts, there are radically different views of what "knowing more" and "teaching better" mean. In other words, there are radically different conceptions of teacher learning, including varying images of knowledge; of professional practice; of the necessary and or potential relationships that exist between the two; of the intellectual, social, and organizational contexts that support teacher learning; and of the ways teacher learning is linked to educational change and the purposes of schooling. Different conceptions of teacher learning - although not always made explicit - lead to very different ideas about how to improve teacher education and professional development, how to bring about school and curricular change, and how to assess and license teachers over the course of the professional life span. (p. 249)

Developing this argument, Cochran-Smith and Lytle (1999a) distinguish between three prominent conceptions of teacher learning that drive reform intended to promote teacher learning: (i) knowledge-for-practice (formal knowledge often delivered to teachers by educational researchers), (ii) knowledge-in-practice

(understood as the practical, sometimes tacit, knowledge that is made explicit through reflection on best practice, and (iii) knowledge-of-practice, respectively. The latter, it is argued, "...unlike the first two... cannot be understood in terms of a universe of knowledge that divides formal knowledge, on the one hand, from practical knowledge, on the other" (p. 266). Rather, it is assumed that:

...The knowledge teachers need to teach well is generated when teachers treat their own classrooms and schools as sites for intentional investigation at the same time that they treat the knowledge and theory produced by others as generative material for interrogation and interpretation. In this sense, teachers learn when they generate local knowledge of practice by working within the contexts of inquiry communities to theorize and construct their work and to connect it to larger social, cultural, and political issues. (Cochran-Smith & Lytle, 1999a, p. 247)

Table 3 (p. 39) combines the work of Cochran-Smith and Lytle (1999a), and Chin and Benne (1969) - specialists in change theory, to highlight the relationships between conceptualisations of teacher learning, professional development design and delivery, typologies of change and learning theory. Summarising the key messages of Table 3, it has been recorded that professional development programmes are frequently didactic rather than constructivist in nature; as a consequence they "...pursue their goals by being directive with teachers in ways that they discourage teachers from being with children..." (Loveless, 1998, p. 188). In turn, this manifests in the perpetuation of a view of "...knowledge as facts and skills, teaching as telling, and learning as remembering..." (Thompson & Zeuilli, 1999, p. 353), ideas that run contrary to progressive theories of learning, and situated, social-constructivism in particular. Adopting this argument to propose that early professional development activities were based on a paradigm that implied a deficit in teacher knowledge - a paradigm that has lately been much criticised in the research literature - Moore (2007) points out that:

...More recently, there has been a fundamental shift in the theories supporting professional development: a switch in emphasis from change as something done to teachers, to teacher ownership and acceptance of change as an active, life-long, learning process.... New models of the process of teacher change have been devised to complement this development.... Foci include: teacher knowledge, skills, strategies, beliefs, attitudes and practice. (p. 3)

Table 3 hints at the nature of these changes, with reference to the cognitive, social-constructivist and distributed perspectives. The importance increasingly attributed to these changes - both by researchers interested in student learning (Black & William, 2006; James, 2006; Gardner, 2006), and those interested in adult learning (Borko, 2004; Cochran-Smith & Lytle, 2002; James & Pedder, 2006b; Lieberman & Pointer Mace, 2008; Thompson & William, 2007) - would suggest that further elaboration is required at this point. Hence, immediately following Table 3 overleaf, consideration is given to emerging theories of learning.

It is not inconsequential in this context that recent research on Irish teachers' experiences of professional learning suggested that provision appeared to combine "... more negative features of 'knowledge for practice' and 'knowledge in practice'..." with "...the absence of support at school/classroom level..." giving rise to a situation in which learning was "...not sustained" in the absence of "...support and context sensitive feedback" (Sugrue, 2002, p.334). While one might argue that these conclusions - at worst, pre-date, and at-best, coincide with - the mushrooming of professional development opportunities in Ireland in recent years, it is noteworthy that, with a few notable exceptions (e.g. Special Education Support Service), models of CPD remain largely unchanged in Ireland. Indeed, as observed, the implications of the continuing the practice of prolonged withdrawal of teaching staff from school contexts to 'attend' CPD are potentially grave. Significantly, Loxley, Johnston, Murchan, Fitzgerald, and Quinn (2007) - reflecting on the seven-year, CPD programme instigated by the DES in 1998 to help teachers mediate the revised primary curriculum - questioned whether:

...A significant long-term implication... may be... (that – *my* insertion) ...it has contributed inadvertently to creating a broad dependency culture in Irish primary education whereby the Department of Education and Science is automatically expected to 'provide' professional development... (rather than – *my* insertion) ...an alternative and perhaps more sustainable approach... in which teachers are expected to identify their own needs and take steps to address them. (p. 283).

Such musings bring into sharp relief the importance of the ideas presented in Table 3.

Table 3. *Conceptualisations of Professional Development, Change Models and Learning Theory*

Interpretations of Teacher Knowledge		
For Practice	In Practice	Of Practice
Associated Interpretations of the Nature of Knowledge		
Empirically verifiable; generated and prescribed by others: "Teachers are knowledge users not generators".	Tacit, contingent, craft knowledge: "Teaching is a wise action in the midst of uncertain and changing situations".	All knowledge is contestable: Teachers generate knowledge by "...making their classrooms and schools sites for inquiry, connecting their work in schools to larger issues, and taking a critical perspective on the theory and research of others".
Associated Conceptions of Professional Learning Opportunities		
Traditional, empirical-rational (assuming the information/policy is uncontroversial); Power-coercive (if information/policy is mandated to ensure compliance); Teachers are rational beings requiring only that knowledge be transmitted; Often decontextualised, remote/off-site, yo-yo delivery reflecting the Balkanized, 'solo practice' of teachers in schools; single session or series focused on range of knowledge and skills teachers need to know/be able to use (<i>disjointed, out of school in-service – education centers/hotels 'delivered by an expert'</i>).	Job-embedded, normative-reductive – teachers require an injection of new knowledge and the opportunity to conceptualise and make sense of it, in context/on-site = school-based; Demands critical reflection – both private (individual) and collective (group/staff collaboration); Teacher learning community - may be contrived and may require external scaffolding, initially; active participation in one's own re-education (<i>In-school days, with/without 'expert' support</i>); on-going/of set duration focused on "artefacts of teaching".	Normative-reductive within multiple contexts and multiple sources of information and data; Action-research, problem-based approach, practitioner-focused and led enquiry; part of a general, on-going commitment to inquiry and reflection and challenging the status quo. Professional Learning Community, frequently internally mediated by teacher leaders/experts = distributed leadership, capacity-building, school culture (<i>on-going, self-started and sustaining; shared/collective responsibility for growth and learning</i>).
Associated Paradigms of Learning		
<i>Cognitive perspective</i> : knowing is results from the individual construction of ever more powerful concepts or logical structures; learning is the acquisition of knowledge, skills and attitudes to be later adapted and applied; individual, solitary pursuit independent of context and intention – horizontal integration of new ideas, "bolt-on" new ideas and strategies to existing repertoire.	<i>Social-constructivist/sociocentric perspective</i> : interaction with others and resources are both the process and the product of learning = learning cannot be analysed without analyzing interactional systems; learning within a socio-cultural context = situated learning; vertical integration of new ideas - challenging and changing established beliefs and practices.	<i>Distributed cognitive and Social-constructivist perspectives</i> – as part of the liberal tradition – one is not seen as distinct from the other but complementary; learning is a process involving the individual <i>both</i> in active individual construction of knowledge and understanding as well as in enculturation into the practices of wider society.

Emergent Theories of Learning

In this section, three clusters of learning theories are described, motivated primarily by the view expressed by Cochran-Smith and Lytle (1999a) that "...different conceptions of teacher learning, although not always made explicit, lead to very different ideas about how to improve teacher education and professional development" (p. 249). However, this work is also spurred by the belief that greater understanding of learning theories will lead, ultimately, to greater alignment between teacher and student education, and classroom assessment, in particular. Before progressing further, a comment on the use of terms is required. As James (2006a) observes, the terms 'behaviorist', 'cognitivist' and 'situated' are frequently used in the US literature, whereas, in the UK, the terms 'behaviourist', 'constructivist' and 'socio-cultural' or 'activist' are more common.

For the purposes of this thesis, ideas from cognitivist, constructivist and socio-cultural theories are combined to inform a general social-constructivist framework that emphasises the nature of situated cognition. By adopting this approach, and introducing this discussion at this point in the chapter, it is intended that subsequent discussions about research on teacher professional development from a situative perspective (i.e. Borko, 2004, next section), and the importance of aligning assessment practices within a social-constructivist frame (i.e. Shepard, 2000, part two), will be more meaningful.

Towards a Social-Constructivist Framework of Situated Learning

The significance of recent developments in understanding regarding learning theories and their implications for teaching and learning cannot be over-emphasised. As Putnam and Borko (2000) note:

The education and research communities are abuzz with new (or at least re-discovered) ideas about the nature of cognition and learning. Terms like "situated cognition", "distributed cognition", and "communities of practice" fill the air.... Some have argued that the shifts in world view... are even more fundamental than the now historical shift from behaviorist to cognitive views of learning. (p. 4)

The shift from the mechanistic theories of knowledge acquisition implicit in the behaviourist perspective - that perpetuated a notion of the learner or subject as completely passive, and the teacher or 'trainer' holding the key to learning success - to the cognitive view of the learner as actively engaged in the learning process, was radical indeed. As Shepard (2000) put it: "...the cognitive revolution reintroduced the concept of the mind" (p. 6) and, specifically, the idea that learning is determined by the mental construction of knowledge within the learner's head. Of particular importance in this reconceptualisation are the ideas that learning construction relies on the use of mental models of the world, that new learning is enabled or impeded by prior beliefs and knowledge structures, and that the components of meta-cognition - including self-monitoring and self-regulation - are instrumental in developing an intelligent approach to learning that raises it above the simple accumulation of facts and isolated pockets of information.

Concomitant with evolving views of the mental processes involved in learning is a growing awareness that learning is not an isolated, individual pursuit. As Lieberman and Pointer Mace (2008) express it:

We are coming to understand that learning rather than being solely individual (as we have taken it to be) is actually also social. In plain terms – people learn from and with others in particular ways. They learn through practice (learning as doing), through meaning (learning as intentional), through community (learning as participating and being with others), and through identity (learning as changing who we are). (p. 227)

This understanding of learning as arising out of the interaction between the individual learner and his/her social environment - that "...what is taken into the mind is socially and culturally determined..." (Shepard, 2000, p. 7) - may be traced back to the work of Vygotsky (1978) who demonstrated the importance of others as learning mediators. James (2006a) elaborates on the implications of this approach for teachers (although, as will be argued in Chapters 3 and 4, this applies equally to the facilitator engaged in teacher professional development):

...Socio-cultural approaches imply that the teacher needs to create an environment in which people can be stimulated to think and act in authentic tasks (like apprentices) beyond their current level of competence (but in what Vygotsky calls their 'zone of proximal development'). Access to, and use of,

an appropriate range of tools are important aspects of such an expansive learning environment. It is important to find activities that learners can complete with assistance but not alone so that the 'more expert other', in some cases the teacher but often a peer, can 'scaffold' their learning... and remove the scaffold when they can cope on their own. Teachers and students jointly solve problems and all develop their skill and understanding. (p. 57)

While not explicit in the quotation above, socio-cultural approaches overlap with situative perspectives on learning, although the roots of the situative view stem more from anthropology, sociology and psychology, and the work of Lave and Wenger (1991) in particular, who conceptualised learning in terms of 'cognitive apprenticeship' in 'communities of learning' (Borko, 2004). Three core themes are associated with the situative perspective:

- First, learning is always context-dependent: "... how a person learns a particular set of knowledge and skills, and the situation in which a person learns, become a fundamental part of what is learned" (Putnam & Borko, 2000, p. 4). Closely associated with this notion of contextualised learning, is the idea that learning is a mediated activity - learners engage with various artefacts, be they physical (e.g. books), or symbolic (e.g. language), and these tools are integral elements of the learning process;
- Second, learning is an inherently social activity with both *what*, and *how*, a person learns being determined largely by his/her active engagement with other learners. This idea extends to include the way in which a person becomes enculturated into the mores and rhythms of groups in which he/she participates and learns; as Black (2001) puts it: "...the process of learning is thus seen as a process of enculturation and one's capacity to learn is seen ... as a capacity to interact and participate effectively in such communities" (p. 79);
- Third, just as learning is not an independent, isolated activity, neither is knowledge owned by any one individual or unit: "...learning involves participation and what is learned is not necessarily the property of an individual but shared within the social group" (James, 2006a, p. 57); hence, the concept of 'distributed cognition' - knowledge and learning that is stretched over people, organisations and communities, where the sum is greater than the individual units.

Black (2001) captures the significance of these themes for one's theory of learning very clearly when he says:

Simply to accept that community participation is a focal element for almost all learning, that knowledge does not exist only in the minds of individuals and that learning proceeds by iteration between the individual and the communal is to accept a fundamental change in one's theory of learning. (p. 79)

Implications for Teacher CPD

It is apparent that, to-date, consideration of these ideas and their implications for educational practice have focused primarily on students; as a consequence, their implications for teacher learning and teacher professional development remain less well-developed and not lacking in challenge. In the context that, by definition, all knowledge from a situative perspective is situated, Putnam and Borko (2000) assert that this perspective "...focuses researchers' attention on how various settings for teachers' learning give rise to different kinds of knowing" (p. 6).

Considerations of this kind have spawned debates about where best to locate professional development; as a consequence, a number of models have been tried, tested and evaluated, including:

- Classroom-based, teacher development, in which researchers or staff development team members work with individual teachers in their classrooms to support the introduction of innovative practices and ideas;
- Ongoing, site-based, professional development workshops/meetings, facilitated by researchers or staff members focusing on teachers' personal classroom experiences, sometimes using videotapes as a catalyst for discussion and critique;
- Off-site, summer courses in which teachers participate in activity-based learning that reflect social-constructivist principles. The primary advantage of this model is that teachers are unencumbered by school and classroom-based responsibilities.

Advantages and disadvantages attend to each model. For instance, site-based models increase the likelihood that what teachers learn is of tangible benefit and support in the classroom; however, these approaches present significant challenges also. For example, the issue of scalability arises in the context of the labour and time-intensity associated with this model. In addition, it might be argued that restricting the context of learning to a particular school or classroom setting effectively straight-jackets teachers' learning opportunities by limiting the breadth of their situational experiences and, hence, their learning. Off-site learning, in contrast, while providing potentially valuable opportunities for teachers to expand the contexts in which they learn, interact and explore new ideas, leaves teachers with the challenge of interpreting and integrating the knowledge gained meaningfully into their classroom practice.

In weighing up the alternatives, Putnam and Borko (2000) suggest that the issue of the location of teachers' professional learning should be decided according to the particular goals of the programme:

Research... suggests that the most appropriate staff development site depends on the specific goals for teachers' learning. For example, summer workshops appear to be particularly powerful settings for teachers to develop new relationships with subject matter and new insights about individual students' learning. Experiences situated in the teachers' own classrooms may be better suited to facilitating teachers' enactment of specific instructional practices. And, it may be that a combination of approaches, situated in a variety of contexts, holds the best promise for fostering powerful, multidimensional changes in teachers' thinking and practices. Further research is needed to better understand the complex dynamics of these multifaceted approaches to teacher learning. (p. 7)

Research on Professional Development: A Review from a Situative Perspective

Given the increased research interest in the potential of situated learning to support teachers' learning, Borko's (2004) research review that adopted a situative perspective to interpret what has been learned about professional development programmes and their impact on teacher learning in order to signpost potential approaches to future research, is significant. Her rationale for adopting a situative perspective is particularly noteworthy, as is her use of the metaphor of multi-focal

contact lenses to explain the potential that this perspective affords the researcher to focus on individual and group development simultaneously:

Multifocal contact lenses provide a useful metaphor for considering situative perspectives on knowing and learning. Researchers use the “near-vision prescription” of a psychological framework to focus on the individual teacher. With this prescription, they collect and analyse data on questions such as how a teacher constructs new knowledge and instructional practices. They use the “distance-vision prescription” of a socio-cultural conceptual framework to focus on the professional development community – to collect and analyse data on norms of communication and patterns of participation in professional development activities. The ability to use multiple frameworks at the same time is a key strength of situative research perspectives. (Borko, 2004, p. 8)

The key point being made here is that when a researcher employs a situative perspective, data may be obtained both on the individual teacher in his/her classroom and on the manner in which that teacher participates with colleagues in the wider professional organisation of the school. Reflecting on how researchers have typically engaged the situative perspective over the years, Borko (2004) argues that research initiatives undertaken typically fall into one of three distinct phases that are distinguished from one another on the basis of the nature of the interaction between four key constituent elements of a professional development system: the programme, the teachers/learners, the facilitator and the context. In turn, Borko (2004) argues, as outlined in Table 4 below, that a pattern of increasing sophistication is detectable, both in professional development and research designs.

Table 4.

Research on Professional Development Programmes from a Situative Perspective

Phases	Number of Sites	Number of Facilitators	Focus/foci of the Research
1	1	1	Interrelationships between PDP and teachers (2 foci)
2	>1	>1	Interrelationships between PDP, Facilitators and teachers (3 foci)
3	Multiple	Multiple	Interrelationships between PDP, facilitators, teachers and the context in which they are working (4 foci)

Phase one research, which reportedly accounts for most of the professional development community’s work so far, focused on the nature of the professional

development programme offered, teachers as learners - and the relationship between these two elements - the goal being to establish “an existence proof”, that is, evidence that a professional development programme can impact positively on teacher learning. Borko (2004) notes that, typically, a phase one professional development programme is:

...Relatively small, and the research on them is labor intensive. In most instances, the designers of the professional development programs are also the researchers. Moreover, the participants are typically “motivated volunteers” – teachers who volunteered to participate and were motivated to try out new ideas.... The resulting existence proofs (evidence that the programme can have a positive impact on teacher learning – *my insertion*) unquestionably are an important contribution to the field... (evoking images of the possible – *my insertion*) not only documenting that it can be done, but also laying out at least one detailed example of how it was organized, developed and pursued. (p. 5)

In contrast, at phase three level, the research focus is more expansive; multiple professional development programmes are provided at a variety of sites which affords the researcher the opportunity to explore multiple relationships among and between each element of the professional development programme: facilitator, professional development program, teachers as learners and context.

Borko's (2004) analysis of research findings in phase one led her to conclude that high quality professional development programmes can indeed help teachers deepen their knowledge and transform their teaching. Using a situative lens, she unpacked this overarching finding with reference to four themes: (1) teachers and their learning, (2) processes and activities of professional development, (3) teacher learning and (4) research methodology and the value of multifocal research lens. Given the potential significance of these insights for this thesis, each of these findings is detailed fully here.

Teachers and their Learning: Individual Focus

Having acknowledged that intensive professional development programmes can effect change in teacher knowledge and practices, Borko (2004) isolates three core characteristics of the programmes reviewed. The key findings associated with each are as follows:

- Subject matter knowledge for teaching: programmes that include an explicit focus on subject matter, particularly those that engage teachers as active learners, can help teachers develop learning significantly;
- Understanding of student thinking: professional development can increase teachers' awareness of the role that children's thinking plays in the learning process, and the importance of listening carefully to students in order to build on their understandings and misconceptions, if this is explicitly incorporated;
- Instructional practices: assuming that a key reason for attempting to deepen teachers' knowledge of subject matter and student thinking is to improve classroom teaching, it follows that professional development programmes that incorporate a focus on instructional practices, and encourage teachers to incorporate ideas into their teaching, optimise the opportunity for students to benefit from teacher learning.

Notwithstanding the beneficial impacts of phase one research, this work also confirms that meaningful learning is a slow, arduous and uncertain process. As observed, not only do teachers differ in the extent to which they change through their participation in professional development but, additionally, some elements of teacher knowledge and practice are more resistant to change than others. Significantly, from an A/L perspective:

...It appears to be easier for teachers to incorporate strategies for eliciting students' thinking into their own teaching than to use what they hear from students to make instructional decisions. (Borko, 2004, p. 6)

Processes and Activities of Professional Development: Group Focus

Phase one research also provides evidence that strong professional learning communities can foster teacher learning and instructional improvement. Leveraging a sociocultural conceptual framework, researchers have used the group as the unit of analysis to extend understanding of teachers' learning by examining the nature and extent of their participation in the processes and activities of professional development. It is noteworthy that Borko (2004) chose to focus, within this context, specifically on teacher learning communities, in recognition of their centrality to the

programmes of professional development and research reviewed. Elaborating on the nature of this particular strand of research, Borko (2004) explained that:

Research on teacher learning communities typically explores features of professional development programs such as the establishment and maintenance of communication norms and trust, as well as the collaborative interactions that occur when groups of teachers work together to examine and improve their practice. This research provides evidence that strong professional development communities are important contributors to instructional improvement and school reform.... (p. 6)

With specific reference to the work of Grossman, Wineburg, and Woolworth (2001), Borko (2004) highlighted a number of key components that impact on community formation, including the establishment of a sense of group identity and interactional and relational norms, the development of a sense of shared responsibility for the maintenance and regulation of these norms, and the assumption by group members of a sense of shared responsibility for colleagues' growth and development. It is of note that no direct links were reported between changes in teachers' interactions and student achievement, however.

In keeping with research on individual change and the concepts of teacher efficacy, motivation and attribution introduced previously, research using the group as the unit of analysis suggests that a number of challenges need to be overcome if the teacher learning community is to serve its purpose. Critical amongst these is the need to engage in conversations about teaching that affirm and support teachers' work, while also allowing for debate and constructive criticism, if and when it is required. As argued: "...such conversations must occur... if teachers are to collectively explore ways of improving their teaching and support one another as they work to transform their practice" (Borko, 2004, p. 7). This places a particular onus on the leader of the professional development programme to foster trust, mutual respect and a sense of safety in disclosure within the group.

Teacher Learning and Research Methodology: Dual Focus

Phase one research reveals that records of practice are powerful contexts for facilitating teacher learning and, hence, teacher change. In keeping with a core tenet

of the situative perspective that the contexts in which people learn, and the activities with which they engage, become a fundamental part of what they learn, research indicates that a teacher's classroom offers powerful learning opportunities. Moreover, artifacts - including lesson plans and classroom tasks or video recordings of lessons that open a window on classroom practice - can be used as a substitute in professional development settings to facilitate teachers' co-examination and review of instructional approaches, and their perceived impact on student achievement.

The Value of Multifocal Research Lens: Simultaneous focus

Borko (2004) makes one overarching comment in relation to the issue of what she terms the multifocal research lens - that this approach is indispensable to rich and inclusive research studies of teacher development:

Studies that focus on either the individual or the group as the unit of analysis can provide valuable insights about teacher learning. However, these insights are limited in scope. To explore the connections among professional development activities and processes on the one hand, and individual teachers' knowledge and instructional practices on the other, researchers must use the multiple frameworks and units of analysis the situative perspectives provide and must coordinate them in a manner that leads to a fuller, deeper explanation of teacher development. (p. 8)

In articulating this view, Borko (2004) in effect draws together the research on teacher change models - introduced in the previous section - with the notion of collective group enquiry in the form of a teacher learning community, to emphasise that if teacher CPD is to be effective, it must attend to the idiosyncratic needs of the individual teacher as an independent unit but also as a participant in the broader professional group. In essence, this echoes the arguments presented by Cochran-Smith and Lytle (1999a) advocating the importance of developing teachers' knowledge of practice by exploiting school-based, situated learning opportunities to encourage and sustain teacher-generated conversations about what, why and how they teach.

The opening paragraph of this section of the literature review promised a selective, judicious review of research and literature on teacher CPD, in the context of

evolving theories of teaching and learning. In drawing this section to a close, brief reference is made to two reviews that stand somewhat apart from the work introduced thus far but are deemed nonetheless to be highly relevant to the work at hand: reviews conducted by Timperley and Alton-Lee (2008), and Marcos and Tillema (2006), respectively.

Best Evidence Synthesis Review

In 2008, the publication of a synthesis of evidence from ninety-seven empirical studies identifying “...what kinds of teacher knowledge, and the conditions under which it was developed, promoted the learning processes that enabled teachers to change their practices in ways that had a positive impact on student outcomes” was published (Timperley & Alton-Lee, 2008, p. 341). One of a series of reviews being undertaken as part of the Iterative Best Evidence Syntheses Programme (BES) in New Zealand, the review led to the discovery of what the authors identified as a “second *black box*” between the provision of professional learning opportunities and teacher outcomes.

As indicated in Figure 6 overleaf, the first ‘black box’ was similar to that introduced by Black and Wiliam in the context of their 1998 review of research on AfL that pointed to the gap in knowledge about what happens in the classroom; the second black box highlighted by Timperley and Alton-Lee (2008) refers to the gaps in knowledge about how students exploit the learning opportunities that present as a result of changes in teaching practices motivated by teacher CPD. In highlighting the impact on students, Timperley and Alton-Lee (2008) bring the issues raised previously in this section full circle, to refocus attention on the original argument articulated by Hattie (2005), that the only true measure of CPD is student achievement.

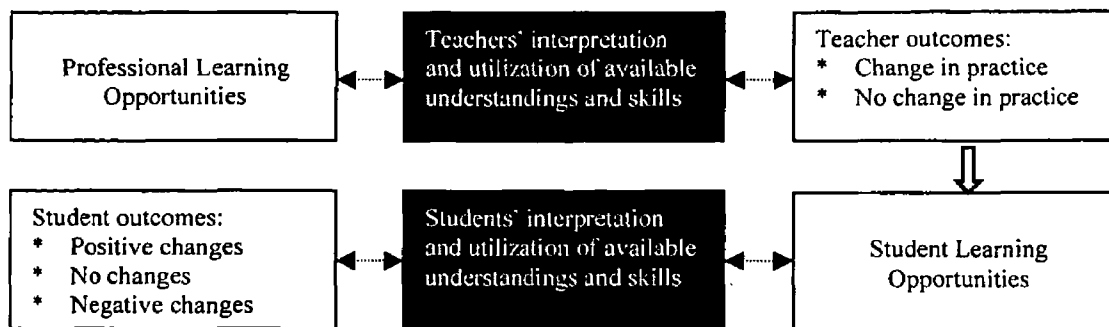


Figure 6.

The Black Boxes of Teacher and Student Learning (Timperley & Alton-Lee, 2008)

While a full report of the findings of this research is beyond the scope of this work, a number of key conclusions are noteworthy. First, as reported, the least effective forms of professional development (measured in terms of student achievement) fall at two extremes of a continuum, ranging from teachers being given too much, to too little, latitude in designing and controlling their own learning. As stated, there is little benefit to be derived from treating teachers as:

...Self-regulating professionals who, if given sufficient time and resources, are able to construct their own learning experiences and develop a more effective reality for their students through their collective expertise... because they typically did not develop teachers' current knowledge and practice and challenge problematic attitudes. (Timperley & Alton-Lee, 2008, p. 347)

At the other extreme, the idea of “outside experts”, developing and presenting “...recipies for teaching (typically based on research about what kinds of pedagogies work for improving student outcomes)” - although appearing initially to be more effective - produce benefits that are either “...short-lived... or relatively limited compared to other kinds of professional development” (Timperley & Alton-Lee, 2008, p. 348).

Among the approaches identified as having impacted positively on student achievement, two are particularly noteworthy. First, where efforts were made to support classroom-based decisions about teaching and learning by developing teachers' foundational knowledge of pedagogy and assessment, broad substantive gains were reported across children's learning. Second, studies that achieved the highest effect sizes were those that employed a variety of approaches including

coaching, workshops and learning communities to assist teachers in translating the professional knowledge gained into changes in their teaching. It might be argued that the strength of this second finding is tempered by the authors' statement that they "...inferred the learning processes from the descriptions provided because they were rarely addressed or described explicitly (in the studies – *my insertion*)..." (Timperley & Alton-Lee, 2008, p. 351).

Reflecting on the findings of the synthesis of evidence, the authors seem to suggest that it is the process - as distinct from the content - aspects of professional development that require rethinking:

...It is not professional development per se that is the problem; rather, it is the way it is typically undertaken. Provision needs to recognise the complexity of professional practice and bring capable and effective expertise to supporting teachers.... Such expertise needs to engage rather than bypass teachers' theories and provide and exemplify alternative visions and practices. Teachers need to have a problem to solve, to have multiple opportunities to learn relevant pedagogical content and assessment knowledge in ways that integrate theory and practice, and to maintain a constant focus on how teaching affects students. Better outcomes for students are sustained when the organizational conditions support ongoing evidence-informed inquiry into the impact of practice on students. Take any of the ingredients out, and its impact is likely to diminish. (Timperley & Alton-Lee, 2008, p. 358-359)

There is growing recognition of the dynamic interaction between curriculum, assessment, pedagogy and learning, leading to a concomitant focus on the crucial role of the teacher (Carless, 2005). As Atkin and Black (2003) note: "...what really counts in education is what happens when teachers and students meet. The wisdom of any decision about education is best judged on the basis of whether or not it raises the quality of these interactions" (p. xi).

One of the difficulties in making judgements of this kind, however, is the gap in understanding and knowledge about what happens in classrooms, as noted in the previous chapter, which is at least partially attributable to the absence of assessment tools to measure the quality of teaching and learning, particularly at scale:

One reason that the quality of instruction has remained a "black box" in many accountability systems and large-scale evaluation designs is because few

assessment tools exist that directly measure the quality of classroom practice on a broad scale. Teacher surveys frequently have been used to indirectly assess the quality of students' learning environments, though this method has limitations as far as accurately describing the interactions between teachers and students, as well as teachers' translations of reform policies (including content standards) into everyday classroom practice.... Likewise, analyses of student work have provided some information about student performance, but have not drawn attention to the opportunities students have in the classroom to produce high-quality work. Classroom observations have been the most direct way to measure instructional quality, but these can be time consuming and expensive to conduct.... New indicators that help ... monitor and support efforts to improve the quality of instruction are clearly needed. (Matsumura & Pascal, 2003, p. 3)

The other major issue relates to what has been termed "the bifurcation of teaching" - a term introduced by Marcos and Tillema (2006) in the context of the final piece of literature on CPD to be reviewed in this section.

Assessing the Impact of Teacher CPD: The Need for Evaluation Tools

In the context of a meta-review of research on teacher reflection and action, Marcos and Tillema (2006) uncovered what they described as a "*blind spot*" in the empirical work reviewed. As reported:

...Studies gather knowledge on constituent elements of teaching but have difficulty in positioning the process as a whole (for example, by treating aspects of teaching as separate objects of study). Studies can range from focusing on levels of ideas to identifying prior beliefs, and from changes in beliefs, modes of reflecting, attitudes toward teaching to personal identity and self-efficacy. But only in rare cases do we find an interrelated study of teacher activity such as teacher planning, teacher intentions to act, teacher action potentials, or activity in classroom performance, or even interaction with pupil learning. Teaching as a process, therefore, is not investigated as an interrelated whole comprised of many functional relationships between thinking and action. By studying only particular aspects, no matter how important each may be, these studies fragment teacher activity, and portray isolated understandings...that can only tell 'half the story'.... This segmentation ultimately signifies an imbalance between the processes studied and the level of grounded knowledge provided through research.... There seems to be a tendency to study the thinking aspect (beliefs and reflections) rather than what teachers actually do (action). As a result, we have a '*blind spot*' in our research efforts. (pp. 113-114)

To highlight this blind spot, Marcos and Tillema (2006) developed an analytic framework to demonstrate how studies on teacher reflection and action typically fall into one of four separate domains: talking the talk; talking the walk; walking the talk and walking the walk; Table 5 offers a distillation of the foci, underlying assumptions and research methods within each domain. Two key findings emerged from this work: first, as one progresses from the domain of talking the talk (which focuses on teachers' beliefs and reflections on action - representational processes), to walking the walk (which focuses on professional performance - presentational processes), the number of empirical studies conducted declines significantly. Second, as currently undertaken, research in this field bifurcates our understanding of teaching.

In this context, the authors concluded:

Instruments to collect data must be designed so that they can fill the gaps between the domains (between reflective thought and action)...that would... at least in part, enable overlapping or linking of the data to develop a full account of beliefs, plans, actions and reflections. (Marcos & Tillema, p. 125)

As will become apparent in Chapter 3, this was a challenge that presented in this study in the context of the need to find research instruments that would capture teachers' needs and concerns in order to offer timely and appropriate professional support, on the one hand, and an accurate record of developments in teachers' classroom practices, on the other.

Table 5.

Teacher Reflection and Action: An Analytic Framework

Domain	Research Focus & Underlying Assumptions	Methodology
Talking the Talk	<p>The teacher's voice: descriptive studies investigating how teachers interpret their work in terms of an existing knowledge base or codified system</p> <p><u>Underlying assumption:</u> Teachers' actions are determined by teacher understandings. Therefore, these studies explain teachers' beliefs and thoughts because they function as filters for making sense of the knowledge and experiences they will encounter.</p>	<p>Questioning:</p> <ul style="list-style-type: none"> - Interviewing - Questionnaires - Checklist surveys
Talking the Walk	<p>Studies on reported action; investigating retrospective accounts of actions to interpret what was done.</p> <p><u>Underlying assumptions:</u> Practitioners can identify and build on their existing understanding to produce new knowledge and make generalizations from particular experiences that will assist them to apply to new practices. Key features of this research: to capture teachers' knowledge construction which holds invaluable promise for developing new understandings and producing new knowledge about teaching and learning.</p>	<p>Narrative inquiry:</p> <ul style="list-style-type: none"> - An approach especially suited to constructing knowledge from actions with the understanding that the knowledge base for teaching resides in the stories or experience of the teacher
Walking the Talk	<p>Studies on teachers' thinking, investigating teachers' plans and intentions in light of their backgrounds and beliefs</p> <p><u>Underlying assumption:</u> The key to success is a purposeful, goal-directed strategy that increases the likelihood of the desired outcome; alternatively described as planned action theory, prospective reflection, anticipatory action research. Expectation: a comparison between the plans and performance to reveal their fit or alignment with the intended outcome.</p>	<p>Written documents/collections of intentions:</p> <ul style="list-style-type: none"> - Notes, plans, diagrams - Diaries, reflective journals, CPD plans
Walking the Walk	<p>Studies on observed action, investigating, in-depth, how action exemplifies teacher knowledge</p> <p><u>Underlying assumption:</u> Situated cognition and distributed practice more adequately represent teacher learning than do deliberate thought and reflection. Action is always performed under local contextual guides, which modify every previous plan. Hence, the renewed interest in collaborative learning and distributed action in teams and teacher learning communities – a low-cost, sustainable, satisfying, and potentially transformative form of teacher professional development.</p>	<p>Participant observation:</p> <ul style="list-style-type: none"> -Field notes -Video recordings of classroom practice and of TLC meetings.

Summary of Literature Review: Part 1

In the context of evolving paradigms of learning, the literature reviewed to this point served to foreground important factors mediating the potential of CPD to effect positive change in teachers' knowledge, skills and attitudes. These need to be borne in mind as the discussion progresses in the next section to focus on the potential of Assessment *for* Learning, as a core pedagogical approach, to improve significantly children's learning experiences and achievements. Three key themes have emerged that are particularly noteworthy.

First, teacher change is not a linear process; it a highly individual, idiosyncratic, iterative process that demands routine reflection and enactment (Clarke & Hollingsworth, 2002) on one's personal beliefs, attitudes and motivations (Smith & Gillespie, 2007), the success of which is mediated in large part by the nature and extent of the support received at critical stages in the process (Hall & Ford, 1987; Mc Kinney et al., 1999; van den Berg et al., 2000). Of importance also, is the degree of congruence between teachers' existing beliefs about teaching and learning and the value attributed to the proposed idea or intervention. The motivation to experiment with new ideas also presupposes an element of dissonance with existing practices that acts as a spur to enquiry. This translates into the need for teachers to experience both support and pressure during the change process (Wiliam, 2007).

Second, teacher professional development is frustrated or progressed by the context, processes and content of the programmes offered (Guskey & Sparkes, 1996), which reflect providers' underlying learning theories, concepts of lifelong learning and their appreciation of the critical importance of social-cognition and situated learning to develop teachers' knowledge of practice, distributed across colleagues and schools (Cochran-Smith & Lytle, 1999a; Guskey, 2002; Putnam & Borko, 2000; Shepard, 2000).

Finally, given the relatively small body of research reporting the impact of CPD on both teachers' instructional practices and student' achievement (Darling-Hammond, 2000), research is required that prioritises both sets of outcomes and their interrelationships (Loucks-Horsley & Matsumoto, 1999), without bifurcating the

integrity of the teaching-learning process (Marcos & Tillema, 2006). Existing research (Borko, 2004; Liebermann & Pointer-Mace, 2008) suggests that this might best be achieved by adopting a situative perspective to provide sustained, site-based learning opportunities for teachers, through the medium of a teacher learning community, and a multifocal lens of inquiry that supports unit and group analyses.

These themes are developed further in the following section which addresses the challenges presenting for researchers, teachers and children alike as they attempt to embrace formative assessment.

Literature Review: Part 2 – Assessment *for* Learning

It was proposed recently that just as “... the advent of formal examinations in nineteenth-century Europe arguably represented a major victory for social justice, so the advent of a focus on assessment for learning in the twenty-first century could be similarly significant for social reform...” (Broadfoot & Black, 2004, p. 22). This second section of the chapter examines the potential of formative assessment to radically change the theory and practice of teaching and learning. Prefaced by a brief commentary on the evolution of assessment terms, the section builds on the insights gained previously in relation to social-constructivist theory to emphasise the influence of enduring, and often conflicting beliefs, about the nature and functions of testing and assessment. This lays the groundwork for a distillation of the findings of a series of prominent literature reviews, including the seminal work of Black and Wiliam (1998a), in order to (a) identify key principles that emerged from interpretations of these findings and (b) examine the content and process challenges that arose when attempts were made to give practical, classroom definition to these principles. Particular reference is made in this context to the current work of the Keeping Learning On Track (KLOT) Project in the US (Wiliam & Thompson, 2006). The chapter concludes by drawing together key messages from sections one and two within a conceptual framework that links the research questions with the study design and the interpretive and evaluation frameworks that guided the study, both of which are detailed in the next chapter.

Towards a Definition of Terms

As noted in the introductory chapter, the term *Assessment for Learning* is used in this study to refer to:

...Any assessment for which the first priority in its design and practice is to serve the purpose of promoting pupils' learning. It thus differs from assessment designed primarily to serve the purposes of accountability, or of ranking, or of certifying competence. An assessment activity can help learning if it provides information to be used as feedback, by teachers, and by their pupils in assessing themselves and each other, to modify the teaching and learning activities in which they are engaged. Such assessment becomes 'formative assessment' when the evidence is actually used to adapt the teaching work to meet learning needs. (Black et al., 2002. p. 1)

As reviews of the history of the growth of formative assessment (used interchangeably with the term *A/L* in this study) attest, definitions of the term have progressed through a number of iterations (Allay & Lopez, 2005; Baroudi, 2007; Black & Wiliam, 2003). While the coining of the term "formative evaluation" is attributed to Michael Scriven, who used the term in the context of the "...ongoing improvement of the curriculum..." (1967, p. 41), Bloom, Hasting and Madaus (1971) introduced the distinction, as understood today, between formative and summative evaluation, describing formative evaluation as "...the use of systematic evaluation in the process of curriculum construction, teaching, and learning for the purpose of improving any of these three processes..." (p. 155). This contrasted with their definition of summative evaluation as "...the type of evaluation used at the end of a term, course, or program for purposes of grading, certification, evaluation of progress, or research on the effectiveness of a curriculum, course of study, or educational plan" (Bloom et al., 1971, p. 155). In considering the evolution of these terms, Black and Wiliam (2003) made the crucial point that "... from their earliest use it was clear that the terms 'formative' and 'summative' applied not to the assessments themselves, but to the functions they served" (p. 623). Over the years, several researchers have revisited this distinction; Harlen (2006a, p.114), for example, has offered a categorization of a possible continuum between formative and summative assessment, as outlined in Table 6 overleaf.

Table 6.

Summative and Formative Assessment (Harlen, 2006a)

	<div style="display: flex; align-items: center; justify-content: space-between;"> Formative ←————→ Summative </div>			
	Informal formative	Formal formative	Informal summative	Formal summative
Major focus	What are the next steps in learning?		What has been achieved to date?	
Purpose	To inform next steps in learning	To inform next steps in teaching	To monitor progress against plans	To record achievements of individuals
How is evidence collected?	As normal part of class work	Introduced into normal class work	Introduced into normal class work	Separate task or test
Basis of judgement	Student referenced	Student and criterion referenced	Criterion referenced	Criterion referenced
Judged by	Student and teacher	Teacher	Teacher	Teacher or external marker
Action taken	Feedback to students and teacher	Feedback into teaching plans	Feedback into teaching plans	Report to student, parent, other teachers etc.
Epithet	Assessment for learning	Matching	Dip stick	Assessment of learning

Arguing that there is no sharp distinction between the two forms of assessment, she suggests that the practices and uses at both extremes (to the left and right of the arrow) typify assessment *for* learning and assessment *of* learning, respectively:

At the purely formative end is assessment that is integral to student-teacher interaction and is also part of the student's role. The teacher and the student consider work in relation to the goals that are appropriate for the particular learner and so the judgements are essentially student-referenced. The central purpose is to enable teachers and students to identify the next steps in learning and to know how to take these. At the purely summative end of the dimension the purpose is to give an account of what has been achieved at certain points. For this purpose, the assessment should result in a dependable report on the achievements of each individual student. (Harlen, 2006a, p. 113)

Hence, Wiliam's (2007) reiteration that AfL is "...any assessment for which the first priority in its design and practice is to serve the purpose of promoting pupils' learning...." (p. 1).

Notwithstanding advances in understanding of the relationship between formative and summative assessment, recent meetings of the American Educational

Research Association (AERA) have witnessed repeated calls for greater clarity in relation to the range of assessment terms being employed generally in the literature (Marion & Pellegrino, 2006; Noonan & Renihan, 2006; Shepard, 2005). At its most basic, what is at issue here is the manner in which, for example, the terms 'assessment *for* learning' and 'formative assessment' are used in the UK, whereas the phrase 'classroom assessment' is typically the term of choice in the US. However, at a more fundamental level is the issue raised by Shepard in her presidential address to the AERA in 2000 in which she argued that "...any attempt to change the form and purpose of classroom assessment to make it more fundamentally a part of the learning process must acknowledge the power of... enduring and hidden beliefs..." (p. 6) about traditional views of testing, teaching and learning. Given the centrality of these issues for an understanding of research on assessment, the key tenets of Shepard's (2000) argument are introduced at this point.

From Testing to Assessment

Shepard (2000) was perhaps the first researcher to draw out the implications for classroom assessment of the shift from the dominant 20th century paradigm of social efficiency curriculum, behaviourist learning theories and scientific measurement to the emergent paradigm, based on social-constructivism and sociocultural theories. Articulating her core argument, Shepard (2000) proposed that:

To be compatible with and to support the social-constructivist model of teaching and learning, classroom assessment must change in two fundamentally important ways. First, its form and content must be changed to better represent important thinking and problem solving skills in each of the disciplines. Second, the way that assessment is used in classrooms and how it is regarded by teachers and students must change. Furthermore, to enable this latter set of changes within classrooms... teachers need help in fending off the distorting and demotivating effects of external assessment. (p. 7)

Shepard (2000) substantiated these remarks with reference to a detailed exposition of the evolution of ideas about the role of the school and education, and the nature of learning and assessment. Chronicling the rejection of the once dominant 20th century paradigm of social efficiency, fixed ideas of intelligence and scientific measurement in favour of an increasing appreciation and advocacy of learning as socially constructed in classrooms where all children are deemed to have the potential

to learn and improve, Shepard (2000) suggested that education, and assessment more particularly, is currently at a crossroads. The current regime embraces ideas of situated, social-constructivism and all that this implies in terms of teaching and learning, yet it has failed to break cleanly from the vestiges of traditional notions of testing, rooted in beliefs about the objectivity of standardised achievement tests and behaviourist learning theories. This manifests most acutely in the conflicts between the increasing promotion of formative classroom assessment in tandem with the upsurge in performativity and externally mandated summative, high-stakes assessment, particularly in the US and the UK.

As originally presented, Shepard (2000) represented these ideas diagrammatically as a series of interlocking relationships within paradigms in relation to beliefs about the role of school, learning theories and measurement/assessment of learning. Table 7 overleaf offers a distillation of these ideas; the greyed boundaries within paradigms are intended to signal the original interconnectivity of the themes; the two horizontal black arrows in the middle of the Table emphasise the current dissonance in beliefs and understanding about the nature of teaching and learning, on the one hand, and assessment/measurement, on the other, with the vertical arrows signalling the paradigm shift which is currently taking place.

Table 7.

The Incompatibility of New Views of Instruction and Traditional Views of Testing

Emergent Paradigm (circa 1990s-2000+)	
<i>Reform Vision of Curriculum</i>	
All students can learn; Challenging subject matter aimed at higher order thinking & problem solving; Equal opportunity for diverse learners; Socialisation into the discourses & practices of academic disciplines; Authenticity in the relationship between learning in and out of school; Fostering of important dispositions and habits of mind; Enactment of democratic practices in a caring community.	
<i>Cognitive & Constructivist Learning Theories</i>	<i>Classroom Assessment</i>
Intellectual abilities are socially & culturally developed; Learners construct knowledge & understandings within a social context; New learning is shaped by prior knowledge & cultural perspectives; Intelligent thought involves "metacognition" or self-monitoring of learning & thinking; Deep understanding is principled & supports transfer; Cognitive performance depends on dispositions & personal identity.	Challenging tasks to elicit higher order thinking; Addresses learning processes as well as learning outcomes; An on-going process, integrated with instruction; Used formatively in support of student learning; Expectations visible to students; Students active in evaluating their own work; Used to evaluate teaching as well as student learning.
Dissolution of Old Paradigm: New Views of Instruction/Old Views of Testing (circa 1980s-2000+)	
<i>Instruction</i>	<i>Traditional Testing</i>
20th Century Dominant Paradigm (circa 1900-2000+)	
<i>The Curriculum of Social Efficiency:</i>	
Scientific management of schools like factories; Carefully specified educational objectives based on job analysis; Utilitarian content, antagonism toward academic content except for elite few; Science of exact measurement, precise standards; Differentiated curriculum based on predicted social roles.	
<i>Hereditarian Theory of Intellect:</i>	<i>Scientific Measurement</i>
IQ as innate, unitary, and fixed.	IQ tests to sort pupils by ability; Objective tests to measure achievement.
<i>Associationist & Behaviourist Learning Theories</i>	
Concepts of mind replaced by stimulus-response associations; Accumulation of atomistic bits of knowledge; Learning tightly sequenced & hierarchical; Limited transfer, each objective taught explicitly; Test-teach-test to ensure learning; Tests isomorphic with learning; Motivation based on positive reinforcement of many small steps.	

The import of Shepard's (2000) message is crucial to the development of AfL, particularly in countries which continue to measure student success by summative test scores, and, increasingly, teachers' effectiveness by the numbers of students receiving high grades. As expressed recently with reference to the current situation in the US:

...The expectation of high reliability and objectivity in the assessment of students' learning within a culture of accountability and litigation when things go wrong, has tended to deflect policy developments from any consideration of improving learning through assessment. (William, 2006b, p. 169)

It is not surprising then, that, more recently, James (2006a) has revisited the challenge of trying to develop consistency between assessment practices and beliefs about learning by comparing and contrasting behaviourist, cognitive-constructivist and socio-cultural/situated, and active theories of learning, in terms of their manifestation in classroom practice. She concluded that, their differences and the challenge of realignment notwithstanding, "...the possibility for a more complete and inclusive theory of learning to guide the practice of teaching and assessment seems a goal worth pursuing" (James, 2006a, p. 60). Indeed, like Shepard (2000) before her, she argued that this seems inevitable given that change in assessment practices or beliefs about learning "...almost always requires a change in the other" (James, 2006a). She concluded, however, that the trajectory is not necessarily clear:

...Assessment practice is sometimes out of step with developments in learning theory and can undermine effective teaching and learning because its washback effect is so powerful, especially in high stakes testing. (p. 58)

Taking another perspective on this theme, Black and William (2006a) cautioned that:

The beliefs of teachers about learning, about their roles as assessors and about the 'abilities' and prospects of their students, will affect their interpretations of their students' learning and will thereby determine the quality of their formative assessment. (p. 23)

It is against this background that the rise in interest in formative assessment is now considered.

Wininger and Norman (2005) record that "...although the term and the processes associated with formative assessment have been around for decades, theories, research, and practice of formative assessment did not surface widely until the late 1980s" (p. 21). Sadler's (1989) introduction of a theory of formative assessment, based on the core principle of student feedback to facilitate student self-assessment, was one of the first to be articulated. In turn, he employed Ramaprasad's (1983) earlier definition of feedback as "...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" (p. 4). In order to avoid the situation of "dangling data" - meaning information on attainment that has not been harnessed to support next steps in learning - Sadler (1989) emphasized the need for students to receive explicit information regarding the standards by which work would be judged (subsequently termed – success criteria):

...The learner has to (a) possess a concept of the standard (or goal or reference level) being aimed for, (b) compare the actual (or current) level of performance with the standard, and (c) engage in appropriate action which leads to some closure of the gap. (p. 121)

In tracing the rise and fall of formative assessment from the late 1970s in the UK, Black and Wiliam (2003) emphasised that the difference in assessment functions advanced attempts to develop a range of tools to support teachers in making the required changes in classroom practice and, more particularly, to aid them in attempting to align both approaches within some 'overall system' of assessment. However, despite a series of interventions at second level and the introduction of a national curriculum for all schools in England and Wales, they reported that "...throughout the 1990s, the debate continued about how evidence from the external tests and teachers' judgments could be combined... with increasing concern that the potential of assessment to support learning was being ignored" (Black & Wiliam, 2003, pp. 623-624). In turn, this concern led to the commissioning of what is now regarded as a seminal, watershed review (Black & Wiliam, 1998a), that drew public attention to the weight of evidence attesting to the power of the effective use of formative assessment to improve learning.

A review by Broadfoot and Black (2004) opened with the lines: "The field of assessment is extensive. It is therefore necessary to be selective" (p. 9). Mindful of this advice, the early literature on formative assessment reviewed in the context of this work is presented here in tabular form initially, in order to highlight core research studies, their primary foci and key findings. The rationale for adopting this approach was two-fold: first, this is an area of research that has been extensively reviewed and reported on in recent years; hence, another iteration was deemed unlikely to add significantly to the body of knowledge; second, and more importantly, the field of assessment is extensive - and notwithstanding the current reviews - getting to grips with this body of literature can be daunting. In this context, it was determined that a visual map of a representative sample of the work undertaken - highlighting key findings - might prove a more useful addition to the field, particularly if the findings were integrated subsequently into a discussion of their practical implications for teaching and learning.

Table 8. A Representative Sample of Research on AfL Pre-Dating 1998

Reviews	Key Foci	Findings
1986: Fuchs & Fuchs	Meta-analysis of experimental/quasi-experimental design studies using formative assessment 2-5 times per week. Key focus on students with SEN, with non-SEN also included.	Effect size of .70 for SEN group; non-SEN group .63. Impact of formative assessment is greater when used within a formally structured programme (.92) rather than in an unstructured way (.42).
1991: Bangert-Drowns et al.	Meta-analysis using 58 effect sizes from 40 reports on feedback, 42 from college-level samples; studies included were undertaken between 1960 and 1990, each consisting of two groups receiving identical instruction.	Overall positive effect sizes (37 out of 58 effect sizes): student achievement raised from 50 th to 59 th PC on standardised tests; 18 out of 58 effect sizes were negative. 'Right/wrong' only feedback impacts negatively on student achievement.
1996: Kluger & DeNisi	Meta-analysis of experimental & quasi-experimental interventions, across academic & non-academic settings, with at least 10 participants. Focused on a broad range of feedback strategies: e.g. grades, written and verbal.	Across all types of formative assessment - moderate effect size of 4.0; in 35% of studies a negative effect was reported. Explanation: the greater the emphasis on individual performance (extrinsic motivation) over learning effort (internal motivation), the worse the impact.
1987: Natriello	Meta-analysis of a range of assessment purposes: certification, selection, direction & motivation within an assessment cycle model.	Research into the effects of evaluation processes conflate key distinctions, e.g. the difference between the quality & quantity of feedback, hence irrelevancy.
1988: Crooks	Review of research in 9 specific areas on the impact of evaluation practices on student learning activities & achievement.	Summative, grading function of assessment is over-emphasised; students require clear learning outcomes, specific, constructive and regular feedback, a strong sense of active involvement & opportunity to set/achieve specific learning goals.
1988: Butler	48, 11-year old Israeli students, half in the top quartile of the class, the other the bottom quartile. Students were given written tasks followed by (a) comments-only feedback, (b) numerical grades only & (c) a combination.	Comments-only led to significant learning gains with both high & low-achieving students - grades & a combination of strategies resulting in reduced achievement. For high achievers, any combination worked.
1991: Bergan et al.	An intervention involving 838, 5-year old kindergarten children, from mainly disadvantaged backgrounds, in 6 different regions of the U.S. Motivated by a belief in the potential of early intervention, focused on the acquisition of basic skills (3-R focus), intervention group teachers received specialist training in formative assessment.	Significant gains in reading, maths & science among the intervention group, though the use of multiple-choice tests jarred with the otherwise child-centred approach employed; special needs placements for the control group were 1 in 3.7 children, with 1 in 5 placed in special education compared with 1 in 17 and 1 in 71, respectively, for the intervention group.
1992: Martinez & Martinez	120 American college students in an introductory algebra course taught by an experienced, highly rated, teacher.	Significant gains for college students tested more frequently & taught by a more experienced teacher; however, the 'exceptionality' of the teacher raised issues re. findings.
1994: Fontana & Fernandes	354 Portuguese, 8-14 year old students taught by 25 maths teachers with specific training in self-assessment methodology & exposed to learning objectives & success criteria.	Mean gain of experimental group was twice that of the control group of younger students; results were less clear for older students.
1995: Whiting et al.	Effects of mastery learning on approx. 7,000 students over an 18-year period taught by one teacher - regular testing & feedback to all with a 90% success requirement to proceed.	Grade point averages of teacher's students were consistently high, & higher, than those of students taught by other teachers.
1996: Schunk	44, 9 and 10-year old students in the U.S. divided into two groups, one stressing learning goals - how to solve problems - and frequent self-evaluation, the other performance goals - how to solve them only.	Using outcome measures of skill, motivation & self-efficacy - effect of self-evaluation outweighed the differential effect of the two types of goals, in favour of the learning goals group.
1997: Frederiksen & White	An inquiry-based, middle school science curriculum module focused on the practical inquiry approach to learning. 12 classes of 30 students, each in two schools, were exposed to the same curriculum plan with all students working in peer groups. Control group engaged in general discussion; intervention group spent the same time period in discussion structured to promote reflective assessment.	Significant gains by the experimental over the control group; differential gains between students of "low" & "high" ability - students with "low" ability students within the intervention group outperforming their peers in the control group by more than 3 SDs, the 'medium ability' students by just over 2SDs and the 'high ability' by just over 1SD.

Table 8 provides an overview of a range of research studies in the field of assessment spanning twenty years, a sample of those included in the Black and Wiliam review (1998a). The breadth and scope of the work ranges from meta-analyses (e.g. Fuchs & Fuchs, 1986), to large-scale reviews (e.g. Crooks, 1988) to individual experimental, or quasi-experimental, intervention studies (e.g. Schunk, 1996). The selection of these particular studies, and the presentation of them here in tabular format, provides an opportunity to stress a number of key points that are of particular relevance to this study.

First, despite the range of studies included in Black and Wiliam's (1998a) extensive review, and with the notable exception of the work of Fuchs and Fuchs (1986), and Bergan, Sladeczek, Schwarz, and Smith (1991), there is a serious under-representation of studies with an explicit, if not exclusive focus, on young students from disadvantaged communities with, or without, special educational needs. This is surprising given the level of investment over many decades - particularly in the US where many of the review studies were undertaken - in early years education (e.g. Head Start), and in programmes targeting educational disadvantage (e.g. Success for All). Second, it is notable that despite the fact that the two studies mentioned above reported significant quantitative evidence of learning gains for children in the intervention groups over their peers in control groups, follow up studies were not conducted. Moreover, across studies of potential interest, where positive effect sizes were reported, no correlations were drawn between the specific formative approaches used and the gains achieved; as a consequence, the potential to build on these findings in further research was weakened. Finally, despite the positive effect sizes derived from studies on 'mainstream' populations that differentiated between 'high' and 'low' ability groups - what Black and Wiliam (1998a) termed the traditional 'low-achievers' - these studies did not give rise to subsequent studies with a specific focus on low achievers either.

As a consequence, in turning now to consider the review conducted by Black and Wiliam (1998a), attention is focused on the composite findings of the eclectic range of studies included, mindful that none of the studies were conducted in an Irish setting, with very few addressing the issue of educational disadvantage specifically.

Driven by a need to establish an evidential basis for advocating the use of Af/L, coupled with an interest in discovering the nature and extent of existing formative practice in classrooms that could potentially guide further dissemination, an extensive review was undertaken by colleagues from King's College, London, in 1998. The Black and Wiliam (1998a) publication extended two previous reviews by Natriello (1987) and Crooks (1988), linking their findings with those of approximately 250 other research sources, culled from an initial yield of approximately 580 articles or book chapters. As indicated in Table 8, in contrast to the work of other researchers (e.g. Bangert-Drowns, Kulik, Kulik, & Morgan, 1991; Fuchs & Fuchs, 1986; Johnson & Johnson, 1990; Kluger & DeNisi, 1996), Black and Wiliam (1998a) undertook a review of the literature organised around six core themes, rather than attempting a meta-analysis. They rationalised that such was the diversity and range of formative assessment practices reported in the literature review - coupled with the contrasting underlying assumptions of the quantitative studies regarding learning theory and the diversity of the participant cohorts - that a meta-analysis was impractical, a view supported by the work of Brookhart (2006; see Appendix A). Reflecting later on this decision, Black and Wiliam (2003) argued that the pre-eminent focus in the review on issues including pupils' self- and peer-assessment and the role of feedback as a pedagogical learning tool, had the positive effect of taking the emphasis in the formative assessment studies reviewed "...away from *systems*..." and "...the formative-summative interface..." (p. 628) to relocate it, more gainfully, within classroom practice.

The Six Sub-Sections of the Review

Black and Wiliam's (1998a) report was presented in six sections. Section one offered a survey of the evidence from approximately thirty experimental or quasi-experimental studies, spanning age groups ranging from 5-year olds to university undergraduates, in different school subjects, across the globe. In each case, significant effect size gains were reported for intervention groups who had been deliberately exposed to teaching practices that incorporated formative assessment methodologies. As reported, the potential of the ideas reviewed to impact positively

on student learning was unequivocal; the research established conclusively "...gains in achievement ... among the highest ever reported for educational interventions...", with an average mean effect size for most studies of between 0.4 and 0.7" (Black & Wiliam, 1998a, p. 61).

The second section reported on existing assessment practices in classrooms which were described as being "depressing" and "weak" (Black & Wiliam, 2006a, p. 11). The difficulties that emerged ranged from teachers' inadvertent reinforcement of rote and superficial learning - as a consequence of inappropriate questioning and feedback styles - to more fundamental problems in relation to the overly bureaucratic, teacher-dominated, nature of teacher-pupil engagement. This led the researchers to conclude that fundamental rather than superficial changes were needed both in relation to teachers' perceptions of their own roles with regard to students and in classroom practice more generally. Section three focused on the role of students in formative assessment including associations between students' beliefs about learning, what it means to learn, their readiness to take risks and to engage in classroom tasks, linked, as appropriate, to students' motivation, attribution and self-efficacy. This led to consideration of proactive strategies that might be used by students to afford them greater control over their learning, including study methods and skills, peer- and self-assessment.

The fourth section sought to identify teaching and learning strategies that might be of practical benefit to teachers in engaging formative assessment methods. Key consideration was given to the nature and choice of classroom tasks, classroom discourse and teachers' use of open questions to maximise student engagement. The focus shifted in the fifth section to examine research that shed light on comprehensive systems of teaching and learning, such as mastery learning, in which formative assessment is a significant component part. The final section zoned in on the centrality of feedback as indicated by a number of the research studies reviewed, notably the work of Kluger and DeNisi (1996), who distinguished between feedback that positively influenced student achievement from that which had the potential to do the reverse (see Table 8). Developing Ramaprasad's (1983) concept of 'closing the gap', and Sadler's (1989) thesis on the 'reference level' of performance, it was argued that learners must understand their current standard of learning, and the reference

point or goal of their learning, in order to close the gap between the two. Closely associated with these ideas were research messages about attribution theory and, particularly, the role of the teacher in inculcating in students an understanding that achievement is attributable to effort rather than luck or fixed intelligence.

Significant Findings

One of the key findings to emerge from this review was that, as a field of research, the area was not well established. It was argued that such was the diversity of approaches reported in the studies reviewed - both in terms of the research designs conceived and the intervention approaches employed - that, collectively, the body of work did not give rise to a coherent programme of formative assessment. For example, the fact that the previous reviews by Fuchs and Fuchs (1986), Natriello (1987), and Crooks (1988) did not refer to one another at all, and rarely overlapped in the citations listed, was indicative of an overall pattern. While the authors argued that the eclecticism of the research, in effect, strengthened the overall conclusion that the constructive use of formative assessment yields significant learning gains (irrespective of the nuances of implementation), they readily acknowledged that this did not resolve the issue of the absence of an underlying theory of AfL that could be used to guide practice. Interesting, some time later, when reflecting on the patchy effects of research on practice, Black (2000) likened researchers to:

...A guerrilla force, engaged in more or less successful attacks on the outposts of the occupying army of public test practice, but rarely able either to take control of any important area or to secure a significant share in the region's administration. (p. 416)

Speculating on why this might be the case, Black (2000) noted that there is little tradition of researchers working collaboratively to formulate coherent programmes aimed at influencing policy, a practice that might be improved if the number of research centers to support such work were to increase.

Of more relevance to this study, perhaps, were two other problems identified by the researchers as requiring particular attention. The first arose from the fact that, although the research indicated clearly that formative assessment can have a

particularly powerful effect on disadvantaged and low-attaining students, this evidence was apparently not supported in the results of other studies, leading Black and Wiliam (1998a) to surmise that:

The apparent contradictions here probably arise because there are some important features of the classrooms that have yet to be recorded and understood. If it is true that the ranges of school achievement might be narrowed by the enhancement of the achievement of those hitherto seen as slow learners, then there are very strong social and educational reasons for giving high priority to sensitive research and development work to see how to understand and tackle the issues involved. (p. 59)

The second problem, or rather “clutch of problems”, identified by Black and Wiliam (1998a) referred to the perennial issue of the need to align assessment to meet formative and summative purposes and, more specifically, the challenges that this would present for teachers and students. They reached the conclusion that “...if an optimum balance (was – *my insertion*) not sought, formative work (would – *my insertion*) always be insecure because of the threat of renewed dominance by the summative” (Black & Wiliam, 1998a, p. 59).

In the absence of an underlying theory of AfL to guide them, and in light of serious challenges requiring further research, rather than suggesting a “single royal road” or “any one optimum model” to policy-makers and practitioners, the authors identified a set of principles to guide practice. These included:

...The setting of clear goals, the choice, range and articulation of appropriate learning tasks, the deployment of these with appropriate pedagogy to evoke feedback... and the use of that feedback to guide the learning trajectory of students. (It was suggested that – *my insertion*)...within and running through any such plan should be a commitment to involving students in the processes of self- and peer-assessment... emphasised by a constructivist approach to learning. (Black & Wiliam, 1998a, p. 61)

Principles of AfL

By way of filling the gap in practical knowledge, at least temporarily, a booklet followed quickly on foot of the publication of the 1998 review that offered a summary of findings, together with recommendations for changes in policy and

practice, that would be required if the potential offered by AfL was to be realised (Black & Wiliam, 1998b). As acknowledged, the recommendations in *Beyond the Black Box* (Black & Wiliam, 1998b) extended beyond the scientific evidence of the review to allow for informed speculation on how best to proceed. In like manner, a subsequent publication in 2002 by the Assessment Reform Group (ARG) argued that "...while assessment of learning has well established procedures, assessment for learning requires some theoretical ideas to be translated into practice in particular ways if the potential benefits are to be gained" (p. 1). This gave rise to the articulation of the 'ten principles of assessment for learning', outlined as follows:

1. AfL should be part of effective planning of teaching and learning.
2. AfL should focus on how students learn.
3. AfL should be recognised as central to classroom practice.
4. AfL should be regarded as a key professional skill.
5. AfL should be sensitive and constructive because any assessment has an emotional impact.
6. AfL should take account of the importance of learner motivation.
7. AfL should promote commitment to learning goals and a shared understanding of the criteria by which they are assessed.
8. AfL should ensure that learners receive constructive guidance about how to improve.
9. AfL should develop the learner's capacity for self-assessment so that they can become reflective and self-managing.
10. AfL should recognise the full range of achievements of all learners.

(ARG, 2002, p. 1)

As stated many times in the literature, AfL is a deceptively simple approach and it really is only when one begins to unpack the research rationale underpinning each of the principles that the complexity of the issues involved, their interconnectivity, and mutual-dependence, begin to emerge. In attempting to draw together the import of these issues for teaching and learning, reference is made at this juncture to Harlen's (2006a) revised conceptualisation of AfL as a cycle of events (see Figure 7). It captures well the non-linearity of the teaching-learning contract and the co-dependence of the teacher-learner interpretation of evidence of performance, leading to decisions on how to close the gap, towards achieving higher goal. Most notably, perhaps, it places the student firmly at the heart of his/her learning, as the fulcrum rather than passive recipient of knowledge creation, with all the power and responsibility that this confers.

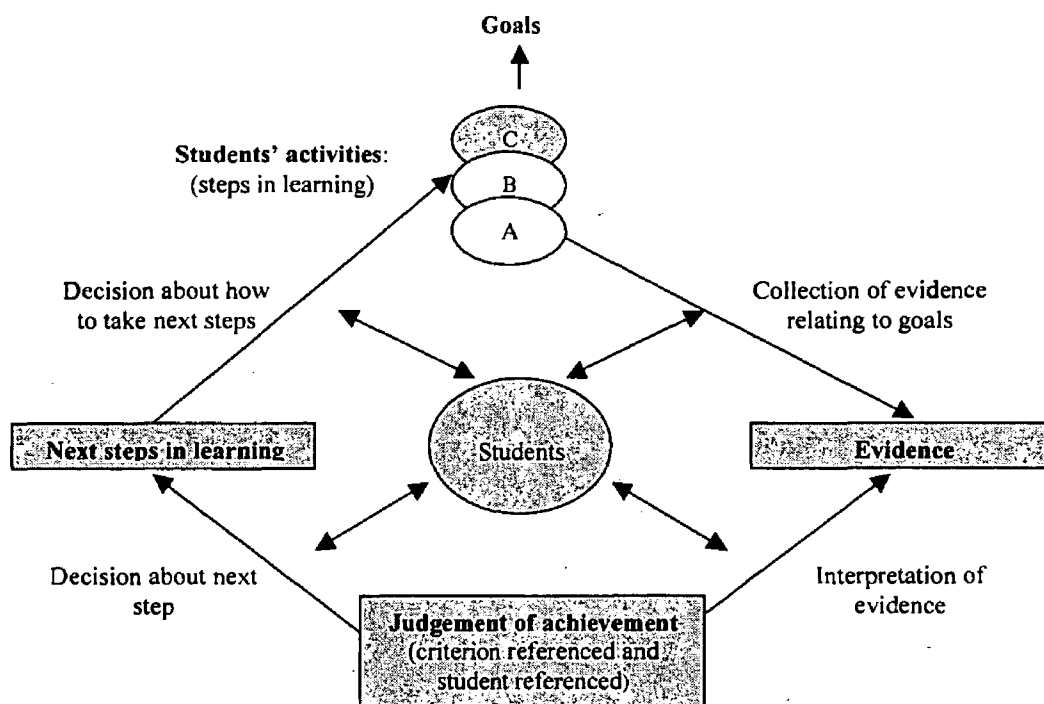


Figure 7.

Assessment for Learning as a Cycle of Events (Harlen, 2006a)

Following the publication of the Black and Wiliam (1998a) landmark review, together with *Beyond the Black Box* (Black & Wiliam, 1998b), and the AfL principles (ARG, 2002), a series of studies were undertaken originating in the UK and spreading to the US, amongst other countries. Each was focused, to a greater or lesser extent, on the content or process elements of the implementation of the research findings. It is towards an examination of a number of the more prominent of these studies that attention now turns; specifically, the King's-Medway-Oxfordshire Formative Assessment Project (KMOFAP) and the Learning How To Learn (LHTL) Programme in the UK, the Assessment is for Learning (AifL) Programme in Scotland, and the Keeping Learning on Track (KLOT) programme in the US, are reviewed. Two key factors determined the inclusion of these programmes over others: first, these studies are recognised within the literature as being well-conceived, pioneering initiatives and second, the continuity of influence of key personnel between these projects - notably members of the King's Project Team - allowed for a tracking of the development of

themes -and indeed changes in direction and emphases - as lessons were learned and areas for further research identified.

The King's-Medway-Oxfordshire Formative Assessment Project

In concluding their 1998 review, Black and Wiliam identified a number of avenues for further research, including the pre-eminent need for 'real world', classroom-based, research that would test the findings of the research studies reviewed. Other issues raised included the need to:

- Redress the absence of attention in previous studies to issues of race, class and gender;
- Unpack assumptions about learning from both teachers' and students' perspectives;
- Examine the effect on practice of teachers' content and pedagogical knowledge;
- Consider the nature of the social setting of the classroom and its impact on the dynamics of teacher-student relationships.

More recently, the need to extend this type of work to include students at both extremes of formal schooling - notably infant children and students in post-16, tertiary and non-statutory assessment contexts - has also been highlighted (Ecclestone, Swann, Greenwood, Vobar, & Eldred, 2004).

Prioritising the first challenge, in 2000, Black and Wiliam spearheaded the King's-Medway-Oxfordshire Formative Assessment Project (KMOFAP), a two-year development project, originally involving four teachers (two science; two mathematics) from six volunteer schools, catering for students in the age range 11 to 18 years. The second year saw the project expand to include two teachers of English and one additional science and maths teacher from each of the participating schools, bringing the total number of teachers involved to 48. As reported, over the period of the project, very deliberate attention was paid to both the content and process aspects of the teachers' professional development.

The content element of the project was informed directly by the reports of strategies employed by the various studies incorporated in the 1998 review. Specifically, the KMOFAP asked participating teachers to pilot a number of key approaches, reported subsequently under four headings: oral feedback in classroom dialogue, feedback through marking, peer- and self-assessment and the formative use of summative tests (Black & Wiliam, 2006a, p. 14).

In contrast, the rationale behind, and, indeed, the focus of the process element of the project reflected many of the ideas addressed in the opening section of this chapter about the constituent elements of effective CPD; as articulated by Black and Wiliam (2006a):

One of the key assumptions of the project was that if the promise of formative assessment were to be realised, a research design in which teachers are asked to test out and perhaps modify a scheme worked out for them by researchers would not be appropriate. We presented them with a collection of ideas culled from research findings rather than a structured scheme. We argued that a process of supported development was an essential next step. In such a process, the teachers in their classrooms had to work out the answers to many of the practical questions which the research evidence we had presented could not answer. The issues had to be reformulated in collaboration with them, where possible in relation to fundamental insights, and certainly in terms that could make sense to their peers in ordinary classrooms. (p. 20)

The model of CPD adopted to support this approach included one, day-long, team meeting every five weeks, supplemented by visits to participating schools by members of the research team to observe classroom practice, offer feedback, collect interview data and identify items for discussion at the team meetings. In order to provide a quantitative measure of achievement, data from tests used customarily by the schools were collated, with no additional tests being introduced.

The findings from the project were such that Black, Harrison, Lee, Marshall, and Wiliam (2003) remarked later that the evidence from the research 1998 review could finally be supplemented by evidence of enhanced performance on the UK national, and on schools' own, examinations. However, as observed subsequently by James and Pedder (2006b), the KMOFAP showed the potential of A/L to extend far beyond enhanced performance on traditional tests; indeed, the KMOFAP

demonstrated that, with appropriate support, teachers and students were willing to risk changing time-honoured classroom dynamics, and traditional teacher-student roles, to the extent that boundaries between assessment and teaching, and teachers and pupils, effectively became blurred:

That was the main message (of AfL – *my insertion*). In its fullest expression it gives explicit roles to learners, not just to teachers, for instigating teaching and learning. Thus students are not merely the objects of their teacher's behaviour, they are animators of their own effective teaching and learning processes. This has its clearest embodiment in processes of peer and self-assessment when students... become autonomous, independent and active learners. When this happens, teaching is no longer the sole preserve of the adult teacher; learners are brought into the heart of teaching and learning processes and decision making as they adopt pedagogical practices to further their own learning and that of their peers. It gives the old expression of being 'self-taught' a new meaning. (p. 28)

Adopting a broader frame of reference, and considering specifically the process element of AfL, Black and Wiliam (2006a) remarked that "... more generally, this work raised questions about the 'application' of research to practice and the links between this and the professional development of teachers... how teachers take on research, adapt it and make it their own...." (p. 24). Significantly, they recalled how team discussions following the completion of the programme revealed that, although the research team explained that the project was fundamentally concerned with turning research ideas into practice – a task that could only be completed through the classroom-based, trialling of strategies and techniques - participating teachers always believed that the research team already knew how to achieve this and were simply providing opportunities for teacher self-discovery. One might speculate that this reflected teachers' earlier experiences of CPD that followed more tradition, empirical-rational approaches (see Table 3). At the very least it highlighted a point raised in the opening chapter of this thesis that the implementation of research, like policy, is neither predictable nor linear but rather, as McLaughlin (2006) puts it, "...implementation is a multi-layered phenomenon, and each layer or level acts on the policy as it interprets intention, resources and regulatory framework...." (p. 6). Hence, the researchers' conclusion that an obvious next step would be to consider more closely the factors that influence teachers' uptake of innovations like formative assessment.

Whereas the KMOFAP tried to portray how the findings from formative assessment research could be translated into classroom practice, the Learning How to Learn (LHTL) project took up the challenge of examining how teachers learn at three, inter-related levels: at the level of the classroom, the school, and across internal, external and virtual learning networks. Aware that the levels and quality of training and collegial interaction experienced during the KMOFAP had been exceptional - and therefore unsustainable – research focus had inevitably shifted to questions about the process of AfL and CPD. As outlined by James, Black, McCormick, Pedder, & William (2006), issues arose that extended beyond the classroom associated with the ‘rolling out’ or ‘scaling up’ of innovations:

Claims for the effectiveness of Assessment for Learning raise a number of questions... which go beyond the classroom. For example, how does one spread knowledge and promote change in these specific practices across teachers and schools? How can one ‘leverage’ using minimum resource for maximum impact? (p. 102)

In response, the four-year (2001-2005) development and research project, LHTL, was undertaken, involving forty schools (infant, primary and second level) across five, UK local education authorities (LEAs) and a virtual education action zone, working collaboratively with researchers from four universities. Building on the experiences of KMOFAP, and working with a number of the researchers from this project, the principal aim of the LHTL project was to bring “...two areas of research together in order to investigate the conditions in classrooms and schools, that would promote better learning by pupils through the development of AfL practices by teachers...” (James et al., 2006, p. 103). The logic underpinning this approach was that extending and scaling up best practice in formative assessment was entirely contingent on gaining greater insight into the kinds of teaching practices that bring about effective learning and the nature of professional development and institutional cultures that support teachers’ learning of these new practices.

Mixed-methods instruments, including questionnaires to measure antecedent and outcome variables, standardised performance measures to gauge student achievement, together with more qualitative approaches, such as narrative accounts,

interviews, document analysis and some video, were employed. No control groups were used, based on the decision that the primary focus was on teacher engagement with formative assessment interventions, and, less so, on the 'impact' of the intervention.

Given the interest in this thesis on classroom-level intervention, and a site-based teacher learning community as the vehicle for CPD, commentary on the findings of this project are limited to the first level of the LHTL project. (For detailed reviews of the overall findings of the project, the reader is referred to a special edition of *Research Papers in Education*, 21(2), 2006: <http://www.citeulike.org/journal/routledg-rred>).

A key question that arose in the course of the school-level work was whether, in order to effect change in classroom practice, teachers actually need to share their underpinning principles of teaching and learning or whether "...altering classroom practice through the application of certain principles is sufficient" (Marshall & Drummond, 2006, p. 135). Noting that although much of the work on formative assessment has developed from what might loosely be called a cognitive constructivist learning perspective, Marshall and Drummond (2006) suggested that it is also feasible to adopt a more situated perspective, as advanced by Putnam and Borko (2002) and Lieberman (2008). In this context, in reflecting on what the teachers in the LHTL project attempted to achieve, they observed that "...the implementation of AfL in the classroom... (is – *my insertion*) much more than the application of certain procedures – questioning, feedback, sharing the criteria with the learner and peer and self-assessment – but about the realization of certain principles of teaching and learning (Marshall & Drummond, 2006, p. 135), a view also expressed by James and Pedder (2006b) in the context of the KMOFAP.

They arrived at a number of important conclusions from reviewing video data, cross-referenced with interview data, key among them being that only about 20% of the teachers observed employed AfL to the extent that it clearly promoted and supported enhanced student autonomy. As a consequence, the researchers proposed that the principles of AfL were no longer sufficient to bring about the nature and

degree of teacher change required and that far greater attention had to be placed on the choice of classroom task:

...The four original headings, under which AfL practice was conceived – questioning, feedback, sharing criteria and self-assessment – need revision.... The spirit of AfL is instantiated in the way teachers conceptualize and sequence the tasks undertaken by pupils in the lesson. The nature of these tasks affects all subsequent interactions within the class. Moreover these tasks tend to demand ‘high organization based on ideas’ if it is going to help pupils become independent learners. (Marshall & Drummond, 2006. p.147)

One of the implications drawn from this conclusion was that teachers’ beliefs about the nature of learning needed to be made explicit through group discussion and debate. This was in keeping with teachers’ requests during the KMOFAP for the researchers to provide input on learning theories (Black & Wiliam, 2006a). In this context, Black and Wiliam (2006a) expressed initial surprise at this request but, upon reflection, acknowledged that if teachers were to feel confident in giving constructive feedback to students, this presupposed teacher-efficacy in their knowledge of learning theories. A second conclusion reached by the LHTL team is particularly interesting in the context of discussion in the first section of this literature review on situated learning. Specifically it was observed that, depending on teachers’ existing/initial beliefs about learning, some teachers embraced the spirit of AfL more easily and readily, in part, because they valued student autonomy more but also because they adopted a situative perspective on learning. The research team reported that the teachers whose practice they categorised as illustrating the spirit of AfL demonstrated a “progressive, rather than fixed, view of what went on in any given lesson”:

Neither circumstances nor the disposition of pupils were beyond change.... Indeed these provided a challenge to be reflected upon and overcome. Such an attitude (gave - *my insertion*) these teachers a far greater sense of agency than those who tended to see constraints in the school culture, the examination system or the ability of the pupils. (Marshall & Drummond, 2006, p. 147)

Further light is shed on this and other issues by research conducted in Scotland since 2001.

The Assessment is for Learning Programme

Described as "...probably one of the most ambitious developments in Scottish education in the past 25 years" (Condie, Livingston, & Seagraves, 2005: 6.1: Discussion and Implications), the Assessment is for Learning Programme (AifL) was initiated in 2001 by the Scottish Executive Education Department (SEED) under the aegis of the newly formed Assessment Reform Group. Involving all sectors, pre-school though post-primary, across thirty-two local authorities in Scotland, the programme sought to build and expand on the work by Black, Wiliam and colleagues in the UK; indeed, members of the team from King's College, London, provided direct support in the initial phases of the work in Scotland. The breadth and scope of the programme is evident in the following description from the external evaluators' report of the overall aim of the programme:

Essentially, the AifL Programme was designed to bring together the various purposes of assessment into a single coherent framework which would answer questions of accountability, standards and monitoring of progress and performance but which emphasised the role of assessment in supporting individual pupils' learning in the classroom. This meant reviewing existing practice, including national testing and monitoring procedures, as well as introducing and developing new ideas and strategies within schools. In the event, 10 projects were identified which, together, addressed the aims of the programme. (Condie et al., 2005:1.3)

The ten component projects are listed in Appendix B, together with brief descriptors of each. While it is beyond the scope of this thesis to review in detail the work of the AifL programme, two of the design features are particularly noteworthy. First, in keeping with the KMOFAP and the LHTL project, the Scottish programme adopted a collaborative approach to the work, engaging the expertise of policy-makers, researchers and practitioners. Although this aspect of the programme reported varied success that was attributed to the attitudes and inclinations of those involved, the inclusion of researchers beyond their customary involvement at the design and/or evaluation stages of such projects, served to highlight the potential which such alliances might offer.

A second interesting feature of the AifL programme was its approach to policy-implementation and the issue of scalability, in particular. The combination of

a top-down, bottom-up model facilitated the conception and instigation of the individual projects at national level, within an overall national framework. Additionally, it allowed for the fielding of the implementation of the projects to interested schools, supported in turn by LEAs, to undertake action-research-like, site-based, projects, intended to produce final case-study reports. Local implementation was supported by the provision of project bursaries from SEED funding to release teachers (substitution costs), purchase resource materials or provide professional development opportunities, as determined and prioritised by individual schools and school clusters.

The evaluation of this programme, which began shortly after it was launched in schools, sought to trace the individual and collective success of the implementation of the ten projects to realise the overall objectives of the initiative. Summarising the overall evaluation conclusions, the evaluators highlighted various aspects of the work that proved successful, many of which reinforced those of previous research. Specifically, the value of small-scale, site-based projects, in which funded teacher learning communities were established to encourage and sustain changes in teaching and learning was noted. Moreover, in keeping with the LHTL project, it was reported that change was more likely to take place where teachers were convinced of the merits and potential of the innovation, where emphasis was placed on the reality of their classroom experiences (i.e. situated learning) and reflected their day-to-day concerns about teaching and learning (Marshall & Drummond, 2006). In this context, the importance of 'hot' information in generating interest and commitment, defined as "...feedback and advice from colleagues and other practitioners..." was deemed to be "...far more relevant and immediate (and trustworthy) than when presented with 'cold' information in printed reports and guidelines" (Condie et al., 2005).

One of the more pertinent projects within this programme, from the perspective of this thesis, was the Support for Professional Practice in Formative Assessment project, the first listed in Appendix B. Involving four groups of eight or nine schools, from primary and post-primary sectors - which received routine support from a development officer, third level personnel and occasional contributions from the KMOFAP staff - this project recorded interesting findings regarding its impact on students. For example, a significant difference was reported in the visibility and

impact of the project between primary, secondary and special schools, with the proportion of participating staff being reportedly higher in primary and special schools. This is an important finding which points to the need to be cautious about making generalisations about the implementation of formative assessment across different school levels.

More generally, staff responses in the 'field visit schools' to an evaluative rating scale attested to the overall success of the project: success in improving pupil learning and motivation (89%), the quality of students' work (88%), attainment (78%), learning skills (94%), concentration (83%) and behaviour (55%), with 77% of staff reporting an enhanced learning climate in their schools (Condie et al., 2005). From teachers' perspectives, greater awareness of children's individual needs was noted, as well as increased personal motivation, confidence and enjoyment of the teaching role. These positive findings were linked to the nature of the supports provided during the project, including time release to engage in the action research cycle of planning, preparing, reflecting and evaluating intervention strategies, in addition to school cultures, organisational and management structures that supported and encouraged innovation.

Yet, challenges similar to those identified by teachers in the KMOFAP emerged, ranging from the nature and extent of change in pedagogical approaches required to implement AfL as intended, to tension arising from the conflicting demands of formative and summative assessment. Indeed, one of the more fundamental challenges to emerge from across the range of projects in the AifL programme was the difficulty encountered by teachers in going beyond surface engagement with strategies and techniques of AfL. As reported:

For some teachers, getting to grips with formative assessment, with its implications for the way children learn and teachers teach, was disconcerting, while others saw only opportunities. Where it was taken on board and fundamental changes in practice occurred, important changes included: better understanding of pupils' learning; improved dialogue with pupils and parents; and better feedback. Some did not engage with the deeper issues and adopted some of the strategies and routines (e.g. traffic lights, two stars and a wish and wait time) more superficially, without really engaging with the underpinning theories and philosophies. (Condie et al., 2005)

However, the evaluators cautioned that challenges - particularly in terms of scalability - remained. In fact, the observation that implementation was 'patchy' served to highlight a potential drawback of engaging a combined bottom-up, top-down, policy implementation approach; specifically, it emerged that the degree of choice afforded made it difficult to generalise across the programme subsequently and/or to derive general principles from the overall findings. As the evaluators noted, what appears relevant and appropriate at the start-up phase of an initiative may not translate easily into longer-term implementation approaches:

Where projects were taken on board in this spirit, they were successful and benefits accrued to pupils, parents and teachers, as relevant. It is a process that is generally considered an appropriate one for small-scale change, for the seeding of new practices within a community which will grow more widely. It has been successful, but at the cost of generalisability across Scotland. However, to revert to more traditional approaches of top-down delivery requiring implementation will run the risk of losing goodwill and commitment. The challenge to SEED and to local authorities is to encourage networks of practitioners who will continue to challenge, to encourage reflection and to promote change in schools. (Condie et al., 2005)

What is at issue here is, in effect, the process of AfL, and specifically, the issue of devising a system of CPD that balances choice and opportunities for site-based, situated learning with the creation of self-sustaining, local action networks that will maintain the momentum once the initial round of intensive support is scaled back. In similar vein, reflecting on the KMOFAP, the AifL programme in Scotland and the LHTL projects, Black and Wiliam (2006a) warned that, such was the specificity of each project, that the replicability of the initiative and the generalisability of their findings were issues that had to be handled with caution. However, they also argued that certain lessons seemed to be universal:

The experience so far of schools basing their own innovations on the existing findings of results from research and from recently developed practice is that a sustained commitment over at least two years is needed, that evaluation and feedback have to be built in to any plan and that any teachers involved need strong support, both from colleagues and from their school leadership. (p. 23)

This is the challenge that was taken up by the Keeping Learning on Track (KLOT) project in the USA, under the steerage of Dylan Wiliam, the final initiative to be considered at length in this review.

Building on the experiences of the KMOFAP project, Black and Wiliam (2006b) articulated a tentative theory of AfL that they hoped would guide further, comparable research and help them conceptualise the dynamic changes that accompany the introduction of formative assessment in classrooms. Leveraging a situative perspective, and, specifically, the concept of a community of practice as advanced by Lave and Wenger (1991), they proposed that effective implementation of AfL in the classroom necessitated changes in:

- The way teachers and students interact with subject matter from a one-way delivery model to a two-way, interactional relationship;
- The teacher's role, from a focus on teaching to a focus on learning;
- The student's role, from receptivity to activity;
- The student-teacher relationship, from adversaries to collaborators.

In turn, Wiliam and Thompson (2006) argued that if the AfL ideas were to be scalable, i.e. introduced widely and with integrity - so that the kinds of changes described could be achieved - then much greater clarity would be required in relation to the kinds of practices that do, and do not, constitute formative assessment. Wiliam (2007) remarked that while the flood of publications that followed the research reviews evidencing the potential of formative assessment - if implemented well - to impact student learning was predictable, in many cases, the assignment of the term 'formative' was not always legitimate or warranted:

It is *not* legitimate... to claim that the existing research indicates that such use of standardized benchmark assessment will raise student achievement.... Although common assessments, benchmark assessments, interim assessments, and the like play an important role in monitoring student progress and providing system-level support information for policymakers, there is no evidence at this time that such assessments increase student achievement.... Instead of putting their faith in such solutions, schools need to implement the kind of formative assessment that research clearly supports. (p. 36)

In seeking to clarify this issue, Wiliam (2007) proposed a typology of formative assessments, categorised according to the type, focus and length or duration of time,

as shown in Table 9 in order to make a key point, which is that, currently, there is no evidence that supports the use of long-cycle ‘formative’ assessment; all of the existing research is based on studies of medium, and short-cycle, assessment.

Table 9.

A Typology of Formative Assessment (Adapted from Wiliam, 2007)

Type	Focus	Length
<i>Long-cycle</i>	Across marking periods, quarters, semesters, years	4 weeks to 1 year
<i>Medium-cycle</i>	Within and between instructional units	1 to 4 weeks
<i>Short-cycle:</i>	Within and between lessons	
✓ <i>Day-by-day</i>		24 to 48 hours
✓ <i>Minute-by-minute</i>		5 seconds to 2 hours

This should not, however, be interpreted simply in temporal terms. As observed in reviews of previous studies: “...the vision of formative assessment... involves more than adding “extra” assessment events to the flow of teaching and learning” (Wiliam & Thompson, 2006, p. 5). Indeed, it is notable that whereas reviews prior to, and/or included by Black and Wiliam (1998a) focused on the functions of feedback and, more particularly, the nature of feedback that supported learning, subsequent work reviewed previously in this chapter evidences an increasing acceptance that AfL demands nothing short of a complete overhaul and renegotiation of the classroom contract:

In a classroom where assessment is used with the primary function of supporting learning, the divide between instruction and assessment becomes blurred. Everything students do, such as conversing in groups, completing seatwork, answering questions, asking questions, working on projects, handing in homework assignments – even sitting silently and looking confused – is a potential source of information about what they do and do not understand. The teacher who is consciously using assessment to support learning takes in this information, analyses it, and makes instructional decisions that address the understandings and misunderstandings that are revealed. In this approach, assessment is no longer understood to be a thing or an event (such as a test or a quiz); rather, it becomes an ongoing, cyclical process that is woven into the minute-to-minute and day-by-day life of the classroom. (Wiliam & Thompson, 2007. 5)

Hence, Thompson and Wiliam's (2007) expression of what they termed the "big idea", reproduced below as in the original:

Students and teachers
 Using evidence of learning
 To adapt immediate learning needs
 Minute-by-minute and day-by-day. (p. 6)

Distinguishing between the roles of teachers, students and their peers, and drawing on Ramaprasad's (1983) concept of 'closing the gap', Thompson and Wiliam (2007, p. 7) 'unpacked the big idea' and the five key strategies of formative assessment as shown in Table 10.

Table 10.

Aspects of Assessment for Learning (Thompson & Wiliam, 2007)

	Where the learner is going	Where the learner is right now	How to get there
Teacher	Clarifying and sharing learning intentions and criteria for success	Engineering effective classroom discussions, questions, activities, and tasks that elicit evidence of learning	Providing feedback that moves learners forward
Peer	Understanding and sharing learning intentions and criteria for success	Activating students as instructional resources for one another	
Learner	Understanding learning intentions and criteria for success	Activating students as the owners of their own learning	

However, collaborating with teachers involved in the Keeping Learning on Track (KLOT) project in the United States led to the realisation that "...these generic strategies offer a necessary but still insufficient framework..." to guide the practical implementation of AfL (Thompson & Wiliam, 2007, p. 7). Arising, in part, from differences between teachers' craft knowledge (know how) and knowledge of universal truths (know why), it was determined that much greater specificity was required in relation to the concrete teaching and learning techniques that manifest the five core strategies (see Appendix C). However, as Wiliam (2008) observed some time later, even with strategies and techniques in place, there are significant hurdles to be overcome:

When you've been doing something one way half a million times, doing it another way is going to be pretty difficult! That's the bad news. The good news is that, if a school understands just how hard effective formative assessment is to implement and puts the right supports in place, the necessary changes are achievable in every classroom. But to bring these changes about, schools need to make sustained investments in a new kind of teacher professional development. (p. 37)

Hence, the KLOT project was designed, chiefly, as a teacher professional development programme with a firm focus "...on the black box of day-to-day instruction as the central axis of capacity building efforts" (Thompson & Wiliam, 2007, p. 3). As outlined, it comprises three core components:

- A content component (*what* we would like teachers to learn about and adopt as a central feature of their teaching practice): minute-by-minute and day-by-day;
- A process component (*how* we support teachers to learn about and adopt assessment for learning as a central part of their everyday practice): an ongoing program of school-based collaborative professional learning;
- An empirical/theoretical component (*why* we expect teachers to adopt assessment for learning as a central part of their everyday practice, and the outcomes we expect to see if they do): intervention's theory of action buttressed by empirical research.

Assuming both the content (*what*), and empirical (*why*), components of the programme to have been detailed sufficiently at this point, focus shifts now to look more closely at the process (*how*) component of KLOT, beginning with the authors' description of the logic behind their 'scaling-up' model.

Delivering AfL at Scale

While it was readily acknowledged that the KMOFAP - in which participants were selected on the basis of personal interest coupled with the support of Local Educational Authorities - was not generalisable to other contexts, the KLOT project embraced this as its principal challenge. Moreover, in describing the challenge of scaling up, due recognition was given to the perennial issues of teacher isolation and the privacy of classroom practice:

Scaling up a classroom-based intervention isn't like gearing up factory machinery to produce more or better cars. Scaling up an intervention in a

million classrooms (roughly the number of teachers in the U.S.) is a different kind of challenge. Not only is the sheer number of classrooms daunting, the complexity of the systems in which classrooms exist, the separateness of these classrooms, and the private nature of the activity of teaching mean that each and every teacher has to “get it” and “do it” right, all on their own. No one else can do it for them, just as no one else can do students’ learning for them. No matter how good the intervention’s theory of action, no matter how well designed its components, the design and implementation effort will be wasted if it doesn’t actually improve teachers’ practices – in all the diverse contexts in which they work, and with a high level of quality. This is the challenge of scaling up. (Thompson & Wiliam, 2007, p. 1)

In responding to these challenges, Thompson and Goe (2006) identified a number of reasons why teacher learning communities seem to be “...particularly functional vehicles...” (pp. 6-7) to support teacher learning about AfL. First, they argued that given teachers’ requirement for high level professional judgement to implement formative assessment, recourse to school-based expert personnel is an attractive option. Second, they suggested that site-based, teacher learning communities, if set up and sustained well, by their very nature offer the kind of sustained, developmental, personalised, situated support that research suggests is needed for real change to occur. Third, they proposed that teacher learning communities provide a safe environment for teachers to disclose gaps in their knowledge and understanding, thereby “...supporting teachers in converting the broad assessment for learning strategies into “lived” practices within their specific subjects and classrooms” (Thompson & Goe, 2006, p. 6).

The genesis of the idea of using school-based, teacher learning communities as the vehicle for delivering effective, scalable, CPD resulted from a series of research initiatives in the US over a period of five to six years and, in particular, the discovery that “...the content of assessment for learning was not the only content that (they – *my insertion*) needed to convey explicitly to teachers and school leaders” (Thompson & Goe, 2006, p. 22). Figure 8 overleaf conceptualises these two content areas within a two-by-two matrix, the argument being that an effective and scalable programme of teacher professional development would need to attend to each of the four quadrants.

		Content Areas	
		Assessment for learning	Teacher Learning Communities
Learning Phases	Initial exposure & motivation	Quadrant A	Quadrant C
	Ongoing learning, practice, reflection, & adjustment	Quadrant B	Quadrant D

Figure 8.

KLOT: Content Areas and Learning Phases (Thompson & Goe, 2006)

In cataloguing their learning since 2003, and their work with hundreds of U.S. schools in multiple-districts across the socio-economic divide, Thompson and Goe (2006) made reference to "...the staggered rate..." (p. 9) of their own learning about the need to attend explicitly to both the content and process elements of A/L. Significantly, they recounted how they progressed from delivering, three-day workshops to groups of teachers, with a singular focus on transferring their knowledge of the A/L principles, strategies and techniques, without any reference to how this might be supported and maintained within individual schools (Quadrant A), to the point where they realised that "...just as there is an implicit classroom contract, there is an implicit school contract, and that contract has not, historically, supported teachers to be learners" (Thompson & Goe, 2006, p. 9), (Quadrant B). This led, ultimately, to the development by the research team of a series of content modules, each comprising the directions and materials for guiding the conduct of school-based, ninety-minute meetings on A/L. The modules were introduced to nominated school-based teacher leaders, who attended regionally based, monthly meetings, facilitated by the research team, as part of a cascaded model of CPD (Quadrant C). The leaders were supported and trained in content and process knowledge, with a view to them facilitating month to six-weekly meetings in their individual schools with teaching colleagues (Quadrant D). Each module followed a uniform design style, the aim being to maintain the fidelity of the A/L message as it filtered down from the originators of the ideas through to the classroom implementation, thereby anticipating one of the key challenges of using TLCs at scale.

The work, as outlined, has been evolving rapidly and the impact of this work on student achievement is currently unavailable. However, clear messages regarding organising, facilitating and sustaining professional learning communities focused on A/L have emerged; these have been distilled and presented in Table 11.

Table 11.

Facilitating and Sustaining TLCs (Adapted from Thompson & Wiliam, 2007)

Practical Suggestions	<ul style="list-style-type: none"> ✓ Plan for the teacher learning community to run for at least two years ✓ Start with volunteers ✓ Meet monthly for at least 75 minutes ✓ Aim for a group size of 8-10 ✓ Try to group teachers with similar assignments ✓ Establish building-based groups ✓ Require teachers to make detailed, modest, individual action plans
Guiding Questions for an Action Plan	<ul style="list-style-type: none"> ✓ What is one thing that you will find easy to change? What difference do you expect it to make to your practice? ✓ What is one thing that you would like to change that will require support? What help would you need? ✓ What other change would you like to make later on in the year? What help might you need? ✓ What will you do differently or stop doing to implement these changes?
Guidelines for running TLC Meetings	<ul style="list-style-type: none"> ✓ Introduction (5-10 minutes): Agree on aims of the meeting & move to agenda ✓ How's It Going (30-50 minutes): the "active ingredient" of the meetings in which each participant summarises their previous month's work and receives the groups' support and advice on how to progress ✓ New learning about formative assessment (25-40 minutes): focusing on the introduction of new ideas to the group ✓ Personal action planning (10-15 minutes): participant think through, in detail, their plan for the next month ✓ Review of the meeting (5 minutes): meeting review in the context of the original aims and objectives.

This work has significant import for this study; consequently, it is revisited in Chapter 3. However, more generally, it is important to note at this point the influence exerted by international research on A/L and CPD on recent Irish-based initiatives.

A/L in the Irish Context

The collective influence of the literature reviews undertaken, and the research studies initiated in their wake, is evidenced in a growing body of research studies being undertaken at both post-graduate (e.g. Cahillane McGovern, 2006; Lee, 2008),

and national levels. An initiative spearheaded by the NCCA, in which an English and a geography teacher from each of ten post primary schools in counties Cork and Sligo were supported by local Education Centres to pilot AfL strategies with their first year students, deserves particular attention at this juncture.

Facilitated by the establishment of two, teacher networks and the development of a web portal, this year-long project offered teachers an initial, three-day, training input on "...the broad outline of assessment for learning in the classroom..." including "...a menu of in-class approaches from which the participants might choose" (NCCA, 2005a, p. 4) to support their use of AfL in the classroom and, in turn, the reporting of student progress to parents.

An initial progress review, based on qualitative data spanning the period 2003-2004, indicated that teachers reported the introduction of a range of AfL strategies, with a notable preference for sharing learning intentions and comment-only marking. In keeping with previous research (e.g. Butler, 1998), variations in feedback yielded mixed results, with "more able students" (NCCA, 2005a, p. 7) pressing for traditional graded feedback. Effects on learners were noted in terms of teachers' perceptions of "...a generally positive impact on the involvement and motivation of learners..." including "...hitherto reluctant learners..." (NCCA, 2005a, pp. 7-8). One of the most significant findings of the study, perhaps, were teachers' perceptions of the ease with which AfL was assimilated into daily practice, leading to the conclusion being drawn that "...the assessment for learning approaches recommended to the teachers who participated in the NCCA initiative (were – *my insertion*) intuitive and flexible and (could – *my insertion*) be adapted easily to a range of subject disciplines in an Irish context" (NCCA, 2005a, p. 12). Given the frequent and over-lapping challenges reported in evaluations of previous AfL studies in the UK and US, detailed previously in this chapter, this conclusion is unexpected and suggests a need for further research, perhaps employing more rigorous research tools, to support in-depth, classroom-based analysis of the nature and degree of the implementation being undertaken.

This chapter was presented in two parts, the first focusing on teacher professional development, in the context of developments in thinking regarding theories of learning, the second offering a selective review of the literature on assessment. Further to the key themes identified in section one, a number of important issues have emerged in relation to the literature and research on assessment.

First, as a warranted approach to effecting change in teaching and learning that has enjoyed considerable success worldwide, AfL provides both an educational imperative and a research rationale for undertaking similar work in Ireland, particularly with young children from disadvantaged backgrounds who may be described as perennial low-achievers.

Second, despite a range of interventions in recent years - in the UK and the US in particular - that adopted a variety of innovative strategies and research approaches to address both the process and content challenges of CPD and AfL, significant challenges remain. These include issues in relation to how teachers variously implement formative assessment (according to the *letter* or *spirit* of AfL) and the extent to which this is determined by teachers' underlying beliefs about teaching and learning, through to issues of the generalisability of research findings and the scalability of CPD that maintains the integrity of AfL, on the one hand, while affording teachers the opportunity to engage in sustained, situated learning, on the other.

Third, there is a notable silence in the literature reviewed in relation to core issues such as the evolving profile and role of the facilitator of CPD; while it is readily acknowledged that the kind of expert support provided by researchers as part of the KMOFAP was unrealistic, the burden placed on teacher-leaders in the KLOT project requires further exploration if the approach is to prove tenable for a wider audience. An associated concern relates to the degree of support and pressure that needs to be applied in order to ensure that top-down and bottom-up initiatives, such as those introduced in Scotland, enjoy more uniform success.

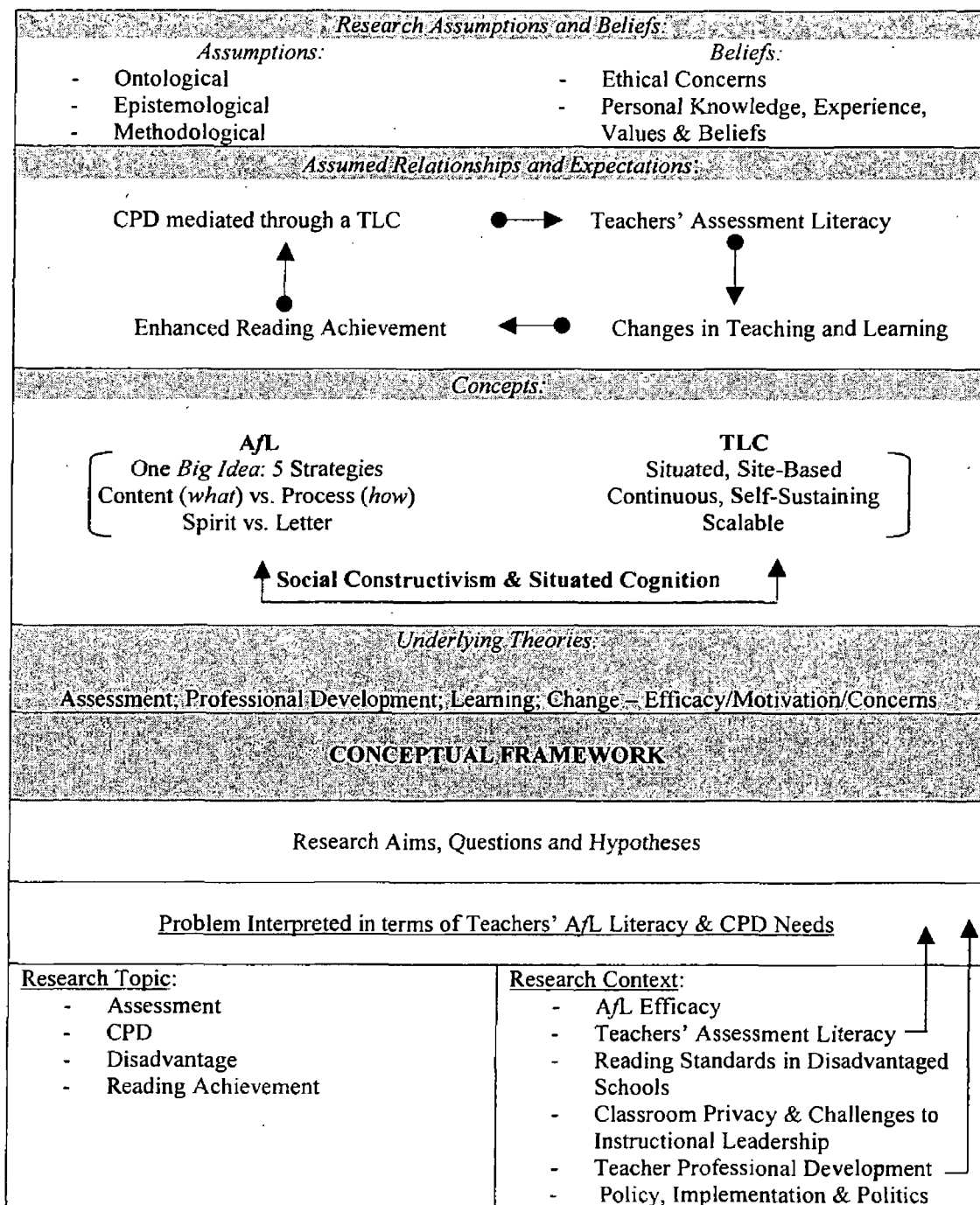
For a further elaboration on these issues, the reader is directed to Appendix D which offers a distillation of a review by Gardner (2008) of the conclusions drawn by the Analysis and Review of Innovations in Assessment (ARIA) project in relation to initiatives and developments in assessment in England, Scotland, Wales and Northern Ireland.

Conclusion: Towards a Conceptual Framework for the Study

As promised at the outset, this chapter concludes by introducing the conceptual framework for the study, defined as “...the system of concepts, assumptions, expectations, beliefs and theories that supports and informs research” (Maxwell, 2005b, p. 33). According to Eisenhart (1991), the primary function of a conceptual framework is to justify “...the concepts chosen for investigation/interpretation, and any anticipated relationships between them... given the research problem under investigation” (p. 209). As employed in this context, the framework depicted in Table 12 overleaf is intended to bridge the research questions identified in the opening chapter, with the literature review conducted here, to inform the research study outlined in the next chapter.

Progressing upwards from the foot of the table, attention is drawn to the contextual issues from which the research topic emerged, through the particular interpretation of the challenges drawn, to the research questions posed. In turn, the underlying theories and guiding concepts, derived in the main from this literature review, informed the assumption of a number of key relationships and expectations about the potential impact of an A/L intervention on teachers’ knowledge, skills and attitudes and children’s reading achievement. These assumptions are based on a range of assumptions and beliefs that are outlined in full in the next chapter, which details the study design and methodology.

Table 12.

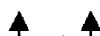
Conceptual Framework for the Study**Note:**

This table is intended to be read from the base, upwards.

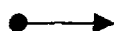
Elements of the Conceptual Framework:



Underpinned by:



Impacts on:



CHAPTER 3

THESIS DESIGN AND METHODOLOGY

Introduction

It has been argued that a frank articulation of the conceptual framework of a research study, defined as "...the system of concepts, assumptions, expectations, beliefs, and theories that supports and informs... research..." (Maxwell, 2005b, p. 33) is often missing in doctoral dissertations and that this is inherently problematic given that it is nigh impossible for any research to be a-theoretical. Referring specifically to doctoral work, Lester (2005) observes that very often there is "...a lack of attention to clarifying and justifying why a particular question is proposed to be studied in a particular way and why certain factors (e.g. concepts, behaviors, attitudes, societal forces) are more important than others" (p. 460). As a consequence, the reader has to infer, amongst other things, the rationale for the research design and choice of methods.

Taking cognisance of these views, the introductory chapter of this thesis sought to contextualise the study, with a view to clarifying the scope of the work and the manner in which the research problem was interpreted. In turn, the previous chapter reviewed and assessed a selection of research on Assessment *for* Learning, teacher professional development and learning theory, with a view to developing a robust conceptual framework for the research. This was presented in the preceding pages. The broad aim of this third chapter, then, is to focus on the study itself and to introduce the research design and methods employed to test the research hypotheses, in the context of an appropriate research paradigm.

Organisation of the Chapter

The chapter is presented in three parts. Following an initial reminder of the research problem and hypotheses, the relationship between three dominant research paradigms - their ontological, epistemological and methodological assumptions - and their associated theoretical frameworks, is considered. This sets the stage for an

introduction to the underlying interpretive/constructivist paradigm that guided this work, and the mixed methods design employed. This is followed by an overview of the research study itself that culminates in a discussion on the issues of quality control, validity, reliability and ethics. The chapter is grounded in contemporary dialogues on the merits and demerits of mixed methods research, in a post-modern, post-positivist tradition.

From Research Problem to Research Hypotheses

As explained in the opening chapter, the research problem was interpreted as being fundamentally one of a lack of research knowledge and practical acumen in the Irish context regarding (a) the potential of AfL to effect improvements in children's learning and (b) how best to develop teachers' assessment literacy through CPD so that they would feel empowered to embrace the *spirit*, and not just the *letter*, of the pedagogy. Following on from the literature review and the exposition of a conceptual framework for the study, a set of research questions emerged which are now presented as research hypotheses:

1. A nine-month, school-based intervention, employing Assessment *for* Learning principles and practice, would make a quantitative difference (i.e. effect size) to the reading achievement of a target group of children when compared to a similar cohort not involved in the intervention;
2. There would be a discernible, positive impact on children's attitudes, motivation and approaches to reading;
3. Using the medium of a site-based, teacher learning community to provide a professional development programme on Assessment *for* Learning would have a positive impact on teachers' knowledge, skills and attitudes of/to AfL and, in turn, how they teach reading.

Research Paradigms

In designing the study, cognizance was taken of the ongoing international debate on what constitutes good research and evidence in education (Cochran-Smith & Lytle, 1999; Eisenhart, 2005; Johnson & Onwuegbuzie, 2004; Leech & Onwuegbuzie, 2007). In this context, careful consideration was given to the role of research paradigms, the contrasting ontological, epistemological and methodological assumptions on which they are premised, and the import of these perspectives for research in the social sciences. These ideas and themes impacted significantly on the design of this study, given the contrasting but complementary nature of the hypotheses: the first two requiring a quantitative response, the third a qualitative one. As a consequence, a mixed methods approach was employed, in keeping with an interpretive/constructivist paradigm, the rationale for which is now presented.

Paradigm Wars and the Incompatibility Thesis

Coined originally by Kuhn (1970) in his seminal work, *The Structure of Scientific Revolution*, the term 'paradigm' has survived in current research literature to describe:

...A set of very general philosophical assumptions about the nature of the world (ontology) and how we can understand it (epistemology), assumptions that tend to be shared by researchers working in a specific field or tradition. Paradigms also typically include specific methodological strategies linked to these assumptions, and identify particular studies that are seen as exemplifying these assumptions and methods. (Maxwell, 2005b, p. 36)

For more than a century, research in the social sciences has been dogged by vociferous debate and dispute between researchers, peaking in the "paradigm wars" (Maxwell, 2005b) of the 1980s. Characteristic of these wars were the polarised views on quantitative and qualitative methods, articulated by positivist and interpretive/constructivist purists, with each side advancing their particular approach as ideal for social research. Johnson and Onwuegbuzie (2004) argue that these purists "...implicitly if not explicitly..." advocated "...the *incompatibility thesis*... which posits that qualitative and quantitative research paradigms, including their associated methods cannot and should not be mixed" (p. 14). Hence the emergence of two

distinct research cultures, one trumpeting the “hard” evidence to be derived from scientific, causal research, the other promising richer, more nuanced knowledge, from approaches such as ethnography and observational studies.

In response to the extreme polarisation of the debate, recent years have witnessed calls for some sort of rapprochement between the two sides; as a consequence, a third research paradigm has been proposed that seeks to shift the focus from metaphysical to methodological concerns, based on the argument that adopting a “...needs-based or contingency approach...” frees researchers to “...create designs that effectively answer their research questions...” instead of following “...completely... either the qualitative or quantitative paradigm” (Johnson & Onwuegbuzie, 2004, p. 20).

In attempting to capture the import of these debates, and their implications for choosing a research paradigm to guide the work of this study, Table 13 was prepared. Drawing from the work of Lather (1992), Sparkes (1992) and Eisenhart (1991), Table 13 is intended to serve a dual purpose. First, it seeks to illuminate the distinctions in philosophical approaches exemplified by the three paradigms regarding the nature of knowledge (epistemology), the relationship between the researcher and knowledge (ontology), and the implications of adopting particular epistemological and ontological perspectives for researchers’ methodological approaches. Second, it aims to link the three paradigms, and their underlying concepts and approaches, to three kinds of research frameworks, thereby demonstrating that a researcher’s choice of framework is not arbitrary, but reflects important personal beliefs and understandings about the nature of knowledge, how it exists (in the metaphysical sense) in relation to the observer, and the possible roles to be adopted, and tools to be employed consequently, by the researcher in his/her work.

Table 13.

Research Paradigms and Theoretical Frameworks

Paradigms		
Empirical/Positivist	Interpretive/Constructivist	Critical/Emancipatory
Associated Terms/Terms Used Interchangeably (Adapted from Lather, 1992)		
Experimental Quasi-experimental Correlational Causal comparative Quantitative	Naturalistic Phenomenological Hermeneutic Symbolic Interaction Ethnographic Qualitative	Neo-Marxist Feminist Race specific Freirean Participatory Transformative
Paradigms: Ontological, Epistemological and Methodological Assumptions (Adapted from Sparkes, 1992)		
1. Ontological Assumptions		
External-Realist	Internal-Idealist, Relativist	External-Realist or Internal Idealist
2. Epistemological Assumptions		
Objectivist, Dualist	Subjectivist; Interactive	Subjectivist; Interactive
3. Methodological Assumptions		
Nomothetic Experimental Manipulative	Ideographic Hermeneutical Dialectical	Ideographic Participatory Transformative
A Summary of Three Kinds of Research Frameworks (Adapted from Eisenhart, 1991)		
1. Theoretical: A skeletal structure of explanation - Relies on formal theory, i.e. it is "...constructed by using an established, coherent explanation of certain phenomena and relationships..." (Eisenhart, 1991, p. 205); - <i>Signing up</i> to an established theory means the researcher is bound by the conventions of argument and experimentation associated with the theory.	2. Conceptual: A skeletal structure of justification - "...An argument that the concepts chosen for investigation/interpretation, and any anticipated relationships among them, will be appropriate and useful, given the research problem under investigation..." (Eisenhart, 1991, p. 209); - In common with Theoretical frameworks, the framework may be "...based on different theories and various aspects of practitioner knowledge, depending on exactly what the researcher thinks (and can argue) will be relevant and important to address about a research problem, at a given point in time and given the state-of-the-art regarding the research question (Eisenhart, 1991, p. 209); - Unlike either theoretical or practical frameworks, these frameworks "...are not constructed of steel girders made of theoretical propositions or practical experiences; instead they are like scaffoldings of wooden planks that take the form of arguments about what is relevant to study and why..." (Eisenhart, 1991, pp. 210 – 211).	3. Practical: A skeletal structure of accumulated knowledge - Research problems that focus on "...problems that really pay off for practitioners" (Eisenhart, 1991, p. 207); - Not informed by formal theory but by the accumulated practical knowledge (ideas) of practitioners and administrators, the findings of previous research and often the viewpoints of politicians and public opinion; - Research hypotheses and questions derived from this knowledge base - Research results used to support, extend and revise the practice.

It should be noted that the linkages made between the three research paradigms and the three research frameworks in Table 13 are intended to be loosely interpreted; they do not represent a 'direct fit' but they do help to indicate broadly how theoretical and conceptual frameworks differ and, in turn, how these differences reflect contrasting philosophical and methodological persuasions. Moreover, they serve to signal the link between current calls for some sort of rapprochement between these paradigms and the early work of Eisenhart (1991), who highlighted the implications of adopting particular paradigmatic approaches for the design of research studies. Indeed, these issues have particular currency in the context of the relentless demand from policy-makers for evidence-based research, premised on the belief that the success of education policies depends on the availability of warranted, scientific knowledge about how to improve instruction. As Raudenbush (2005) explains:

...Causal effects of instructional interventions belong at the heart of the current policy research agenda in education and... randomized experiments are the best way to warrant these effects.... However, ... such experiments are insufficient to achieve the aims of this agenda.... A knowledge gap needs to be addressed so that educators can act on incentives and use resources in ways that will supply students with coherent and effective instruction. It follows that identifying, testing, and warranting the effectiveness of strategies for instruction is currently the central task of applied research in education. (pp. 26-27)

The search for warranted interventions and/or evidence-based teaching and learning strategies, however, presents education researchers with formidable design challenges that pivot around the issue of how evidence is to be constructed and understood. As Cochran-Smith (2006) warns, there is a lot at stake:

...The danger is a too narrow version of evidence grounded in "scientific research" as causal studies only... (when, what is required is – *my insertion*) ...a perspective on evidence that includes but is not limited to clinical trials and a broad and inclusive view of science that includes but is not limited to the investigation of causal questions.... Perhaps this could be along the lines of... a "postpositivist" perspective on science that incorporates experimental research as well as qualitative research, with the latter accepted on its own terms rather than forced into an overarching framework governed by the assumptions of the former. (p. 11)

This argument complements the call by Eisenhart (2005) for a "science plus" approach to education research, based on the rationale that "...qualitative studies are

likely to offer more, not less, that quantitative studies to scientific research in education” (p. 55). In response, Cochran-Smith (2006) introduced the term “evidence plus” recently to highlight the need for teacher education to be informed “...by a wealth of critical and theoretical enquiry... in particular, ...the large bodies of work that now exist about teacher learning in communities....” (p. 11).

Calls for evidence/science plus inquiry sit well with the conclusion drawn in the previous chapter of this thesis highlighting the need for research that would prioritise the impact of CPD on teachers’ classroom practices as well as student learning (Darling-Hammond, 2000; Loucks-Horsley & Matsumoto, 1999), without bifurcating the integrity of the teaching-learning process (Marcos & Tillema, 2006). Specifically, these calls helped to rationalise the view taken in this thesis that a purely empirical/positivist paradigm - although appropriate to the quantitative inquiry undertaken in response to Hypotheses 1 and 2 - would not have provided a useful frame for Hypothesis 3, given that its primary focus was neither to predict nor test teachers’ behaviour. In like manner, a critical/emancipatory approach was ruled out on the basis that no attempt was being made to investigate, or influence, the power relationships within the TLC. In this context, an interpretive/constructivist framework - endorsing the view that each individual would construct his/her own reality, and assign different meaning to their experiences of participation in the TLC on the basis of personal/professional life experiences - was considered more suited to this element of the research. However, the need to apply a rigorous quantitative approach to test Hypotheses 1 and 2 necessitated the employment of a mixed methods approach, as outlined in Figure 9, which is inherently compatible with an interpretive/constructivist paradigm and the delivery of rich data anticipated by a science-plus/evidence philosophy.

Mixed Methods Research Typologies

Recent work by Leech and Onwuegbuzie (2007), and Johnson and Onwuegbuzie (2001), serve to reinforce the belief that “...epistemological and paradigmatic ecumenicalism is within reach” (Johnson & Onwuegbuzie, 2001, p. 15) by proposing a mixed methods classificatory systems highlighting how the criteria outlined by Driscoll, Appiah-Yeboah, Salib, & Rupert (2007) may be variously

configured. As explained, mixed methods designs are described and classified according to "...the level of prioritisation of one form of data over another, ...the combination of data forms in the research process (such as during the collection or analysis phases) and by the timing of data collections, such as whether the quantitative and qualitative phases take place concurrently or sequentially, and if so, in what order" (Driscoll et al., 2007, p. 19). Researchers employ mixed methods designs, variously termed multi-method, combined or hybrid, primarily because they are appropriate to the research questions raised but, additionally - as in the case of this work - because they have the potential to expand the breadth and scope of the research undertaken, and to mitigate the limitations of relying on one method.

Distinguishing between two major types of mixed methods research, namely mixed-model and mixed-method designs, Leech and Onwuegbuzie (2007) have suggested a typology of mixed methods research that proved very useful in framing the particular approach adopted in this study as shown in Figure 9: a partially mixed, concurrent, equal status design (P1), with the qualitative data assuming equal status with the quantitative data (QUAN + QUAL), reflecting the espousal of a science/evidence plus philosophy.

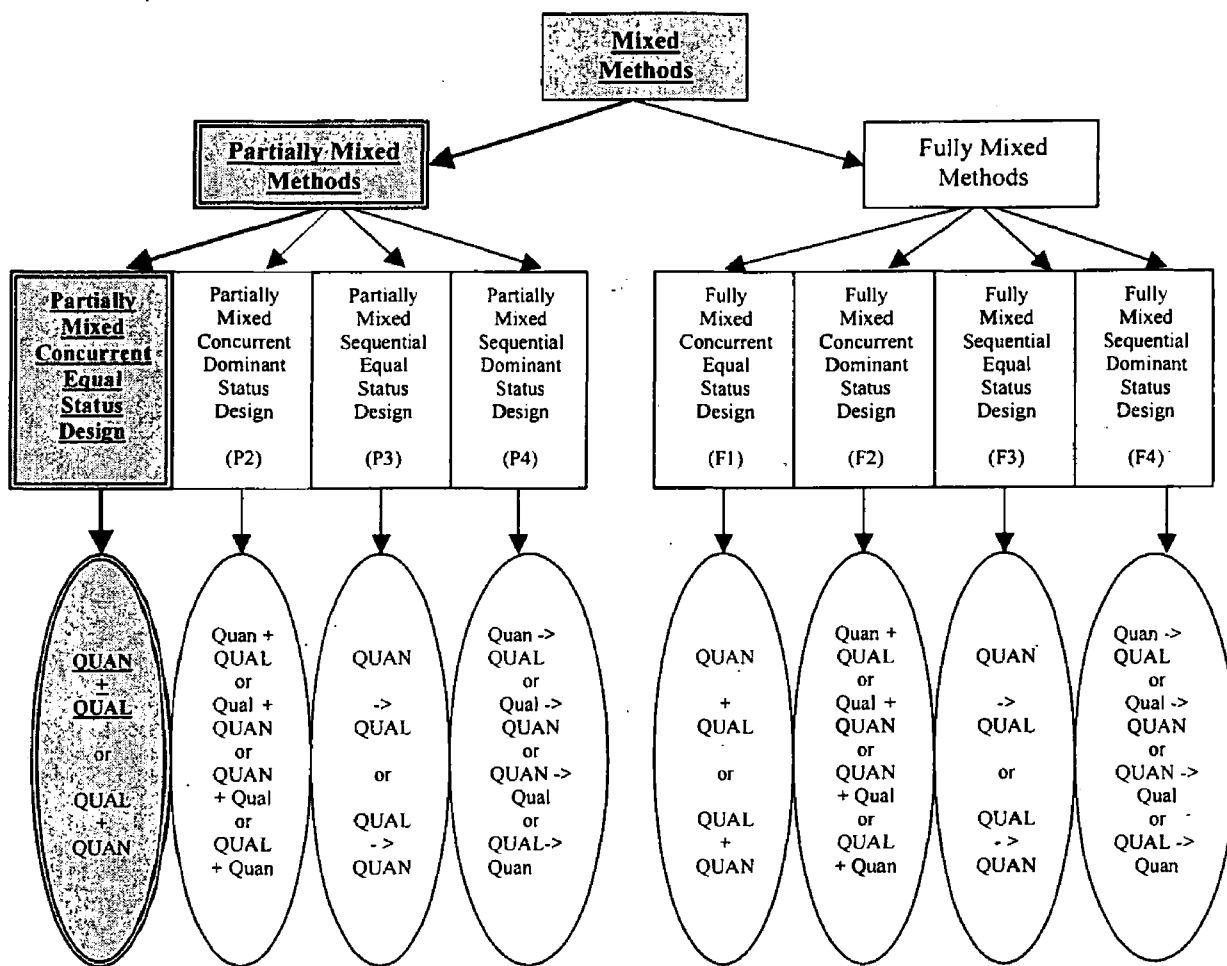


Figure 9.

Mixed Methods Research Typologies (Adapted from Leech & Onwuegbuzie, 2007)

Note:

Leech and Onwuegbuzie's typology (2007) has been adapted in Figure 9. Items highlighted in gray represent the options within the typography chosen for this research.

Creswell (2002) suggests that concurrent, mixed methods designs - such as the one described here - frequently serve a number of purposes including addressing different kinds of questions, validating and/or transforming data for comparison purposes. It should be noted, in this context, that the term 'mixed methods' as distinct from 'mixed model' is being used intentionally to signal that, although the research design employed both quantitative and qualitative approaches within and across stages of the research process, it did not seek to blend the data or to *qualitize* or *quantitize* data, subsequently.

Drawing together the ideas introduced so far, Leech and Onwuegbuzie's (2007) work on typologies of mixed methods research, together with the concepts of "science-plus" and "evidence-plus", as conceived by Eisenhart (2005) and Cochran-Smith (2006), led the researcher to develop a framework for the study shown in Table 14 overleaf. Justification for this approach was derived from Eisenhart's (2005) advice that:

In education, research must be practically relevant as well as scientifically proficient. The most elegant, sophisticated research designs can easily lead to naught if the results cannot be understood by practitioners, are not relevant to practice, or cannot be put into practice. (p. 57)

Table 14, then, presents the research undertaken, identifying the paradigms that framed each hypothesis, the variables considered prior to, and during, the intervention period and, finally, the outputs and insights gained. Closer examination of these elements, with respect to the hypotheses, shows that the input and output variables used in Hypotheses 1 and 2 were readily quantifiable, in keeping with the quantitative approach employed, and yielded evidence associated with a scientific approach to research. In contrast, Hypothesis 3 focused on the process elements of the work that demanded a less *scientific* but equally rigorous set of instruments to provide complementary data (science/evidence plus).

Table 14 extends the ideas introduced in the Conceptual Framework (Table 12) outlined at the end of the preceding chapter, in particular the section at the top: Research Assumptions and Beliefs. As indicated, Table 14 attests to the researcher's belief that different types of research questions demand different kinds of paradigmatic and methodological responses, which, in turn, suggests an affiliation to a non-purist, philosophical ideology.

Table 14. *Towards an Interpretive, Mixed Methods, Research Framework*

Hypotheses 1 and 2: Changes in children's Reading Ability, Motivation and AfL Practices			
Science/Evidence: Quantitative, Pre-Test/Post-Test Quasi-Experimental Design			
Theoretical Frame	Input	Process	Output
<u>A Skeletal structure of explanation</u> Relies on formal theory, i.e. it is "...constructed by using an established explanation of certain phenomena and relationships..." (Eisenhart, 1991, p. 205). <i>Signing up</i> to an established theory means the researcher is bound by the conventions of argument and experimentation associated with the theory.	<u>Pre-Intervention</u> Children's reading achievement; ✓ Standardised reading tests (MICRA-T; DPRT); Children's motivation to read: ✓ ERAS; Children's AfL practices: ✓ SA/LQ Children's characteristics: ✓ SEN ✓ Age ✓ Gender	<u>Not Evaluated</u> Children's classroom engagement with AfL strategies and techniques (video data)	<u>Post-Intervention</u> Children's reading achievement: ✓ Standardised reading tests (MICRA-T; DPRT); Children's motivation to read: ✓ ERAS; Children's AfL practices: ✓ SA/LQ
Hypothesis 3: Impact of TLC-mediated CPD on Teachers' AfL Knowledge, Skills & Attitudes			
Science/Evidence Plus: Qualitative, Situated Exploration and Evaluation			
Conceptual Frame	Context	Process	Insights
<u>A Skeletal structure of justification</u> "...An argument that the concepts chosen for investigation/interpretation, and any anticipated relationships among them, will be appropriate and useful, given the research problem under investigation...". Unlike either theoretical or practical frameworks, these frameworks "...are not constructed of steel girders made of theoretical propositions or practical experiences; instead they are like scaffoldings of wooden planks that take the form of arguments about what is relevant to study and why..." (Eisenhart, 1991, pp. 209 – 211).	Micro: School Culture ✓ Leadership ✓ Innovation Challenges ✓ Reading ✓ Achievement ✓ AfL Competence Teachers' characteristics: ✓ Teachers' assessment knowledge, skills & attitudes ✓ Age profile ✓ Teaching Experience ✓ Qualifications ✓ Efficacy ✓ Motivation ✓ Concerns	TLC-mediated CPD: ✓ Knowledge <i>in practice</i> (rather than <i>for/of</i>) ✓ Situative Perspective ✓ Dual focus: Group = TLC; Individual = Adapted CBAM Group focus: ✓ 'Closing the gap' reviews of practice; videos Individual Focus: ✓ Teacher monthly reviews – Learning Logs	Dual Focus of Analysis: 1. Group: Teachers' independent review of CPD ✓ Content ✓ Process ✓ Context (including facilitator) 2. Individual: ✓ Completion of & Reflection on personal learning logs - Efficacy ✓ Completion of & Reflection on personal SA/LAI

Note:

A number of acronyms are introduced in Table 14 which are referred to routinely within the study from this point forward. It should be noted, therefore, that a full list of these acronyms is provided in Appendix E.

Having detailed the framework that guided this research and the mixed methods approach employed, this section addresses the core elements of the study itself, outlining the data obtained at each stage of the project (pre-, during-, and post-intervention), the participants involved and research instruments employed. Although more detailed information is provided in relation to the samples of children and teachers who participated in this research in the following chapter, some initial contextual information is provided here.

Host School

The opening chapter of this thesis detailed the need for research to explore the potential of A/L to improve the reading levels of perennial low achievers in an Irish context. Hence, criteria were developed to guide the choice of a host school(s); these included:

- Band One, DEIS designation in an urban area, i.e. one of the 300 urban schools nationwide considered to be teaching children experiencing the most extreme levels of socio-economic disadvantage;
- A minimum of three cohorts at each class level, to ensure sufficient numbers of children for the control and experimental groups;
- Documented records of previous achievement in reading, as evidenced by the availability of standardised test results, that could be used to establish baseline standards;
- A school culture and organisation that was positively disposed to innovation, i.e. a school that would be interested in, and supportive of, trying something new but which, in other respects, would not be considered exceptional.

Initial approaches were made by phone to a number of school principals known to the researcher, indicating the nature of the project and seeking initial expressions of interest. Following presentations to the staff of two schools, an agreement was reached to work with one junior school in accordance with the research programme and timeline presented in Table 15.

Table 15.

Quasi-Experimental Research Programme and Timeline

Stage	Data	Participants	Instruments
<u>Pre- Intervention</u> May/June 2007	Baseline data	Children of: a. Control Group b. In-coming Experimental Group	*Mary Immaculate College Reading Attainment Test (MICRA T) *Drumcondra Primary Reading Test (DPRT) *Elementary Reading Attitude Survey (ERAS) *Scaled questionnaire on current A/L practices during reading lessons (SA/LQ) *Samples of children's work
		Teachers: a. Control Group b. In-coming Experimental Group	*Audits of attitudes to, and perceptions of, current use of A/L practices, within the school generally, and specifically, within individual teacher's classrooms (A/LAI)
<u>During- Intervention</u> Sept/May 2007/2008	Intervention data	Teachers of Experimental Group	<u>Arising from CPD days:</u> *Copies of teachers' monthly action plans for the application of A/L *Teachers' Learning Logs (TLL) detailing their reflections on the CPD provided and the challenges/opportunities presenting, from professional and personal perspectives <u>Arising from Review meetings:</u> *Video recordings of aspects of monthly review meetings focusing on teachers' perceptions of the intervention project as it unfolded *Teachers' records of written critiques from colleagues, arising from the viewing of teaching videos from each of the four classes, showing teachers' efforts at piloting a range of A/L strategies and techniques (Teachers' Video Review Sheets – TVRS) <u>Arising from classroom practice:</u> *Video recordings of teachers' use of A/L methodologies during reading lessons
		Children: Experimental Group	*Portfolios of children's work *Video capture of children's responses to/use of A/L methodologies during reading lessons
<u>Post Intervention</u> May/June 2008	Post- Intervention data	Teachers of Experimental Group	*Audit of attitudes to, and perceptions of, current use of A/L practices, within the school generally and, specifically, within individual teacher's classrooms (A/LAI)
		Children: Experimental Group	*MICRA T *Drumcondra Primary Reading Test *Elementary Reading Attitude Survey *Scaled questionnaire on current A/L practices during reading lessons *Samples of children's work

Note:

The Control Group refers to the children who completed Second Class in June 2007, without receiving any additional support; the Experimental Group refers to the children who were in Second Class during the academic year, 2007-2008, and participated in the intervention project; CPD is shorthand for Continuous Professional Development and refers to the two days spent at the school every month to six weeks which used the medium of a teacher learning community to review video practice of the previous month's teaching before focusing on the introduction of new A/L strategies and techniques.

As outlined in Table 15, a quasi-experimental design which according to Singleton and Straits (1999) attempts to incorporate elements of a “true experimental” design without maintaining the same level of experimental control - for example, not employing randomization - was employed in response to Hypotheses 1 and 2. In order to bolster the design and reduce the impact of non-randomization, pre-intervention sample matching was undertaken with the researcher choosing to locate both the control and experimental cohorts within the same school environment. Participants were drawn from assembled collectives, namely, pre-formed classes, with the experimental group consisting of all of the children enrolled in Second Class for the duration of the academic year, 2007-2008 (N = 85); the control group comprised the outgoing Second Classes (N = 85).

Pre- and Post-Intervention Data: Control and Experimental Classes

As required by the quasi-experimental design employed to test the first two hypotheses, pre- and post-intervention data were obtained by requesting children in both the control and experimental classes to complete four, group-administered, assessments, beginning in May-June, 2007. These included:

- Two reading tests, standardised in the Republic of Ireland:
 - Mary Immaculate College Reading Attainment Test (MICRA-T);
 - Drumcondra Primary Reading Test (DPRT);
- An attitudinal survey on reading motivation, developed in the United States:
 - Elementary Reading Attitude Survey (ERAS);
- A scaled questionnaire on the use of A/L approaches in reading, developed by the researcher:
 - Scaled A/L Questionnaire (SA/LQ).

Standardised Reading Tests

With the exception of the MICRA-T, which was used routinely in the school, the assessments chosen were unfamiliar to the children. The DPRT was introduced as a second standardised reading measure because its focus on reading for comprehension was considered complementary to that of the MICRA-T which, at Levels 1 and 2, focuses mainly on the assessment of reading skills. Employing independent standardised tests in this way is not unusual in this kind of research; as noted by Black and Wiliam (2003) in the context of the KMOFAP, it is possible "...that gains could be sustained over time, even when measured using externally mandated standardized tests" (p. 633).

Elementary Reading Attitude Survey

While there is some evidence to suggest a positive correlation between a child's motivation to read, the frequency with which he/she might read and overall reading achievement (Guthrie, Wigfield, & VonSecker, 1999), Stiggins (2001) argues that the manner in which teachers use assessment determines children's overall desire to learn more than any other classroom factor. Described as the conditions and processes that account for "...the arousal, direction, magnitude and maintenance of effort..." (Katzell & Thompson, 1990, p. 144), motivation to read is deemed to be central to learning. Hence, as a method of triangulating the data obtained from the standardised reading tests and, more importantly, to widen the lens of enquiry sufficiently to capture data that might suggest changes in children's learning otherwise missed by standardised tests, baseline data on children's motivation to read were collected in May/June 2007, using the Elementary Reading Attitude Survey (ERAS) (McKenna & Kear, 1990).

Instrument Reliability

Developed in the US, the ERAS was administered to a stratified sample (in terms of race and gender) of 18,138 children in grades 1-6 from 95 school districts, representing 38 US states (McKenna & Kear, 1990; McKenna, Kear, & Ellsworth, 1995). The test comprises twenty items that seek to determine children's attitudes to

(a) recreational reading: "...reading for fun outside the school setting..." and (b) academic reading: "...reading aloud in class; reading workbooks and worksheets; and reading schoolbooks..." (Kush & Watkins, 1996, p. 316). Children are required to respond to each of ten items in two subscales, by circling one of four pictures of the cartoon character Garfield, that depicts him in a range of moods from very happy (4 points) to very unhappy (1 point), as shown in Appendix F. The test yields three scores, one for each of the subscales, and a total composite score.

It is notable that although a recent publication by Kazelskis, Thames and Reeves (2004) questioned the reliability of the instrument across gender and race - specifically with regard to potential factor invariance for African American and European American respondents - no definitive conclusions were reached that cast doubt on the reliability of the instrument for use with Irish students. Nonetheless, the Statistical Package for Social Sciences (SPSS) was used by the researcher in this study to obtain reliability values for the scales by measuring the internal consistency of the two attitude scales. The Cronbach's alpha coefficient values were reported as ranging from .74 to .89 - values well above the recommended 0.7 level (DeVellis, 2003) - thereby justifying its use in this study as detailed in the next chapter.

Scaled AfL Questionnaire

Additional insights into the children's attitudes to reading and, more specifically, their familiarity with, and use of, AfL approaches when reading, were sought by adapting the content of the National Assessment of English Reading Test (NAER) (Educational Research Centre, 2004), which is the most recent in a series of national assessment instruments. As designed by the researcher, the Scaled AfL Questionnaire (SAfLQ) has two sections; the first was modeled closely on the original NAER test; the second comprises twenty-four items across three subscales (before, during and after reading) that seeks to determine to what extent children use formative assessment in support of their reading (see Appendix G). Like the ERAS, the test provides three scores, one for each of the subscales, and a total score.

Instrument Reliability

Given the non-standardised nature of this instrument, statistical analyses were undertaken to determine the reliability of the scale as a whole and the interreliability of the subscales, details of which are reported in full in Appendix H. Pallant (2007) suggests that, ideally, Cronbach alpha values should be above 0.7 if the internal consistency of a research instrument is to be assumed. The values reported in Appendix H range from 0.81 to 0.89 for the instrument as a whole, with inter-item correlations reporting values well within the optimal range for inter-item correlation of 0.2 to 0.4, thereby verifying the robustness of the instrument overall, and its use in this study.

Test Administration

With regard to the administration of each of these tests, Pallant's (2007) advice on designing a study was heeded, with clear guidelines and protocols being issued to each teacher in order to ensure parity and consistency in test administration.

Research Instruments and Data Collection: Hypothesis 3

In contrast to the quantitative strategies employed to measure changes in children's reading achievements, the research approach adopted in response to Hypothesis 3 of the study drew on fundamentally different philosophical and epistemological assumptions and beliefs. According to Gasson (2002), an interpretive approach allows the researcher to make the implicit explicit. As outlined (Gasson, 2002, p. 3), the cycle of research associated with this genre typically includes the follows stages and elements:

- Understanding the "problem situation" objectively and determining appropriate situations in which phenomena relevant to the research problem can be observed;
- Engaging in those situations inter-subjectively with actors who regularly participate in such situations, to obsessively collect data on phenomena whether or not they seem significant at the time of data collection;

- Disengaging from the situation at relevant points, to analyze the data, to question which phenomena are significant (from a pattern-recognition and utility perspective, rather than a frequency or quantity perspective), then re-engaging if data saturation has not been achieved. Disengagement also requires *regular*, objective questioning of the value-systems and assumptions that are brought to bear in interpreting the situation;
- Disengaging from the situation, to present the subjective as objective and to interpret the situation for others.

In keeping with this approach, a range of research instruments were used to collect “rich data” (Alton-Lee, 2006, p. 622) concerning teachers’ individual and collective experiences of involvement in the TCL and their evolving knowledge, skills, attitudes and practices of AfL, as outlined below.

Pre- and Post-Intervention Data: Teachers of Control and Experimental Classes

The research instruments used with respect to Hypothesis 3 included:

- An AfL Audit Instrument (AfLAI);
- Teachers’ Learning Logs (TLL);
- Video footage of teaching and Teachers’ Video Review Sheets (TVRS).

AfL Audit Instrument

Teachers were invited to complete an AfLAI before and after the intervention period in order to establish a baseline against which to determine any changes in their knowledge, skills and attitudes of/to AfL, both within their own classrooms and within the school as a whole (based on individual knowledge and perception). Adapted from an instrument designed by Inspectors of the Department for Education and Science (DfES) in the UK to encourage dialogue between staff members about the strengths and weaknesses of their assessment policies and practices (<http://www.aaia.org.uk>), the AfLAI presents a range of questions across eleven subscales. In response, teachers are requested to choose one option from a four-point

scale indicating the degree to which individual statements reflect formative practices within the school/classroom, as shown in Appendix I.

It is noteworthy that, as used in the UK, teachers grouped at similar class levels discussed and completed one audit, having aired any differences of opinion regarding school practices. In the case of this research study, teachers completed the audit independently; hence, the findings regarding school practices represented individual opinions and perceptions of school assessment policies and practices. This is why teachers of the control classes were also requested to complete the audit; although they only did so once - before the intervention project began - auditing their views provided an opportunity to (a) gain some insight into the kinds of classroom assessment practices to which the control group had been exposed (albeit according to teachers' self-reports) and (b) gauge the degree of agreement or dissonance between teachers about school-wide assessment practices. These are important points that are revisited in the next chapter.

Teachers' Learning Logs

In keeping with one of the key principles of the KLOT programme (Thompson & Wiliam, 2007), following each meeting of the TLC, teachers were required to complete a Learning Log (TLL, Appendix J). Whereas the opening section of the log required teachers to critically reflect and review their learning following a two-day period of CPD, the second part was designed by the researcher to provide a window into the evolving needs and concerns of the teachers. In turn, this information was used, in accordance with the philosophy of the CBAM (Hall & Ford, 1997) to guide the researcher in preparing for, and facilitating, subsequent meetings of the TLC.

Piloting of Instruments

Before concluding this review of the research instruments, and moving to consider the professional development programme offered during the intervention period, the issue of the piloting of the research instruments arises. In this context, it must be reported that although a pilot phase was envisaged when this study was originally being considered, difficulties in identifying a host school, coupled with the

need to design a range of research instruments within a very short timeframe, resulted in the opportunity to conduct a pilot being missed. As will become evident in the next chapter, with the exception of one instrument, this did not compromise the findings of the study. However, it is fully acknowledged that this outcome was fortuitous and might very easily have resulted in more serious consequences.

Professional Development Programme and Timeline

As noted in the previous chapter, international research has found the establishment of teacher learning communities in schools to be a highly efficient way of supporting the delivery of school-based professional development (Ellsworth, Martinez, Lyon, & Wylie (2007), although many questions remain regarding how best to initiate, sustain and scale TLCs (William, 2008). In describing the programme of professional development provided as part of this study, consideration is given first to the content of the programme, before turning to consider how it was mediated through the TLC.

The 'Content' of the AfL Professional Development Programme

Over the period of the academic year, 2007-2008, a programme of professional development - that commenced with an introductory workshop on the findings of current research on the potential of AfL to improve children's learning achievement - was provided by the researcher; it exposed participating teachers, incrementally, to the five, research-based, AfL strategies, to which gains in student achievement have been linked. As outlined previously (Leahy et al., 2005), these include:

1. Clarifying and sharing the learning intentions and criteria for success with students;
2. Engineering effective classroom, discussions, questions and learning tasks;
3. Providing feedback that moves learners forward;
4. Activating students as the owners of their own learning;
5. Encouraging students to be instructional resources for one another.

Taken together, "...these strategies define the territory of assessment for learning" (Thompson & Goe, 2006, p. 5); they are the "...nonnegotiable... things that ...teachers must... do..." (Wiliam, 2008, p. 37); hence, as outlined in Appendix K, they were all included in the professional development programme offered.

The core content of the AfL programme was taken from a commercially available Scottish product called *The Learning Set* (Learning Unlimited, 2004). This content was deemed appropriate because it incorporates key lessons from the research evidence on AfL, as well as practical experience derived from the Scottish - Assessment is for Learning Programme - reviewed in the previous chapter. Adaptations to the content and recommended mode of delivery of the programme were made throughout the year, in response to the perceived needs of the teachers and the dynamics of the TLC. As a result, a significant range of supplementary materials were developed and/or sourced by the researcher, a selection of which are included as Appendix L.

The 'Process' of the AfL Professional Development Programme: The TLC

Although the structure and organisation of the monthly meetings of the TLC were modeled closely on the approach adopted, with success, by the Keeping Learning On Track (KLOT) project (Thompson & Wiliam, 2007), as detailed in the previous chapter, there were three notable differences. First, the site-based, TLC was facilitated by the researcher rather than a teacher-leader trained remotely with other leaders from different schools. Second, the duration and intensity of the TLC meetings differed significantly; in the KLOT model, it is recommended that TLC meetings are scheduled to take place at four-to-six week intervals, for a period of approximately one to two hours. Meetings of the TLC of this project met at similar intervals but stayed together for a period of two days, during which time the classes of participating teachers were taught by internal school colleagues - who knew the children because they routinely co-taught in the classes - with paid substitutes assuming the roles of the stand-in teachers. This arrangement was made possible by the financial assistance of the DES, negotiated through the NCCA. Third, the TLC

meetings of the KLOT project follow a closely scripted, prescribed agenda and timetable, which Wiliam (2008, p. 38) outlines as follows:

- How's it going (30-50 minutes – based on teachers' self-reports);
- New learning about formative assessment (25-40 minutes);
- Personal action planning (10-15 minutes);
- Review of the meeting (5 minutes).

In contrast, meetings of the project TLC began by negotiating a timetable for the two days. This negotiation was not entirely fluid, however, given that there was a number of non-negotiables that included the need to allow time to:

- Watch each teacher's video and allow for discussion, review and recommendations;
- Introduce a new A/L strategy and/or revisit previous strategies and techniques;
- Complete the Learning Logs and identify action steps for the next period.

These differences were significant in practical terms; for example, the amount of work involved for the researcher in sustaining and facilitating the programme over the period of the project was formidable. However, the most significant differences were in terms of the opportunities presented by the variations in the model employed in this project, specifically, the use of video to access teachers' classrooms, and the opportunity for the role of facilitator to be unpacked. In order to maximise these opportunities, a range of research instruments were carefully designed.

Video Footage of Teaching

As identified in previous chapters, one of the key challenges to be overcome in trying to improve teaching practice and educational standards is gaining access to the 'Black Box' of the classroom (Black & Wiliam, 1998a; Conway, 2002). The use of video, as an alternative to the more common research approach of requesting teachers to report on their practice and/or negotiating access to observe classroom practice, presents significant learning opportunities, not least for the teachers themselves.

Hence, in this research study, consent was sought at the outset from teachers, parents and children, to allow the placement of a video camera in each classroom for the duration of the intervention programme, for use by the teachers to capture their trialing of AfL strategies and techniques when teaching reading. As employed in the project, the use of video served two purposes: first, it allowed the TLC direct access to individual classrooms, thereby facilitating personal and group reflection on both the teaching and learning being observed. Second, it offered the opportunity, yet to be exercised, of sharing these experiences with others engaged in professional development focused on AfL. The use of video in this way presented a range of ethical issues that are detailed later in this chapter.

This section presented the research and professional development programmes, respectively, their associated timelines and the research tools used to collect data on (a) children's reading achievement and (b) teachers' knowledge, skills and attitudes of/to AfL and their classroom practice. The final section addresses issues of quality control, validity, reliability and ethics.

Data Analysis

In keeping with the mixed methods approach employed in this study, different approaches were used to analyse the data with respect to each of the research hypotheses. In the context that the statistical analyses conducted in relation to Hypotheses 1 and 2 was limited to the use of SPSS, and is detailed in full in the next chapter, comment is restricted at this point to a number of key observations relating to the implications for the researcher of adopting an interpretive approach to understanding the teachers' work.

First, donning the interpretive hat necessitated that the researcher engage in some self-reflexivity (Angen, 2000), recognizing that - in contrast to the detached, objective stance demanded by positivist and empiricist traditions - she invariably had to assume one of two roles in the TLC: that of outside observer or that of involved researcher. However, as Walsham (1995) observes, in either case, the analysis of data inevitably involved the researcher's own subjectivity which brings to the fore the issue of 'double-hermeneutics':

...Particularly with reference to in-depth case studies carried out over a period of time, researchers inevitably influence the interpretations of those people who are being researched, a process referred to as 'double hermeneutics'.... So, even if researchers view themselves as outside observers, they are in some sense conducting action research by influencing what is happening in the domain of action, if only by the sharing of concepts and interpretations with the personnel in the field site. (p. 77)

This raises two key issues: (a) the issue of self-disclosure by the researcher, both to the community with which she participated and the research community to whom the final report is directed and (b) the strategies to be employed in analysing data to minimize subjectivity. With regard to the first, it should be noted that, while information of this kind was readily shared within the TLC, it was deemed inappropriate to disclose biographical information for its own sake in this thesis, opting instead to share such information when it was immediately relevant to the study, as arises in the final chapter. With regard to the second and more crucial issue however, the concept of disciplined subjectivity, as advocated by McMillan & Schumacker (1997), was embraced to minimise the impact of researcher' subjectivity by engaging (a) multi-method strategies, as described previously, (b) participant language and verbatim accounts in recounting teachers' opinions, included in the next chapter and (c) low-inference descriptors to ensure that "...the interpretations and concepts (had – *my insertion*) mutual meanings between the participants and the researcher" (p. 407).

Second, in analysing the data with respect to Hypothesis 3, the aim of research generalisability was reframed in terms of the contribution that an in-depth, situated exploration of this kind could offer teachers in similar circumstances, in the event that the data obtained was deemed to be sufficiently 'rich' and 'saturated' to allow for general patterns and themes to emerge. This is an issue revisited in the context of the recommendations of this research in the final chapter.

Quality Control

Concepts of quality control - and validity and reliability in particular - are central to research. As generally interpreted in the literature on social science methodology, the issue of reliability is understood in terms of replicability (Shaxson,

2005), the key question to be considered being: were this research study to be replicated under the same conditions elsewhere, would the results be the same or similar?

In the context of the quantitative element of this research employed in response to Hypotheses 1 and 2, the issue of replicability obliged the researcher to meet two main criteria. First, it had to be established that the children assigned to the experimental group would have exhibited the same average outcome as the children assigned to the control group, had the children been alternatively grouped. Second, it had to be apparent that the experimental group assignment of children was unrelated to their potential to benefit from the intervention (Raudenbush, 2005).

Given that threats to reliability in experimental research are generally dealt with in advance of the study being implemented - as part of the design process (Robson, 2006) - the use of a quasi-experimental design to assess children's reading achievement, rather than a true experimental design, presented challenges, not least in terms of the need to control for confounding variables. Guided by Raudenbush's (2005) remark that "... much of the association between a potential confounder (e.g. an aspect of home environment) and the outcome is removed once one has controlled for a reliable pretest of achievement" (p. 28), the researcher administered a range of pre- and post-intervention tests to both the control and experimental groups. In turn, where non-standardised instruments were used, SPSS analyses were undertaken to examine their reliability, as outlined previously. In addition, the researcher attempted to minimise the impact of bias in the administration of the tests by ensuring that the same teachers were involved on each occasion and that test conditions were matched as closely as possible.

Although issues of quality control arise in undertaking qualitative research also, the issues present differently. As reported previously in the context of data analysis, a number of strategies were employed to obtain "thick descriptions" (McMillan & Schumacher, 1997, p. 416) - both of the teachers' emerging understanding of A/L and also of their concerns and needs in attempting to introduce practical changes in classroom instruction. However, the desire to collect rich data was tempered by a recognition of potential threats to the reliability of the findings

including reactivity, respondent' biases and researcher' bias, in particular (Lincoln & Guba, 1985). Consequently, in an effort to reduce these threats, four types of triangulation were employed systematically over the course of the project. These included:

- Data triangulation, supported by the use of a variety of tests before and after the intervention, coupled with video analysis during meetings of the TLC;
- Observer triangulation, facilitated by the presence of an NCCA representative as a participant observer at the meetings of the TLC which allowed for the sharing of ideas and interpretations;
- Methodological triangulation, enabled by the use of standardized and non-standardised tests and video observations of classroom engagement;
- Theory triangulation, supported by engagement in the TLC.

Notwithstanding the potential of triangulation as a strategy for minimizing researcher' bias in particular, two additional approaches were employed. First, a systematic and detailed audit trail of the activities of the research, including minutes of meetings, resources compiled for the TLC and field notes, was maintained. Second, member checking, which involved providing participating teachers with an opportunity to review and respond to transcripts of their Learning Logs and the findings of their A/L audit instruments, as part of an iterative approach to data analysis, was employed. It should be noted that this approach, in particular, proved very useful, as indicated in the next chapter.

Ethical Considerations

According to Robson (2006), ethics "...refer to rules of conduct, typically, to conformity to a code or set of principles" (p. 65). In undertaking this study, the researcher complied assiduously with the code of ethics of St. Patrick's College, undertaking the study only after written consent had been received from the College Ethics Committee. This sanction, in turn, came on foot of a series of meetings with prospective, study-participants. This process began with a presentation on the project to the full staff body of the host school, subsequent to which a letter of information, together with letters of consent, were sent to the Principal who, in turn, sought and

received the sanction of the Board of Management, thereby granting the researcher access to the school.

Upon receipt of the Board's imprimatur, two public meetings were held in the school, the first for the parents/guardians of the children in the control and experimental groups, held in May 2007. In both cases, the research study and the associated aims, objectives and methodology were outlined in full before participants' consent was sought. Furthermore, participants' rights to absent themselves and/or withdraw, subsequently, were clarified in full, with a commitment being given that no adverse consequences would attend to any such decisions. Parents/guardians who were not present at the meetings were visited by the Home-School-Community Liaison Teacher, and/or met with the researcher, to discuss the project and any fears or concerns that they had.

In addition to adherence to the College ethical code, the researcher was also vigilant in ensuring - insofar as she was capable and aware - that she avoided the "Ten questionable practices in social research":

1. Involving people without their knowledge or consent;
2. Coercing them to participate;
3. Withholding information about the true nature of the research;
4. Otherwise deceiving the participant;
5. Inducing participants to commit acts diminishing of their self-esteem;
6. Violating rights of self-determination (e.g. in studies seeking to promote individual change;
7. Exposing participants to physical or mental stress;
8. Invading privacy;
9. Withholding benefits from some participants (e.g. in comparison groups);
10. Not treating participants fairly, or with consideration, or with respect.

(Kimmel, 1998, quoted in Robson, 2006, p. 69)

Some of these practices were easily avoided by ensuring that informed consent, together with the individual assent of each of the children in the experimental group, was obtained - in writing - before the study began. The procedures adopted to obtain this consent/assent were such that issues such as coercing people to participate, withholding information from them and/or attempting to deceive them in any way, were consciously tackled and removed. Other issues required more creative consideration, however. For example, in requesting teachers to video their teaching

practices, with a view to vignettes of their work being shared during the monthly meetings of the TLC, and subsequent to the completion of the study edited versions being used for professional development purposes, issues of privacy and participant anonymity arose. Equally, the fact that one cohort of children was the focus of the intervention, with the anticipation that their exposure to A/L pedagogy would enhance their potential to read, whilst another cohort received no obvious, immediate benefit-in-kind, raised other questions.

Regarding the use of video, while the potential it would offer the TLC to critique practice was stated clearly, it was emphasised that the final arbiters were the teachers, parents/guardians and children. Two assurances were made in advance of an agreement being sought to employ video in the work. First, an undertaking was given that any stills or photographs taken would be of children's work, omitting faces where possible. The second issue, that of children's anonymity on video footage, proved more challenging, particularly in light of current research on the use of video (Arafey & McLaughlin, 2002) which suggests that attempts to blur faces present technical difficulties that cannot be overcome without seriously compromising the education value of the data ultimately. Hence, no guarantee of anonymity was given although it was agreed that teachers and parents/guardians would be invited to preview video vignettes intended CPD use following the completion of the project. Despite these caveats, full sanction was obtained from all parties to use video in the research.

Conclusion

As indicated at the outset, the main aim of this chapter was to introduce the research design and mixed methods approach adopted, in light of the research paradigms chosen. Towards achieving this, the chapter traced the link between research paradigms and their underlying philosophical and methodological assumptions, in order to justify the interpretive/constructive framework for the study. A subsequent review of mixed methods research led to discussion of the research methodologies employed, and the instruments used, at each phase of the research. Towards the end of the chapter, issues of data analysis, quality control, validity, reliability and ethics were raised in preparation for the analysis of data in the next chapter.

CHAPTER 4

DATA ANALYSIS AND RESEARCH FINDINGS

Introduction

This chapter aims to offer a coherent and critical analysis of the research data so that the findings will be of value to others. Reflecting on the process of research evaluation, Thompson, Ponte, Pack, & Goe (2004) suggest that empirically based social science research is particularly vulnerable to "...logistical challenges, unforeseen methodological difficulties, and disruptions that threaten both the completion and the validity of the research" (p. 7). Moreover, it is argued that despite the substantial debate within the academic literature on this issue (outlined in the previous chapter in the context of the 'paradigm wars'), the tendency to over-sanitize the research process so that it appears as a rational, clean, linear process remains, with the result that "...an idealized version of method over the practical reality and content of research" becomes privileged (Thompson et al., 2004, p. 7).

In an attempt to integrate a focus on disruptions into evaluative work, Thompson et al. (2004) proposed a five-point continuum to capture the recursive process by which research moves "...from idealized conception to "pock-marked" conclusion" (p. 7). The five stages outlined include: conceiving the idealized research, creating the initial research plan, becoming aware of the existence of problematic realities, designing and implementing adaptations, and producing outcomes. This continuum serves as a useful tool to understanding the approach adopted in organising this study. Reflecting on the work undertaken so far: the principal arguments presented in Chapters 1 and 2 led to the development of the conceptual framework for the study (Table 12). Based on this framework, Chapter 3 distinguished between the various research approach options, with reference to key differences between their philosophical, ontological and epistemological underpinnings (Table 13), to legitimise the decision to employ a partially mixed, concurrent, equal-status, research design (Figure 9), within an interpretive/constructivist frame (Table 14). This approach corresponds loosely with the first three steps of the Thompson et al., (2004) continuum. Steps four and five are

addressed in this chapter, ensuring, as recommended, that the story told is as real as possible, and as a consequence, as “pock-marked” as required.

Organisation of Chapter

In keeping with the order in which the research questions were posed, the first part of this chapter presents the findings from the quantitative data derived from pre- and post-test measures of children’s performance on two standardised reading tests, a reading motivation test and a reading approaches audit. This is followed by an analysis, in part two, of data from teachers’ pre- and post-intervention AfL audits that sought to determine their knowledge, skills and attitudes in relation to assessment, coupled with qualitative data derived from their individual learning logs and their group analysis of the process, content and context elements of the CPD offered. Each section begins with a statement of the research hypothesis investigated.

Pallant (2007) suggests that in studies involving human subjects, it is important to collect information regarding the key characteristics of the sample, in addition to running a number of analyses to ensure that the assumptions underpinning individual statistical tests are being met. Hence, the testing of Hypothesis 1 is prefaced by information derived from descriptive statistics regarding (a) the sample and (b) the research variables.

Characteristics of the Children’s Sample

Table 16 overleaf provides important information on key characteristics of the cohort of one hundred and seventy children who participated in this research. Organised by group, Control (C) and Experimental (E), the data detail the ratio of boys to girls in each group, the numbers of New Irish, i.e. children who immigrated to Ireland in recent years mainly from African or European countries, and the percentage of children who received additional educational support within the school in First and/or Second Class as a result of the identification of a Special Educational Need (SEN) of some kind. It should be noted that the school operated a policy that prioritised the support of children’s literacy (over maths, for example) when providing additional teaching support to children with SEN.

Table 16.

Descriptive Statistics: Control and Experimental Groups

Group	Gender		Nationality		Special Educational Needs (SEN)	
	Boys	Girls	Irish	New Irish	End 1 st Class	End 2 nd Class
Control (C)	60%	40%	80%	20%	24%	23%
Experimental (E)	56%	44%	67%	33%	43%	42%

The sample data highlighted some significant differences between the control and experimental groups; specifically, the number of children with SEN in the experimental group was almost twice that of the control group, in both First and Second Class. Furthermore, the number of New Irish children was more than 13% higher in the experimental group. These figures would suggest, at the very least, that the complexity of educational needs was consistently greater within the experimental group. An important factor regarding the SEN subgroup that needs to be borne in mind is that the constituency of this cohort changed from First to Second Class, for both the control and experimental groups, with some, but not all, of the same children appearing at both times. As will become evident, this issue had particular import for the interpretation of data relating specifically to this subgroup of children, which Black and Wiliam (1998a) referred to as the traditional low achievers. More specific information regarding the comparability of these groups emerges in the context of the various statistical tests undertaken which are fully reported later in this chapter.

Research Variables

In order to avoid the violation of assumptions underpinning the various statistical tests employed in this study, descriptive statistics on the research variables were obtained, using both frequency and descriptive statistics (Pallant, 2007, pp. 53-64). Many statistical techniques assume that the distribution of scores of the dependent variable (in respect of Hypothesis 1, the dependent variable is reading achievement) is normally distributed in the population. As employed in this context, 'normal' describes a symmetrical, bell-shaped curve, with the greatest frequency of scores in the middle, and a smaller frequency of scores towards the extremes (Robson,

2006, p. 414). Tables 17 and 18 show the histograms for the sampling distributions in this study.

Table 17.

Mary Immaculate College Reading Attainment Test Pre- and Post-Intervention Histograms for Standard Scores

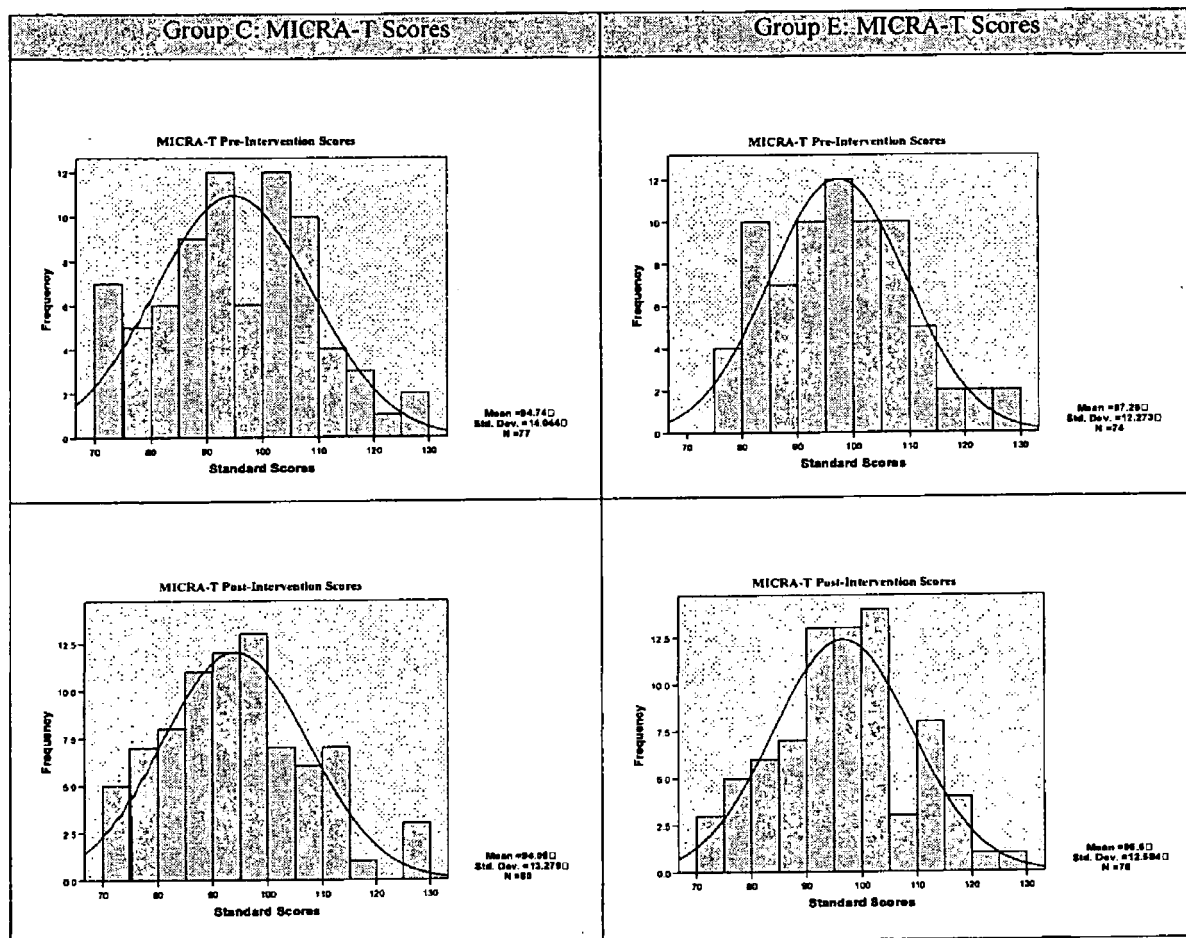
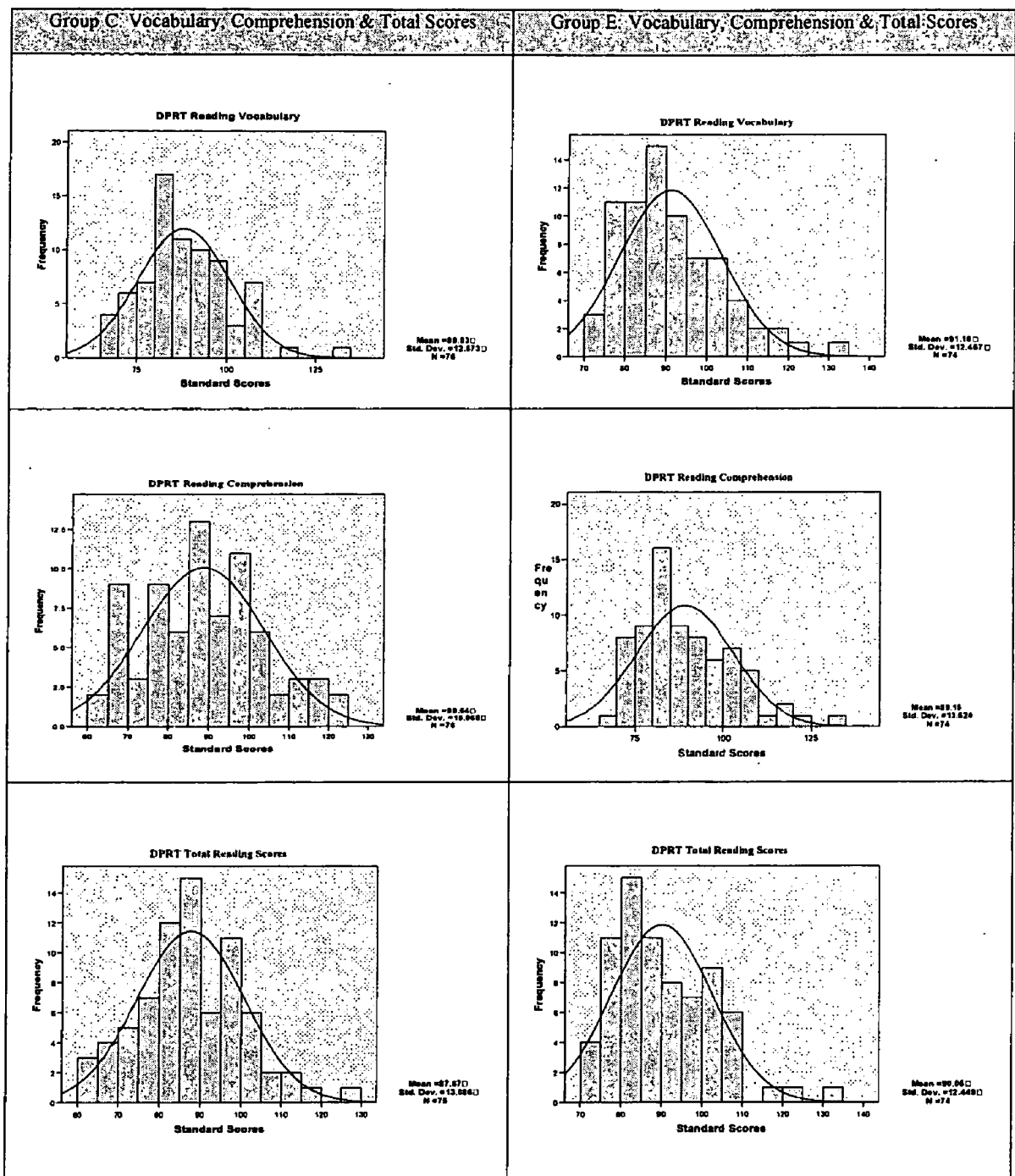


Table 18.

Drumcondra Primary Reading Test Post-Intervention Histograms for Standard Scores



As the histograms indicate, the sampling distributions for the two instruments (Mary Immaculate College Reading Attainment-Test, MICRA-T, and Drumcondra Primary Reading Test, DPRT) looked reasonably normally distributed in each case.

Moreover, the majority of Kolmogorov-Smirnov (KS) normality tests revealed non-statistical results for both groups, before and after the intervention programme. The one exception was the DPRT Vocabulary scores for both groups, which had a significant KS statistic, and therefore these distributions had to be investigated further. Closer inspection of the histograms, however, revealed that the distributions did not seriously diverge from the normal distribution. Kurtosis and skewness values were also acceptable (Tabachnick & Fidell, 2007, p. 80). Given the nature of the distributions and the satisfactory results from the statistical tests of normality, a decision was made to proceed with the use of parametric tests with respect to the analyses conducted on data pertaining to the first hypothesis.

Hypotheses and Findings

Hypothesis 1: A nine-month, school-based intervention, employing Assessment for Learning principles and practice, would make a quantitative difference (i.e. effect size) to the reading achievement of a target group of children when compared to a similar cohort not involved in the intervention.

The Statistical Package for Social Sciences (SPSS) was used to conduct a number of independent-samples t-tests to compare the standard reading scores for the control group (C) with the experimental group (E). Due to the application of multiple comparisons, a Bonferroni adjustment was made to the Alpha level, as recommended by Pallant (2007, p. 206), which is noted in the reporting of each comparison. In the case of the MICRA-T, the number of comparisons was three and for the DPRT, it was two.

The first t-test was conducted on data obtained from the standard scores of groups C and E on the MICRA-T at the end of First Class (Control: May/June, 2006; Experimental: May/June, 2007), prior to the commencement of the intervention programme. An independent-samples t-test was used to compare the mean score on a continuous variable, for two different groups. Table 19 shows the mean standard scores and standard deviations derived from these data.

Table 19.

Descriptive Statistics for MICRA-T Pre- and Post-Intervention Standard Scores

Instrument	Administered	Group	Number	Mean	Standard Deviation
MICRA-T	End 1 st Class	C	77	94.74	14.04
		E	74	97.26	12.27
MICRA-T	End 2 nd Class	C	80	94.06	13.28
		E	78	96.60	12.58

The t-tests revealed that there was no significant difference in the pre-intervention reading scores (MICRA-T) for the control group ($M = 94.75$; $SD = 14.04$) and experimental group, $M = 97.26$, $SD = 12.27$; $t(149) = -1.17$, $p = .24$ (two-tailed, adjusted for multiple comparisons).

The effect size (ES) refers to a group of indices that quantify numerically the magnitude of a treatment effect. In contrast to significance tests, ES indices are independent of sample size and may be calculated either in terms of the standardized difference between two means or as the correlation between the independent variable classification and the individual scores on the dependent variable, known as the "effect size correlation" (Rosnow & Rosenthal, 1996). In this thesis, three effect sizes are reported: *Eta squared* effect size statistics which indicate the proportion of variance in the dependent variable that is explained by the independent variable, with values ranging from 0 to 1; *Cohen's d*, which presents difference between groups in terms of standard deviation units, and *r values*, which are one of a range of possible values used in reporting difference between groups and/or matched pairs, as determined by the application of non-parametric tests to the data. According to Cohen (1994), the strength of the effect size statistics may be variously interpreted, as outlined in Table 20 overleaf.

Table 20.

Interpreting Effect Sizes (Adapted from Pallant, 2007)

Size interpretation	Eta squared	Cohen's d (standard deviation units)	r Values (non-parametric test)
Small	.01	.2	.1
Medium	.06	.5	.3
Large	.10 +	.8	.5

In this context, the magnitude of the differences in the means (mean difference = -2.52, 95% CI: -6.77 to 1.73) can be described as small (eta squared = 0.01). Following the intervention programme, no significant differences were detected in standard reading scores (MICRA-T) for the control ($M = 94.06$; $SD = 13.28$) and experimental group, $M = 96.06$, $SD = 12.58$; $t(156) = -1.23$, $p = .22$ (two-tailed, adjusted for multiple comparisons). The magnitude of the differences in the means (mean difference = -2.54, 95% CI: - 6.61 to 1.53) was again very small (eta squared = .01).

Hence, the findings from the t-test suggest that the differences between the two groups were not statistically significant at the outset and that the intervention programme did not have a significant impact, as determined statistically and with reference to effect sizes, on children's reading achievement.

In order to triangulate the findings from the independent samples test, a paired-samples t-test was undertaken, in this case, comparing the pre- and post-mean scores within each group, as indicated in Table 21.

The Paired-Samples T-Test

A paired-samples t-test, also referred to as repeated measures, was conducted to evaluate the impact of the intervention on the experimental group's reading achievement relative to that of the control group, who received no additional help. In keeping with the findings of the independent-samples t-test, no significant difference in reading attainment was found. Mean scores from Time 1 (end 1st class) for the

control group ($M = 94.10$, $SD = 12.52$) were not statistically different for those obtained at Time 2 (end 2nd class) ($M = 94.90$, $SD = 14.09$), $t(72) = -1.16$, $p < .25$ (two-tailed, adjusted for multiple comparisons). The mean increase in reading scores was .81 with a 95% confidence interval ranging from -2.20 to $.58$. The eta squared statistic (.009) indicated that there was no effect size difference between the mean scores of the control group between First and Second Class.

Table 21.

Descriptive Statistics for the MICRA-T Pre- and Post-Intervention Standard Scores

Instrument	Administered	Group	Number	Mean	Standard Deviation
MICRA-T	End 1 st Class	C	73	94.10	12.52
	End 2 nd Class	C	73	94.90	14.09
MICRA-T	End 1 st Class	E	69	97.64	12.03
	End 2 nd Class	E	69	97.42	12.45

Scores from Time 1 (end 1st class) for the experimental group ($M = 97.64$, $SD = 12.03$) were not statistically different either for those obtained at Time 2 (end 2nd class) ($M = 97.42$, $SD = 12.45$), $t(68) = .29$, $p < .78$ (two-tailed, adjusted for multiple comparisons). The mean decrease in reading scores was .23 with a 95% confidence interval ranging from -1.30 to 1.73 . The eta squared statistic (.001) indicated that there was almost no effect size difference between the mean scores of the experimental group between First and Second Class. It should be noted, however, that what is being referred to here is the relative performance of the children rather than absolute improvement. In other words, the fact that the children succeeded in maintaining their existing reading levels, albeit at a consistently lower standard than the national norm, is noteworthy.

The mean standard scores and standard deviations from the DPRT for the control and experimental groups are presented in Table 22. It should be noted that unlike the MICRA-T, the DPRT provides sub-test scores in vocabulary and comprehension as well as a total reading score. In addition, unlike the MICRA-T test, the DPRT was administered once to each group, i.e. when the children reached the

end of Second Class. Hence, there were no pre-intervention scores against which to compare reading achievements.

Table 22.

Descriptive Statistics for the Drumcondra Primary Reading Test Post-Intervention Standard Scores

Instrument	Administered	Group	Number	Mean	Standard Deviation
Drumcondra Primary Reading Test (vocabulary)	End 2 nd Class	C	76	88.03	12.67
		E	74	91.18	12.47
Drumcondra Primary Reading Test (comprehension)	End 2 nd Class	C	76	88.64	15.07
		E	74	89.16	13.62
Drumcondra Primary Reading Test (total reading score)	End 2 nd Class	C	75	87.67	13.09
		E	74	90.05	12.45

Again, there was no significant difference in scores at the end of 2nd class in the reading vocabulary scores for the control group ($M = 88.03$; $SD = 12.67$) and experimental group, $M = 91.18$; $SD = 12.47$; $t(148) = -1.53$, $p = .127$ (two tailed, adjusted for multiple comparisons) or in the reading comprehension scores ($M = 88.64$; $SD = 15.07$) and experimental group $M = 89.16$; $SD = 13.62$; $t(148) = -.22$, $p = .82$ (two-tailed, adjusted for multiple comparisons). Not surprisingly, the total reading score for the control group ($M = 87.67$; $SD = 13.09$) and experimental group, $M = 90.05$; $SD = 12.45$; $t(147) = -1.14$, $p = .26$ (two-tailed, adjusted for multiple comparisons) were not statistically significantly different. The magnitude of the differences in the means (mean difference = -2.39 , 95% CI: -6.52 to 1.75) was small (eta squared = .01).

It is important to note that while the effect size for the total test was not significant, there was a small effect size of .02 recorded for the reading vocabulary element of the test. This suggests that vocabulary may be more susceptible to change than other elements of the test. Alternatively, teachers' reports that the introduction of A/L strategies, and sharing learning intentions and sharing success criteria techniques in particular, resulted in significant additional time investment to facilitate oral discussion with the children, in relation to both the purpose and outcomes of their

learning, might also have contributed to changes in children's vocabulary. In this context, further testing - specifically in relation to children's oracy - might usefully be considered should a further study of this kind be undertaken.

In summary, then, results from all of the statistical tests conducted led to the conclusion that the first hypothesis should be rejected. In light of the recommendations of previous research that - in order to effect change - interventions using A/L need to be maintained for a minimum period of eighteen months (Gardner, 2008), this finding is not altogether surprising. However, given the observation in the literature that research on CPD rarely extends its lens of analysis beyond examining changes in instructional practices, teachers' knowledge and beliefs - to include impact on student learning (Darling-Hammond, 2000; Houcks-Horsley & Matsumoto, 1999) - it was deemed important to investigate the possibility of potential gains in children's reading.

The large numbers of children within the experimental group with SEN and the reported potential of A/L to narrow the ranges of school achievement by raising the standards of traditional low achievers – an issue highlighted by Black and Wiliam (1998a) as warranting closer investigation - led to further analysis being undertaken to determine if there were any notable changes in the reading patterns of children with SEN, before and after the intervention. In this context, an independent-samples t-test was conducted; the relevant descriptive data are reported in Table 23.

Table 23.

Descriptive Statistics for the MICRA-T Pre- and Post-Intervention Standard Scores (SEN Groups)

Instrument	Administered	Group	Number	Mean	Standard Deviation
MICRA-T	End 1 st Class	SEN - C	17	85.59	14.49
		SEN - E	32	88.66	8.78
MICRA-T	End 2 nd Class	SEN - C	18	82.00	8.77
		SEN - E	32	89.59	9.02

As indicated in Table 23, there was no significant difference in scores in the pre-intervention reading scores (MICRA-T) for the SEN control subgroup ($M = 85.59$; $SD = 14.49$) and the SEN experimental subgroup, $M = 88.66$; $SD = 8.78$; $t(47) = -.924$, $p = .36$ (two-tailed, adjusted for multiple comparisons). The magnitude of the differences in the means (mean difference = -3.07 , 95% CI: -11.24 to -1.713) was small (eta squared = $.02$). However, following the intervention programme, a significant difference was detected in standard reading scores (MICRA-T) for the control ($M = 82.00$; $SD = 8.77$) and the SEN experimental subgroup, $M = 89.59$ $SD = 9.02$; $t(48) = -2.89$, $p = .006$ (two-tailed, adjusted for multiple comparisons). The magnitude of the differences in the means (mean difference = -7.59 , 95% CI: -12.88 to -2.31) was large (eta squared = $.15$).

This finding suggests that, although the difference between the two groups was not significant at the outset, the intervention programme did have a significant impact on the SEN experimental subgroup as evidenced by the fact that their reading scores did not decrease as much as those of the SEN control subgroup. That said, it should be noted that a paired samples t-test revealed that there was no statistical difference between the MICRA-T standard scores for the SEN experimental group before, and after, the intervention ($t(27) = .78$; $p = .44$). In other words, the children maintained their reading levels although, relative to the national norm, their scores remained depressed.

Nonetheless, the finding that children with SEN in the experimental group succeeded in maintaining their reading standards to the end of Second Class is an important one, particularly in light of an earlier Irish study by Martin (1979) that reported evidence of a widening achievement gap in reading during primary schooling, with socio-economically advantaged children increasing their lead over their more disadvantaged peers. Furthermore, the relative achievement of this SEN subgroup is underscored by the fact that the tests used at Time 1 (end of First Class) and Time 2 (end of Second Class) were at different levels (from MICRA-T Level 1 to MICRA-T Level 2), which demanded more complex reading skills. That the children managed to maintain their standardized reading scores under these conditions can be considered noteworthy.

Hypothesis 2: There would be a discernible, positive impact on children's attitudes, motivation and approaches to reading.

As reported previously, the KMOFAP demonstrated that the potential of A_{FL} extends far beyond enhancing performance on traditional, standard tests to engender change in the traditional teacher-student roles to the point where "...students are not merely the objects of their teacher's behaviour, they are the animators of their own effective teaching and learning processes" (James & Pedder, 2006b, p. 28). The decision to include the Elementary Reading Attitude Survey (ERAS) and the Scaled A_{FL} Questionnaire (SA_{FLQ}), as instruments to assess alternative impacts of the intervention on children's learning, was taken precisely because they had the potential to capture changes that more traditional reading tests might fail to detect.

Due to the nature of the scores (rank order) obtained from the ERAS, no assumptions were made about the sampling distributions of the population and therefore a decision was taken to use non-parametric tests to aid the data analysis in this case, as recommended by Pallant (2007, pp. 210-231).

The Mann-Whitney *U* test is a non-parametric alternative to the t-test for independent samples. It tests for differences between two independent groups (the control and experimental) on a continuous measure (motivation to read). Rather than comparing the means of the two groups, as in the case of the t-test, the Mann-Whitney *U* Test compares median scores; it converts the scores on the continuous variable to ranks across the two groups and then evaluates whether the ranks for the two groups differ significantly. As outlined in the previous chapter, the ERAS measures children's motivation to reading according to the choices they make on a four-point scale (from 4 = happy to 1 = very unhappy) for each of ten items in two subscales (recreational and academic reading). The test yields three scores, one for each of the subscales, and a total composite score. Again, adjustments were made to the Alpha level applied due to the application of multiple comparisons (six).

Table 24.

Descriptive Statistics for the Elementary Reading Attitude Survey Post-Intervention Standard Scores

Instrument	Administered	Group	Number	Median Rank
Elementary Reading Test (recreational)	End 2 nd Class	C	76	29.00
	End 2 nd Class	E	70	28.00
Elementary Reading Test (academic)	End 2 nd Class	C	76	33.50
	End 2 nd Class	E	70	31.00
Elementary Reading Test (total score)	End 2 nd Class	C	76	61.50
	End 2 nd Class	E	70	58.50

The Mann-Whitney *U* test revealed no significant difference in the overall motivation to read levels of children from the control group ($Md = 61.50$, $n = 76$) and children from the experimental group ($Md = 58.50$, $n = 70$), $U = 2398$, $z = -1.03$, $p = .30$. In addition, the effect size derived from the test, $r = .09$, was small (Pallant, 2007, p. 223). Further analyses revealed similar findings for the subtests.

In keeping with the commitment made at the beginning of this chapter to present the findings of this research frankly, it must be noted at this point that the Elementary Reading Attitude Survey did not prove as useful or sensitive an instrument as originally anticipated, based on the recommendations for its use (McKenna et al., 1995). This comment is made in light of the *patterns* of responses made by some children, detected when inputting the raw data into SPSS (continuous 4s or 1s being marked), which appeared to suggest that some children either didn't understand, or were not sufficiently motivated, to answer the questions conscientiously. One might speculate that since children from disadvantaged homes do not have the same access to books as their more privileged peers (Eivers et al., 2004), that asking them questions about recreational reading was inappropriate. Similarly, given the literacy challenges such children face (Shiel, 2007), being requested to answer questions about their motivation to read for academic purposes might have elicited responses that reflected more their perceived efficacy than their actual motivation, factors which are intimately connected (Harlen, 2006b), as outlined in Chapter 2. However, the fact that these difficulties were not detected in advance

may also be accounted for by the fact that no pilot of this instrument took place. Although this runs contrary to recommended practice (Robson, 2006), it resulted from the restricted timeframe that presented between the identification of the school and the receipt of ethical clearance to conduct the project, which, in turn, was delayed by the need to obtain individual consent from each child's parent/guardian, in addition to individual children's assent. In this context, one can only recommend close examination of the suitability of this instrument in the Irish context in the event that it is proposed for use in a similar study, and/or that consideration might be given to the development of a more robust instrument.

As stated, the second research hypothesis anticipated that the intervention project would have a discernible, positive impact on children's motivation, attitudes and approaches to reading. The second instrument employed to test this hypothesis was the Scaled A/L Questionnaire (SA/LQ), designed by the researcher in an effort to capture any changes that might have occurred in children's use of A/L approaches before, during, and after reading as a consequence of their exposure to A/L strategies and techniques during reading lessons.

Again, no assumptions were made in relation to the distribution of the population; therefore, the Mann-Whitney *U* test, that compared the reading approaches of the control group and the experimental group at the end of Second Class, was conducted initially. It revealed statistically significant findings for each subscale, as indicated by the median rank scores reported in Table 25.

Table 25.

Descriptive Statistics for the Scaled AfL Questionnaire Post-Intervention

Instrument	Administered	Group	Number	Median Ranks
Scaled Assessment For Learning Questionnaire (before reading)	End 2 nd Class	C	76	27.00
	End 2 nd Class	E	60	32.00
Scaled Assessment For Learning Questionnaire (during reading)	End 2 nd Class	C	76	10.00
	End 2 nd Class	E	60	11.00
Scaled Assessment For Learning Questionnaire (after reading)	End 2 nd Class	C	76	16.00
	End 2 nd Class	E	60	20.00
Scaled Assessment For Learning Questionnaire (total score)	End 2 nd Class	C	76	58.00
	End 2 nd Class	E	60	66.00

Taking the findings of each Mann-Whitney *U* test in turn, it may be reported that:

- The Mann-Whitney *U* test for the *before reading* sub-scale revealed a significant difference in the use of AfL approaches to support their reading between the control group ($Md = 27$, $n = 76$) and children from the experimental group ($Md = 32$, $n = 60$), $U = 1226.5$, $z = -4.62$, $p = .00$, $r = .40$ – indicating a medium effect size in favour of the experimental group;
- The Mann-Whitney *U* test for the *during reading* sub-scale revealed a significant difference in the use of AfL approaches to support their reading between the control group ($Md = 10$, $n = 76$) and children from the experimental group ($Md = 11$, $n = 60$), $U = 1850$, $z = -1.90$, $p = .06$, $r = .16$ – indicating a small effect size in favour of the experimental group;
- The Mann-Whitney *U* test for the *after reading* sub-scale revealed a significant difference in the use of AfL approaches to support their reading from the control group ($Md = 16$, $n = 76$) and children from the experimental group ($Md = 20$, $n = 60$), $U = 1529$, $z = -3.30$, $p = .001$, $r = .28$ – indicating a small to medium effect size in favour of the experimental group;

- The Mann-Whitney U test for the SAfLQ - *total score* - revealed a significant difference in the use of AfL approaches to support their reading between the control group ($Md = 58, n = 76$) and children from the experimental group ($Md = 66, n = 60$), $U = 1529, z = -3.30, p = .001, r = .31$ - indicating a medium effect size in favour of the experimental group.

Subsequent analyses using a Wilcoxon Signed Rank Test, which is designed for use with repeated measures, i.e. when subjects are measured on two occasions or under two conditions, was conducted on the data. A non-parametric alternative to the repeated measures t-test, the Wilcoxon converts scores to ranks and compares them at Time 1 (pre-intervention) and Time 2 (post-intervention), rather than comparing mean scores. These mean rank data are presented in Table 26 and are organised by means for the subscales (before, during, and after reading) and for the total reading score for the experimental group at the end of First and Second Class.

Table 26.

Descriptive Statistics for the Scaled AfL Questionnaire Pre- and Post-Intervention

Instrument	Administered	Group	Number	Mean Rank	Standard Deviation
Scaled Assessment For Learning Questionnaire (before reading)	None	C	-	-	-
	End 1 st Class	E	80	25.68	8.89
	End 2 nd Class	E	60	31.52	6.63
Scaled Assessment For Learning Questionnaire (during reading)	None	C	-	-	-
	End 1 st Class	E	80	9.51	3.65
	End 2 nd Class	E	60	11.02	3.20
Scaled Assessment For Learning Questionnaire (after reading)	None	C	-	-	-
	End 1 st Class	E	80	17.79	6.43
	End 2 nd Class	E	60	19.08	5.62
Scaled Assessment For Learning Questionnaire (total score)	None	C	-	-	-
	End 1 st Class	E	80	57.30	16.69
	End 2 nd Class	E	60	65.25	14.21

As a comparison of each pair of means would suggest, it was not surprising to discover that the Wilcoxon Signed Rank test revealed a statistically significant difference for the experimental group from Time 1 (pre-intervention) to Time 2 (post-intervention) for each subscale and for the total score ($z = -3.65$, $p < .000$ with a large effect size, $r = .48$, in the latter case). The median score for the SAfLQ total reading increased from $MD = 57.3$ (pre-intervention) to $MD = 65.25$ (post-intervention).

As indicated in the previous chapter, given the non-standardised nature of this instrument, analyses were undertaken to determine the reliability of the scale as a whole - and the interreliability of the subscales - the outcome of which verified that it was robust (see Appendix H).

Consequently, with regard to the second hypothesis, it may be reported that while there were no changes detected in children's reading motivation, analyses suggested important and significant changes in the use by the experimental group of formative assessment approaches in support of their reading, following their exposure to a range of AfL strategies and techniques during the intervention project. Hence, the conclusion is drawn that the second hypothesis is neither rejected nor affirmed, given the contrasting findings in respect of the two elements of Hypothesis 2, i.e. children's motivation to read and children's use of AfL when reading, respectively. In turning now to examine the findings of the teachers' data, it would be anticipated that changes in children's approaches to reading would be reflected in teachers' comments about developments in teaching practices and their observations of children's work, with the latter acting as a motivation for teacher change, as suggested in the literature (Guskey, 2002; Clarke & Hollingsworth, 2002).

Hypothesis 3: Using the medium of a site-based, teacher learning community to provide a professional development programme on Assessment for Learning would have a positive impact on teachers' knowledge, skills and attitudes of/to AfL and, in turn, how they teach reading.

As outlined in Chapter 3, there were four class teachers immediately involved in the implementation of this project, with an additional teacher attending each of the meetings of the Teacher Learning Community (TLC) as the Principal's delegate, with responsibility for liaising with all parties to ensure the smooth running of the TLC. In addition, a representative of the National Council for Curriculum and Assessment (NCCA), the statutory body with responsibility for advising the Minister for Education and Science on matters relating to curriculum and assessment, attended as a participant observer. This concession was made at the request of the NCCA in the context of their provision of funding to pay substitution costs for the release of the teachers during the period of the intervention and for providing video recording facilities. The researcher facilitated all meetings of the TLC.

In keeping with a commitment given to the teachers to respect their individual identities, and, more critically, to deter from reporting comments and 'progress' on a case-by-case or class-by-class basis, a profile of the group rather than individuals is provided here, and subsequent commentaries seek to maintain the teachers' anonymity.

Group Profile

As outlined previously, the number of teachers who sought to participate in this project exceeded the numbers required; as a consequence the Principal nominated a number of people on the basis that, collectively, they would bring a range of interests and experiences to the project, which would enrich the collective experience without creating any element of competition. The profile of the group reflected this thinking. The teachers, all female, ranged in age from mid-twenties to mid-fifties; they had followed different professional pathways into teaching, one having been qualified as a National Teacher, others as Bachelors of Education, and a fourth, initially as a post-primary teacher, before doing a post-graduate degree in the UK to

qualify as a primary teacher. In keeping with their age profiles, their years of teaching experience and the extent of their career advancement varied: one teacher had taught in the host school for more than twenty years, another less than five; three held recognised positions of responsibility within the school. Of more immediate significance, perhaps, was the fact that for two teachers, this was their first experience of teaching Second Class in many years; consequently, familiarity with the curriculum at this level was uneven within the group.

Research Findings: Teachers' Data

As outlined in the previous chapter, three instruments were used to monitor and gauge changes in teachers' knowledge, skills and attitudes of/to AfL and the way in which this impacted on their teaching practices: pre- and post-intervention assessment audits (AfLAI), individual teacher's learning logs (TLL) and a group analysis by the teachers of their experiences of involvement in the teacher learning community. A fourth instrument that was pivotal to the intervention project was the use by teachers of video to record reading lessons. For the purposes of this study, video recording served two key purposes: (a) it provided evidence that the AfL strategies and techniques introduced by the researcher at the monthly meetings of the TLC were being trialled in the classrooms and (b) it facilitated collective review and critique of teacher's individual efforts to implement AfL by colleagues within the TCL (Teachers' Video Review Sheets). However, video evidence is not considered here in response to Hypothesis 3; rather, it is intended to form the basis for post-doctoral analysis of teacher-child engagement in AfL, in addition to being used for professional development purposes. This decision was taken on the basis that the literature has reported incidents of teachers withdrawing their initial consent to participate in video recording of their work, subsequent to the initiation of an intervention (Sherin, 2004). In the event that video had been the primary source of data, and such action had taken place during this project, it would have jeopardized the work. Hence, it was deemed more prudent to employ the range of instruments described, with the explicit intent of revisiting the video data subsequent to the completion of the study.

Analyses of data in this section, therefore, follows the order in which it was collected, beginning with a review of teachers' assessment audits, their reflections on the findings of their individual learning logs, and concluding with reference to their group review of the TLC. It should be noted that, as outlined in Chapter 3, an initial review of the data was undertaken by the researcher at the end of May 2008 and this information, together with the original data, was returned to each teacher for their further consideration and commentary. Hence, the findings reported underwent a number of iterations and are therefore judged to accurately represent the voices of the teachers involved. Moreover, it should be noted that teachers' comments have been reported as received, with particular emphases such as underlining of words and exclamations having been added by the teachers themselves.

Teachers' Assessment Audits: Pre- and Post-Intervention

Table 27 shows the results of each teacher's average responses to a range of questions within the eleven subscales of the AfL audit instrument (AfLAI), before and after the intervention. The mean differences on the four-point scale in responses are shown in the last column on the right in the table, with the data being organised by mean change from high to low, as scored by the individual teachers. In reviewing these data, attention is drawn to the fact that, as presented, reported gains appear to be modest and restricted to a number of key statements including Fundamental Principles, Marking and Providing Feedback, Assessment for Learning and Assessment of Learning, respectively, with very small and/or negative changes being associated with other statements. In interpreting these data, due consideration must be given to the AfL and change literature, as reviewed in Chapter 2, attesting to (1) the tendency for teachers to experiment with a limited number of AfL principles in the initial stages of exposure to AfL and (2) the non-linearity of the change process in general. Hence, the importance of inviting teachers' considered reflections on the data as reported immediately following Table 27 overleaf.

Table 27. *Teachers' Pre- and Post-Intervention AfL Audit Responses*

Sub-Categories of Statements	Pre- Intervention Mean Score	Post- Intervention Mean Score	Post- Intervention Mean Change
Teacher A			
Fundamental Principles	1.4	3.4	2.0
Marking and Providing Feedback	1.7	3.3	1.7
Assessment of Learning	1.6	3.2	1.6
Assessment for Learning	1.7	3.2	1.6
Planning Learning	2.2	3.6	1.3
Recording and Evidence	1.2	2.2	1.0
Management and Monitoring	1.6	2.4	0.8
Using Assessment Information to Monitor Progress	1.2	1.8	0.6
Transfer and Transition	1.0	1.4	0.4
Reporting to Parents and Carers	1.7	2.0	0.3
Evaluation as Part of School Improvement	1.7	2.0	0.3
Teacher B			
Fundamental Principles	1.7	3.0	1.3
Marking and Providing Feedback	1.6	2.5	0.8
Assessment of Learning	1.4	2.0	0.6
Assessment for Learning	1.8	2.3	0.6
Planning Learning	2.0	2.4	0.4
Recording and Evidence	1.6	2.0	0.4
Management and Monitoring	2.2	2.5	0.3
Using Assessment Information to Monitor Progress	1.8	1.8	0.0
Transfer and Transition	1.9	1.8	-0.1
Reporting to Parents and Carers	1.9	1.7	-0.2
Evaluation as Part of School Improvement	2.1	1.9	-0.3
Teacher C			
Fundamental Principles	0.8	4.0	3.2
Marking and Providing Feedback	2.3	3.5	1.2
Assessment of Learning	2.6	3.4	0.9
Assessment for Learning	3.5	4.0	0.5
Planning Learning	2.2	2.6	0.3
Recording and Evidence	2.9	3.2	0.3
Management and Monitoring	3.1	3.2	0.1
Using Assessment Information to Monitor Progress	2.8	2.6	-0.2
Transfer and Transition	2.0	1.8	-0.3
Reporting to Parents and Carers	3.0	2.7	-0.3
Evaluation as Part of School Improvement	3.3	2.9	-0.4
Teacher D			
Fundamental Principles	1.4	3.3	1.9
Marking and Providing Feedback	1.8	3.3	1.6
Assessment of Learning	3.7	5.2	1.5
Assessment for Learning	2.3	3.6	1.4
Planning Learning	2.0	3.3	1.3
Recording and Evidence	2.2	2.8	0.6
Management and Monitoring	2.7	3.0	0.3
Using Assessment Information to Monitor Progress	2.6	2.9	0.3
Transfer and Transition	3.2	3.4	0.2
Reporting to Parents and Carers	3.0	3.1	0.1
Evaluation as Part of School Improvement	3.4	2.2	-1.2

As acknowledged, Table 27 reports mixed results ranging from very positive to negative changes in response to which teachers were invited to reflect and respond. In reviewing their responses, a number of themes emerged; these are listed initially here before being revisited in turn:

- Teachers expressed different ideas about the accuracy of the mean changes reported, with each one for whom a negative response featured proposing that this did not reflect their experience of change in these areas, and those whose data showed positive changes endorsing the findings;
- Each teacher commented on the extent of the change that had occurred in their knowledge, skills and attitudes as a result of their involvement with the project, reflecting not only on the progress made but, more critically, on the limitations of their understanding of assessment at the outset;
- More than any other issue, teachers made reference to the extent to which giving and receiving feedback, as a consequence of sharing learning intentions and success criteria, had assumed centre-stage in the teaching and learning in their classrooms.

The quantified changes in teachers' responses to the AfLAI shown in Table 27 elicited different responses from the teachers concerned, with many expressing dissatisfaction, both with the instrument and, more particularly, the initial ratings they gave to the individual statements. In general, the initial ratings were deemed inflated for a variety of reasons, as suggested below:

I think that the change is not properly reflected as I think that in the initial audit, I didn't truly understand the statements and what AFL was really about. I think that the first audit scores should be lower.

Elaborating on this point, the teacher attributed her early scores both a lack of understanding of what was being asked and a reluctance to 'let the side down':

I was "ambitious" by putting the high scores in the first audit. I think that my understanding is greater now and perhaps I'm not as insecure or worried about revealing deficits in my practice!! (Or areas which require further development).... I feel a little disappointed that I don't think that I've been helpful enough in the completion of the first audit. I think I was overwhelmed

by it, particularly the part which refers to the whole school. I felt that I was lacking in knowledge about the whole school but felt that I would be letting everyone down by revealing low scores!! I then felt that I had to show a bit of understanding and try to show that I was at least trying to aspire to some of the statements!

Interestingly, these comments echo concerns raised by Elmore (2000), as referenced in the opening chapter of this thesis, in relation to the problem of loose coupling in schools and, specifically, the tendency to try to protect any perceived internal school difficulties from the critical glare of public scrutiny, often to the detriment of instructional leadership. Additionally, it explains the decision taken not to include reports by teachers of their perceptions of assessment practices within the school as a whole in this review. Indeed, such was the variation in teachers' perceptions, as reported by both the control and experimental class teachers, that further interrogation of the findings would have been required – including, at the very least, a series of focus interviews – before any clarity in relation to the findings could have been reached.

Another teacher whose data suggested a regression in her learning, discounted the findings, citing equally personal reasons:

Firstly, I'm a harsh marker anyway (I never give a child "excellent" on a report). But the main thing is that I didn't answer 2 questions in May 08's audit for reasons unknown (but I was unwell that day). Had I done so they'd (the final scores – *my* insertion) be a 2 or 3 which I think puts the average at 2.33 which is up 0.4. Also I think now that some of my 2s in June '07 should have been 1s, given what I now know.

Overall, this teacher felt that she had:

...Come a long way in developing knowledge, skills and attitudes in AFL and although the audit generally shows a small upward trend, it may be that my harsh ratings do not fully reflect the progress I think I have made. I believe that my teaching of reading has improved hugely through AFL and more than these small changes imply.

The comments of other teachers were more forceful in their criticism of the audit instrument, *per se*, in particular to the terminology it used, which reflected their

unfamiliarity at the pre-intervention stage with basic AfL concepts. As one teacher whose data had recorded very positive changes remarked:

The biggest change for me was in the fundamental principles section and it demonstrates how far I have come in terms of understanding AFL – last June, I didn't understand the questions asked. It was all theory and words e.g. sharing of learning intentions, pupils assessing work etc.... I had no real knowledge or understanding of what AFL was....

Subsequent comments by this teacher revealed the extent of her transition, moving from a position where assessment conjured up feelings of negativity and reduced efficacy in the context of teaching in a disadvantaged school, to one where assessment acts as a vehicle for success:

I know that there has been a sea-change in my attitudes to the very word assessment – it had a negative connotation – MICRA-T and ways to make all of us working in a disadvantaged school to feel – less. Now I see it as an integral part of teaching – involving the teacher and children working together to see where they are in their learning and to go forward.... This year has changed my day-to-day teaching and interaction with the children to an amazing extent. The whole atmosphere, ethos and learning of the children has changed. It is an amazement to me how this has happened – when did the change happen, was there a tipping point of change and when did this occur in the course of the school year?

This teacher's opening comments emphasise the significance of the findings of previous research particularly in relation to feedback and the impact on low-achieving children of receiving grades-only feedback (Butler, 1988), both on their motivation to learn and their feelings of attribution and control (Bandura, 1993a; 1993b). However, it also raises concerns about DES/NCCA policy (2004; 2007) as highlighted by Hall, Conway, Rath, Murphy, and McKeon (2008), that requires teachers to share the findings of standardised test results with parents, and more particularly, the absence of guidelines on how sensitive assessment information of this kind may be used to empower rather than disable struggling learners, their parents and teachers.

It is significant, in this context, that in reflecting on the changes they witnessed in their teaching, each of the teachers referred to changes in the nature and extent of their interactions with the children in their care and the purposes served by this communication:

Prior to AFL, I didn't communicate enough with the children. I was teaching about their learning and didn't provide an opportunity for them to communicate to me on their learning. The biggest change for me is getting the children to reflect and talk about their learning and progress on a regular basis. This has been invaluable to me and to the children. This is now part of our school day; it's embedded and the children are really enjoying it. Also, now when I assess their work, it has a purpose; I'm not just doing it for the sake of it. Children are now more aware of what they are learning, how they can achieve success and assess their own learning and set their own goals.

These findings contrast sharply with those of another Irish study, which included a focus on teachers' assessment practices and beliefs, and determined that although the teachers interviewed seemed to endorse assessment for learning for teaching purposes, the descriptions they provided of their assessment practices did not indicate core involvement by the learners themselves. As reported:

...Teachers did not talk about sharing success criteria with pupils or helping pupils become aware of how their work is judged and there was little or no emphasis on pupil self-assessment or peer assessment or ipsative assessment – features considered essential by the theoretical research on formative assessment. (Hall & Kavanagh, 2002, p. 265)

In light of these earlier observations, this research would seem to indicate that the reason teachers may not talk about formative assessment strategies is because they are either unaware of them or they confuse their existing assessment practices with what is recommended. Certainly, the following comments would suggest that this is an issue:

Before the intervention I had my own opinion, ideas and views on what assessment was. I thought that marking copies, correcting worksheets and chatting to pupils, tests... was keeping up to date with assessment so I therefore thought the aspects of recording, marking and providing feedback were ongoing in my teaching. Now I understand what feedback is – it needs to inform the children and the teacher about where they are at and how they can move forward. Before the intervention, I said that the statement Pupils are provided with opportunities to assess their own and others work was mirrored in my classroom. But what I thought that meant was children swapping copies and sheets and correcting each other's work. Now I see that this has totally new meaning for me. Pupils in my classroom now do that but in a different way – orally and give suggestions as to how to move on and progress. Feedback for my pupils now has a purpose; it has its own worth. If I knew then what I know now about feedback, I certainly wouldn't have given myself a 4.

Intervention teachers' reports of sharing information with children in this way, and the information being used by teachers and children alike to inform the next step in learning, signifies a significant departure in practices from that reported in the Irish study referred to previously. In contrast, comments made by teachers interviewed for that study led the authors to conclude that:

...Teachers place more emphasis on furnishing information that informs their teaching decisions than information that informs individual pupils' learning decisions. Adding weight to this claim that their assessments are mainly for their own teaching decisions is their approach to assessment evidence. The collection, recording and reporting of assessment evidence, though not entirely absent, did not feature strongly in their discourse of assessment. (Hall & Kavanagh, 2002, p. 266)

In the absence of formal training in AfL in Ireland, as noted by (Eivers et al., 2004), coupled with epistemological confusion in policy text (Hall, 2000), the findings reported here are hardly surprising and reinforce the urgency of the case for providing Irish teachers with extensive and sustained CPD in assessment, as recently recommended (O'Leary, 2008). It is reassuring, in this context, that as the findings of this study show, quite radical changes in knowledge, skills and attitudes can take place in a relatively short period of time (albeit with significant support), with far-reaching consequences:

Using assessment information to monitor progress – I think that it is something that I've always done informally but I think that now I'm more aware of the importance of steering children towards learning goals and "closing the gap". I think that I'm more in control of helping the children to learn and in helping them become owners of their own learning.

Teachers' Learning Logs

While teachers' commentary in response to their assessment audits was generally positive, this should not be interpreted as suggesting that involvement in the project was unchallenging. In keeping with the literature on change (Smith & Gillespie, 2007), teachers experienced a range of emotions and concerns, that were both very personal and, at times, unexpected. In general, the concerns raised in the learning logs related to the issues of time, personal- and teacher-efficacy, the

challenge of inter-personal critique (particularly on video) and the unpredictability of the change process. The central mediated support to which they attributed their success in navigating these concerns was the teacher learning community.

The issue of time is a perennial bugbear in the literature on change (Cuban, 1984; Hargreaves & Goodson, 2006). However, significantly, the time concerns raised by the intervention teachers related more to the amount of time required to come to grips with the language, strategies and techniques of formative assessment – and writing lesson plans in particular – than actually finding opportunities to undertake this kind of work, which was not entirely unexpected given their routine release from teaching to attend CPD. The nuances of their concern are evident in the contrasting commentaries that follow, with one teacher expressing concern about the project becoming too theoretical and the other accepting, in retrospect, that learning is challenging and requires a significant investment of time, effort and perseverance, and a large dollop of humour:

In the first term I had huge concerns around lesson plans and the time needed to write them. I was concerned about my understanding of WALTs (acronym for ‘we are learning to’, i.e. learning intentions – *my insertion*) and WILFs (acronym for ‘what I am looking for is’, i.e. success criteria – *my insertion*) and the project moving towards a theoretical base that has nothing to do with classroom practice.

My concerns about time diminished as I became more familiar with AFL and planning became easier. This was due to all the time we spent talking about strategies, techniques and even getting bogged down on language and minutia was part of the learning curve. So all these handouts, videos and talk were very useful Zita and that time meant that my concerns about time at the end are only that it's over and there's still more learning to do. (But you can't hold our hands forever).

While the issue of time was an important one, there were other concerns that teachers raised which were of a much more personal nature, reflecting both the literature on the kinds of change that teachers are likely to experience when involved in an initiative of this kind (Hall & Ford, 1987), and the interrelatedness of teachers' feeling of efficacy, motivation and attribution (McKinney et al., 1999). In requesting teachers to reflect on their learning logs and the nature of the concerns they expressed over time – how these evolved or stagnated, and why – teachers were provided with

some contextual information on the Concerns Based Adoption Model, and the Stages of Concern (Hall & Ford, 1987), as shown in Appendix M. As a consequence, teachers' statements made frequent reference to the stages at which they associated their concerns. The following commentary from three different teachers reflects Hall and Ford's (1987) view that personal or self-concerns dominate in the early stages of an intervention:

Yes, my concerns early on were more the fear of the unknown as such and once I moved on, my concerns and needs changed as earlier concerns were dealt with. Time was always going to be a big need – planning time, preparation and reflection. I probably started on part 3 of scale –Personal – I was my biggest concern – could I do this, am I capable of doing this project justice? Part of me was at point 2 – I wanted to learn and I was interested in this new innovation.

During the intervention period, my biggest concern was “Am I able to do this and do justice to AFL for Zita? Can I change my current style of teaching and adopt new practices?

I began very much at 3. Initially my energy and the demands concerned me as I have a demanding class and a child at home who wakes me every night.

The literature also suggests that navigating the change process is an idiosyncratic and non-linear process (van den Berg et al., 2000), that advancement and progression depends, in large part, on timely and appropriate support (Hall & Ford, 1987; Stein et al., 1988) and that, as reported by Condie et al. (2005) in relation to the LHTL project, ‘hot’ information and support from colleagues is prized far more than the ‘cold’ information of research. These themes seem to be borne out by the following comments from different teachers:

I think that I started at 3 – the personal stage of concern. I think that I have reached 6 – the collaboration stage because I have spent some time at each of the intervening stages but in some ways I feel I haven't actually moved away from stage 3. I always seem to return to my own personal needs and concerns. In other ways I am very much excited about starting a full school year (from September) using all that I've learned this year to help the children learn more and more successfully!

I'd say a part of me is now at point 5 on the scale. My concern is now how can I take my teaching and feedback to inform the next lesson; what consequences is my teaching having on their future learning? How can I

improve in order to enhance their performance in a lesson? Am I going forward and are they?

There are signs of progression in my learning from the intervention stage to now. I'm not as concerned now about myself, my confidence and ability, working and sharing with my peers, getting feedback from them and the honesty among the group was fantastic. Watching other teachers teach was invaluable – I learned so much more that I would have reading notes from a page. Also when we worked together to create WALTs and WILFs - that was extremely helpful. Overall, sharing our learning was what supported me most.

As indicated in the last statement, although some teachers were initially very fearful of inviting colleagues to watch videos of their teaching, which is not unsurprising given the “legendary autonomy” among Irish teachers (OECD, 1991) and the inevitable isolation of teaching in Ireland as elsewhere (Hindin et al., 2007; Sugrue, 2004c), the medium proved an enormous success:

I began by being concerned about myself, being “watched” while I taught and being “criticized” (now I know it’s “critiqued” or reviewed).... The idea of others watching me teach and reviewing me was daunting. I became very at ease with the video and better able to hear what others were saying in review and while I did (and do) continue to have concerns about me they are more about fine-tuning and embedding... because I am trying for more children’s talk/involvement, more meaningful lessons for children, more self and peer assessment again in a meaningful (for the children) way...

The video and watching my colleagues was a great learning experience. I would love to continue to have time out of classroom to meet my colleagues and review and plan (not 2 days but an hour would help).

While there may have been differences in the nature of teachers’ concerns, there was no ambiguity in the feedback in relation to the role played by the teacher learning community:

The biggest support was the group experience – it was a genuine safe space for complete sharing and learning. Everyone put themselves out there to help, support and share both the good and the bad without egos or personalities getting in the way. The feedback received from my lessons shared made me continue to try different ideas – to adapt and refine what I was doing. I also liked getting the background information from Zita – I found it really useful to look into my AFL files on feedback, questioning etc. to get ideas or to transfer my actual practice back to the theory and compare/contrast both.

I can't seem to over-emphasize the benefits of the Teacher Learning Communities. I've learned so much – both through others' review of my practice and through reviewing others' practice! I was initially terrified of showing the videos of my lessons and worried that I would be unable to take the criticism (albeit constructive and supportive at all times!).

The heterogeneous group was great as everybody brought something different to the table and this was a great group for affirming, accepting, comforting. It was easy to review video with this group. No element of competition intruded.

These comments seem to endorse the situative perspective adopted in this study that site-based, teacher learning communities offer a particularly rich environment in which to undertake CPD because of the immediacy of the classroom contexts as a focal point for teachers' critical review and reflection. In this context, video footage of classroom practice presented as a unique artefact to focus teachers' learning because of the opportunity it offered to revisit the complexity of the classroom after the event, thereby affording teachers the opportunity "...to develop a different kind of knowledge for teaching – knowledge not of 'what to do next' but rather, knowledge of how to interpret and reflect on classroom practice" (Sherin, 2004, p. 14). This echoes closely the view as expressed that "...learning 'to do' assessment for learning requires the development of *expertise*, not just the rote application of declarative or procedural knowledge" (Thompson & Wiliam, 2007, p. 16).

In keeping with the commitment to offer a 'warts-and-all' report on this intervention, it should be noted that there were times when the project faltered, as teachers' comments evidence:

I felt in January – a system overload and I found the February sessions were just what I needed – overview and review of all that had gone before. I feel that in the last few months the AFL methods are completely embedded. I couldn't see myself teaching any other way.

I got stuck early on with the language of writing WALTs and WILFs and now I feel that this happened because it was all new to us. Now writing WALTs and WILFs comes more freely – I don't get too hung up about it now. Also the children's vocabulary and language has improved so much that they have a greater understanding of the language used.

I don't feel stuck and I think I plodded along and maybe around April I had a small eureka moment and it all came together for me. But I think it was the plodding that got me there. At the February review I came in feeling a bit low and our discussion on the 19th February was great as we all seemed to start off feeling down and through talking we moved towards more positive feelings. TLC really was great and got us through!

What is particularly interesting from the teachers' comments in relation to the challenges cited is that no specific reference was made to the role played by the facilitator in trying to anticipate and respond, in timely fashion, to the individual and collective concerns that arose along the way. It is noteworthy in this context, that on foot of their research into the relationship between teachers' conversations and collaboration outside the classroom and their actual classroom teaching of literacy, Hindin et al. (2007, p. 372) identified the need "...to explore the role of the facilitator in teacher learning groups..." as a priority; hence, this is a point that is revisited in the next section in the context of the teachers' group response to their experience of participation in this project.

In concluding this section of the review, two final comments - both by the same teacher - are recorded that serve to reinforce the point made frequently in the research literature that "...if the teaching practices one is aiming to change are recurrent, central, and entrenched within everyday teaching and school culture, then teachers will need sustained support to change them" (Thompson & Wiliam, 2007, p. 15). Reflecting on the success of the project, it was remarked that "...the children in 2nd class this year have had a wonderful experience of reading and surely it must feed into their attitudes and understanding". However, this articulation of joy was countered with an expression of concern, not only in relation to the scalability of what had been achieved within the school but, more critically, about the feasibility of an AFL approach in general, despite all the evidence of its success:

...I would love ideas about how the philosophy of AFL could be disseminated throughout the school. I wonder about how realistic it is to have AFL in an infant classroom. I know it has been done in England but in our particular environment – DEIS Band 1 – I see difficulties.

In some ways, this comment runs contrary to expectations. For example, in light of their experiences with teachers in the KLOT programme, Thompson and

Wiliam (2007, p. 18) pointed to the fact that when teachers heard colleagues recount stories of the success enjoyed, both by themselves and their students following the introduction of AfL strategies and techniques, these stories served as “existence proofs” that the changes required were both feasible and worthwhile, spurring teachers to experiment. As reported, the “existence proofs” thereby acted to counterpoint the common lament that: “...that’s all well and good for teachers at *those* schools, but that won’t work here with the kinds of students we get at this school” (Thompson & Wiliam, 2007, p. 18). That this kind of lament was voiced by a teacher in this study, on foot of the successes reported by all of the teachers involved, highlights both the fragility of the change process and the need to ensure that the support teachers receive is sufficient to sustain their initial enthusiasm and belief in the intervention when the initial “bootstraps” of support (Thompson & Wiliam, 2007) have been removed and they must carry the mantle alone. This is very important if teachers are to feel sufficiently empowered to continue to prioritise and sustain the efforts made when faced with the inevitable challenges of the status quo within their schools, and more specifically, “...school cultures that do not easily align with the needs of sustained, school-embedded, collegial work with colleagues” (Thompson & Wiliam, 2007, p. 20).

Teachers’ Group Response to the CPD Programme

This final element of the analysis of teachers’ data considers the findings of the teachers’ independent review of the programme of professional development offered, with specific reference to the content, process and context of the work. Framed by the work of Smith and Gillespie (2007), teachers were encouraged to reflect critically on what they had experienced and record as “pock-marked” an account as they deemed fit (Thompson et al., 2004). Their review was scaffolded by the provision of information on issues that are considered within the literature to impact on the success of CPD, together with some prompt questions to facilitate their navigation of these issues (see Appendix N).

It should be noted, at this point, that although sufficient time was given to this task, the findings shared with the researcher were sparse, relative to the amount of

information derived from the other research instruments employed. This may have been because the direction given to the group was to discuss the issues liberally and record the key findings in writing thereafter. While the rationale for adopting this approach was to afford the teachers the latitude to express their opinions unimpeded by any feelings of guilt or fear of offence - particularly to the researcher - in hindsight, had the meeting been facilitated by an independent party, more information may have been recorded in relation to the issues discussed.

In light of the condensed nature of the responses given, the findings have been included in full in Appendix O, thereby affording the reader full access to the data. Rather than offering a commentary on each of the responses made by the group, a number of important themes are considered which have either not emerged in the discussion thus far or require further comment, including critically, teachers' content knowledge and the role of the facilitator in the CPD process. However, this analysis is prefaced by a brief overview of the main points raised by the teachers in their review.

As Appendix O shows, the teachers reviewed the CPD under three categories: content, process and context. In contrast to the individual reviews analysed in the previous section, the tone of this data was to some extent more critical. For example, in addition to the teachers' re-articulation of their concerns about the amount of time spent in the initial stages of the intervention and the suitability of the approach for children so young, they questioned the role played by government bodies and educational institutions in supporting their learning in this area. In addition, they made specific reference (for the first time) to the issue of content knowledge, commenting, as the teachers who participated in the Assessment is for Learning project in Scotland had before them (see Condie et al., 2005), that they learned more from themselves (in this case, their videos) than from any "AFL research/handouts" (Teachers' Independent Review, Appendix O). In addition, the group made a number of specific references to the contribution of the researcher, in her role as facilitator. The observations were generally very positive, with some interesting comments being made about the facilitator's personal mastery of AFL content and the nature of the observations made in response to observations of teaching.

These commentaries raise a number of important issues, including that addressed by Black and Wiliam (2006a), following their experience with the KMOFAP "...about the notoriously difficult problem of turning research into practice" (p. 18) and the potential role of academic researchers in supporting school-based initiatives of this kind, an issue revisited in the final chapter. More pressing issues that arise at this point, however, include (a) the potential pitfalls of the 'knowledge of practice' (Cochran-Smith & Lytle, 1999) approach to CPD, as outlined in Table 3, Chapter 2 and (b) the nature and contribution of the facilitator to the TLC.

Speculating on the challenges for teachers of engaging in independent reviews of their own practice, Sugrue (2003) makes the point that it is important to recognise that:

...There are pitfalls to practitioner-focused inquiry, sometimes disparagingly referred to as a 'pooling of ignorance' rather than a building of expertise.... Espousal of this approach, therefore, needs to avoid exclusive focus on teachers' craft knowledge. (p. 23)

The point raised here is that teachers' ability to undertake peer-to-peer analysis, particularly in AfL - which is premised on the idea of closing the gap between students' current levels of attainment and the next step in their learning - rests not just on the degree to which they feel free to give and receive feedback to colleagues but, at least as critically, on teachers' underlying content knowledge. Hence, the observation by Grossman et al. (2001) that "...learning from colleagues requires both a shift in perspectives and the ability to listen hard to other adults, especially as these adults struggle to formulate thoughts in response to challenging intellectual content" (p. 993).

It is significant in this regard that distinctions have been drawn in the literature between the various categories of knowledge that teachers require, including subject matter knowledge, pedagogical knowledge and pedagogical content knowledge (Grossman, Wilson, & Shulman, 1989). Subject matter knowledge, as the term suggests, refers to the depth and breadth of knowledge in a particular curricular content area (in this case, literacy and the teaching of reading), and the concepts, facts, rules and conventions on which a teacher draws to teach. Pedagogical knowledge, that includes knowledge about learning theory, cognitive and affective

development is more generic in nature, and supports teachers more generally in their day-to-day decision-making and interactions with all students. This contrasts with pedagogical content knowledge, which is content/subject specific and serves to inform teachers' pedagogical decisions about instruction and learning in relation to the individual subject areas, including what should be taught, in what sequence, using what methodologies and media.

Interestingly, while there is little disagreement within the research community in relation to the importance of aiming to develop teachers' content knowledge as part of CPD (Shulman, 1986; Smith & Gillespie, 2007), what happens when teachers - with varying degrees of expertise - engage in "practitioner-focused inquiry" (Sugrue, 2003, p. 23) with a view to developing 'knowledge of practice' (Cochran-Smith & Lytle, 1999) is less clear. Recent remarks by researchers involved in the KLOT project are noteworthy:

This issue of deficiencies in teachers' subject matter knowledge raises a question for the model we describe: are there limits to the effectiveness of teacher learning communities focused on assessment for learning in transforming teacher practice, given pre-existing limits on teachers' subject matter knowledge? We do not have a definitive answer to this question, though we can report that we have repeatedly observed groups of teachers improve their pedagogical practice, even when no teacher in the group has had strong content knowledge. (Thompson & Wiliam, 2007, p. 18)

Moreover, Thompson and Wiliam (2007) argue that because teachers are centrally engaged in the same learning voyage as their colleagues, this affords them a unique insight into the dynamics of the change process, and, as a consequence, the nature of the dilemmas presenting and supports required. Hence, the belief that "...teachers themselves can provide effective leadership for their peers..." provided that they do not "...assume an "extra" expertise just because they are in the role of facilitator and advocate for the teacher learning community" (Thompson & Wiliam, 2007, p. 18), and receive some external guidance on how to mediate a group.

In responding to these views, it is necessary to consider the role adopted by the researcher in this study, who acted as facilitator of the TLC for the duration of the project. In many ways, the description of the "legitimate peripheral participant", (Thompson & Wiliam, 2007, p. 19), seems fairly appropriate. Arguing from the

premise that, to be a fully-fledged member of a teacher learning community, a participant must be attempting to change his/her own practice, Thompson and Wiliam (2007) suggest that provided participants "...recognize, and accept, their peripherality, they can be of substantial help to the community, brokering ideas, acting as advocates, and facilitating the community's learning" (p. 19). Certainly, the experience of the researcher would resonate with the notion that it is very difficult to participate fully in a TLC when one is not in fact a member of the teaching community, *per se*. However, as this researcher learned, mediation of the role of facilitator is contingent on the ability to draw both on theoretical knowledge of AfL and incremental learning in relation to the application of strategies and techniques. Hence, the decision was made very early on by the researcher to trial any practices being recommended to the teachers during meetings of the TLC in her own context, albeit this was at third level and in the context of teacher education. Given the "deceptively simple" (Black & Wiliam, 2006b) nature of AfL - in retrospect - it would be hard to conceive of a situation in which this work could have been undertaken in the absence of this developing knowledge because it would not only have jeopardised the credibility of the researcher as guide to the group, but would also have seriously undermined her potential to empathise with the teachers in the context of the challenges that presented as they struggled to implement AfL.

Two related points need to be made at this juncture. First, reflecting on earlier views expressed in relation to the impact of competencies in content knowledge on the development of AfL skills (Thompson & Wiliam, 2007), it is noteworthy that, at times, gaps in teachers' understanding of core elements of the reading process, including, for example, how to teach phonics, certainly threatened to hamper progress. Given that the concept of 'closing the gap' on children's learning is premised on teachers' ability to diagnose - with speed and acuity - individual zones of proximal development (Vygotsky, 1978) across a range of subject areas at primary level, in order to scaffold the student's next learning step, it is hardly surprising that hiatuses in knowledge emerged. However, rather than expecting this issue to be resolved through on-going, internal debate amongst the teachers, the experience of this project would suggest that, having highlighted the need, it presented a further opportunity for focused CPD, that would most likely require an injection of specialist guidance and support.

The second and final point to be made in this context relates back to the role of the facilitator. Although Thompson and Wiliam (2007) suggest that teachers can assume the role of facilitation with relative ease, provided they receive some kind of reassurance in relation to the expectations associated with the role and some external support, the experience of this project would suggest that, in the absence of consistent modelling and the development of - what might be termed - situational artefacts (Borko, 2004), teachers are extremely reticent to participate in any kind of genuine critique of their colleagues' work. The key artefact developed to support teachers during this project was an adaptation of the P.E.N. (Prompt/Error/Next) AfL technique to the P.I.N. (Prompt/Improvement/Next), as shown in Appendix P. Prior to its introduction, the degree to which teachers engaged in critical review was minimal, with the task falling almost exclusively on the researcher/facilitator. However, its introduction acted as a *de facto* instrument to apply, in alternative measures, the pressure and support deemed to be required to bring about changes in teaching practices (Wiliam, 2008).

Consideration of these findings would suggest the need for further review, both of the role and skills-base of facilitators, and teachers' content knowledge, if consideration were being given to extending the model of site-based, TLCs to develop teachers' knowledge, skills and attitudes of/to formative assessment. However, this is not to detract in any way from the potential which this approach offers; indeed, as teachers' comments indicate, this is a model of very high leverage.

Before concluding this section, it is important to note teachers' comments regarding the context in which the CPD took place, and specifically their reiteration of the views expressed previously about the unqualified value of sustained, school-based learning, as recorded in Appendix P.

In contrast to Hypotheses 1 and 2 then, analyses and review of the data led to the conclusion being drawn that the third hypothesis should be affirmed, despite the apparently equivocal nature of the data reported in Table 27. Moreover, as teachers' reflections and commentary evidenced, the use of a TLC provided a suitable medium to support the development of their knowledge, skills and attitudes of/to AfL and concomitant changes in their classroom practices.

Summary

This chapter provided a detailed description of the results from the analyses undertaken in respect of three research hypotheses that sought to determine the impact of a one-year intervention project in A/L on a cohort of children and their teachers. The final chapter reflects on these findings and the limitations of the work, before making a number of recommendations that relate to research, policy and practice.

CHAPTER 5

CONCLUSIONS AND RECOMMENDATIONS

Introduction

This thesis examined the potential of a teacher learning community to bring about changes in teachers' understanding and use of Assessment *for* Learning, and the extent to which these changes would impact on the reading achievement of a cohort of children in a disadvantaged school. Arguing that Irish education is currently witnessing a fundamental shift in focus from an emphasis on curriculum revision, professional development and policy implementation to one of review, the case was made that, despite the significant progress enjoyed in recent decades, there are perennial difficulties that require urgent attention, including the problem of persistently low levels of literacy among vulnerable Irish populations. A selective review of international literature on student achievement pointed to the pivotal role played by teachers - second only to demographic variables - in determining children's academic success at school. However, it was also acknowledged that instructional influence is subject to the 'dilution effect' of other variables including teachers' access to, and experience of, sustained professional development. In turn, this raised issues about the nature of CPD, and the contribution played by evolving theories of learning in aiding understanding of the factors that influence its success.

The catalytic idea that drove the work of this thesis, however, was the potential of formative assessment or AfL, reported by previous international research, to effect significant improvements in children's learning, and traditional low-achievers in particular. A selective review of literature highlighted how the early promise of quantifiable effect-size improvements in students' learning, following exposure to the judicious employment of a range of AfL strategies and techniques, motivated a diversity of follow-up studies focused variously on developing practical classroom strategies, and developing financially-viable models of CPD that would maintain the fidelity of the policy message when taken to scale. However, the literature reviewed, both in respect of AfL and CPD, indicated that many questions remain unanswered, especially (a) how AfL might act as a lever to level out the

achievement gap between traditional high and low achievers and (b) how the potential of site-based CPD, mediated through a teacher learning community (TLC), could provide teachers with the professional support required to stimulate and sustain radical changes in teaching, learning and assessment practices. Against this background, the research project described in this thesis was conceived and implemented in accordance with an interpretive/constructivist paradigm that facilitated the adoption of a mixed methods approach, in response to three hypotheses. As described, the study employed a quasi-experimental design to test the first two hypotheses and a qualitative approach with respect to the third. In light of the extensive report in the previous chapter of the findings of the analyses conducted, only selective highlights are reiterated at this point.

Research Hypotheses and Findings

Hypothesis 1: A nine-month, school-based intervention, employing Assessment for Learning principles and practice, would make a quantitative difference (i.e. effect size) to the reading achievement of a target group of children when compared to a similar cohort not involved in the intervention.

The application of a series of statistical tests to the study led to the conclusion that the first hypothesis should be rejected and that the effect size differences between the groups were small. However, comparative analyses of the reading standards of children with SEN in the control and experimental groups suggested that although the differences between the two groups were not significant at the outset, the intervention programme did have a significant impact on the SEN experimental subgroup, evidenced by the fact that their reading scores did not decrease to the same extent as those of the SEN control subgroup. That said, it was noted that there was no statistical difference between the test scores for the SEN experimental group before and after the intervention. These findings point to the need for further exploratory research – perhaps with closer examination of the work of Fuchs and Fuchs (1986) and Bergan et al. (1991), whose work also focused on disadvantaged communities (see Table 8, p. 66) - particularly in relation to the suitability of existing measurement tools, such as standardised tests, to measure subtle changes in children's learning that

may occur after a short intervention period. In light of the experimental class teachers' anecdotal reports that the children's oral language developed significantly as a result of engaging in classroom discussion on learning intentions, success criteria and providing oral feedback to each other, there is a clear need to ensure that such evidence is not lost by virtue of the bluntness of the research instruments employed.

Hypothesis 2: There would be a discernible, positive impact on children's attitudes, motivation and approaches to reading.

The analyses of data conducted using non-parametric statistical tests suggested that while there were no detectable changes in children's attitudes or motivation to reading, significant changes in the use by the experimental group of formative assessment approaches in support of their reading, following their exposure to a range of AfL strategies and techniques during the intervention project, were found. Hence, the conclusion was drawn that the second hypothesis could be neither fully rejected nor affirmed, signalling the need for further research and the need to develop instruments with sufficient design-sensitivity to capture subtle changes in children's attitudes, motivation and approaches to reading.

Hypothesis 3: Using the medium of a site-based, teacher learning community to provide a professional development programme on Assessment for Learning would have a positive impact on teachers' knowledge, skills and attitudes of/to AfL and, in turn, how they teach reading.

As anticipated, the reported changes in children's approaches to reading were reflected in changes in teachers' pre- and post-intervention responses to an AfL audit, together with data from Learning Logs and the group review of the CPD offered, indicating overall positive developments in teachers' knowledge, skills and attitudes to AfL. Among the key findings to emerge from these data were the following:

- In keeping with the findings of previous Irish research (DES, 2005a; 2005b; 2005c; Hall, 2000; NCCA, 2005a; 2005b; 2005c) participating teachers lacked baseline understanding of the principles of formative assessment;

consequently, their awareness of deficits in pre-intervention classroom assessment practices was initially limited;

- During the course of the intervention, as a consequence of participation in a TLC characterised by trust and mutual respect, teachers embraced the use of video as a mediating artefact to facilitate critical review of practice. As reflected in individual Learning Logs, despite the idiosyncratic, non-linear and highly individual nature of the learning experiences, teachers progressed from initial pre-occupation with self-concerns to consider higher order issues, such as how to exploit the teaching-learning process to maximise student engagement and achievement. In this context, specific reference was made to the sharing of learning intentions, success criteria and formative feedback to close the gap on children's learning;
- Teachers prized learning derived from critical peer-review of video over research-based input and questioned the roles played, to-date, by academic and DES agencies in guiding teachers' implementation of AfL;
- Despite the reported success achieved, concerns regarding the suitability of AfL as a key pedagogical approach to teaching young children in a severely disadvantaged context persisted. One of the primary challenges in undertaking an intervention of this kind is that, as indicated in Chapter 2, there is a notable absence of international research on the impact of AfL on very young children and/or children from disadvantaged communities. In light of Black and Wiliam's (2006a) observation that teachers' willingness to experiment with AfL is influenced by knowledge of colleagues who face similar teaching challenges and yet manage to implement AfL with success, this finding is not altogether unexpected.

Notwithstanding these concerns, the conclusion drawn was that the third hypothesis should be affirmed.

Limitations of the Study

Reflecting on these findings, and the study generally, served to highlight the limitations of the work as well as potential avenues for further research. Beginning

with the limitations of the study, the following key issues are acknowledged and should be borne in mind when evaluating the research conclusions:

- This was a project of modest proportions. It involved a small sample of children and teachers, in a very specific context, using a non-randomised design, over a timescale that research suggests is insufficient to bring about quantifiable changes in children's learning. Moreover, by virtue of the cohort of children included, and the fact that they transferred to a senior school at the end of the project, the potential to document the extent to which exposure to this intervention might influence the learning achievements of this group overtime was truncated;
- The design features of the project were compromised by (a) the absence of a pilot phase to test the proposed research instruments and (b) the instruments used to capture changes in children's learning and to audit teachers' assessment literacy and post-intervention review of the group experience of CPD. As a consequence, one might argue that aspects of participants' learning may not have been detected. As previously noted, the practice of employing national standardised tests as a mechanism of measuring developments in children's learning achievements has been established in the A/L literature (e.g. KMOFAP, Black & Wiliam, 2006a). However, as remarked previously in this chapter, the findings of this study suggest that, even though a number of reliable tests may be used, modest, but important gains in children's learning may go undetected because the instruments used are not sufficiently refined (e.g. oracy, see Table 22, p. 132);
- Although video played a key role in mediating teachers' learning in this study, its potential as a source of rich data to inform a wider audience of the classroom-based challenges which teachers and children face in adapting to A/L was lost, by virtue of the decision to exclude its analysis as a consequence of the short timeframe of the project. In addition, the opportunity to formally triangulate these data with other findings of the study was not exploited;
- It might also be suggested that teachers' mastery of A/L may have been compromised by the fact that the researcher had neither professional training nor practical expertise in formative assessment prior to the intervention, and

was depending on the findings of research and experimentation with strategies and techniques in her own teaching at third level, to support her in guiding the teachers effectively;

- Finally, an intervention of this kind is time-consuming and expensive, and while it can legitimately claim elements of success, arguably, it has raised as many questions as it has answered. Specifically, the issue of how this particular model of CPD, which was judged so highly, could be leveraged to inform a research and policy agenda in formative assessment, that would begin to address the immediate need for quality CPD - while also feeding into a medium-to-long-term agenda of national reform in assessment - requires careful consideration.

By way of response to some of the limitations outlined, four sets of recommendations are offered with the intention of signposting the implications of the study and how they may be used to progress research, policy and practice in this area. In keeping with the attempt made throughout this work to anchor the conceptual framework and research design firmly within the parameters of current research literature and research paradigms, these recommendations are also conceived within these frames.

Research and Policy Implications and Recommendations

Recommendation 1: Replicate the mixed-methods research design employed in this study in a number of sites around the country, serving similar populations of children, with a view to generating patterns of research findings that could guide more broad-scale experimentation. In turn, use the research data derived from these initial research trials to inform a review of assessment policy in Ireland.

As discussed in Chapter 3, the epistemological and methodological approaches of positivist/empirical research differ fundamentally to those of interpretivism/constructivism, not least in relation to generalisability. Whereas positivists and empiricists adopt the view that, in every aspect of existence, there are essential, independently-observable laws that may be abstracted and generalised sufficiently to have broad application in a range of situations, interpretivists take the

polarised view that social meaning is contextualised, fluid and heterogeneous. Hence, the belief that an “insider” perspective is required if the complexity and variability within the phenomena being studied are to be fully understood. Moreover, interpretivists propose that rich data yield insights, which, although not generalisable in scientific terms, are transferable to individuals in comparable situations.

It is of some significance in this context that the historical schism between quantitative and qualitative research seems to parallel the shift, described in the second chapter of this thesis, from a 20th century paradigm of social efficiency curriculum and measurement - construed in terms of summative assessment against normative expectations - to an emerging paradigm, promoting formative assessment, social-constructivist and socio-cultural theories of learning. This raises two challenges. On the one hand, there is a need to avoid ‘either-or’ standoffs, within both research and assessment communities. Indeed, as Broadfoot and Black (2004) argue: “...if formative assessment is to prosper, initiatives aimed at supporting a positive link between formative and summative work are sorely needed” (p. 17). Second, for this to be achieved, mixed methods approaches to research in assessment will be required.

In this context, it is proposed that the challenges facing the educational community in Ireland at this time, outlined in the introductory chapter, in many ways echo those that stimulated Black and Wiliam’s initiation of the KMOFAP almost a decade ago. Specifically, the need for research to establish Irish-based empirical evidence of the efficacy of assessment in general, and A/L in particular, remains, and this is paralleled by a requirement for multiple, small-scale studies, exploring site-based interpretations of the research recommendations on A/L, by Irish teachers. In keeping with international calls for a multifaceted research agenda, as an essential component of systematic experimentation, and in light of Black’s observation (2000, p. 416) (as noted in Chapter 3) of the “guerrilla force” response of researchers to collaborative enquiry, the logic behind this recommendation is that:

By consulting expert knowledge, attempting to implement novel programs on small scales, and making preliminary (nonrandomized) assessments, we can determine which interventions to discard or refine.... The synthesis of research from a variety of methods conducted at different scales ought to be a

prerequisite for the construction of a large-scale randomized field trial.
(Raudenbush, 2005, p. 29)

An interpretive approach has the potential to facilitate this kind of foundation-level inquiry. Assuming progress in formative assessment to be linked to understanding how teachers interpret assessment generally, its role in the teaching and learning process and the supports they need to risk changing time-honoured practices to implement the ideas as intended, in-depth investigation in a selection of carefully chosen schools - e.g. those with Band I DEIS categorisation initially - would yield data that could be generalised across this stratum of the educational community. Furthermore, if research of this nature were pursued to the point of "saturation" (as understood in interpretivist terms) when some general patterns began to emerge - notwithstanding the variability across individuals and social contexts - findings from one context or population could be generalised to comparable contexts or populations.

Marrying the data derived from interpretive research of this kind could be used, in turn, as a yardstick against which to critique assessment policy documents, such as the recently published guidelines on assessment (NCCA, 2007), and professional development opportunities, as well as providing a method of triangulating data from more frequent research methods, such as national surveys. Of course, such an approach would be contingent on good will and, more particularly, on the sustained participation by representatives of the research, policy making and practitioner communities alike which, in turn, would necessitate the successful mediation of traditional boundaries and domains. Given the firmly entrenched, traditionally 'balkanised' affiliations of specific interest groups (Hargreaves, 1994), and notwithstanding the espousal/rhetoric of a partnership approach to Irish education (Sugrue, 2004) outlined in the introductory chapter, it is argued that much of the existing literature on learning has been framed within 'a set of binaries', that serve to demarcate and separate domains from one another. As Edwards argues (2005):

Each of these binaries identifies that learning is occurring across a range of domains and sites, but that this learning is in some senses situated or contextualized. The range of learning contexts may therefore be extended, along with what can be identified as learning. However, their very situatedness and pedagogical approaches that assume domains to be discrete

... mean that learning from one site is not necessarily realized as a resource in other sites by either teachers or learners. (Para. 11)

Hence, as an extension of the iterative call in this thesis for site-based, constructivist and socio-cultural approaches to teacher-learning, a second recommendation is made in response to the challenges which the bifurcation of knowledge, implicit in the demarcation of domains of learning, presents.

Recommendation 2: Explore further the potential of TLCs as a model of site-based CPD, through the judicious use of video as a boundary object, to motivate and sustain inter-agency participation.

Recent work by Tsui and Law (2007) is informative in this regard. Arguing that issues of globalization and lifelong learning force educators to cross community boundaries if they are to engage in collective knowledge generation, they investigated the learning afforded participants in a school–university partnership mediated by lesson study. They concluded that learning was transformed from helping student teachers learn to teach, into learning for all participants, including university tutors and teacher mentors.

Their study was rooted in the work of Star (1989) who conceived of the concept of ‘boundary crossing’ from one organisation or domain to another by employing ‘boundary tools’, that, like a blackboard, would ‘sit in the middle of a group of actors with divergent viewpoints’ (p. 46). As depicted in Figure 10 overleaf, boundary-crossings typically involves journeying into unfamiliar territories or ‘boundary zones’ (denoted by number 1). A boundary zone represents “...a hybrid, polycontextual, multi-voiced and multi-scripted context... where it is possible to extend the object of each activity system and to create a shared object between them. In that way, the activity itself is reorganized, resulting in new opportunities for learning” (Tuomi-Gröhn & Engeström, 2003, p. 5). These zones are frequently viewed as areas of potential conflict or competing discourses; hence, the need for (a) someone to broker the ‘crossing’ by initiating communication between the independent parties and (b) support ‘cognitive retooling’ by developing artefacts - such as diaries, video and/or lesson plans - that would be shared and fully understood

within the boundary zone, but without the requirement for full understanding of the independent context of each party' usage.

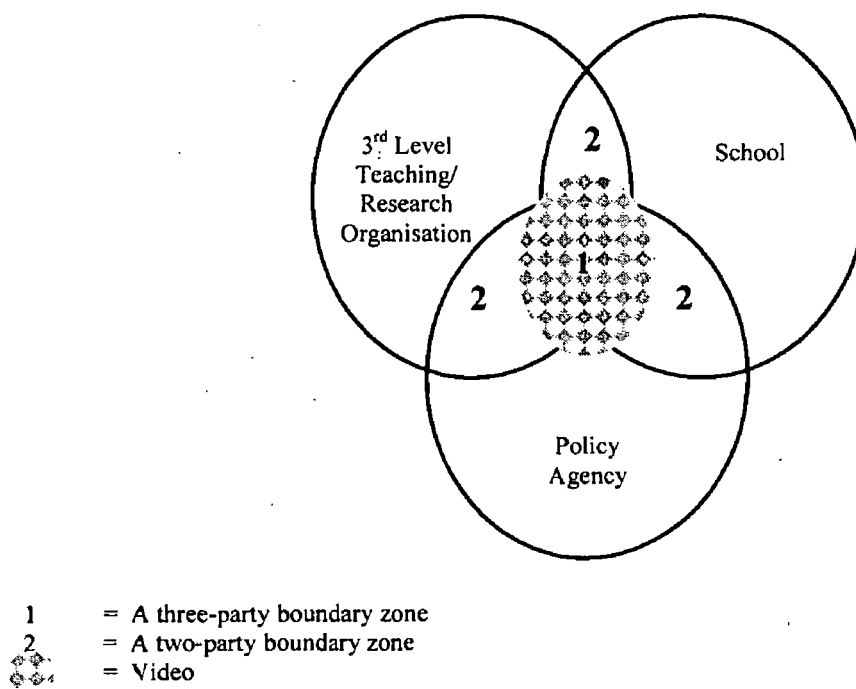


Figure 10.

From Balkanisation to Boundary Crossing - Video as a Boundary Object

Notwithstanding the fact that video analysis was not formally undertaken as part of this study, the intervention did provide insight into the potential of video as a relatively simple, yet potentially highly effective, mechanism to challenge participants to look afresh at their long-standing practices and assumptions, as suggested by Marcos and Tillema (2006) in their call for research on Walking the Walk (Table 5, p. 55). Moreover, the introduction of a TLC, and the brokering by the researcher of video as the primary artifact or boundary object to mediate teachers' experience of implementing AfL in their classrooms - in keeping with Borko's (2004) analysis of phase one research on CPD from a situated perspective (Table 4, p. 45) - transformed the traditional experience of CPD from knowledge *for*, or *in*, practice to knowledge *of* practice (Cochran-Smith & Lytle, 2002). More, crucially, the opportunity for professional development extended to all participants, including the researcher - and by default, the undergraduate and post-graduate practitioners whom she teaches - thereby exemplifying how "...both the fuzzy social differentiations that develop

between groups within the school, and the clearer borders that separate the school's members from those in the community and in other schools..." (Stoll & Louis, 2007, p. 4) might be blurred.

Recommendation 3: It is recommended that any programme of CPD in formative assessment - in accordance with developments in conceptualizations of professional development, change models and learning theory (Table 3, p. 39) - would incorporate an explicit focus on teachers' foundational understanding of subject and pedagogical content knowledge, as well as recognition of the need to consider existing theories of learning.

As argued previously, expertise in formative assessment is closely allied to, if not contingent on, teachers' ability to draw on tacit content and pedagogical knowledge in order to respond with acuity to children's evolving learning needs, as well as a renegotiation both of the teacher's and student's role. In line with other researchers (e.g. Borko, 2004; Timperley & Alton-Lee, 2008), and further to their work on the KMOFAP, Black and Wiliam (2006a) identified "...the effect on practice of the content knowledge, and the pedagogical content knowledge, that teachers deploy in particular school subjects..." including "...the way in which these resources underlie each teachers' composition and presentation of the learning work, and the interpretative frameworks that they use in responding to the evidence provided by feedback from students..." (p. 24), as issues requiring further exploration. As indicated in the previous chapter, notwithstanding the contribution made by the KLOT project (Wiliam, 2008), there is still considerable uncertainty regarding the impact of teachers' content and pedagogical knowledge on their mediation of A/L. What is not in dispute, however, is the complexity of the challenge that faces teachers, in Ireland as elsewhere as signaled in Chapter 1, in trying to draw – almost instantaneously – on internalized tacit knowledge in order to make minute-by-minute decisions to guide children's learning. As expressed by Thompson and Goe (2006):

An expert in assessment for learning is able to rapidly note essential details of the complex social and psychological situation of a lesson (especially the state of the students' learning), while disregarding distracting, yet non-essential details. She is then able to swiftly compare that situation with her intended goals for the lesson, her knowledge of the content being taught, her

developmental knowledge of students in general and these students in particular, and other relevant schema. Guided by the results of these comparisons, she then selects her next instructional move from a wide array of options – most well-rehearsed, some less familiar, and some invented on the spot, such that these next steps address the students' immediate learning needs in real time. (p. 4)

In the context of the research evidence presented in the opening chapter of this thesis, pointing to the paucity of research in the Irish context on teaching and learning in disadvantaged schools (Conway, 2002) and, more particularly, the qualitative difference in pedagogical practices in such schools (Oakes & Lipton, 1999; Shiel et al., 1996), efforts to support teachers' use of formative assessment would have to take cognizance of these issues as well as models of, and factors influencing, teacher change, as outlined in Chapter 2. Indeed, as Black and Wiliam (2006a) acknowledged when attempting to set out a theory of formative assessment, any effort to improve teacher-student interaction, if it is to accord with the *spirit* and not just the *letter* of AfL as distinguished previously in this thesis, will necessitate changes both in teachers' roles, those adopted by their students, and, in turn, reconsideration of theories and models of learning. Viewed in this way, like video, AfL may be reconceived as a boundary object, albeit not holding communities together, but as a mechanism by, and through, which teachers and students renegotiate and co-construct the *what*, *how* and *why* of learning and teaching.

Given the complexity of the challenges involved, notably the epistemological confusion linked to the absence of robust and sustained policy analysis signposted in the discussion of the contextualization of this study, there is a requirement to streamline the existing range of policy texts (e.g. Revised Curriculum Guidelines; NCCA Assessment Guidelines) and support materials (e.g. The Drumcondra Primary Curriculum Profiles; First Steps) available to school – all of which contain assessment recommendations. This would provide a mechanism both to support teachers' planning for AfL and to maintain a clear policy message in relation to what teachers are expected to do and how this might be achieved efficiently and effectively.

Recommendation 4: Give careful consideration both to the suitability of existing tests to support the development and promotion of formative assessment approaches in schools and the need for alternative assessment instruments that will command the same level of credibility as those traditionally used for summative purposes.

Reflecting Matsumura and Pascal's (2003) position that the quality of classroom instruction has remained a "black box" due to the absence of classroom-focused assessment tools, this study has indicated that there is a dearth of research instruments in Ireland and, indeed, internationally, to gauge developments both of teachers' assessment literacy and children's achievements in learning, following exposure to AfL. One of the potential benefits of engagement by researchers in a research agenda along the lines of that proposed previously in this chapter would be an opportunity to design and pilot such instruments in conference with teachers and pupils.

As alluded to previously, the current absence of a range of tried and tested research instruments, with proved effectiveness in detecting subtle changes in children's learning, before, during and after exposure to AfL strategies and techniques - particularly after a short period of time - may have limited the potential of this study. One might argue that this is not altogether unexpected given that current interest in AfL is often attributed to the Black and Wiliam review of 1998a, the evidence base of which was restricted to quantitative studies of effect size changes in controlled experiments. However, if, as Wiliam (2007) argues, currently there is no evidence to support the use of long-cycle 'formative' assessment, with improvements being linked to 'minute-by-minute, day-by-day' assessment cycles, this would raise questions about the employment of traditional summative assessment tests (e.g. MICRA-T; DPRT) to capture short-cycle learning gains. This issue is part of the bigger challenge outlined in Chapter 2, with reference to the work of Shepard (2000), regarding the current dissonance in beliefs and understanding about the nature of teaching and learning, on the one hand, and assessment and measurement, on the other. Indeed, as indicated in Table 7 (p. 62), there is an incompatibility between the growing espousal of a socio-constructivist model of teaching and learning and scientific measurement, which, as noted by James (2006a), may be explained by the fact that assessment practices are not always aligned with developments in learning

theory. Of course, there are deep-seated reasons why standardised tests continue to hold sway in the measurement of children's learning; as Wiliam (2006b) noted in the US context, "... the expectation of high reliability and objectivity in the assessment of students' learning within a culture of accountability and litigation when things go wrong, has tended to deflect policy developments from any consideration of improving learning through assessment" (p. 169).

In making this recommendation for the development of alternative assessment tools, then, it is readily acknowledged that such assessment instruments need to be both highly-sensitive to subtle changes in teaching and learning, while also commanding the respect of researchers, policy-makers and teachers alike, by virtue of their trustworthiness. The key point, however, is that these assessments need to focus on learning, to capture the *spirit* of AfL, understood in terms of changes in children's understanding of what they are learning and why, and their motivation to engage in tasks, evidenced, for example, by the nature, extent and quality of child to teacher, and child-to-child interactions, as the tenth principle of AfL articulated by the ARG (2002) holds. In other words, the assessments need to reflect not just what children learn, but why and how and to what end, so that what is valued as learning is clearly reflected by the assessments.

Epilogue

One might argue that countries like Ireland, that tend to follow rather than herald international policy reforms, have the advantage of being able to learn from the experiences of the reform pioneers. In the case of standards-based reform, this could prove a wise course of action given the view of many eminent education commentators that increased performativity and accountability pose, not just issue-related problems about data analysis, warehousing, 'teaching to the test' and so forth, but challenges of the deepest and most fundamental kind about how schools are organised and led (Ball, 2003; Elmore, 2000; Lyotard, 1984). Moreover, it has been argued that as school systems try to operate in increasingly performance and accountability-driven environments, they need to learn "...not just new ways of doing things, but very different ways of thinking about the purposes of their work, and the skills and knowledge that go with those purposes" (Elmore, 2000, p. 35).

Against this backdrop, the publication in 1998 of the Black and Wiliam review of formative assessment represented a watershed moment in teaching and learning, not least because it presented the educational community with a formidable challenge: to shift the focus from teaching to learning and to place assessment in its service. As evidenced by this study, and other more large-scale research initiatives, the implications of the prioritization of learning over assessment has far-reaching implications for teachers and students, but it also has major implications for researchers and policy-makers, who invariably must reconcile their, perhaps conflicting, allegiances to high-stakes and formative assessment agendas.

In Ireland, the jury is out on this key issue. There are mixed signals and inconsistencies in policy texts and implementation programmes and, given the dramatic downturn in the global economy, the auguries are of increasing pressures to compete internationally on all fronts, not least education. In such a climate, vulnerable communities, who have traditionally benefited least from education, stand to loose out even more, both in the immediate and long-term. Educators, armed with the knowledge that A/L offers one potential to change these trajectories, have important decisions to make at this critical time.

Appendix A

Brookhart's Review of Classroom Assessment from 1982 - 2002

The impracticality, noted by Black and William (1998a), of undertaking a meta-analysis of the research literature on A/L - given the absence of an overall theoretical framework on assessment coupled with differences in learning theories underpinning research initiatives - was brought to the fore in a recent review of classroom assessment literature by Brookhart (2004). Spanning the twenty-year period, 1982-2002, the literature review undertaken focused on classroom-based, assessment at primary level, in the US, UK and elsewhere. In explaining the work, Brookhart (2004, p. 450) noted that a large proportion of the articles "...consisted of inventories of practice or studies based in one theoretical tradition, often pedagogy..." with some studies being informed by more than one theoretical tradition, or discussing assessment practices in conjunction with instructional and classroom management practices. This piece of work is significant because it maps the research undertaken from the early 1980s, with particular reference to the three main research frameworks identified: psychology, sociology and measurement, as condensed in Table A1 overleaf.

Brookhart (2004) argued that the underlying theoretical concepts of the literature reviewed conflicted at (at least) three intersections: at the intersection between measurement theory and the group nature of learning, measurement theory and educational psychology, and among methods conventionally required for the different theoretical approaches. Positioning classroom assessment at the first intersection - between measurement theory and the group nature of learning - she made a plea (as interpreted by the researcher) for situative, multi-methods inquiry - which has particular import for this study:

Studies taking seriously a theory that situates the actual construct measured in group settings, and studies of assessment practices that are appropriately multidimensional, are needed for a complete understanding of classroom assessment at this intersection. (Brookhart, 2004, p. 451)

Table A1.

Classroom Assessment Literature 1982-2002 (Brookhart, 2004)

Practical Bases	Studies Included:
Largely inventories, without a theoretical framework, i.e. reports of practice	<i>Inventories of classroom assessment patterns or practices</i> - (Barnes, 1985; Gullickson, 1985; Schmidt & Brosnan 1996; Nicholson & Anderson 1993; Mavrommatis 1997).
Each study used a different framework	<i>Inventories of classroom assessment patterns or practices based in a theoretical framework</i> - (Stiggins & Bridgeford, 1985; Wilson, 1990; Adams & Hsu, 1998; McMillan, 2001; Gipps et al., 2000).
An area under-estimated/represented in the literature	<i>Classroom management</i> - (Higgins, Harris & Kuehn, 1994; Tunstall & Gipps, 1996; Mavrommatis, 1997; Barnes, 1985; Mooreland & Jones, 2000; Kusch, 1999).
Theoretical Bases	Studies Included:
<i>Psychology:</i> The majority of classroom assessment literature used psychological theory as the statement that drove the research questions and selection of methods	<i>The study of individual differences</i> - (Dassa, 1990; Butler & Nisan, 1986; Butler, 1987; Shepard (1996); Stiggins, Griswold & Wikelund, 1986; Crooks, 1988; Natriello, 1987; Stiggins, 1999). <i>Feedback theory</i> - (Elawar & Corno, 1985; Banger-Drowns, Kulik, Kulik, & Morgan, 1991; Banger-Drowns, Kulik & Kulik, 1991; Kluger & DeNisi, 1996; Tunstall and Gipps, 1996; Ross, Rolheiser & Hogaboam-Gray, 2002). <i>Studies of teachers' beliefs and practices</i> - (Niemi, 1997; McMillan, 2001; Thomas & Oldfather, 1997; Kusch, 1999; Shepard, 2001; Barnes, 1985; Pryor & Akweisi, 1998; Johnson, Wallace & Thompson, 1999; Wilson, 1999; Wilson & Martinussen, 1999; Anderson, 1999; Shulha, 1999).
<i>Theories about groups:</i> Focus: The cultural nature of learning and assessment	<i>Classroom assessment environment theory</i> - (Stiggins & Bridgeford, 1985; Stiggin & Conklin, 1992; Tittle, 1994; Brookhart et al., 1997a, 1997b; Brookhart & deVoge, 1999). <i>Social constructivism</i> - (Thomas & Oldfather, 1997). <i>Culture and sociology</i> - (Wolf, 1993; Torrance & Prior, 1998; Stiggins & Conklin, 1992; Rosenholtz & Rosenholtz, 1998; Simpson, 1981; Natriello, 1996).
<i>Measurement theory:</i> Focus: Constructs and how they are quantified	<i>Validity and reliability</i> - (Frisbie, Miranda & Baker, 1993; Cicmanec & Viechnicki, 1994; Impara, Plake & Fager, 1993; Frary, Cross & Weber, 1993; Mertler, 2001; Traub, 1990; Campbell & Evan, 2000; Bultermann-Bos, Terwel, Verloops & Wardekker, 2002; Messick, 1986; Shepard et al., 1996; Goldberg & Roswell, 2000). <i>Theory of formative assesement</i> - (Scriven, 1967; Sadler, 1983, 1989; Black and Wiliam 1996, 1998a, 1998b; Barnes, 1985; Thomas & Oldfather, 1997; Pryor & Akwesi, 1998; Torrance & Prior, 1998; Gipps et al., 2000; Rea-Dickens, 2001; Brookhart, 2001).

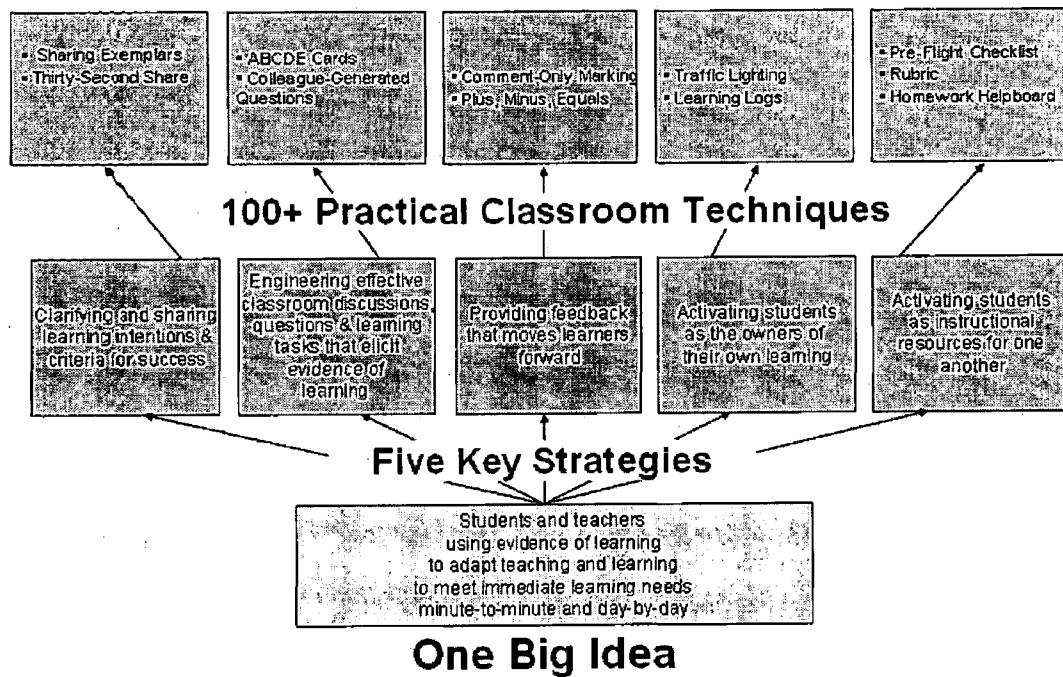
Appendix B

The Ten Projects of the AifL Programme

Project	Descriptor
1. Support for Professional Practice in Formative Assessment	To investigate effective classroom approaches to formative assessment & produce guidance for schools & teachers about assessment policies & practice to raise attainment.
2. Personal Learning Plans	To draw together existing work on Personal Learning Plans (PLPs) & other existing records & design a single recording framework.
3. Support for Management of Personal Learning Plans	To produce guidance for schools & teachers on management of the record-keeping process & PLPs, taking into account work on Progress Files & Individualised Educational Programmes (IEPs).
4. Gathering & Interpreting Assessment Evidence	To produce guidance & exemplification for teachers on the range of assessment evidence to be gathered & retained for each aspect of the curriculum at each stage, & how to evaluate it.
5. Local Moderation	To investigate ways of 'sharing the standard' with other teachers, & to produce local portfolios of examples of assessed work across the various aspects of the curriculum.
6. New National Assessment	To create an online 'bank' of assessment materials, based on Assessment of Achievement Programme (AAP) tests & tasks, to replace the current National Tests.
7. Assessment of Achievement Programme	To continue & update the Scottish Executive Education Department's (SEED) existing AAP in order to improve the monitoring of standards in the 5-14 Programme & give increased importance to the professional development of teachers.
8. ICT Support for Assessment	To produce a framework document outlining the key features/requirements of software to be used in support of PLPs & to guide LEAs in determining how to progress this element.
9. Reporting to Parents & Others	To link reporting to the PLP framework & propose a common format for reports and to include guidance for teachers on ways of working with parents, guidance for parents on the assessment process, and exemplars of good practice (video).
10. Meeting the needs of Pupils with Additional Support Needs	To ensure that all the projects in the programme are inclusive of pupils with the whole range of educational & social needs.

Appendix C

100+ Practical Classroom Techniques (Thompson & William, 2007)



Appendix D

Formative Assessment in the UK: Review by Gardner (2008)

The recent publication edited by Gardner (2008), entitled *Changing Assessment: Processes, Principles and Standards*, is timely. It reports on the conclusions drawn by the Analysis and Review of Innovations in Assessment (ARIA) project, following their review of recent initiatives and developments in assessment, in England, Scotland, Wales and Northern Ireland. Arguing that "...changes in assessment practice have been notoriously difficult to sustain" (Gardner, 2008, p. 1), the ARIA study took as its main focus the teacher's role in bringing about change through the use, at school level, of assessment for formative and summative purposes. In offering a synthesis of the lessons learned, Gardner (2008, pp. 2-3) drew two initial conclusions:

1. Education systems must fully commit to all of the necessary ingredients for sustainable development if their objective is to promote and embed changes in assessment practice. This is because, all too frequently when initiatives aimed at changing assessment practice are being conceived, insufficient attention is paid to the key planning and design process.
2. A common language of principles and standards is needed to guide the development of effective assessment practice. In turn, principles and standards should be designed to enable any stakeholder group to assess the extent to which they are effectively promoting and sustaining desirable changes in assessment and its use.

Towards meeting these challenges, a change-process model was proposed, together with a set of standards for classroom assessment practices, aimed at each of the four stakeholder groups identified (Table D1 overleaf). Given the focus of this study, the discussion of lessons learned in relation to initiating and sustaining change in assessment practices drew particular attention. Given the currency of this report, and the fact that it reflects the findings of such a broad base research inquiry across four countries, a summary of the key comments are reproduced in Table D2.

Table D1.

Standards for the Use of Formative Assessment (Adapted from Gardner, 2008)

Standards for Classroom Assessment Practice	Standards for Use by School Management Teams	Standards for Use in National and Local Inspection and Advice Arrangements	Standards for Use in National Policy Formulation
Formative Use of Assessment			
<p>1. Teachers gather evidence of their students' learning through questioning, observation, discussion and study of products relevant to the learning goals.</p> <p>2. Teachers involve students in discussing learning goals and the standards to be expected in their work.</p> <p>3. Teachers use assessment to advance students' learning by:</p> <ul style="list-style-type: none"> - Adapting the pace, challenge and content of activities - Giving feedback to students about how to improve - Providing time for students to reflect on and assess their own work. <p>4. Students use assessment to advance their learning by:</p> <ul style="list-style-type: none"> - Knowing and using the criteria for the standards of work they should be aiming for - Giving and receiving comments from their peers on the quality of their work and how to improve it - Reflecting on how to improve their work and taking responsibility for it. 	<p>Teachers collaborate in developing their practice in:</p> <ul style="list-style-type: none"> - Communicating goals and quality criteria to students - Helping students to take part in self- and peer-assessment - Providing feedback to help learning - Enabling students to take responsibility for their work. 	<p>1. The use of assessment to support learning is included as a key factor in evaluating the effectiveness of schools.</p> <p>2. Help is available for schools to ensure that all areas of achievement benefit from the formative use of assessment.</p> <p>3. Schools are encouraged to develop their formative use of assessment.</p>	<p>1. Assessment to support learning is at the heart of government programmes for raising standards of achievement.</p> <p>2. Initial teacher education and professional development courses ensure that teachers have the skills to use assessment to support learning.</p> <p>3. School inspection frameworks give prominence to the use of assessment to support learning.</p> <p>4. Schools are encouraged to evaluate and develop their formative use of assessment.</p>

Table D2.

Formative Assessment Practices in the UK (Gardner, 2008)

Key Processes	Key Findings
Innovation	<ul style="list-style-type: none"> - Innovation may fail in the face of workload issues or in simply not being a convincing enough change for teachers to adopt.
Warrant	<ul style="list-style-type: none"> - Any innovation in assessment requires solid research evidence that it works before its rollout can be justified.
Dissemination	<ul style="list-style-type: none"> - A failure to recognise that the learning process has to be the same for the first person as the last has meant that many approaches to rolling out innovations in assessment have had only limited success. - The success of involving teachers in new assessment practices in pilot developments often falters in the large-scale roll-out phase as engagement of teachers switches to telling teachers. - For real change, as with all learning, the individuals involved need to take more control of what they are being asked to do. They need to make sense of it through reflection and sharing it with others until new ideas and processes become internalised. - Successful dissemination cannot rely on any single strategy and assumptions cannot be made that what works in one culture will work in another. In planning the scaling up of innovations in assessment practice, the context, the nature of the innovation and its impact in relation to (or competition with) other initiatives are important considerations.
Agency	<ul style="list-style-type: none"> - Self-agency is a powerful device in fostering change because it draws on self-motivation. Unless teachers are committed through self-agency to changing their assessment practice, the prospects for successful dissemination and professional learning, leading to its embedding and sustainable development, are likely to be slim.
Professional Learning	<ul style="list-style-type: none"> - Using the term professional learning recognises that attempts to change practice in education must aim for a change in understanding rather than merely superficial change in teaching techniques. - When decisions need to be made about when and how assessment techniques can be used, the lack of a fundamental understanding of the purposes may lead to confusion and ultimately to rejection of the techniques. - Teachers need: 1. Time to reflect and to adjust their teaching to take on new practices; 2. Professional development activities which are best spread over time with opportunities for trying out new assessment ideas between sessions; 3. To reflect on and share their experiences with others in order to promote effective professional learning and the ownership and understanding needed for successful implementation of new assessment procedures.
Impact	<ul style="list-style-type: none"> - One of the most common reasons for 'no-difference' or even negative findings for the impact of innovation in education is that the intended changes are not properly in place. A key reason for proposing principles and standards, therefore, is to offer a framework to assist key stakeholders in ensuring that quality assessment practices are in place.
Sustainable Development	<ul style="list-style-type: none"> - In assessment the notion of sustaining new practices as fixed and unchanging is inappropriate. Such classroom activities as the A/L traffic-lights, wait time and no marks, which were at one time new to many teachers, can quickly become drab routine if they do not develop in tandem with changing needs of teachers and pupils. - Criteria used by schools and inspectors in school evaluation must endorse and reflect the importance of the changes being pursued. Where policy arrangements and official endorsements are not consistent with the demands on schools, the potential for innovation to be sustained is compromised. <p>Sustainable development is the necessary condition and the desired endgame of any assessment innovation of value. An essential ingredient will be the readiness of teachers to scan the horizon continuously for ways to improve their assessment practices.</p>

Appendix E

List of Acronyms

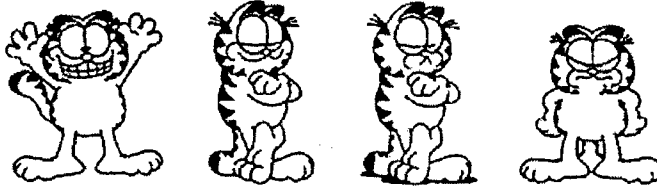
AfL	Assessment <i>for</i> Learning
AfLAI	Assessment <i>for</i> Learning Audit Instrument
C	Control
CBAM	Concerns Based Adoption Model
CI	Confidence Interval
DEIS	Delivering Equality of Opportunity in Schools
DPRT	Drumcondra Primary Reading Test
E	Experimental
ERAS	Elementary Reading Attitude Survey
ES	Effect Size
M	Mean
MD	Median
MICRA-T	Mary Immaculate College Reading Attainment Test
N	Number
SAfLQ	Scaled Assessment <i>for</i> Learning Questionnaire
SD	Standard Deviation
SEN	Special Educational Needs
SPSS	Statistical Package for Social Sciences
TLC	Teacher Learning Community
TLL	Teacher's Learning Log
TVRS	Teachers' Video Review Sheets

Appendix F

Elementary Reading Attitude Survey

Name _____ Class _____ Date _____

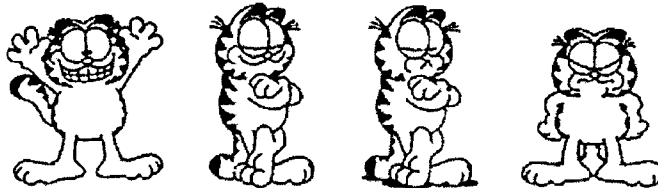
1. How do you feel when you read a book on a rainy Saturday?



2. How do you feel when you read a book in school during free time?



3. How do you feel about reading for fun at home?



4. How do you feel about getting a book for a present?



5. How do you feel about spending free time reading?



6. How do you feel about starting a new book?



7. How do you feel about reading during summer vacation?



8. How do you feel about reading instead of playing?



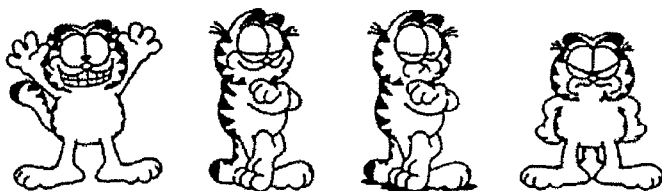
9. How do you feel about going to a bookstore?



10. How do you feel about reading different kinds of books?



11. How do you feel when the teacher asks you questions about what you read?



12. How do you feel about doing reading workbook pages and worksheets?



13. How do you feel about reading in school?



14. How do you feel about reading your school books?



15. How do you feel about learning from a book?



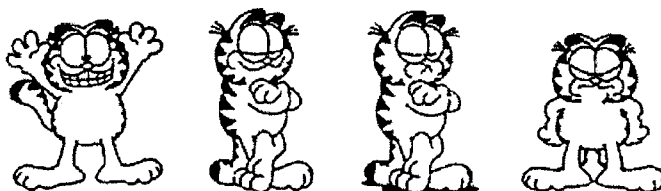
16. How do you feel when it is time for reading class?



17. How do you feel about the stories you read in reading class?



18. How do you feel when you read out loud in class?



19. How do you feel about using a dictionary?



20. How do you feel about taking a reading test?



GARFIELD: © 1978 United Feature Syndicate, Inc.

Kear, D.J. & McKenna, M. C. (1999). Measuring attitude toward reading: A new tool for teachers. In S. J. Barrentine (Ed.), *Reading assessment: principles and practices for elementary teachers. A collection of articles from "The Reading Teacher."* p. 199-214. Newark, DE: International Reading Association. Garfield Images approved for use in this document only.

Questionnaire for 2nd Class Pupils

What is your name? _____

What is your date of birth? _____ / _____ / _____
day / month / year

Here are some questions about reading. Tick the box beside the answer that is best for you. Please tick just one answer for each question on this page.

1. DO YOU THINK YOU ARE GOOD AT READING?

- a) Yes, excellent ☐₁
- b) Yes, very good ☐₂
- c) Yes, ok ☐₃
- d) No, not so good ☐₄
- e) No, poor ☐₅

2. IF YOU COME TO A NEW WORD IN YOUR READER, WHAT DO YOU USUALLY DO?

- a) Skip the word ☐₁
- b) Try to sound it out ☐₂
- c) Use the words around it to figure it out ☐₃
- d) Ask someone for help ☐₄

Please put a tick in one the boxes for each of the short questions asked
 – a (never), b (sometimes), c (often) and d (always).

3. BEFORE I READ...	<i>Never</i>	<i>Sometimes</i>	<i>Often</i>	<i>Always</i>
a) ...I talk with my teacher or friends about what makes a good reader.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
b) ...I talk with a partner about what I need to do to be a better reader.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
c) ...I tell a partner what I think they need to do to be a better reader.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
d) ...I decide to work on one or two things that will make me a better reader.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄

4. BEFORE I READ...	<i>Never</i>	<i>Sometimes</i>	<i>Often</i>	<i>Always</i>
a) ...I understand what I am going to learn from my reading.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
b) ...I think about what I did very well the last time I was reading.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
c) ...I think about what I found difficult to do the last time I was reading. ...	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
d) ...I know exactly what my teacher wants me to work especially hard at during the lesson.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄

5. BEFORE I READ...	<i>Never</i>	<i>Sometimes</i>	<i>Often</i>	<i>Always</i>
a) ...I make some guesses about what will happen.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
b) ...I ask questions I would like answered.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
c) ...I read the title to see what the story is about.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
d) ...I think about where the story might be taking place.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄

6. WHILE I READ...	<i>Never</i>	<i>Sometimes</i>	<i>Often</i>	<i>Always</i>
a) ...I check to see if I understand the story so far.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
b) ...I think about what the characters in the story are doing.....	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
c) ...I try to work out what will happen next.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
d) ...I re-read parts of the story I don't understand.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
 7. AFTER I READ...	 <i>Never</i>	 <i>Sometimes</i>	 <i>Often</i>	 <i>Always</i>
a) ...I ask myself questions to see if I understood the story.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
b) ...I think about how I would have acted if I were the main character....	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
c) ...I compare the story to other stories I have read.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
d) ... I talk to other children about what happened in the story	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
 8. AFTER I READ...	 <i>Never</i>	 <i>Sometimes</i>	 <i>Often</i>	 <i>Always</i>
a) ...I ask myself how well I did at my reading.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
b) ...I think about what I did well when I was reading.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
c) ...I think about what I found difficult when I was reading.....	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
d) ... I discuss with my teacher what I think I need to do next time to become better at reading.	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄

Thank you for completing this questionnaire. ©

Appendix H

Scaled AfL Questionnaire Reliability Analyses

In making claims about changes in the attitudes to reading of the experimental group, the issue of the reliability of the scales used arose, in particular, the scale's internal consistency that determines the degree to which the individual items within the scale 'hang together'. In this context, Cronbach alpha coefficient values were sought for (a) the scale as a whole, for the control group (post-intervention), and the experimental group (pre- and post-intervention) and (b) within scales for both groups (pre, during and after reading). Table H1 shows the coefficient values for the total scale.

Table H1.

Scaled AfL Questionnaire Pre- and Post-Intervention Total Scale

Instrument	Administered	Group	Number	Number of Items	Cronbach's Alpha
Scaled Assessment For Learning Questionnaire (total score)	End 1 st Class	C	-	-	-
	End 1 st Class	E	55	24	0.89
Scaled Assessment For Learning Questionnaire (total score)	End 2 nd Class	C	72	24	0.81
	End 2 nd Class	E	57	24	0.88

Note:

The number of children included in Tables 11 and 12 above differ from that reported in previous tables for the same test. This discrepancy is explained by the fact that, when undertaking this kind of analysis, SPSS only includes cases for which full sets of data are returned.

Pallant (2007) suggests that, ideally, Cronbach alpha values should be above 0.7 if the internal consistency of a research instrument is to be assumed. Given that the values reported in Table H1 range from 0.81 to 0.89, the internal consistency of the SAfLQ appears high. However, in the context that Cronbach alpha values are sensitive to the number of items, particularly short scales of fewer than ten items, and that two of the three sub-sections of the SAfLQ had four and eight items, respectively, the decision was taken to conduct further analysis to establish the inter-item correlation for the items in question. Table H2 shows the Cronbach alpha values for each of the sub-scales (before, during and after reading), once for the control group at the end of Second Class, and before and after the intervention programme for the

experimental group. It is not surprising that the second sub-scale, during reading, reported relatively lower values given that it comprised just four items. Nor was it unexpected that the reliability of the scale seemed to differ between the groups, and, indeed, within the experimental group, when tested before and after the intervention programme.

Table H2.

Scaled AfL Questionnaire Pre- and Post-Intervention Within Sub-Scales

Instrument	Administered	Group	Number	Number of Items	Cronbach's Alpha
Scaled Assessment For Learning Questionnaire (before reading)	End 1 st Class	C	-	-	-
	End 1 st Class	E	62	12	0.81
Scaled Assessment For Learning Questionnaire (during reading)	End 1 st Class	C	-	-	-
	End 1 st Class	E	74	4	0.55
Scaled Assessment For Learning Questionnaire (after reading)	End 1 st Class	C	-	-	-
	End 1 st Class	E	68	8	0.73
Scaled Assessment For Learning Questionnaire (before reading)	End 2 nd Class	C	72	12	0.62
	End 2 nd Class	E	58	12	0.74
Scaled Assessment For Learning Questionnaire (during reading)	End 2 nd Class	C	72	4	0.58
	End 2 nd Class	E	58	4	0.59
Scaled Assessment For Learning Questionnaire (during reading)	End 2 nd Class	C	72	8	0.73
	End 2 nd Class	E	57	8	0.76

Additional analysis was undertaken subsequently to determine the mean inter-item correlation for items reporting values lower than 0.7 for both of the sample groups. These are reported in Table H3. Particularly significant is the fact that, although there were small number of minus values from the 4-item subscale *during reading*, many of the reported values are well within the optimal range for inter-item correlation of 0.2 to 0.4, which suggests that the scale as a whole is robust. This would suggest that further research might be required to raise the reliability of this subsection in particular.

Table H3.

Inter-Item Correlation Matrices for the Scaled Assessment for Learning Questionnaire, Control and Experimental Groups

Scale Item	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Experimental Group During Reading Sub-Scale: End 1st Class					
6a	7.34	7.076	.312	.107	.504
6b	7.54	7.896	.233	.057	.562
6c	7.39	6.899	.389	.185	.439
6d	7.61	6.488	.424	.213	.406
Experimental Group During Reading Sub-Scale: End 2nd Class					
p6a	8.14	6.998	.242	.109	.624
p6b	8.16	6.168	.502	.270	.423
p6c	8.00	6.456	.426	.248	.480
p6d	8.40	6.454	.346	.159	.541
Control Group Before Reading Sub-Scale: End 2nd Class					
p3a	25.07	29.051	.224	.265	.604
p3b	25.04	28.435	.300	.411	.592
p3c	25.00	28.873	.256	.324	.599
p3d	24.65	27.779	.283	.450	.593
p4a	24.18	27.925	.223	.196	.606
p4b	24.18	26.404	.401	.236	.569
p4c	24.72	29.696	.135	.149	.619
p4d	23.86	26.910	.301	.339	.589
p5a	24.50	26.789	.285	.295	.593
p5b	24.71	26.942	.393	.322	.573
p5c	23.50	27.972	.223	.232	.605
p5d	24.38	27.843	.204	.154	.611
Control Group During Reading Sub-Scale: End 2nd Class					
p6a	7.57	6.418	.351	.142	.495
p6b	7.61	6.156	.368	.157	.480
p6c	7.85	6.216	.348	.153	.497
p6d	7.31	6.272	.332	.138	.510

Appendix I

Teachers' AfL Audit Instrument

How important/relevant is the statement for your establishment?	Fundamental principles	How does your current practice match the statement?
4: essential 3: very important 2: quite important 1: of limited importance		4: mirrors the statement 3: room for minor improvements 2: elements require development 1: requires re-thinking
1 2 3 4		1 2 3 4
	Assessment offers all pupils an opportunity to show what they know, understand and can do	
	Assessment practice helps pupils to understand what they can do and where they need to develop further	
	The key learning outcomes of each subject or learning experience (early years) have been identified so that assessments made against these can be used to help develop children's learning	
	Assessments are not restricted to national curriculum subjects	
	Sharing of learning intentions is routine practice, which enables the pupils to understand their role in the lessons	
	Assessment practice in the school enhances the learning process	
	Assessments made by the teachers inform daily and weekly planning and allow learning to be matched to the needs of the pupils	
	Assessment of pupils' learning is reported to parents in a way which identifies achievements and what the child needs to do to improve	
	Pupils are involved in assessing their own work and that of their peers	
	Pupils and teachers work together identifying targets for learning and ways of achieving these	
	Core assessment data on each child is updated each year and passed to the receiving teacher or school to aid future planning	

Note:

Teachers were advised that the term 'Establishment' should be interpreted as 'School' and 'Current Practice' as 'In your own teaching'. In addition, a number of the statements were omitted from the audit completed by the teachers in this study, given the differences in educational systems between the UK and Ireland.

How important/relevant is the statement for your establishment?

- 4: essential
- 3: very important
- 2: quite important
- 1: of limited importance

1 2 3 4

Planning learning

How does your current practice match the statement?

- 4: mirrors the statement
- 3: room for minor improvements
- 2: elements require development
- 1: requires re-thinking

1 2 3 4

Long term planning

Involves the whole staff and reflects our school's aims

Reflects the whole-school curriculum framework taking into consideration the foundation curriculum, the programmes of study, schemes of work, time available and any planned thematic work

Maps progression throughout the school in terms of level of demand with reference to the programmes of study and level descriptions

Medium term planning

Identifies key learning intentions for assessment

Identifies intended progression and curriculum targets for classes

Takes into consideration a range of teaching techniques and assessment approaches which reflect our pupils' different learning styles

Short term planning

Contains clear specific learning intentions for assessment

Takes into account the prior knowledge, skill and understanding that the pupils bring to the learning situation

Takes into account a range of learning styles

Recognises that all learning does not need to be assessed

Recognises that unanticipated learning occurs

Identifies what will be assessed for pupils or groups, how this will be done, and who will do it

How important/relevant is the statement for your establishment?

- 4: essential
- 3: very important
- 2: quite important
- 1: of limited importance

1 2 3 4

Assessment for learning

How does your current practice match the statement?

- 4: mirrors the statement
- 3: room for minor improvements
- 2: elements require development
- 1: requires re-thinking

1 2 3 4

Pupils know what they are learning, what they have achieved and how they can improve

Pupils are provided with regular opportunities to reflect and talk about their learning, progress and goals

Teachers use a range of assessment methods confidently and appropriately

The organisation of the classroom enables planned assessments to occur and unexpected achievements to be assessed as well

Strategies are in place which reveal when pupils have difficulties or are not making progress

Assessments are used to decide what to do next with individuals, groups or the class

Shared ways forward are agreed between pupils and teachers which focus on how pupils will achieve their goals

Other adults working in the classroom are clear about their role in assessment and communicate significant information about pupils

The portfolios agreed within the school are consistently used to confirm assessment judgements

It is recognised that all learning need not be assessed

How important/relevant is the statement for your establishment?

- 4: essential
- 3: very important
- 2: quite important
- 1: of limited importance

1 2 3 4

Marking and providing feedback

How does your current practice match the statement?

- 4: mirrors the statement
- 3: room for minor improvements
- 2: elements require development
- 1: requires re-thinking

1 2 3 4

Prompt and regular marking occurs in all classes and all subjects

The marking process includes both verbal and written feedback

Marking focuses on the learning intentions as the criteria for success

Pupils are provided with opportunities to assess their own and others' work

Marking strategies help the pupils understand what they have achieved and what they need to do next

The outcomes of marking, along with other information, are used to adjust future teaching plans

The policy for marking is reviewed regularly, making sure new members of staff understand it, so that practice can continue to reflect the school policy

How important/relevant is the statement for your establishment?

- 4: essential
- 3: very important
- 2: quite important
- 1: of limited importance

1 2 3 4

Assessment of learning

How does your current practice match the statement?

- 4: mirrors the statement
- 3: room for minor improvements
- 2: elements require development
- 1: requires re-thinking

1 2 3 4

Assessment of learning is not just tests

Assessment of learning is always undertaken for a specific purpose

The current requirements and guidance for statutory assessment are understood and followed

A holistic and 'best-fit' approach is used when working with the level descriptions

In making judgements against level descriptions a range of assessment information is used

Consistent judgements are reached through activities that promote shared understanding of standards

Item analysis is used as a tool to identify gaps in teaching and learning

The outcomes of assessment of learning activities provide feedback and 'feed-forward' for pupils

Assessment of learning information is used to evaluate teaching and for monitoring progress

How important/relevant is the statement for your establishment?

- 4: essential
- 3: very important
- 2: quite important
- 1: of limited importance

1 2 3 4

Recording and evidence

How does your current practice match the statement?

- 4: mirrors the statement
- 3: room for minor improvements
- 2: elements require development
- 1: requires re-thinking

1 2 3 4

Information from record of ongoing assessment records, together with records of summative assessment, contributes to the agreed attainment information used to track progress

There is a whole-school agreed set of attainment information, which is recorded

The agreed attainment information is updated at least once a year and used to track progress

Beyond whole school records teachers decide what to record

Each teacher uses a range of recording strategies for additional records

Progress against key learning intentions is observed, noted and where significant is recorded

Progress against key learning intentions feed forward into future planning

Pupils are involved in recording comments on their work

Samples of assessed work are kept to exemplify agreed standards and to model success

Records, which are passed on, are useful, clear and easy to interpret

Records enable reports to be written easily

Whole school records provide the information required for the National Transfer Form

Information from the previous teacher is used to plan work in a new class

How important/relevant is the statement for your establishment?

4: essential
3: very important
2: quite important
1: of limited importance

1 2 3 4

Reporting to parents and carers

Efforts are made to ensure that parents understand the information which is given to them about their child's attainment and progress

Reports outline strengths in all aspects of school life and indicate areas that need to be developed

Children are actively involved and contribute to the reporting process

There are opportunities for teachers, pupils and parents to talk together

Parental involvement in pupils' learning is encouraged

Targets are set, shared with parents and reviewed with pupils

Statutory requirements for reporting are met

The relationship between individual attainment and comparative results is clearly explained

Timing of reports allows appropriate discussion and action to take place

Parents are helped to understand that teacher assessment and test levels are equally important and provide different and complementary information

How does your current practice match the statement?

4: mirrors the statement
3: room for minor improvements
2: elements require development
1: requires re-thinking

1 2 3 4

How important/relevant is the statement for your establishment?

- 4: essential
- 3: very important
- 2: quite important
- 1: of limited importance

1 2 3 4

Transfer and transition

How does your current practice match the statement?

- 4: mirrors the statement
- 3: room for minor improvements
- 2: elements require development
- 1: requires re-thinking

1 2 3 4

Curriculum liaison and trust is well established with partner schools and settings

Pupils' work is used as a focus for discussions on progression and approaches to learning as pupils transfer

There are opportunities for teachers to observe learning in partner schools and settings

There is agreement within and across schools about what information should be passed on

Information transferred identifies pupils' strengths and areas for development

The information which is transferred to the next school covers the statutory requirements

When pupils stay in the same school, information gets to the next teacher in time for effective planning to occur

When pupils move schools there are procedures in place to ensure that the right information gets to the right people in good time

When new pupils arrive the information that comes with them is always used to pitch the curriculum appropriately

How important/relevant is the statement for your establishment?

- 4: essential
- 3: very important
- 2: quite important
- 1: of limited importance

1 2 3 4

Management and monitoring

The policy for assessment, recording and reporting is developed together and is understood by all

There are agreed guidelines for implementing our policy, which are communicated to the staff and are placed in writing

Procedures are in place to monitor and evaluate what is happening in all aspects of assessment, recording and reporting

Curriculum co-ordinators monitor the planning and delivery of their subjects through scrutiny of short-term plans and pupils' work

There is a member of staff with overall responsibility for assessment, recording and reporting who co-ordinates the overall assessments made

How does your current practice match the statement?

- 4: mirrors the statement
- 3: room for minor improvements
- 2: elements require development
- 1: requires re-thinking

1 2 3 4

How important/relevant is the statement for your establishment?

4: essential
3: very important
2: quite important
1: of limited importance

1 2 3 4

Using assessment information to monitor progress

How does your current practice match the statement?

4: mirrors the statement
3: room for minor improvements
2: elements require development
1: requires re-thinking

1 2 3 4

When new pupils arrive the information that comes with them is always used to pitch the curriculum appropriately

The whole school agreed set of attainment information about each pupil is used to track the individual's progress

Attainment information is used to inform curriculum planning and to identify key learning objectives

Year on year trends are monitored

Attainment is analysed for each cohort and is used to set appropriate targets

Progress towards targets is checked regularly for individuals and year groups

The performance of specific groups of pupils monitored: e.g. gender, ethnic groups, EAL, SEN, looked after and talented and gifted pupils etc.

Performance management strategies provide information about pupil progress in different teaching groups and subjects

National and local data is used to provide a realistic comparison with other schools

Benchmark information is used to identify other similar schools in order to learn from their experience

Analysis of information informs decisions about what to include in subject action plans and school improvement plans

How important/relevant is the statement for your establishment?

- 4: essential
- 3: very important
- 2: quite important
- 1: of limited importance

1 2 3 4

Evaluation as part of school improvement

All staff are involved in the review and evaluation process

Assessment recording and reporting practices and policy are systematically reviewed and evaluated, in terms of the impact upon both teaching and pupils' learning

Decisions are made about how the practice of assessment, recording and reporting can be improved and targets for development are identified and agreed

The records of pupil's progress are reviewed regularly to ensure information held is used effectively and when necessary adjustments are made

The review and evaluation of assessment practice and policy informs school development planning

The school improvement plan is the vehicle for describing how the strategies for further developing assessment recording and reporting practice will be implemented

How does your current practice match the statement?

- 4: mirrors the statement
- 3: room for minor improvements
- 2: elements require development
- 1: requires re-thinking

1 2 3 4

Teachers' Learning Log

Please complete **at least three** of the following sentence starters with your thoughts on today's sessions.

- [illegible]

At this stage in the project, I am most concerned about

My greatest need is

What would help and support me most would be

Appendix K

Professional Development Programme

September 19 September 26	General introduction to AfL Methodologies <i>AfL Strategy: <u>Sharing Learning Intentions</u></i>
October 23 October 24	Video Analysis and Review <i>AfL Strategy: <u>Tuning into Learners' Minds: Success Criteria</u></i>
November 27 November 28	Video Analysis and Review <i>AfL Strategy: <u>Stepping Forward with Feedback</u></i>
January 8 January 9	Video Analysis and post Christmas Review <i>AfL Strategy: <u>Marking Less to Achieve More</u></i>
February 19 February 20	Video Analysis and Review <i>AfL Strategy: <u>Promoting Assessment by Pupils</u></i>
April 1 April 2	Video Analysis and post Easter Review <i>AfL Strategies: <u>Revisiting the Big Idea: Clozing the Gap</u></i>
May 6 May 7	Video Analysis and Review <i>AfL Strategies: <u>Revisiting the Big Idea: Clozing the Gap</u></i>
June 3, 4, 5	Three-day Review and Post-Intervention Data Collection

Appendix L

Supplementary Teaching Materials

Literacy:	Reading	Writing	Oral Language
Learning Intention (Bit or Application) WALT: 1. Be able to read and show we understand what we are reading.			
Context (distinct and separate from the LI)			
Success Criteria (what you need to do to achieve the learning intention) WILF: Is that you/for you to: <ol style="list-style-type: none"> 1. Try to read the passage. 2. Use the title of the (story/letter/poem – <u>specify the genre</u>) to help you guess what the (<u>specify genre</u>) is about. 3. Use the picture to help you understand what's happening in a (story/letter/poem... - specify the genre). 4. Ask yourself questions quietly as you read. 5. Use all the steps you can learned if you find your reading hard and get stuck. 			
Notes for future Planning (if necessary)			

Note:

LI: This states what the student is learning to do (skill), know (knowledge) and understand (concept). NOT the context. It must be shared with the students both visually and orally. It means that the student knows what the purpose is and helps to transfer the responsibility for learning from teacher to student.

SC: This is a set of criteria that the student can check back against to see if they have been successful. NOT the task. It is more effective if generated by a discussion between the teacher and student. This should also be written down.

Learning Intentions (LIs)		Context	Success Criteria	Feedback
Remember: the LI, Context/Activity and SC need to be stated explicitly and separately so as not to confuse the LI and activity.				
<ul style="list-style-type: none">➤ Clarify at the planning stage; share at the beginning of the lesson➤ Two categories of LIs: - Bit LIs (Knowledge, Skills, Attitudes: which are very specific) Or - Application LIs (Application of Bits: more general/broad in nature)➤ Should use both types, punctuating Bits with Apps➤ The same LIs should be given to the whole class; differentiation may occur in the context of the SC (must/could/should)➤ Always try to use words associated with learning not doing (the activity/product):<ul style="list-style-type: none">- Knowledge: WALT know/understand/explain- Skills: WALT develop/be able to/get better at- Understanding: WALT understand how/why/that- Attitudes/values: WALT be aware of/think more about/consider➤ Make them SMART and maximum 2 or 3➤ Display prominently, same place, get kids to remind you if you forget➤ Make continuous reference throughout the lesson to them – making sure the kids are focused not on the activity they are engaged in but the skills/knowledge they are learning➤ It might be useful to get the kids to write the LI as the title to their work➤ Give the kids a nose for quality by showing eggs and/or modeling the process➤ Use them to give feedback, mark work, praise/reward➤ Research shows if self evaluation is linked to the LI, children's progress, persistence and self-esteem improves	<ul style="list-style-type: none">➤ Except for Bit LIs focusing on knowledge development, the context for the learning should be independent from the LI➤ In this way the kids learn that it is not the particular context that matters but the mastery of the basic/core skill➤ This means that there are usually any number of contexts in which a LI might be taught➤ It is possible that having identified the LI and the SC, you will notice that the context will not support the child's mastery of the LI and will need to be changed	<ul style="list-style-type: none">➤ Should not be expressed as behaviour but as learning➤ SC can be thought of as falling into 4 types: form, content, process and impact➤ Process or Product?➤ A process focus emphasizes <i>how</i> the child can achieve the LI and is more useful to the child; in contrast a product focus indicates <i>what</i> the end result will be but gives no clue as to how to achieve it➤ Visible or invisible?➤ 90% of what the brain remembers is from visual images so visual reminders are very useful.➤ SC should echo/reflect the key steps/teaching points in the lesson➤ SC should be elicited from the kids just before they begin a task/activity or just after they have had a trial run but they <u>must</u> be planned by the teacher before the lesson begins. This ensures that the teacher's focus is on the learning not the teaching/activities of the lesson and gives the teacher a chance to check that the SC identified will actually ensure the child's mastery of the LI➤ In this way, the SC set the activity agenda for the lesson➤ SC may be differentiated by using the terms <i>must</i> (minimum requirement for all children), <i>could</i> (additional elements for more able/advanced children) and <i>should</i>. The kids succeeding at <i>shoulds</i> can then be encouraged to move to <i>coulds</i>.➤ The Aside should be given <u>orally</u> before the kids start work	<ul style="list-style-type: none">➤ Written (PEN)/oral/plenary➤ Highlighters: traffic lights➤ have been found to be most successful for young children➤ Applaud effort as much as achievement➤ Summative feedback is of little use in shaping the learning experience for the child so prompt continuously during the lesson: remember – minute-by-minute, day-by-day assessment!➤ Remember, feedback is only formative if it is used by the learner to improve➤ The plenary is important: what did you learn; what was most difficult/easy; what would you do differently; what do you need to do next to improve	
Bit LIs: <ul style="list-style-type: none">➤ Aim to teach knowledge, skills or attitudes➤ Remember: a knowledge Bit must include the context➤ Bit LIs subsequently become SC for Application LIs		<ul style="list-style-type: none">➤ Easy to create➤ Chronological and step-by-step	Instant/Immediate. Use prompts: <ul style="list-style-type: none">- Reminder- Scaffolded- Example.	
Application LIs: <ul style="list-style-type: none">➤ Aim to give the kids an opportunity to apply the Bits they have learned➤ Make sure the Apps LI is not too broad!		<ul style="list-style-type: none">➤ Comprise the a list of previous Bit LIs➤ Read like a list of ingredients	Feedback is more Reflective <ul style="list-style-type: none">Gives rise to individualized targets for kids	

Skill vs. Strategies Approaches

1. **Bottom-up model:** readers acquire the ability to read by learning a hierarchy of skills in both word recognition and comprehension. Instruction concentrates on the acquisition of separate subskills in decoding and comprehension, such as phonics, context clues and main idea identification
2. **Top-down model:** reading is perceived as sampling, selecting, predicting, comparing and confirming what the reader sees and expects to see. Instruction emphasises the use of prior knowledge to develop hypotheses and make predictions.

An effective reader uses skills as well as appropriate strategies.

1. The reader who comes across an unfamiliar work may apply a strategy to unlock the pronunciation and/or meaning. However, if a reader has a limited number of skills in word recognition, he/she may be hampered in applying a strategy.
2. Thus to hold an either/or approach can be limiting in instruction and unfair to students. The two approaches (**interactive model**) can work in concert to provide maximum success in reading instruction practices

Questions to guide assessment and diagnosis of children's reading include:

1. At what reading level is the child functioning?
2. What are the child's reading interests?
3. What is the child's attitude toward reading and/or school?
4. What are the child's strengths and abilities?

A Diagnosis model for a child who is an early or transition reader

1. Sight words (graded word list, Informal Reading Inventory)
2. Comprehension (Informal Reading Inventory)
 - a. Analyse comprehension levels and errors
 - b. Check vocabulary
 - c. Check comprehension on narrative vs. expository text
 - d. Check comprehension on oral vs. silent reading
 - e. Examine oral reading rate
 - f. Writing Sample
3. Decision: What are the child's patterns of strengths and weaknesses?
4. Decision: What are the appropriate instruction strategies?

Closing the gap

1. Do we know what standard we want (or our children are expected to achieve according to standardised tests) by the end of second class?
2. Do we know where they are currently relative to this standard?
3. Have we clear incremental steps in place to support their achievement of this standard?

The National Standard of Reading in 2nd Class

MICRA-T: Level 2: 2nd and 3rd class.

There are three parts to Level 2 of the test:

1. 20 items, assesses pupils' word recognition and decoding skills.
2. Comprehension section one: which requires pupils to demonstrate their understanding of individual sentences by responding in a variety of ways.
3. Comprehension section 2: consists of 36 sentences, graded according to difficulty, with each sentence containing a superfluous word which must be deleted in order to render the sentence meaningful.

While this level of the test is highly similar to the original Level 2, the difficulty level has been extended upwards so that **it remains challenging even at the end of Third Class.**

Total administration time, including sample items for each part, is approximately one hour.

The Drumcondra Primary Reading Test-Revised

There are three parts to Level 2 has three parts to the test. For the purposes of this project, we are only asking them to complete sections one and two. Section three, which relates to word analysis is deemed to have been examined sufficiently in the MICRA T.

1. Reading Vocabulary: 36 items. A short sentence is given with one word underlined and the child has to tick one word from a list of four words which has a similar meaning. **Estimated time required: 40 minutes.**
2. Reading Comprehension Stories: 3 stories approximately one A4 sheet in length, size 12 font approximately with a picture in black and white at the top of each but without any title to the story. The child has to answer 10 multiple-choice questions relating to each story, i.e. 30 items. The phrasing of the questions tends to demand higher order thinking of the children, i.e. it is not sufficient to be able to locate the answer because as the questions are phrased they do not echo exactly the language used in the story. **Estimated time required: 45 minutes.**
3. The Reading Vocabulary and Reading Comprehension subtests at Level 2 take about **1.5 hours to administer**; this estimate include time for completing identification information and sample items.

So: Apart from the reading acumen required, children are expected to

1. Sit a test of between one and one and a half hours duration with short breaks
2. Follow directions – understand the teacher's instructions
3. Be internally motivated to achieve
4. Persevere when faced with the unknown
5. Working in isolation and in silence without recourse to teacher assistance of any kind.

Appendix M

Teachers' Learning Log Review

Learning Logs:

Mapping your personal journey through Professional Development during the year

In reviewing your logs, it might be helpful to note that there are two key elements to each one:

- **Part A** offers you a series of prompts about your learning which relate to
 - What you liked/found most useful
 - What you learned/were interested in/were not sure about/wanted to find out more about
 - How you felt/what surprised you/might have got more from if...
- **Part B** focuses on your personal concerns and needs and asked specifically what would help and support you most at a particular juncture.

Your responses to these prompts are very important because research shows that the impact of professional development is affected by factors that support or hinder teachers from making changes, including your:

- **Concerns**
(Particularly those expressed in Part B above; what were they and most importantly, were they acknowledged and responded to during the course of the professional development programme, thereby helping you to move on?)
- **Self-efficacy**
(Your perception that you are capable of undertaking the challenges presenting with success; specifically, was there a significant change in your knowledge and skills as you progressed through the year?)
- **Cognitive styles**
(Your particular way of learning; did the processes used in delivering the professional development meet your learning needs – e.g. were the approaches sufficiently varied; was there opportunity for group discussion and debate; did you feel you were being “given the truth” by an expert or constructing knowledge and understanding together; did you get a chance to build on previous knowledge and skills or was the focus running in parallel with your current interests and needs?)
- **Reflectiveness**
(To what extent have you had an opportunity to reflect on your learning/teaching practices during the year? In general, have your skills of critical reflection been heightened by this involvement and, if so, is this something that you think has become sufficiently embedded in your practice for it to be sustained?)
- **Formal education and years of experience**
(As a professional group, we brought different educational backgrounds, number of years in teaching and a range of different experiences to the table. Do you think that given your personal background, the professional development you received worked for you and, if so, was it enhanced or hindered by the heterogeneous nature of the group?)

Note: Not all of these issues will arise as you review your learning logs and that's fine; it's just useful to be aware of them. Others will come to the fore when you look at your AfL audits and when you participate in the focus group discussion on your experiences of involvement with the TLC.

Suggested Approach

1. Start by reviewing Part A of the logs. It might be useful to use highlighters to categorise the kinds of responses you made over the period of the year
 - What you liked/found most useful (green)
 - What you learned/were interested in/were not sure about/wanted to find out more about (orange)
 - How you felt/what surprised you/might have got more from if.... (blue).
2. Then take a look at the concerns, needs and supports you identified in Part B; these have been printed out on individual sheets for ease of analysis.
3. Now compare sections A and B. Remember, you are trying to map the changes in your concerns and needs over time, in the context of the supports you received. The Concerns-Based Adoption Model shown in Table 1 below (and other developmental models of its type) holds that people considering and experiencing change evolve in the kinds of questions they ask and in their use of whatever the change is. In general, early questions are more self-oriented: What is it? and How will it affect me? If, and when, these questions are resolved, questions emerge that are more task-oriented: How do I do it? How can I use these materials efficiently? How can I organize myself? and Why is it taking so much time? Finally, if, and when, self- and task- concerns are resolved, the individual can focus on impact: How will this affect the children...?

Table 1: Concerns-Based Adoption Model

Stages of Concern	Expressions of Concern
7. Refocusing	<i>I have some ideas about something that would work even better.</i> (The focus is on exploration of more universal benefits from the innovation, including the possibility of major changes or replacement with a more powerful alternative. Individual has definite ideas about alternatives to the proposed or existing form of the innovation.)
6. Collaboration	<i>How can I relate what I am doing to what others are doing?</i> (The focus is on collaboration and cooperation with others regarding the innovation.)
5. Consequence	<i>How is my use affecting learners? How can I refine it to have more impact?</i> (Attention focuses on impact of the innovation on students in her/his immediate sphere of influence. The focus is on relevance of the innovation for students, evaluation of student outcomes, including performance and competencies, and changes needed to increase student outcomes.)
4. Management	<i>I seem to be spending all my time getting materials ready.</i> (Attention is focused on the processes and tasks of using the innovation and the best use of information and resources. Issues related to efficiency, organizing, managing, scheduling and time demands are utmost.)
3. Personal	<i>How will using it affect me?</i> (Individual is uncertain about the demands of the innovation, her/his inadequacy to meet those demands, and her/his role with the innovation. This involves analysis of her/his role in relation to the reward structure of the organization, decision-making, and consideration of potential conflicts with existing structures or personal commitment. Financial or status implications of the programme for self and colleagues may also be reflected.)
2. Informational	<i>I would like to know more about it.</i> (A general awareness of the innovation and interest in learning more detail about it is indicated. The person seems to be unworried about herself/himself in relation to the innovation. She/he is interested in substantive aspects of the innovation in a selfless manner such as general characteristics, effects, and requirements for use.)
1. Awareness	<i>I am not concerned about it</i> (Little concern about or involvement with the innovation is indicated.)

Use your colour-coding of Part A and the pages listing the concerns, needs and supports you identified to answer the questions overleaf.

What kinds of concerns and needs did you express during the intervention period?

1. Did they evolve over time? Where did you start on the CBAM scale and where are you currently?

2. If your progression was steady and incremental, what supported you most in this?

3. If you seem to have become stuck somewhere along the line, where did this happen and why? More importantly, why do you think you remain stuck at this point and what support(s) do you think would be required to move you on?

Any other comments. Please add overleaf.

Appendix N

Teachers' Independent Review of the Professional Development Provided

So far you have engaged in two complementary reviews – the Learning Logs and the AfL Audits. In reviewing the Learning Logs, the focus was on reviewing your personal responses to the individual professional development sessions over the year and, specifically, on the concerns, needs and supports you identified as being important to you. In contrast, the AfL audits gave you an insight into the development of your knowledge, skills and attitudes in relation to AfL - having engaged with the project for a year - both in the TLC and in the classroom. This session seeks to build on this previous work by offering you an opportunity to review -collectively and anonymously - the professional development, including the content, process and context elements. Researchers categorise the factors that impact on the success of professional development in different ways. For ease of categorisation, three main areas may be identified:

1. **Content characteristics:** the “what” of professional development including: the credibility and scope of the practice or concept being conveyed; the ideas, practices, and strategies taught or suggested during the professional development etc.
2. **Process variables:** the “how” of professional development, the model(s) used (e.g. traditional = one/two-day withdrawal vs. sustained site-based/job-embedded, Teacher Learning Community) and their key characteristics including -
 - Quality of facilitation (e.g. on Day One - input days: nature and range of methods used - video, PowerPoint, Hand outs, Think Pieces, Brainstorming...; personality/style/accessibility/approachability of facilitator...; opportunities for discussion/debate/reflection...)
 - Planning and organization of sessions (e.g. on Day Two – video review days: protocols used to support “disciplined conversations” e.g. learning logs, video footage, video review sheets, lesson planning sheets...)
3. **Context characteristics:** the “who” (people involved/personalities...), “when” (timing/duration...), “where” (school-based vs. hotel/education center...), and “why” of the professional development (motivation/need – personal, school, national contexts); the organizational or school/system culture (specific to St. John’s); the expectations and incentives for using new practices/ideas etc.

Suggested Approach

1. Consider each of the three categories in turn: content, process and context. In the case of each, use the prompts above to help you identify and discuss the issues. It might be helpful to brainstorm the categories initially, record all your responses and then agree on the top two or three that, as a group, you believe warrant detailed discussion.
2. Appoint a rapporteur to note the final comments of the group; it is important that these are read back to the group so that you are satisfied that they reflect the opinions of all accurately and completely. Do this for each of the three categories.
3. Finally, review the three sets of comments together and on a separate sheet make your final overall comments/remarks. The tables attached may aid you in your discussions.

<i>Content: Key Issues Raised</i>	<i>Final Comments/Conclusions re. Content</i>
<i>Process: Key Issues Raised</i>	<i>Final Comments/Conclusions re. Process</i>
<i>Context: Key Issues Raised</i>	<i>Final Comments/Conclusions re. Context</i>
Final Overall Comments/Conclusions	

<i>Content: Key Issues Raised</i>	<i>Final Comments/Conclusions re. Content</i>

<i>Process: Key Issues Raised</i>	<i>Final Comments/Conclusions re. Process</i>

<i>Context: Key Issues Raised</i>	<i>Final Comments/Conclusions re. Context</i>

Final Overall Comments/Conclusions

Appendix O

Teachers' Independent Review Findings

<i>Content: Key Issues Raised</i>	<i>Final Comments/Conclusions re. Content</i>
Learning around Learning Intentions and Success Criteria	Too much time spent – got bogged down with language – was it necessary? But everything fell into place afterwards!
Feedback and Peer/Self assessment	We felt initially that the children were too young but found it worked.
Theoretical vs. practical introduction	A bit overwhelming! Very theoretical to begin with – maybe practical examples/videos better.
Nothing done on AFL re. Irish curriculum	Surprised it was such a shot in the dark for us all. St. Pat's/NCCA lacking work done on AFL in Irish context.
Our own videos!	Content from our own videos as important as AFL research/handouts.
Reading comprehension	Content outside AFL strategies/techniques. New approach to reading comprehension which came from working on Success Criteria/Learning Intentions.
Strategies	How we adapted them – they became personal and we could pick and choose from a range.

<i>Process: Key Issues Raised</i>	<i>Final Comments/Conclusions re. Process</i>
Facilitators	Debate with regard to expert versus facilitator = prefer non-expert but slightly more versed in content; Quality of facilitator = open, driven, understanding, motivated and motivating; Very accessible.
Range of Methods	Volume of handouts was overwhelming; English videos = only useful at end of process for purpose of affirmation.
Planning of Sessions	Review day (videos) should have happened on first day of session.
Video Review Sheets	Difficult to begin with but it was good to have a written record and feedback; Zita's modelling was very good; however, she moved beyond A/L remit at times and was a little unrealistic of what could be done in a single lesson.
Lesson Planning Sheet	Felt like a long process but final draft was ok; Felt like these were more for Cliona's benefit; Still not sure how useful they are from teacher's perspective – own personal notes may be more useful.
Learning Logs	Arduous but necessary; More reflective time needed (earlier time); Suggestion – “soft background music”; Sentence starters sometimes weren't relevant to session.

Context: Key Issues Raised	Final Comments/Conclusions re. Context
Who?	<p>Zita's enthusiasm and motivation were brilliant but at times too theoretical and too much jargon used;</p> <p>Great to have Jen as liaison person, organisation of tea/coffee, room, subs – liaising with Zita;</p> <p>Principal and other staff always very supportive and willing to help out;</p> <p>Importance of group moral among teachers involved. Discussions always took place in a safe, honest and trusting environment.</p> <p>Constructive criticism given to each other.</p>
When?	<p>The scheduling of days after Bank Holidays not ideal;</p> <p>Having 2 days together for review and in-service very helpful.</p>
Where?	<p>School ideal location to maintain focus and accessibility to resources and materials being used;</p> <p>Initially, there was difficulty finding available rooms – frequent interruptions and also the inconvenience experienced by staff members who had to move from their rooms.</p>
<p>Why?</p> <p>Rejuvenation of our teaching skills</p>	<p>Presented us with a personal, professional challenge and a chance for professional development, substitute-supported and based in school.</p>

Appendix P

Teachers' Video Review Sheet – P.I.N.

Focus feedback on the learning objective/success criteria of the task.

In the past, feedback has often focused on four main elements:

1. Presentation
2. Quantity
3. Surface features of any writing (especially spelling)
4. Effort.

While these aspects are important, we have *overemphasised* them so that the main focus of the lesson has been marginalized. We need to focus on these elements every now and again rather thereby selling the message to the children that what matters most is the learning intention of the lesson which is what we will be judging their work on and basing our feedback on.

3 conditions for effective feedback:

The learner has to

1. Possess a concept of the standard (or goal, or, reference level) being aimed for
2. Compare the actual (or current) level of performance with the standard and
3. Engage in appropriate action which leads to some closure of the gap (Sadler, 1989).

Feedback only leads to learning gains when it includes guidance about how to improve.

The greatest motivational benefits will come from focusing feedback on:

1. The qualities of the student's work, and not on comparison with other students
2. Specific ways in which the student's work could be improved
3. Improvements that the student has made compared to his/her earlier work.

So:

1. Give specific and concrete strategies to help the student improve; don't simply reiterate the LI
2. Allow time for the improvements to be made: for assessment to be formative, the feedback information has to be **used**.

Students can be trained to self- and peer-assess their work.

Assuming that the criteria are clear – LI and associated SC, students can learn to reflect on their own work and that of another leading to more independent, self-regulated and motivated work.

Good Guide: Making Assessment for Learning Work in Your Classroom: Practical Strategies for Enhancing Student Learning pg. 8 – 9.

Closing the Gap: Review of Teaching Videos: PIN Strategy (Praise-Improvement-Next)

Teacher Observed: _____

Observer: _____

Review Date: _____

Praise (specific to what the teacher was trying to achieve as per LIs and SC)

Improvements (improvements that could be made)

Next (suggest the most immediate step(s) forward to help this person improve)

Possible prompts:

- Reminder (*Think more about; consider why you did ...*)
- Scaffold (*Can you explain more about; try one of these options, ideas/suggestions*)
- Example (*Suggest a particular strategy or idea: You might try this...*)

BIBLIOGRAPHY

- Adey, P.S. (1995). *The effects of a staff development program: The relationship between the level of use of innovative science curriculum activities and student achievement*, (1-11). Retrieved February, 2009 from http://eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/13/fb/1d.pdf
- Akiba, M., LeTendre, G. K., & Scribner, J. P. (2007). Teacher quality, opportunity gap, and national achievement in 46 countries. *Educational Researcher*, 36(7), 369-387.
- Alexander, K., Entwisle, D., & Olson, L. (2001). Schools, achievement, and inequality: A seasonal perspective. *Educational Evaluation and Policy Analysis*, 23(2), 171-191.
- Alexander, R. (2004). Still no pedagogy? Principle, pragmatism and compliance in primary education. *Cambridge Journal of Education*, 34(1), 7-33.
- Allal, L., & Lopez, L. M. (2005). Formative assessment of learning: a review of publications in French. In: J. Looney (Ed.), *Improving learning through formative assessment: cases, policies, research* (1-25). Paris, France: Organisation for Economic Co-operation and Development.
- Alton-Lee, A. (2006). How teaching influences learning: Implications for educational researchers, teachers, teacher educators and policy makers. *Teaching and Teacher Education*, 22(5), 612-626.
- Amrein-Beardsley, A. (2008). Methodological concerns about the education value-added assessment system. *Educational Researcher*, 37(2), 65-75.
- Anderson, J. R., Greeno, J. G., Reder, L. M., & Simon, H. A. (2000). Perspectives on learning, thinking, and activity. *Educational Researcher*, 29(4), 11-13.
- Anderson, S. E. (1997). Understanding teacher change: Revisiting the concerns based adoption model. *Curriculum Inquiry*, 27(3), 331-367.

- Angen, M. J. (2000). Peals, pith, and provocation: Evaluating interpretive inquiry: Reviewing the validity debate and opening the dialogue. *Qualitative Health Research*, 10(3), 378-395.
- Arafeh, S., & McLaughlin, M. (2002). *Legal and ethnical issues in the use of video in education research*. (NCES 2000-2001, U.S. Department of Education). Washington, DC: National Centre for Educational Statistics.
- Archer, P., & O'Flaherty, B. (1991). Literacy Problems Among Primary School Leavers in Dublin's Inner City. *Studies in Education*, 7, 7-13.
- Arkalgud, R. (1983). On the definition of feedback. *Behavioral Science*, 28(1), 4-13.
- Armstrong, T. (2000). Two types of shock in modernity. *Critical Quarterly*, 42(1), 60-71.
- Ashcroft, K., & Foreman-Peck, L. (1995). Assessment and evaluation. In K. Ashcroft & L. Foreman-Peck (Eds.), *The lecturer's guide to quality and standards in colleges and universities* (pp. 109-136). UK: Routledge.
- Assessment Reform Group (ARG) (2002). *Assessment for Learning: 10 Principles*. Retrieved January, 2009 from <http://gtcni.openrepository.com/gtcni/bitstream/2428/4623/1/Assessment%20for%20Learning%20-%2010%20principles.pdf>
- Atkin. M., & Black, P. (2003). *Inside Science Education Reform: a history of curricular change*. New York: Teachers' College Press.
- Au, K. H. (1998). Social constructivism and the school literacy learning of students of diverse backgrounds. *Journal of Literacy Research*, 30(2), 297-319.
- Avramidis, E., & Smith, B. (1999). An introduction to the major research paradigms and their methodological implications for special needs research. *Emotional and Behavioural Difficulties*, 4(3), 27-36.
- Ball, S. (2003). The teacher's soul and the terrors of performativity, *Journal of Education Policy*, 18(20), 215-288.

- Ball, D. L., & Cohen, D. K. (1999). Developing practice, developing practitioners: Toward a practice-based theory of professional education. In L. Darling-Hammond and G. Sykes (Eds.), *Teaching as the learning profession* (pp. 3–31). San Francisco, CA: Jossey-Bass.
- Bandura, A. (1993a). The effects of teachers' behaviour on pupils' attributions: A review. In C. Antake & C. Brewin (Eds.), *Attributions and psychological change: Applications of attributional theories to clinical and education practice* (pp. 311-320). London: Academic Press.
- Bandura, A. (1993b). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist*, 28, 117-148.
- Bangert-Drowns, R. L., Kulik, C-L. C., Kulik, J. A., & Morgan, M. T. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research*, 61 (2), 213-238 (1991).
- Bangert-Drowns, R. L., & Rudner, L. M. (1991). Meta-analysis in educational research. *Practical Assessment, Research & Evaluation*, 2(8). Retrieved February, 2009 from <http://PAREonline.net/getvn.asp?v=2&n=8>
- Barab, S. A., Barnett, M., & Squire, K. (2002). Developing an empirical account of a community of practice: Characterizing the essential tensions. *The Journal of the Learning Sciences*, 11(4), 489-542.
- Barker, B. (2007). The leadership paradox: Can school leaders transform student outcomes? *School Effectiveness and School Improvement*, 18(1), 21-43.
- Baroudi, Z. (2007). Formative assessment: Definition, elements and role in instructional practice. *Post-Script: Postgraduate Journal of Education Research*, 8(1); 37-48.
- Barry, M. J. (2005). A School's Journey into Data-Driven Decision Making: A Case-study of Data's Impact on Instruction in an urban school, *Unpublished Doctoral Thesis*, University of Southern California, pp. 1-72, Retrieved February, 2009 from <http://www.usc.edu/dept/education/inside/FC/docs/ch1->

- Bazeley, P. (2003). Teaching mixed methods. *Qualitative Research Journal*, 3(Special Issue), 117-126.
- Bazeley, P. (2004). Issues in mixing qualitative and quantitative approaches to research. In R. Buber, J. Gadner, & L. Richards (Eds.), *Applying qualitative methods to marketing management research* (pp. 141-156). UK: Palgrave Macmillan.
- Benjamin, S., Nind, M., Hall, K., Collins, J., & Sheehy, K. (2003). Moments of inclusion and exclusion: Pupils negotiating classroom contexts. *British Journal of Sociology of Education*, 24(5), 547-458.
- Bergan, J. R., Sladeczek, I. E., Schwarz, R. D., & Smith, A. N. (1991). Effects of a measurement and planning system on kindergartners' cognitive development and educational programming. *American Educational Research Journal*, 28, 683-714.
- Bernhardt, V. L. (2004). Continuous Improvement: It takes more than test scores, *ACSA Leadership*, pp. 16-19. Retrieved February, 2009 from <http://principalsoffice.osu.edu/ACSAArticle.Bernhardt.pdf>
- Bernhardt, V. L. (2005). Data Tools for School Improvement, *Educational Leadership*, 62(5), 66-69. Retrieved February, 2009 from <ftp://eff.csuchico.edu/Downloads/ArticlesOfNote/DataTools.pdf>
- Biggs, J. (1998). Assessment and classroom learning: A role for summative assessment. *Assessment in Education: Principles, Policy & Practice*. Retrieved February, 2009 from <http://search.ebscohost.com/login.aspx?direct=true&db=tfh&AN=725614&site=ehost-live>
- Biggs, J. (2001). The reflective institution: Assuring and enhancing the quality of teaching and learning. *Higher Education*, 41(3), 221-238.
- Black, P. (2000). Research and the development of educational assessment. *Oxford*

- Black, P. (2001). Dreams, strategies and systems: Portraits of assessment past, present and future. *Assessment in Education: Principles, Policy & Practice*, 8(1), 65-85.
- Black, P. (2005). Formative assessment: Views through different lenses. *Curriculum Journal*, 16(2), 133-135.
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2002). *Working inside the black box: Assessment for learning in the classroom*. London: King's College, Department of Education and Professional Studies.
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning: Putting it into practice*. Maidenhead, UK: Open University Press.
- Black, P., McCormick, R., James, M., & Pedder, D. (2006a). Learning how to learn and assessment for learning: A theoretical inquiry. *Research Papers in Education*, 21(2), 119-132.
- Black, P., Swann, J., & Wiliam, D. (2006b). School pupils' beliefs about learning. *Research Papers in Education*, 21(2), 151-170.
- Black, P., & Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-74.
- Black, P., & Wiliam, D. (1998b). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139-148.
- Black, P., & Wiliam, D. (2003). In praise of educational research: Formative assessment. *British Educational Research Journal*, 29(5), 623-637.
- Black, P., & Wiliam, D. (2005). Lessons from around the world: How policies, politics and cultures constrain and afford assessment practices. *Curriculum Journal*, 16(2), 249-261.
- Black, P., & Wiliam, D. (2006a). Assessment for learning in the classroom. In J. R. Gardner (Ed.), *Assessment and learning* (pp. 9-26): Thousand Oaks, CA: Sage.

- Black, P., & Wiliam, D. (2006b). Developing a theory of formative assessment. In J. R. Gardner (Ed.), *Assessment and learning* (pp. 81-100): Thousand Oaks, CA: Sage.
- Black, P., & Wiliam, D. (2007). Large-scale assessment systems: Design principles drawn from international comparisons. *Measurement: Interdisciplinary Research and Perspectives*, 5(1), 1-53.
- Blanchard, J. (2003). Targets, assessment for learning, and whole-school improvement. *Cambridge Journal of Education*, 33(2), 257-271.
- Bloom, B. S., Hasting, J. T., Madaus, G. F. (1971). *Handbook on formative and summative evaluation of student learning*. New York: McGraw-Hill.
- Boote, D. N., & Beile, P. (2005). Scholars before researchers: On the centrality of the dissertation literature review in research preparation. *Educational Researcher*, 34(6), 3-15.
- Boote, D. N., & Beile, P. (2006). On "literature reviews of, and for, educational research": A response to the critique by Joseph Maxwell. *Educational Researcher*, 35(9), 32-35.
- Borko, H. (1997). New forms of classroom assessment: Implications for staff development. *Theory Into Practice*, 36(4), 231-238.
- Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher*, 33(8), 3-15.
- Borko, H., Mayfield, C., Marion, S. F., Flexer, R., & Cumbo, K. (1997). Teachers' developing ideas and practices about mathematics performance assessment: Successes, stumbling blocks, and implications for professional development. *Teaching and Teacher Education*, 13, 259-278.
- Boud, D. (2000). Sustainable assessment: Rethinking assessment for the learning society. *Studies in Continuing Education*, 22(2), 151-167.
- Brandom, A., Carmichael, P., & Marshall, B. (2005). Learning about assessment for learning: A framework for discourse about classroom practice. *Teacher*

- Brighton, A. (2000). Art: Francis bacon's modernism. *Critical Quarterly*, 42(1), 137-142.
- Brindley, G. (2007). Editorial. *Language Assessment Quarterly*, 4(1), 1-5.
- Broadfoot, P. (1998). Records of achievement and the learning society: A tale of two discourses. *Assessment in Education: Principles, Policy & Practice*, 5(3), 447-477.
- Broadfoot, P., & Black, P. (2004). Redefining assessment? The first ten years of assessment in education. *Assessment in Education: Principles, Policy & Practice*, 11(1), 7-26.
- Brookhart, S. M. (2004). Classroom assessment: Tensions and intersections in theory and practice. *The Teachers College Record*, 106(3), 429-458.
- Brookhart, S. M. (2005). Research on formative classroom assessment. Paper presented at the Formative classroom assessment: Research, theory, and practice, Symposium conducted at the annual meeting of the *American Educational Research Association*, Montreal, Canada.
- Brown, G. (2003). Teachers' instructional conceptions: Assessment's relationship to learning, teaching, curriculum, and teacher efficacy, *Joint Conference of the Australian and New Zealand Associations for Research in Education (AARE/NZARE)*. Auckland, NZ.
- Bryman, A. (2007). Barriers to integrating qualitative and quantitative research. *Journal of Mixed Methods Research*, 1(1), 8-22.
- Buber, R., Gadner, J., & Richards, L. (Eds.). (2004). *Applying qualitative methods to marketing management research*: UK: Palgrave Macmillan.
- Buhagiar, M. A. (2007). Classroom assessment within the alternative assessment paradigm: Revisiting the territory. *Curriculum Journal*, 18(1), 39 -56.
- Butler, R. (1987). Task-involving and ego-involving properties of evaluation: Effects

- of different feedback conditions on motivational perceptions, interest and performance. *Journal of Educational Psychology*, 79(4), 474-482.
- 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.
- Cahillane McGovern, S. (2006). *Assessment messages: consumerism and connoisseurship in communities of practice*. Unpublished doctoral dissertation, UK: Leeds Metropolitan University.
- Cantrell, S. C., & Hughes, H. K. (2008). Teacher efficacy and content literacy implementation: An exploration of the effects of extended professional development with coaching. *Journal of Literacy Research*, 40(1), 95-127.
- Carless, D. (2005). Prospects for the implementation of assessment for learning. *Assessment in Education*, 12(1), 39-54.
- Carless, D. (2006). Differing perceptions in the feedback process. *Studies in Higher Education*, 31(2), 219 -233.
- Carless, D. (2007). Conceptualizing pre-emptive formative assessment. *Assessment in Education: Principles, Policy & Practice*, 14(2), 171-184.
- Carmichael, P., Fox, A., McCormick, R., Procter, R., & Honour, L. (2006). Teachers' networks in and out of school. *Research Papers in Education*, 21(2), 217-234.
- Chan, C. K. K., & Pang, M. F. (2006). Teacher collaboration in learning communities. *Teaching Education*, 17(1), 1-5.
- Charlton, S., Hannan, B., Herrick, C., Landy, M., & Mahar, S. (2005). *Research into learning: Implications for teaching: Edited and abridged extracts* (No. 2). State of Victoria: Department of Education and Training.
- Chin, R., & Benne, K. D. (1969). General Strategies for Effecting Changes in Human Systems. In G. Warren, & K. D. Bennis (Eds.), *The Planning of Change* (pp. 32-59). New York: Holt, Rinehart and Winston.

- Christie, D., Cassidy, C., Skinner, D., Coutts, N., Sinclair, C., Rimpilainen, S., et al. (2007). Building collaborative communities of enquiry in educational research. *Educational Research and Evaluation*, 13(3), 263-278.
- Clancy, P. (2005). Education Policy. In S. Quin, S., P. Kennedy, A. Matthews, & G. Kiely (Eds.), *Contemporary Irish social policy*. (pp. 80-114). Dublin: University College Dublin Press.
- Clarke, D. J., & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teacher and Teacher Change*, 18, 947-967.
- Claxton, G. L. (1995). What kind of learning does self-assessment drive? Developing a 'nose' for quality: Comments on Klenowski. *Assessment in Education: Principles, Policy & Practice*, 2(3), 339-343.
- Clucas, S. (April 2000). Cultural phenomenology and the everyday. *Critical Quarterly*, 42(1), 8-34.
- Clymer, J. B., & Wiliam, D. (2006-2007). Improving the way we grade science. *Best of Educational Leadership*, 64(4), 36-42.
- Cobb, P. (1994). Where is the mind? Constructivist and socio-cultural perspectives on mathematical development. *Educational Researcher*, 23(7), 13-20.
- Cobb, P., & Bowers, J. (1999). Cognitive and situated learning perspectives in theory and practice. *Educational Researcher*, 28(2), 4-15.
- Cochran-Smith, M. (2003). Learning and unlearning: The education of teacher educators. *Teaching and Teacher Education*, 19(1), 5-28.
- Cochran-Smith, M. (2005). The new teacher education: For better or for worse? *Educational Researcher*, 34(7), 3-17.
- Cochran-Smith, M. (2006). Taking stock in 2006: Evidence, evidence everywhere. *Journal of Teacher Education*, 57(1), 6-12.
- Cochran-Smith, M., & Lytle, S. L. (1999a). Relationships of knowledge and practice: Teacher learning in communities. *Review of Research in Education*, 24, 249-

- Cochran-Smith, M., & Lytle, S. L. (1999b). The teacher research movement: A decade later. *Educational Researcher*, 28(7), 15-25.
- Cochran-Smith, M., Feiman-Nemser, S., McIntyre, D., & Demers, K. (2008). *Handbook of research on teacher education: Enduring questions in changing contexts*. New York, NY: Routledge.
- Codd, J. A. (1988). The construction and deconstruction of educational policy documents, *Journal of Education Policy*, 3(3), 235-247.
- Coffey, J. E., Sato, M., & Thiebault, M. (2005). Classroom assessment up close and personal. *Teacher Development*, 9(2), 169-184.
- Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist*, 49, 997-1003.
- Cohen, D. K., Raudenbush, S. W., & Ball, D. L. (2003). Resources, instruction, and research. *Educational Evaluation and Policy Analysis*, 25(2), 119-142.
- Condie, R., Livingston, K., & Seagraves, L. (2005). *Evaluation of the Assessment is for learning programme: Final report and appendices – February 2005*. Retrieved January, 2009 from <http://www.scotland.gov.uk/Publications/2005/12/0792641/26428>
- Connor, S. (2000). Making an issue of cultural phenomenology. *Critical Quarterly*, 42(1), 2-6.
- Connor, S., & Trotter, D. (April 2000). Introduction. *Critical Quarterly*, 42(1), 1-1.
- Conway, P. F. (2002). Learning in communities of practice: Rethinking teaching and learning in disadvantaged contexts. *Irish educational studies*, 21(3), 61-91.
- Coolahan, J. (2003). Attracting, Developing and Retaining Effective Teachers: Country Background Report for Ireland, pp. 1-65. Retrieved February, 2009 from <http://www.oecd.org/dataoecd/62/22/19196740.pdf>
- Cowie, B. (2005). Pupil commentary on assessment for learning. *Curriculum Journal*,

- Creswell, J. (2002). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Upper Saddle River, NJ: Merrill Prentice Hall.
- Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research*, 58, 438-481.
- Crossouard, B. (2008). A socio-cultural reflection on formative assessment and collaborative challenges in the states of Jersey. *Research Papers in Education*(1), 1-16.
- Cuban, L. (1984) *How Teachers Taught: Constancy and Change in American Classrooms*, New York: Longman.
- Cuban, L. (1998). How schools change reforms: Redefining success and failure. *Teachers College Record*, 99(3), 453-477.
- Darling-Hammond, L. (1999). *Teacher quality and student achievement: A review of state policy evidence*. Seattle WA: Center for the Study of Teaching and Policy, University of Washington.
- Darling-Hammond, L. (2000). How teacher education matters. *Journal of Teacher Education*, 51(3), 166-173.
- Darling-Hammond, L., & McLaughlin, M. (1995). Policies that support professional development in an era of reform. *Phi Delta Kappan*, 76, 597-604.
- Daugherty, R. (2007). Mediating academic research: The assessment reform group experience. *Research Papers in Education*, 22(2), 139-153.
- David, C. (2005). Prospects for the implementation of assessment for learning. *Assessment in Education: Principles, Policy & Practice*, 12(1), 39-54.
- Davies, P. (2000). Differentiation: Processing and understanding in teachers' thinking and practice. *Educational Studies*, 26(2), 191-203.

- Davies, P. (2006). Peer assessment: Judging the quality of students' work by comments rather than marks. *Innovations in Education and Teaching International*, 43(1), 69-82.
- Davies, B., Browne, J., Gannon, S., Honan, E., Laws, C., & Mueller-Rockstroh, B. (2004). The ambivalent practices of reflexivity. *Qualitative Inquiry*, 10(3), 360-389.
- Delandshere, G. (2001). Implicit theories, unexamined assumptions and the status quo of educational assessment. *Assessment in Education: Principles, Policy & Practice*, 8(2), 113-133.
- Department of Education and Science (2004). *Minister for Education and Science, Noel Dempsey T.D., to make Standardised Testing of Pupils a Requirement, 16th July 2004*, Department of Education and Science: Press Room. Retrieved February, 2009 from <http://www.irlgov.ie/educ>
- Department of Education and Science (1999). *Revised Primary Curriculum*. Dublin: Stationery Office.
- Department of Education and Science (2005a). *An Evaluation of Curriculum Implementation in Primary Schools: English, Mathematics and Visual Arts*. Dublin: Stationery Office.
- Department of Education and Science (2005b). *Chief Inspector's Report, 2001-2004*. Dublin: Stationery Office.
- Department of Education and Science (2005c). *Literacy and Numeracy in Disadvantaged Schools: Challenges for Teachers and Learners: An Evaluation by the Inspectorate of the Department of Education and Science*. Dublin: Stationery Office.
- Department of Education and Science (2005d). *DEIS: Delivering Equality of Opportunity in Schools – An action plan for educational inclusion*. Dublin: Stationery Office.
- Derry, S. J. (Ed.). (2007). *Guidelines for video research in education*:

- DeVellis, R.F. (2003). *Scale Development: Theory and Applications*. Thousand Oaks: Sage Publications.
- Dillon, D., O'Brien, D., & Heilman, E. (2000). Literacy research in the next millennium: From paradigms to pragmatism and practicality. *Reading Research Quarterly*, 35(1), 10-26.
- Dillon, D. R. (2005). There and back again: Qualitative research in literacy education. *Reading Research Quarterly*, 40(1), 106-110.
- Dixon, H., & Williams, R. (July 2003). Teachers' understanding and use of formative assessment in literacy learning. *New Zealand Annual Review of Education*. Retrieved February, 2009 from http://www.tki.org.nz/r/assessment/research/research20_e.php
- Doolaard, S. (2007). Effective schools. *School Effectiveness and School Improvement*, 18(2), 241-244.
- Dooner, A., Mandzuk, D., & Clifton, R. A. (2008). Stages of collaboration and the realities of professional learning communities. *Teaching and Teacher Education*, 24(3), 564-574.
- Downes, P., & Gilligan, A. (Eds). (2007). *Beyond educational disadvantage*. Dublin: Institute of Public Administration.
- Doyle, L., & Godfrey, R. (2005). Investigating the reliability of the key stage 2 test results for assessing individual pupil achievement and progress in England. *London Review of Education*, 3(1), 29-45.
- Driscoll, D., Appiah-Yeboah, A., Salib, P., & Rupert, D. J. (2007). Merging qualitative and quantitative data in mixed methods research: How to and why not. *Ecological and Environmental Anthropology*, 3(1), 19-28.
- DuFour, R. (2004). What is a "professional learning community"? *Educational Leadership*, 61(8), 6-11.

- DuFour, R., & Eaker, R. (1998). *Professional learning communities at work*. Bloomington, IN: National Education Service.
- Duncombe, R., & Armour, K. M. (1999). Collaborative professional learning: From theory to practice. *Journal of In-Service Education*, 30(1), 141-166.
- Dweck, C. (2000). *Self-theories: Their role in motivation, personality and development*. Philadelphia: Psychology Press.
- Dwyer, C. A. (1998). Assessment and classroom learning: Theory and practice. *Assessment in Education: Principles, Policy & Practice*, 5(1), 131-137.
- Earl, L. (2005). From accounting to accountability: Harnessing data for school improvement, *Research Conference: Using Data to Support Learning*, pp. 6-10, Retrieved February, 2009 from <http://www.acer.edu.au/workshops/documents/Earl.pdf>
- Ecclestone, K., & Pryor, J. (2003). Learning careers or assessment careers? The impact of assessment systems on learning. *British Educational Research Journal*, 29(4), 471- 488.
- Ecclestone, K., Swann, J., Greenwood, M., Vobar, J., & Eldred, J. (2004). Improving Formative Assessment in Vocational Education and Basic Skills. Research Project in progress. Retrieved January, 2009 from www.exeter.ac.uk
- Editorial. (2006). Learning-oriented assessment: Principles and practice. *Assessment & Evaluation in Higher Education*, 31(4), 395-398.
- Edwards, R. (2005). Contexts, boundary objects and hybrid spaces: theorising learning in lifelong learning. Paper presented at the *35th Annual SCUTREA Conference* July 5-July 7 2005. England, UK: University of Sussex. Retrieved February, 2009 from <http://www.leeds.ac.uk/educol/documents/142037.htm>
- Edwards, R., & Miller, K. (2007). Putting the context into learning. *Pedagogy, Culture & Society*, 15(3), 263-274.
- Eisenhart, M. (1991). Conceptual frameworks for research circa 1991: Ideas from a cultural anthropologist; implications for mathematics education researchers.

Paper presented at the Proceedings of the *Thirteenth Annual Meeting North American Chapter of the International Group for the Psychology of Mathematics Education*, Blacksburg, Virginia, USA.

Eisenhart, M. (2005). Science plus: A response to the responses to scientific research in education. *The Teachers College Record*, 107(1), 52-58.

Eisenhart, M., & Shepard, L. (June 1999). Annual meeting 2000. *Educational Researcher*, 28(5), 42-43.

Eisenhart, M., & Towne, L. (2003). Contestation and change in national policy on "scientifically based" education research. *Educational Researcher*, 32(7), 31-38.

Eisner, E. W. (1993). Reshaping assessment in education: Some criteria in search of practice. *Journal of Curriculum Studies*, 25(3), 219-233.

Eivers, E., Shiel, G., & Shortt, F. (2004). *Reading literacy in disadvantaged primary schools*. Dublin: Educational Research Centre.

Eivers, E., Shiel, G., & Shortt, F. (2005). *Literacy in disadvantaged primary schools: Problems and solutions*. Dublin: Educational Research Centre.

Ellsworth, J., Martinez, C., Lyon, C., & Wylie, E. C. (2007). Leading while learning: Supporting the leaders of keeping learning on track. Paper presented at the annual meetings of the *American Educational Research Association and the National Council on Measurement in Education*. Chicago.

Elmore, R. (2000). Building a new structure for school leadership. *The Albert Shanker Institute*. Retrieved February, 2009 from http://www.leadspace.govt.nz/leadership/pdf/Building_a_New_Structure.pdf

Elmore, R. (2002). *Bridging the gap between standards and achievement: Report on the imperative for professional development in education*. Washington, DC: Albert Shanker Institute.

Elmore, R. (2004). *School reform from the inside out: Policy, practice, and performance*. Cambridge, MA: Harvard Educational Review.

- Elwood, J. (2006). Formative assessment: Possibilities, boundaries and limitations. *Assessment in Education: Principles, Policy & Practice*, 13(2), 215-232.
- Elwood, J., & Klenowski, V. (2002). Creating communities of shared practice: The challenges of assessment use in learning and teaching. *Assessment & Evaluation in Higher Education*, 27(3), 243-256.
- Educational Testing Service (ETS) (2007). *Keeping learning on track: TLC leader kits*: Princeton, NJ: ETS.
- Eva, K. W. (2004). Issues to consider when planning and conducting educational research. *Journal of Dental Education*, 68(3), 316-323.
- Fallon, J. (2005). Targeting disadvantage among young children in the Republic of Ireland: An overview. *Child Care in Practice*, 11(3), 289-311.
- Firestone, W. A., & Corbett, H. D. (1981). Schools versus linking agents as contributors to the change process. *Educational Evaluation and Policy Analysis*, 3(2), 5-17.
- Firestone, W. A., Mangin, M. M., Martinez, M. C., & Polovsky, T. (2004). *Content and coherence in district professional development: Three case studies. Publication Series No. 3.*
- Fontana, D., & Fernandes, M. (1994). Improvements in mathematics performance as a consequence of self-assessment in Portuguese primary school pupils. *British Journal of Educational Psychology*, 64(4), 407-417.
- Forness, S. R. (2001). Special education and related services: What have we learned from meta-analysis? *Exceptionality*, 9(4), 185-197.
- Fraser, C., Kennedy, A., Reid, L., & McKinney, S. (2007). Teachers' continuing professional development: Contested concepts, understandings and models. *Journal of In-Service Education*, 33(2), 153-169.
- Frederiksen, J. R., & White, B. J. (1997). Reflective assessment of students' research within an inquiry-based middle school science curriculum. Paper presented at *the Annual Meeting of the AERA*. US: Chicago, 1997.

- Freeman, M., deMarrais, K., Preissle, J., Roulston, K., & St. Pierre, E. A. (2007). Standards of evidence in qualitative research: An incitement to discourse. *Educational Researcher*, 36(1), 25-32.
- Fuchs, L.S., & Fuch, D. (1986). Effects of systematic e evaluation: a meta-analysis. *Exceptional Children*, 53, 199-208.
- Fullan, M. (1991). *The new meaning of educational change*: London: Cassell.
- Fullan, M. (2001). *Leading in a culture of change*: San Francisco: Jossey-Bass.
- Fullan, M. (2006). Leading professional learning: think 'system' and not 'individual' if the goal is to fundamentally change the culture of schools. *School Administrator*, 63(10), 10-14.
- Fullan, M., Hill, P. T., & Crevola, C. (2006). *Breakthrough*: Thousand Oaks, CA: Corwin Press.
- Gardner, J. (Ed.). (2006). *Assessment and Learning*. London: Sage Publications.
- Gardner, J. (2007). Is teaching a 'partial' profession? *Make the Grade, Summer*, 18-21.
- Gardner, J., Harlen, W., Hayward, L., & Stobart, G. (2008). *Changing assessment practice process: Principles and standards*: Assessment Reform Group.
- Gasson, S. (2002). Position paper: Phenomenological approaches and inter-subjectivity. Paper presented at *The First International Workshop on Interpretive Approaches to Information Systems and Computing Research, Special Interest Group Interpretive Approaches and Methods*. Brunel University, UK.
- Gijbels, D., & Dochy, F. (2006). Students' assessment preferences and approaches to learning: Can formative assessment make a difference? *Educational Studies*, 32(4), 399-409.
- Gipps, C. (1999). Socio-cultural aspects of assessment. *Review of Research in Education*, 24, 355-392.
- Glatthorn, A. (1995). Teacher Development. In L. Anderson (Ed.), *International*

Encyclopedia of teaching and teacher education (pp. 41-46). London: Pergamon Press.

- Goe, L., & Mardy, D. (2007). A teacher-driven implementation of assessment for learning in New Jersey. Paper presented at the *AERA/NCME Conference: Tight but Loose: Scaling Up Teacher Professional Development in Diverse Contexts*. Chicago: ETS.
- Graue, M. E. (1993). Integrating theory and practice through instructional assessment. *Educational Assessment*, 1(4), 283-309.
- Greaney, V., & Mulryan, C. (1991). Elementary education in Ireland. *The Elementary School Journal*, 92(1, Special Issue), 97-108.
- Green, J. (2000). Language: Language maketh man. *Critical Quarterly*, 42(1), 100-104.
- Griffin, G. A. (1983). Introduction: the work of staff development. In G. A. Griffin (Ed.), *Eighty-Second Yearbook of the National Society for the Study of Education* (1-12). Chicago: University of Chicago Press.
- Grossman, P., Wilson, S. M., & Shulman, L.S., (1989). Teachers of substance: Subject matter knowledge for teaching. In M. C. Reynolds (ed.). *Knowledge Base for the Beginning Teacher*, pp. 23-36. Oxford: Pergamon Press.
- Grossman, P., Wineburg, S., Woolworth, S. (2001). Toward a theory of teacher community. *Teachers College Record*, 103, 942-1012.
- Guskey, T. R. (1998). Teacher efficacy, self-concept and attitudes toward the implementation of instructional innovation. *Teachers and Teaching, Theory and Practice*, 4(1), 63-69.
- Guskey, T. R. (2000). *Evaluating Professional Development*. Thousand Oaks, CA: Corwin Press.
- Guskey, T. R. (2002). Professional development and teacher change. *Teachers and Teaching, Theory and Practice*, 8(3), 381-391.

- Guskey, T. R. (2005). Formative classroom assessment and Benjamin S. Bloom: Theory, research, and implications. Paper presented at the *Annual Meeting of the American Educational Research Association*, Montreal, Canada.
- Guskey, T. R., & Sparks, D. (1996). Exploring the relationship between staff development and improvements in student learning. *Journal of Staff Development*, 17(4), 3-14.
- Guthrie, J. T., Wigfield, A., & VonSecker, C. (2000). Effects of integrated instruction on motivation and strategy use in reading. *Journal of Educational Psychology*, 92(2), 331-341.
- Hackley, C. E. (1998). Social constructivism and research in marketing and advertising. *Qualitative Research Journal: An International Journal*, 1(3), 125-131.
- Haigh, M., & Dixon, H. (2007). Why am I doing these things? Engaging in classroom-based inquiry around formative assessment. *Journal of In-Service Education*, 33(3), 359-376.
- Hall, G. E., & Ford, S. M. (1987). *Change in schools: Facilitating the process*. Albany: Sunny Press.
- Hall, K. (2000). A conceptual evaluation of primary assessment policy and the education policy process in the Republic of Ireland. *Compare: A Journal of Comparative Education*, 30(1), 85-101.
- Hall, K. (2002). Primary assessment policy in the Republic of Ireland. *Educational Management Administration & Leadership*, 30(3), 261-274.
- Hall, K., & Burke, W. (2004). *Making formative assessment work: Effective practice in the primary classroom*. UK: Open University Press.
- Hall, K., & Harding, A. (2002). Level descriptions and teacher assessment in England: Towards a community of assessment practice. *Educational Research*, 44(1), 1-16.

- Hall, K., & Kavanagh, V. (2002). Primary assessment in the Republic of Ireland: Conflict and consensus. *Educational Management and Administration*, 30(3), 261-274.
- Hall, K., Conway, P., Rath, A., Murphy, R., & McKeon, J. (2008). Reporting to Parents in Primary School: Communication, Meaning and Learning. Report Commissioned by the National Council for Curriculum and Assessment (NCCA). Retrieved February 2009 from http://www.ncca.ie/uploadedfiles/Primary/UCC_researchReporting.pdf
- Hamilton, L. (2003). Assessment as a policy tool. *Review of Research in Education*, 27, 25-68.
- Hanushek, E. A. (2004). *Some simple analytics of school quality*. Washington, DC: National Bureau of Economic Research.
- Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (1998). Teachers, school and academic achievement. Washington, DC: National Bureau of Economic Research. Retrieved February, 2009 from <http://www.mccsc.edu/~curriculum/teachers,%20schools,%20and%20acheivement.pdf>
- Hargreaves, A. (1992). Balkanized secondary schools and the malaise of modernity. Paper presented at the *Annual Meeting of the American Educational Research Association*, San Francisco, CA.
- Hargreaves, A. (1994). *Changing Teachers, changing times: Teachers' work and culture in the postmodern age*. NY: Teachers' College Press.
- Hargreaves, A., & Goodson, I. (2006). Educational change over time? The sustainability and nonsustainability of three decades of secondary school change and continuity. *Educational Administration Quarterly*, 42(1), 3-41.
- Hargreaves, E. (2005). Assessment for learning? Thinking outside the (black) box. *Cambridge Journal of Education*, 35(2), 213-224.
- Hargreaves, E. (2007). The validity of collaborative assessment for learning.

- Harlen, W. (2006a). On the relationship between assessment for formative and summative purposes. In J. Gardner (Ed.), *Assessment and learning* (103-118). London: Sage Publications.
- Harlen, W. (2006b). The role of assessment in developing motivation for learning. In J. Gardner (Ed.), *Assessment and learning* (61-80). London: Sage Publications.
- Harrington, P., Gillam, K., Andrews, J., & Day, C. (2006). Changing teaching and learning relationships through collaborative action research: Learning to ask different questions. *Teacher Development*, 10(1), 73-86.
- Harris, A. (2008). Leading innovation and change: knowledge creation by schools for schools. *European Journal of Education*, 43(2), 119-228.
- Harris, A., Chapman, C., Muijs, D., Russ, J., & Stoll, L. (2006). Improving schools in challenging contexts: Exploring the possible. *School Effectiveness and School Improvement*, 17(4), 409-424.
- Harris, L. (2007). Employing formative assessment in the classroom. *Improving Schools*, 10(3), 249-260.
- Harrison, C. (2005). Teachers developing assessment for learning: Mapping teacher change. *Teacher Development*, 9(2), 255-263.
- Hattie, J. (2005). What is the nature of evidence that makes a difference to learning? *Research Conference: Using Data to Support Learning*, Melbourne, 11-21. Retrieved February, 2009 from http://www.acer.edu.au/documents/RC2005_Hattie.pdf
- Hattie, J., & Jaeger, R. (1998). Assessment and classroom learning: A deductive approach. *Assessment in Education: Principles, Policy & Practice*, 5(1), 111-122.
- Hawley, W. D., & Valli, L. (2001). The essentials of effective professional development: A new consensus. In D. Boesel (Ed.), *Continuing professional*

- development* (pp. 1-17). Washington, DC: U.S. Department of Education, National Library of Education.
- Hayward, E. L. (2007). Curriculum, pedagogies and assessment in Scotland: The quest for social justice: 'Ah kent yir faither'. *Assessment in Education: Principles, Policy & Practice*, 14(2), 251-268.
- Hayward, L., & Hedge, N. (2005). Traveling towards change in assessment: Policy, practice and research in education. *Assessment in Education: Principles, Policy & Practice*, 12(1), 55-75.
- Hayward, L., Priestley, M., & Young, M. (2004). Ruffling the calm of the ocean floor: Merging practice, policy and research in assessment in Scotland. *Oxford Review of Education*, 30(3), 397-415.
- Higgins, R., Hartley, P., & Skelton, A. (2002). The conscientious consumer: Reconsidering the role of assessment feedback in student learning. *Studies in Higher Education*, 27(1), 53-64.
- Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal*, 42(2), 371-406.
- Hindin, A., Morocco, C. C., Mott, E. A., & Aguilar, C. M. (2007). More than just a group: Teacher collaboration and learning in the workplace. *Teachers and Teaching*, 13(4), 349-376.
- Hord, S. M., Meehan, M. L., Orletsky, S., & Sattes, B. (1999). Assessing a school staff as a community of professional learners. Retrieved February, 2009 from <http://eric.ed.gov/ERICWebPortal/custom/portlets/recordDetails/detailmini.jsp?nfpb=true&ERICEstSearch SearchValue 0=ED439490&ERICEstSearch SearchType 0=no&accno=ED439490>
- Hogan, P., & Smith, G. (2006). A new model of CPD: The TL21 Project (ROI). In SCoTENS (The Standing Conference on Teacher Education North and South) (2006). Teacher education and schools together towards improvement (79-81). *Report by the Centre for Cross Border Studies*.

- Irish National Teachers' Organisation (INTO) (1997). *Teaching and learning: Issues in assessment*. Dublin: INTO.
- Irish Times Newspaper. (November, 22, 2007). Teachers to get assessment guidelines. Retrieved February, 2009 from <http://www.irishtimes.com/search/index.html?rm=listresults&filter=datedesc&keywords=Dr.+Sarah+Fitzpatrick>
- Ingvarson, L. (2005). Getting professional development right, *Using Data to Support Learning: Conference Proceedings, Melbourne*, pp. 63-71. Retrieved February, 2009 from <http://www.acer.edu.au/workshops/documents/Ingvarson.pdf>
- Ingvarson, L., Meiers, M., & Beavis, A. (2005). Factors affecting the impact of professional development programs on teachers' knowledge, practice, student outcomes and efficacy. *Education Policy Analysis Archives*, 13(10), 1-28.
- InPraxis Group, I. (2006). *Professional learning communities: An exploration*. Canada: Alberta Education.
- James, M. (2006a). Assessment, teaching and theories of learning. In J. Gardner (Ed.), *Assessment and Learning* (47-60). London: Sage Publications.
- James, M. (2006b). Balancing rigour and responsiveness in a shifting context: Meeting the challenges of educational research. *Research Papers in Education*, 21(4), 365-380.
- James, M., Black, P., McCormick, R., Pedder, D., & Wiliam, D. (2006). Learning how to learn, in classrooms, schools and networks: Aims, design and analysis. *Research Papers in Education*, 21(2), 101-118.
- James, M., & Pedder, D. (2006a). Beyond method: Assessment and learning practices and values. *Curriculum Journal*, 17(2), 109-138.
- James, M., & Pedder, D. (2006b). Professional learning as a condition for assessment for learning. In J. Gardner (Ed.), *Assessment and Learning* (27-44). London: Sage Publications.

- John, T. (2005). The formative use of assessment information in planning: The notion of contingent planning. *British Journal of Educational Studies*, 53(1), 54-65.
- Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, 33(7), 14-26.
- Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. (2007). Towards a definition of mixed methods research. *Journal of Mixed Methods Research*, 1(2), 112-133.
- Johnston, K., Murchan, D., Loxley, A., Fitzgerald, H., & Quinn, M. (2007). The role and impact of the regional curriculum support service in Irish primary education. *Irish Educational Studies*, 26(3), 219-238.
- Johnston, P., & Costello, P. (2005). Theory and research into practice: Principles for literacy assessment. *Reading Research Quarterly*, 40(2), 256-267.
- Johnston, P. H., Afflerbach, P., & Weiss, P. B. (1993). Teachers' assessment of the teaching and learning of literacy. *Educational Assessment*, 1(2), 91-117.
- Katzell, R., & Thompson, E. (1990). Work motivation: Theory and practice. *American Psychologist*, 45(2), 144-153.
- Kazelskis, R., Thames, D., & Reeves, C. (2004). The elementary reading attitude survey: Factor invariance across gender and race. *Reading Psychology*, 25(2), 111-120.
- Kennedy, M. M. (2007). Defining a literature. *Educational Researcher*, 36(3), 139-147.
- Killion, J. (1999). *What works in the middle: Results-based staff development*. Oxford, OH: National Staff Development Council.
- Kirkup, C. (2006). Using assessment information to inform teaching and learning. *Education 3-13*, 34(2), 153-162.
- Kirton, A., Hallam, S., Peffers, J., Robertson, P., & Stobart, G. (2007). Revolution, evolution or a trojan horse? Piloting assessment for learning in some Scottish primary schools. *British Educational Research Journal*, 33(4), 605-627.

- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254-284.
- Knapp, M. S. (2003). Professional development as a policy pathway. *Review of Research in Education*, 27, 109-157.
- Knight, P. (2002). A systemic approach to professional development: Learning as practice. *Teaching and Teacher Education*, 18(3), 229-241.
- Koellner, K., Jacobs, J., Borko, H., Schneider, C., Pittman, M. E., & Eiteljorg, E. (2007). The problem-solving cycle: A model to support the development of teachers' professional knowledge. *Mathematical Thinking and Learning*, 9(3), 273-303.
- Korthagen, F. A., & Kessels, J. P. (1999). Linking theory and practice: Changing the pedagogy of teacher education. *Educational Researcher*, 28(4), 4-17.
- Kuhn, T. S. (1970). *The structure of scientific revolutions*. Chicago: University of Chicago Press.
- Kuhn, T. (2002). Negotiating boundaries between scholars and practitioners: Knowledge, networks, and communities of practice. *Management Communication Quarterly*, 16(1), 106-112.
- Kush, J. C., & Watkins, M. W. (1996). Long-term stability of children's attitudes toward reading. *Journal of Educational Research*, 89, 315-319.
- Labaree, D. F. (2003). The peculiar problems of preparing educational researchers. *Educational Researcher*, 32(4), 13-22.
- Lather, P. (1992). Critical frames in educational research: feminist and post-structural perspectives. *Theory into Practice*, XXXI(2), 87-99.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. New York: Cambridge University Press.
- Leahy, S., Lyon, C., Thompson, M., & Wiliam, D. (2005). Classroom assessment: Minute-by-minute and day-by-day. *Educational Leadership*, 63(3), 18-24.

- Learning Unlimited, (2004). *The Learning Set*, www.learningunlimited.co.uk
- Leat, D., Lofthouse, R., & Wilcock, A. (2006). Teacher coaching; connecting research and practice. *Teaching Education*, 17(4), 329-339.
- Lee, A. (2008). *Introducing assessment for learning to children in an Irish special school: An exploration of the challenges involved*. Unpublished master's thesis, St. Patrick's College, Dublin.
- Lee, C., & William, D. (2005). Studying changes in the practice of two teachers developing assessment for learning. *Teacher Development*, 9(2), 265-283.
- Leech, N., & Onwuegbuzie, A. (2007). A typology of mixed methods research designs. *Quality and Quantity: International Journal of Methodology*, 43(2), 265-275.
- Leitch, R., Gardner, J., Mitchell, S., Lundy, L., Odena, O., & Galanouli, D. (2007). Consulting pupils in assessment for learning classrooms: The twists and turns of working with students as co-researchers. *Educational Action Research*, 15(3), 459-478.
- Lester, F. (2005). On the theoretical, conceptual, and philosophical foundations for research in mathematics education. *ZDM: The International Journal of Mathematics Education*, 37(6), 457-467.
- Leung, C., & Mohan, B. (2004). Teacher formative assessment and talk in classroom contexts: Assessment as discourse and assessment of discourse. *Language Testing*, 21(3), 335-359.
- Leung, C., & Rea-Dickins, P. (2007). Teacher assessment as policy instrument: Contradictions and capacities. *Language Assessment Quarterly: An International Journal*, 4(1), 6-36.
- Leusner, D. M., Ellsworth, J., & Goe, L. (2007). Scaling up across diverse contexts: Lessons learned from five implementations of the keeping learning on track program, *Paper presented at the annual meetings of the American Educational Research Association and the National Council on Measurement in Education*.

Chicago.

- Lewis, C., Perry, R., & Murata, A. (2006). How should research contribute to instructional improvement? The case of lesson study. *Educational Researcher*, 35(3), 3-14.
- Lieberman, A., & Pointer Mace, D. H. (2008). Teacher learning: The key to educational reform. *Journal of Teacher Education*, 59(3), 226-234.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic Inquiry*. Newbury Park, California: Sage Publications.
- Lingard, B. (2007). Pedagogies of indifference. *International Journal of Inclusive Education*, 11(3), 245-266.
- Lingard, B., & Mills, M. (2007). Pedagogies making a difference: Issues of social justice and inclusion. *International Journal of Inclusive Education*, 11(3), 233-244.
- Lingard, B., Mills, M., & Hayes, D. (2006). Enabling and aligning assessment for learning: Some research and policy lessons from Queensland. *International Studies in Sociology of Education*, 16(2), 83-103.
- Linn, R. L. (2000). Assessments and accountability. *Educational Researcher*, 29(2), 4-16.
- Little, J. W. (1993). Teachers' professional development in a climate of educational reform. *Educational Evaluation and Policy Analysis*, 15(2), 129-152.
- Loucks-Horsley, S., & Matsumoto, C. (1999). Research on professional development for teachers of mathematics and science: The state of the scene. *School Science and Mathematics*, 99(5), 258-271.
- Loucks-Horsley, S., & Stiegelbauer, S. (1991). Using knowledge of change to guide staff development. In A. Lieberman & L. Miller (Eds.), *Staff development for education in the '90s: New demands, new realities, new perspectives*, Second Edition. (pp. 37-60). New York: Teachers College Press.

- Looney, A. (2001). Curriculum as policy: some implications of contemporary policy studies for the analysis of curriculum policy, with particular reference to post-primary curriculum policy in the Republic of Ireland. *The curriculum journal*, 12(2), 149-162.
- Looney, A. (2006). Assessment in the Republic of Ireland. *Assessment in Education: Principles, Policy & Practice*, 13(3), 345-353.
- Lortie, D. (1975). *Schoolteacher: a sociological study*. Chicago: University of Chicago Press.
- Louise, H., Mark, P., & Myra, Y. (2004). Ruffling the calm of the ocean floor: Merging practice, policy and research in assessment in Scotland. *Oxford Review of Education*, 30(3), 397-415.
- Loveless, T. (1998). The use and misuse of research in educational reform. In D. Ravitch (Ed.), *Brookings papers on educational policy: 1998*. (pp. 279-317). Washington, DC: Brookings Institute.
- Loxley, A., Johnston, K., Murchan, D., Fitzgerald, H., & Quinn, M. (2007). The role of whole-school contexts in shaping the experiences and outcomes associated with professional development. *Journal of In-Service Education*, 33(3), 265-285.
- Lyon, C., Wylie, E., & Goe, L. (2006). Changing teachers, changing schools. Paper Presented at the *American Educational Research Association and the National Council on Measurement in Education*. San Francisco, CA.
- Lyotard, J. F. (1984). *The Postmodern Condition: A Report on Knowledge*. University of Minnesota Press.
- MacCabe, C. (2000). Editorial. *Critical Quarterly*, 42(1), ii-ii.
- McKinney, M., Sexton, T., & Meyerson, M. J. (1999). Validating the efficacy-based change model. *Teacher and Teacher Education*, 15, 471-485.
- MacShane, D. (2000). Politics: Political johnnies. *Critical Quarterly*, 42(1), 128-136.

- Maher, J., & Wiliam, D. (2007). Keep learning on track in new teacher induction. Paper presented at the annual meetings of the *American Educational Research Association and the National Council on Measurement in Education*. Chicago.
- Marchant, P., & Hall, K. (2003). Explaining differences in key stage 2 pupil attainments: A multilevel analysis. *London Review of Education*, 1(2), 142-151.
- Marcos, J. J. M., & Tillema, H. (2006). Studying studies on teacher reflection and action: An appraisal of research contributions. *Educational Research Review*, 1(2), 112-132.
- Martin, M. (1979). Reading Attainment in Irish Primary Schools-A Progressive achievement gap? Paper presented at the *Annual Meeting of the Reading Association of Ireland*, September 1979. Retrieved February, 2009 from http://eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/32/fc/d2.pdf
- Martinez, J. G. R., & Martinez, N. C. (1992). Re-examining repeated testing and teacher effects in a remedial mathematics course. *British Journal of Educational Psychology*, 62, 356-363.
- Marion, S., & Pellegrino, J. (2006). A validity framework for evaluating the technical quality of alternate assessments. *Educational Measurement: Issues and Practice*, 25 (4), 47-57.
- Maroussia, R. (2004). Assessment in French and English infant schools: Assessing the work, the child or the culture? *Assessment in Education: Principles, Policy and Practice*, 11(2), 193-211.
- Marshall, B., & Drummond, M. J. (2006). How teachers engage with assessment for learning: Lessons from the classroom. *Research Papers in Education*, 21(2), 133-149.
- Marshall, B., & Wiliam, D. (2005). Editorial. *Teacher Development*, 9(2), 165-167.
- Matsumura, L. C., & Pascal, J. (2003). *Teachers' assignments and student work*:

Opening a window on classroom practice. C.S.E. Report. Los Angeles: National Centre for Research on Evaluation, Standards and Student Testing.

Maxwell, J. A. (2004). Causal explanation, qualitative research and scientific inquiry in education. *Educational Researcher*, 33(2), 3-11.

Maxwell, J. A. (Ed.). (2005a). *Qualitative research design: An interactive approach* (2nd ed.). London: Sage Publications.

Maxwell, J. A. (2005b). Conceptual framework: what do you think is going on? In J. A. Maxwell (Ed.). (2005). *Qualitative research design: An interactive approach* (pp. 33-63). London: Sage Publications.

Maxwell, J. A. (2006c). Literature reviews of, and for, educational research: A commentary on Boote and Beile's scholars before researchers. *Educational Researcher*, 35(9), 28-31.

McDonald, B., & Boud, D. (2003). The impact of self-assessment on achievement: The effects of self-assessment training on performance in external examinations. *Assessment in Education: Principles, Policy & Practice*, 10(2), 209-220.

McDonnell, L., & Elmore, R. (1987). Getting the job done: Alternative policy instruments. *Educational Evaluation and Policy Analysis*, 9(2), 133-152.

McDonnell, P. (2003). Developments in special education in Ireland: Deep structures and policy making. *International Journal of Inclusive Education*, 7(3), 259-269.

McKenna, M. C., & Kear, D. J. (1990). Measuring attitude toward reading: A new tool for teachers. *Reading Teacher*, 43, 626-639.

McKenna, M. C., Kear, D. J., & Ellsworth, R. A. (1995). Children's attitudes toward reading: A national survey. *Reading Research Quarterly*, 30, 934-955.

McKernan, J. (1991). Rationality and intention in Irish education: A grid for the design and evaluation of a core curriculum. *European Journal of Education*,

- McLaughlin, M. (2006). Beyond "misery research" - new opportunities for implementation research, policy and practice, (1-34). Draft of paper presented at an *International Invitational Symposium*, St. Patrick's College, Dublin.
- McLaughlin, M., & Talbert, J. (1993). *Contexts that matter for teaching and learning: Strategic opportunities for meeting the nation's educational goals*. CA: Palo Alto: Stanford University Center for Research on the Context of Secondary School Teaching.
- McLaughlin, M., & Talbert, J. E. (2006). *Building school-based TCLs: Professional strategies to improve student achievement*. New York: Teachers College, Columbia University.
- McMillan, J. H., & Schumacher, S. S. (1997). *Research in education: A conceptual introduction*. New York: Longman.
- McMunn, N., McColskey, W., & Butler, S. (Winter 2003-2004). Building teacher capacity in classroom assessment to improve student learning. *International Journal of Educational Policy, Research, & Practice*, 4(4), 25-48.
- McNaughton, S. (2007). Dame Professor Emerita Marie Clay. *Journal of Early Childhood Literacy*, 7(3), 263-266.
- Meirink, J. A., Meijer, P. C., & Verloop, N. (2007). A closer look at teachers' individual learning in collaborative settings. *Teachers and Teaching*, 13(2), 145-164.
- Miles, M. B., Saxl, E. R., & Lieberman, A. (1988). What skills do educational "change agents" need? An empirical view. *Curriculum Inquiry*, 18(2), 157-193.
- Millar, R., & Morton, M. (2007). Bridging two worlds: Special education and curriculum policy. *International Journal of Inclusive Education*, 11(2), 163-176.
- Miller, D., & Lavin, F. (2007). But now I feel I want to give it a try: Formative

- assessment, self-esteem and a sense of competence. *Curriculum Journal*, 18(1), 3-25.
- Miller, D., & Moran, T. (2007). Theory and practice in self-esteem enhancement: Circle-time and efficacy-based approaches: A controlled evaluation. *Teachers and Teaching*, 13(6), 601-615.
- Moore, M. (2007). Teacher professional development as a transformative experience. In P. L. Jeffery (Ed.). *Proceedings of Australian Association for Research in Education (AARE). Engaging Pedagogies*, 1, 1-22. Adelaide.
- Morgan, D. L. (2007). Paradigms lost and pragmatism regained: Methodological implications of combining qualitative and quantitative methods. *Journal of Mixed Methods Research*, 1(1), 48-76.
- Morgan, W., & Wyatt-Smith, C. M. (2000). Im/proper accountability: Towards a theory of critical literacy and assessment. *Assessment in Education: Principles, Policy & Practice*, 7(1), 123-142.
- Morine-Dersheimer, G. (1990). Book review: Research perspectives on the graduate preparation of teachers. *Journal of Teacher Education*, 41(3), 73-76.
- Muijs, D., Campbell, J., Kyriakides, L., & Robinson, W. (2005). Making the case for differentiated teacher effectiveness: An overview of research in four key areas. *School Effectiveness and School Improvement*, 16(1), 51-70.
- Muijs, D., Harris, A., Chapman, C., Stoll, L., & Russ, J. (2004). Improving schools in socio-economically disadvantaged areas: A review of research evidence. *School Effectiveness and School Improvement*, 15(2), 149-175.
- Muijs, D., & Lindsay, G. (2008). Where are we at? An empirical study of levels and methods of evaluating continuing professional development. *British Educational Research Journal*, 34(2), 195-211.
- Munns, G., & Woodward, H. (2006). Student engagement and student self-assessment: The real framework. *Assessment in Education: Principles, Policy & Practice*, 13(2), 193-213.

- Murchan, D., Johnson, K., Loxley, A., & Fitzgerald, H. (2006). Teacher learning and policy intention: Same coin, different sides? Paper prepared for presentation at the *European Conference on Educational Research*, Geneva.
- National Council for Curriculum and Assessment (NCCA) (2005a). *Interim Report on the developmental initiative for learning in junior cycle*. Dublin: NCCA.
- National Council for Curriculum and Assessment (NCCA) (2005b). *Primary Curriculum Review Phase 1: Summary of Findings and Recommendations*, Dublin: NCCA.
- National Council for Curriculum and Assessment (NCCA) (2005c). *The Primary Curriculum in Schools: Joint DES/NCCA Seminar: The Proceedings*, Dublin: NCCA.
- National Council for Curriculum and Assessment (NCCA) (2007). *Assessment in the Primary School Curriculum: Guidelines for Schools*, Dublin: NCCA.
- Natriello, G. (1987). The impact of evaluation processes on students. *Educational Psychologist*, 22(2), 155-175.
- Nelson, R. R. (2003). On the uneven evolution of human know-how. *Research Policy*, 32(6), 909-922.
- Newton, P. E. (2003). The defensibility of national curriculum assessment in England. *Research Papers in Education*, 18(2), 101-127.
- Newton, P. E. (2005). The public understanding of measurement inaccuracy. *British Educational Research Journal*, 31(4), 419-442.
- Newton, P. E. (2007). Clarifying the purposes of educational assessment. *Assessment in Education: Principles, Policy & Practice*, 14(2), 149-170.
- Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199-218.
- Nolen, A. L., & Putten, J. V. (2007). Action research in education: Addressing gaps in ethical principles and practices. *Educational Researcher*, 36(7), 401-407.

- Noonan, B., & Renihan, P. (2006). Demystifying assessment leadership. *Canadian Journal of Educational Administration and Policy*. Retrieved January, 2009 from <http://www.umanitoba.ca/publications/cjeap/articles/noonan.html>
- Nyquist, J. B. (2003). *The benefits of reconstructing feedback as a larger system of formative assessment: A meta-analysis*. Unpublished master's thesis, Vanderbilt University, Nashville, TN.
- Oakes, J., & Lipton, M. (1999). *Teaching to change the world*. Boston: McGraw-Hill.
- OECD (1991). Ireland: Reviews of National Policies for Education. Paris: OECD.
- OECD (2004). Additional Country Papers: Ireland, International Schooling for Tomorrow Forum, Canada: June 6-8th. Retrieved January, 2009 from <http://www.oecd.org/dataoecd/41/1/32507342.pdf>
- Ottoson, J. M. (1997). After the applause: Exploring multiple influences on application following an adult education program. *Adult Education Quarterly*, 47, 92-107.
- O'Leary, M. (2006). Towards a balanced assessment system for Irish primary and secondary schools. *Oideas*, 52, 7-24.
- O'Leary, M. (2008). Towards an agenda for professional development in assessment. *Journal of In-Service Education*, 34(1), 109-114.
- O'Leary, M., & Shiel, G. (1997). Curriculum profiling in Australia and United Kingdom: Some implications for performance-based assessment in the United States. *Educational Assessment*, 4(3), 203-235.
- O'Sullivan, D. (1992). Shaping education debate: a case study and an interpretation. *The economic and social review*, 23(4), 423-438.
- Onwuegbuzie, A. J., Witcher, A. E., Collins, K. M. T., Filer, J. D., Wiedmaier, C. D., & Moore, C. W. (2007). Students' perceptions of characteristics of effective college teachers: A validity study of a teaching evaluation form using a mixed-methods analysis. *American Educational Research Journal*, 44(1), 113-160.

- Oshima, J., Horino, R., Oshima, R., Yamamoto, T., Inagaki, S., & Takenaka, M. (2006). Changing teachers' epistemological perspectives: A case study of teacher researcher collaborative lesson studies in Japan. *Teaching Education*, 17(1), 43-53.
- Pallant, J. (2007). *SPSS survival manual: A step-by-step guide to data analysis using SPSS for Windows (Version 15)*, 3rd Edition. UK: Allen and Unwin.
- Pedder, D. (2006a). Are small classes better? Understanding relationships between class size, classroom processes and pupils' learning. *Oxford Review of Education*, 32(2), 213-234.
- Pedder, D. (2006b). Organizational conditions that foster successful classroom promotion of learning how to learn. *Research Papers in Education*, 21(2), 171-200.
- Pedder, D. (2007). Profiling teachers' professional learning practices and values: Differences between and within schools. *Curriculum Journal*, 18(3), 231-252.
- Pedder, D., James, M., & MacBeath, J. (2005). How teachers value and practise professional learning. *Research Papers in Education*, 20(3), 209-243.
- Perrenoud, P. (1998). From formative evaluation to a controlled regulation of learning: Towards a wider conceptual field. *Assessment in Education: Principles, Policy & Practice*, 5(1), 85-102.
- Pirie, S. E. B. (1996). Classroom video-recording: When, why and how does it offer a valuable data source for qualitative research? *Paper presented at the Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. Panama City.
- Plowman, L., & Stephen, C. (2008). The big picture? Video and the representation of interaction. *British Educational Research Journal*, 34(4), 541-565.
- Poulson, L., & Avramidis, E. (2003). Pathways and possibilities in professional development: Case studies of effective teachers of literacy. *British Educational Research Journal*, 29(4), 543-560.

- Price, M., Donovan, B., & Rust, C. (2007). Putting a social-constructivist assessment process model into practice: Building the feedback loop into the assessment process through peer review. *Innovations in Education and Teaching International*, 44(2), 143-152.
- Priestley, M. (2007). *The social practices of curriculum making*. Unpublished Doctor of Education, University of Stirling, Scotland.
- Priestley, M., & Sime, D. (2005). Formative assessment for all: A whole-school approach to pedagogic change. *Curriculum Journal*, 16(4), 475-492.
- Primary Curriculum Review Body. (1990). *Primary Curriculum Review Body Report*. Dublin: Department of Education.
- Primary Curriculum Support Programme (PCSP) (2008). Retrieved February, 2009 from <http://www.pcsp.ie/html/index.php>
- Pryor, J., & Crossouard, B. (2008). A socio-cultural theorisation of formative assessment. *Oxford Review of Education*, 34(1), 1-20.
- Putnam, R., & Borko, H. (1997). Teacher learning: Implications of new views of cognition. In B. J. Biddle, T. L. Good, & I. F. Goodson (Eds.), *The international handbook of teachers and teaching* (pp. 1223-1296). Dordrecht, The Netherlands: Kluwer.
- Putnam, R., & Borko, H. (2000). What do new views of knowledge and thinking have to say about research on teacher learning? *Educational Researcher*, 29(1), 4-15.
- Ramaprasad, A. (1983). On the definition of feedback. *Behavioral Science*, 28(1), 4-13.
- Raudenbush, S. W. (2005). Learning from attempts to improve schooling: The contribution of methodological diversity. *Educational Researcher*, 34(5), 25-31.
- Reeves, J., McCall, J., & MacGilchrist, B. (2001). Change leadership: Planning, conceptualization, and perception. In J. Macbeath & P. Mortimer (Eds.),

- Improving school effectiveness* (pp. 122-137). Buckingham, UK: Open University Press.
- Reiman, A. J. (1999). The evolution of the social role-taking and guided reflection framework in teacher education: Recent theory and quantitative synthesis of research. *Teaching and Teacher Education*, 15(6), 597-612.
- Robson, C. (2006). *Real World Research, 2nd Edition*. UK: Blackwell Publishing.
- Rodriguez, M. C. (2004). The role of classroom assessment in student performance on TIMMS. *Applied Measurement in Education*, 17(1), 1-24.
- Rosenholtz, S. (1989). *Teachers' workplace: The social organization of schools*. New York: Longman.
- Rosnow, R. L., & Rosenthal, R. (1996). Computing contrasts, effect sizes, and counternulls on other people's published data: General procedures for research consumers. *Psychological Methods*, 1, 331-340.
- Roos, B., & Hamilton, D. (2005). Formative assessment: A cybernetic viewpoint. *Assessment in Education: Principles, Policy & Practice*, 12(1), 7-20.
- Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General and Applied*, 80, 1-28.
- Ruddell, M. R. (2007). Literacy research, diverse perspectives, and the big tent. *Journal of Curriculum Studies*, 39(2), 235-252.
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18(2), 119-144.
- Sadler, D. R. (1998). Formative assessment: Revisiting the territory. *Assessment in Education: Principles, Policy & Practice*, 5(1), 77-84.
- Sato, M., Coffey, J., & Moorthy, S. (2005). Two teachers making assessment for learning their own. *Curriculum Journal*, 16(2), 177-191.
- Schunk, D. H. (1996). Goal and self-evaluative influences during children's cognitive skill learning. *American Educational Research Journal*, 33, 359-382.

- SCoTENS (The Standing Conference on Teacher Education North and South) (2006). *Teacher education and schools together towards improvement*. Report by the Centre for Cross Border Studies.
- Scriven, M. (1967). The methodology of evaluation. In R. W. Tyler, R. M. Gagne & M. Scriven (Eds.), *Perspectives of curriculum evaluation* (pp. 39-83). Chicago: Rand McNally.
- Sebatane, E. M. (1998). Assessment and classroom learning: A response to Black and William. *Assessment in Education: Principles, Policy & Practice*, 5(1), 123-130.
- Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. *Educational Researcher*, 27(2), 4-13.
- Shaxson, L. (2005). Is your research robust enough? Questions for policy makers and practitioners. *Evidence & Policy: A Journal of Research, Debate and Practice*, 1(1), 101-112.
- Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4-14.
- Shepard, L. A. (2005). Formative assessment: Caveat emptor. Paper presented at the ETS Invitational Conference: *The future of assessment: shaping teaching and learning*. New York: October 10-11, 2005.
- Sherin, M. G. (2004). New perspectives on the role of video in teacher education. In J. Brophy (Ed.), *Using video in teacher education* (pp. 1-27). New York: Elsevier Science.
- Sherin, M. G. (In press). The development of teachers' professional vision in video clubs. In R. Goldman, R. Pea, & S. Derry, (Ed.s). *Video research in the learning sciences*. Hillsdale: NJ: Lawrence Erlbaum.
- Shiel, G. (2007). Reading standards in Ireland: Moving forward, looking back. In B. Dwyer & G. Shiel (Eds.), *Literacy at the crossroads: Moving forward, looking back* (pp. 1-18). Dublin: Reading Association of Ireland.

- Shiel, G, Cosgrove, J., Sofroniou, N., & Kelly, A. (2001). *Ready for Life: The literacy achievement of Irish 15 year olds*. Dublin: Educational Research Centre.
- Shiel, G., Forde, P., & Morgan, M. (1996). International studies of educational achievement: what do they tell us? *Oideas*, 44, 80-97.
- Shulman, L.S. (1986). Those who understand: Knowledge growth in teaching. *Educational Research Review*, 57(1), 4-14.
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153-189.
- Singleton, R. A., & Straits, B.C. (1999). *Approaches to Social Research*. New York and Oxford Oxford: University Press.
- Slavin, R. E. (2008). Perspectives on evidence-based research in education-what works? Issues in synthesizing educational program evaluations. *Educational Researcher*, 37(1), 5-14.
- Sloane, F. (2008). Comments on Slavin: Through the looking glass: Experiments, quasi-experiments, and the medical model. *Educational Researcher*, 37(1), 41-46.
- Smardon, R. (2005). Where the action is: The micro-sociological turn in educational research. *Educational Researcher*, 34(1), 20-25.
- Smith, C., & Gillespie, M. (2007). Research on professional development and teacher change: implications for adult basic education (205-244). Retrieved February, 2009 from http://www.ncsall.net/fileadmin/resources/ann_rev/smith-gillespie-07.pdf
- Smith, E. (1995). Where is the mind? Knowing and knowledge in Cobb's constructivist and socio-cultural perspectives. *Educational Researcher*, 24(7), 23-24.
- Soodak, L. C., & Podell, D. M. (1996). Teacher efficacy: Toward the understanding of a multi-faceted construct. *Teaching & Teacher Education*, 12(4), 401-411.

- Sparkes, A. C. (1992). The paradigms debate: an extended review and a celebration of difference. In C. A. Sparkes (Ed.), *Research in Physical Education and Sport: Exploring Alternative Versions* (9-60). London: Falmer Press.
- Spillane, J. P., Diamond, J. B., Sherer, J. Z., & Coldren, A. F. (2005). Distributing Leadership. In M. J. Coles and G. Southworth, (Eds.). *Developing Leadership, Creating the schools of tomorrow*, pp. 37-49. Open University Press,
- Star, S. L. (1989). The structure of ill-structured solutions: Boundary objects and heterogeneous distributed problem solving. In *Distributed artificial intelligence* (pp. 37-54). UK: Morgan Kaufmann Publishers Inc.
- Stein, M. K., & Wang, M. C. (1988). Teacher development and school improvement: The process of teacher change. *Teacher and Teacher Education*, 4(2), 117-187.
- Stiggins, R. J. (2001). *Student-involved classroom assessment* (3rd ed.). Upper Saddle River, NJ: Prentice-Hall.
- Stiggins, R. J., Arter, J. A., & Chappuis, J. (2004). *Classroom assessment for student learning: Doing it right-using it well*. Portland, OR: Assessment Training.
- Stobart, G. (2001). The validity of national curriculum assessment. *British Journal of Educational Studies*, 49(1), 26-39.
- Stobart, G. (2006). Influencing classroom assessment. *Assessment in Education: Principles, Policy & Practice*, 13(3), 235-238.
- Stoll, L., & Louis, K. (2007). *Professional learning communities*. Milton Keynes: Open University Press.
- Stone, D. (2002). *Policy paradox*. London & New York: W.W. Norton.
- Sugrue, C. (1997). *Complexities of teaching: Child-centred perspectives*. London: Falmer Press.
- Sugrue, C. (2002). Irish teachers' experiences of professional learning: Implications for policy and practice. *Journal of In-Service Education*, 28(2), 311-338.

- Sugrue, C. (2003). Principals' Professional Development: Realities, Perspectives and Possibilities. *Oideas*, 50, 8-39.
- Sugrue, C. (2004a). Introduction. In C. Sugrue (Ed.), *Curriculum and ideology: Irish experiences, international perspectives* (pp. 1-15). Dublin: The Liffey Press.
- Sugrue, C. (2004b). Signposts and silences: Situating the local within the global. In C. Sugrue (Ed.), *Curriculum and ideology: Irish experiences, international perspectives* (pp. 269-308). Dublin: The Liffey Press.
- Sugrue, C. (2004c). Whose curriculum is it anyway? Power, politics and possibilities in the construction of the revised primary curriculum. In C. Sugrue (Ed.), *Curriculum and ideology: Irish experiences, international perspectives* (pp. 167-208). Dublin: The Liffey Press.
- Sugrue, C. (2006). Three decades of teacher education reform: 1973-1999. In J. Kelly (Ed.), *A history of St. Patrick's College* (pp. 225-265). Dublin: Four Courts Press.
- Sugrue, C. (2008). The future of educational change? In C. Sugrue (Ed.), *Future of educational change: International perspectives*. London and New York: Routledge.
- Sugrue, C., & Day, C. (Eds.). (2002). *Developing teachers and teaching: International research perspectives*. London: Routledge/Falmer Press.
- Sugrue, C., & Morgan, M., Devine, D., & Raftery, D. (2001). *The quality of Irish teachers' professional learning: A critical analysis*. A report commissioned by the Department of Education and Science, Dublin: Irish Stationery Office.
- Sugrue, C., & Ui Thuama, C. (1997). Lifelong learning for teachers in Ireland: Policy, provision and vision. *Journal of In-Service Education*, 23(1), 55-70.
- Swaffield, S., & MacBeath, J. (2006). Embedding learning how to learn in school policy: The challenge for leadership. *Research Papers in Education*, 21(2), 201-215.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Experimental Designs Using ANOVA*.

Mass: Pearson, Allyn & Bacon.

Taras, M. (2007). Machinations of assessment: Metaphors, myths and realities. *Pedagogy, Culture & Society*, 15(1), 55-69.

Tharp, R., & Gallimore, R. (1988). *Rousing minds to life: teaching, learning, and schooling in social context*. Cambridge, England: Cambridge University Press.

The Evidence-Based Intervention Work Group. (2005). Theories of change and adoption of innovations: The evolving evidence-based intervention and practice movement in school psychology. *Psychology in the Schools*, 42(5), 475-494.

Thompson, S. (2001). The authentic standards movement and its evil twin. *Phi Delta Kappan*, 82(5), 358-362.

Thompson, C. L., & Zeuli, J. S. (1999). *The frame and the tapestry: Standards-based reform and professional development*. San Francisco: Jossey-Bass.

Thompson, M., & Goe, L. (2006). Models for effective and scalable teacher professional development. Paper presented at the annual meeting of the *American Educational Research Association*. San Francisco, CA.

Thompson, M., Ponte, E., Paek, P., & Goe, L. (2004). *Study of the Impact of the California Formative Assessment and Support System for Teachers: Methodological considerations and recommendations for future research*, 1-47. *CFASST Report 4*. Retrieved February, 2009 from <http://ets.org/Media/Research/pdf/RR-04-33.pdf>

Thompson, M., & Wiliam, D. (2007). Tight but loose: A conceptual framework for scaling up school reforms. Paper presented at the annual meeting of the *American Educational Research Association*. Chicago, IL.

Thrupp, M. (2001). Recent school effectiveness counter-critiques: Problems and possibilities. *British Educational Research Journal*, 27, 443-457.

Tierney, R. D. (2006). Changing practices: Influences on classroom assessment. *Assessment in Education: Principles, Policy & Practice*, 13(3), 239-264.

- Timperley, H., & Alton-Lee, A. (2008). Reframing teacher professional learning: An alternative policy approach to strengthening valued outcomes for diverse learners. *Review of Research in Education*, 32(1), 328-369.
- Tocci, C., & Taylor, G. (2007). *A state-sponsored pilot project in selected schools in Vermont*. Paper presented at the *Tight but loose: Scaling up teacher professional development in diverse contexts*. Princeton, NJ.
- Topping, K. J., Samuels, J., & Paul, T. (2007). Does practice make perfect? Independent reading quantity, quality and student achievement. *Learning and Instruction*, 17(3), 253-264.
- Torrance, H. (2007). Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning. *Assessment in Education: Principles, Policy & Practice*, 14(3), 281-294.
- Trotter, D. (2000). The new historicism and the psychopathology of everyday modern life. *Critical Quarterly*, 42(1), 36-58.
- Tsui, A. B. M., & Law, D. Y. K. (2007). Learning as boundary-crossing in school-university partnership. *Teaching and Teacher Education*, 23(8), 1289-1301.
- Tuomi-Grohn, T., & Engestrom, Y. (Eds.). (2003). *Between work and school: new perspectives on transfer and boundary-crossing*. London, Pergamon.
- van den Berg, R., Sleepers, P., Geijssel, F., & Vandenberghe, R. (2000). Implementation of an innovation: Meeting the concerns of teachers. *Studies in Educational Evaluation*, 26(4), 331-350.
- Vygotsky, L. (1978). *Mind and society*. Cambridge, Mass: Harvard University Press
- Walsh, J. F. (1997). From consultation to legislation: a review of recent developments in Irish education. *Unpublished master's thesis*, University College, Cork.
- Walsham, G. (1995). Interpretive case studies in is research: Nature and method. *European Journal of Information Systems*, 4(2), 74-81.

- Watkins, C. (2005). Classrooms as learning communities: A review of research. *London Review of Education*, 3(1), 47-64.
- Watson, A. (2006). Some difficulties in informal assessment in mathematics. *Assessment in Education: Principles, Policy & Practice*, 13(3), 289-303.
- Weaver-Hightower, M. B. (2008). An ecology metaphor for educational policy analysis: A call to complexity. *Educational Researcher*, 37(3), 153-167.
- Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, 92, 548-573.
- Westheimer, J. (1999). Communities and consequences: An inquiry into ideology and practice in teachers' professional work. *Educational Administration Quarterly*, 35(1), 71-105.
- Whiting, B., Van Burgh, J. W., & Render, G. F. (1995). Mastery learning in the classroom. Paper presented at the *Annual Meeting of the AERA*. US: San Francisco.
- William, D. (2000). Education: The meanings and consequences of educational assessments. *Critical Quarterly*, 42(1), 105-127.
- William, D. (2006a). Assessment for learning: why, what and how. *Critical Quarterly*, 42(1), 105-127.
- William, D. (2006b). Assessment for learning: Why no profile in US policy. In J. Gardner (Ed.), *Assessment and learning* (169-184). London: Sage Publications.
- William, D. (2007). Five "key strategies" for effective formative assessment. Retrieved January, 2009 from http://www.nctm.org/uploadedFiles/Research_News_and_Advocacy/Research/Clips_and_Briefs/Research_brief_04_-_Five_Key%20Strategies.pdf
- William, D. (2008). Changing Classroom Practice. *Informative Assessment*, 65(4), 36-42.

- Wiliam, D., Lee, C., Harrison, C., & Black, P. (2004). Teachers developing assessment for learning: Impact on student achievement. *Assessment in Education: Principles, Policy & Practice*, 11(1), 49-65.
- Wiliam, D., & Thompson, M. (2007). Integrating assessment with instruction: What will it take to make it work? Paper presented at the annual meeting of the *American Educational Research Association*. Chicago, IL.
- Williams, J., & Ryan, J. (2000). National testing and the improvement of classroom teaching: Can they co-exist? *British Educational Research Journal*, 26(1), 49-73.
- Willis, D. (1993). Learning and assessment: Exposing the inconsistencies of theory and practice. *Oxford Review of Education*, 19(3), 383-402.
- Wilson, S. M., & Berne, J. (1999). Teacher learning and the acquisition of professional knowledge: An examination of research on contemporary professional development. *Review of Research in Education*, 24, 173-209.
- Windschitl, M. (2002). Framing constructivism in practice as the negotiation of dilemmas: An analysis of the conceptual, pedagogical, cultural, and political challenges facing teachers. *Review of Educational Research*, 72(2), 131-175.
- Wininger, S. R., & Norman, A. D. (2005). Teacher-Candidates' exposure to formative assessment in educational psychology textbooks: A content analysis. *Educational Assessment*, 10, 19-37.
- Wylie, E. C., Thompson, M., Lyon, C., & Snodgrass, D. (2007). Keeping learning on track in an urban district's low performing schools. Paper presented at the annual meetings of the *American Educational Research Association and the National Council on Measurement in Education*. Princeton, NJ.
- Wylie, E. C., Thompson, M., & Wiliam, D. (2007). Keep learning on track: Impact on student learning. Paper presented at the annual meetings of the *American Educational Research Association and the National Council on Measurement in Education*. Chicago.

- Yorke, M. (2003). Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher Education*, 45(4), 477-501.
- Zientek, L. R., Capraro, M. M., & Capraro, R. M. (2008). Reporting practices in quantitative teacher education research: One look at the evidence cited in the AERA panel report. *Educational Researcher*, 37(4), 208-216.