TITLE PAGE

Using Assessment for Learning to Enhance the Teaching and Learning of Mathematics in one Primary School: A Lesson Study Approach

By Ann Marie Gurhy

(B.Ed., H.Dip. in Ed., M.Ed. in School Leadership)

Thesis Submitted in Partial Fulfilment of the Requirements for the Award of the Degree of

Doctor of Education

Institute of Education, Saint Patrick's Campus

Dublin City University

Co-Supervisors: Dr. Zita Lysaght

Professor Michael O'Leary

September 2017

DECLARATION

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of Doctor of Education is entirely my own work, and that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge breach any law of copyright, 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.

Signature: ____________________________

Date: __________________

ID Number: 59275537

ABSTRACT

Using Assessment for Learning to Enhance the Teaching and Learning of Mathematics in one Primary School. A Lesson Study Approach

Ann Marie Gurhy

In recent decades the merits of assessment for learning (AfL) have been particularly lauded, nationally and internationally. Scholars have linked effective use of AfL with improved student learning and achievement, increased student motivation and self-esteem, enhanced self-regulation and metacognition, improved teacher professional and organisational learning and better student-teacher relationships. In the Irish context, while government policy emphasises the centrality of AfL in teaching and learning few teachers have received assessment-related continuing professional development (CPD). The Department of Education and Skills (DES, 2011a) has highlighted that AfL is not used sufficiently widely in our schools and concerns have also been raised about teacher assessment literacy. Regarding mathematics, data from the 2009 National Assessments of Mathematics and English Reading (DES, 2010b), school inspections (DES, 2010a), international reports (PISA, 2009), and the *Literacy and Numeracy Strategy* (DES, 2011a) have suggested Irish students are underperforming.

This practitioner action research case study aimed to address these issues. Operating within the pragmatic paradigm, it utilised a convergent parallel mixed methods design. Over the course of one academic year, it investigated the impact of AfL practices on the teaching and learning of mathematics at fourth-class level in one primary school. Specifically, it explored how the use of AfL principles, strategies and techniques affected students' attainment on standardised mathematics tests and their dispositions towards mathematics. Additionally, the research investigated the potential of lesson study (LS) as a vehicle of collaborative professional learning in AfL and considered the impact engaging in LS had on teachers' skills, knowledge, and use of AfL, and their beliefs towards AfL as a form of assessment. This study also provided unique insights into learners' perspectives of using AfL in mathematics, both teachers and students. Findings revealed significant effect size gains in children's confidence, motivation and attitudes regarding mathematics, although there was no appreciable difference in students standardised mathematics scores when compared to the comparison group. Additionally, indications are that teachers found LS to be a very effective model of CPD in AfL. Finally, implications for conducting further research are discussed.

This thesis is dedicated to my mother Anne who has always been an inspiration to me, and whose love, support and affirmation I value more than words

ACKNOWLEDGEMENTS

Throughout the period of researching and writing this thesis I have been extremely fortunate to have benefited from the advice and support of my family, supervisors, colleagues and friends. Therefore, there are many people that I want to acknowledge.

I wish to sincerely thank both my supervisors, Professor Michael O'Leary and Dr. Zita Lysaght, for their indispensable advice and unstinting support throughout the research and writing process. I am indebted to you both for your guidance, thoughtful critical suggestions, interest and encouragement throughout this journey, which has undoubtedly improved this dissertation. Each of you in your own way has challenged and extended my thinking and helped hone my research skills. In particular I'd like to thank Professor Michael O'Leary for his unfailing kindness, patience and constant encouragement, and particularly for his assistance with SPSS and all the quantitative data. I could not have asked for a better mentor and would have been lost without him. Thanks also to Dr. Zita Lysaght whose doctoral thesis helped inspire this one and who provided many insights and constructive comments that were invaluable and greatly appreciated. She generously shared her wealth of knowledge and academic experience and I feel fortunate that she was my supervisor.

Thank you sincerely to Mairin Lambe and Meliosa Walshe without whom this research would not have been possible. I am indebted to you both for giving so generously of your time in our collaborative endeavour and for participating so enthusiastically throughout the research process over one full academic year. It was a privilege to observe your teaching and I have learned so much from you both. Not only that but your engagement and support undoubtedly contributed to the success of the intervention. Additionally, your wisdom and wit made learning together fun.

Special thanks are due to all the children who took part in this case study and who made it such an enjoyable experience. Thanks also to members of the SEN team who facilitated the research lessons. Thanks to Catriona, Mairín, Marian, Mary and Melíosa for reading various chapters of this thesis along the way and especially to Bernie, Peter and Patricia for proofreading the whole document.

I am extremely fortunate to work closely with my principal, Bernie Murphy, and consider her a mentor and friend. She has facilitated, encouraged and supported me throughout this research journey and I am forever in her debt. Her work ethic, leadership and love of education are an inspiration to me and I thank her for all she has done, including supervision of classes and proofreading the final document.

Special thanks are due to my family, especially my mother Anne who has always been a constant source of inspiration and support to me, particularly over the last few years – it is to her that this dissertation is dedicated. Thanks also to Catríona who always helps at times of academic crisis, Patricia whose love of education inspires me and to Ruth and James for their encouragement.

To all the different scholars who answered my various queries and offered me advice and encouragement: Dr. Heidi Andrade, Professor Paul Black, Dr. Dolores Corcoran, Dr. Peter Dudley, Dr. Mairéad Hourigan, Professor Catherine Lewis, Dr. Siobhan Leahy, Dr. Aisling Leavy, and Professor Dylan Wiliam, Thank you all most sincerely. Thanks also to all my friends in the EdD. class for their insights and support along the journey and to Ms. Maria Thornbury for all her kindness and professionalism.

Without the help of all the people mentioned above this dissertation would not have been possible. Thank you all most sincerely.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

LIST OF ACRONYMS

AfL	Assessment for Learning
AfLAi	Assessment for Learning Audit Instrument
AifL	Assessment is for Learning
AoL	Assessment of Learning
ARG	Assessment Reform Group
ARM	Audit Reflection Meeting
ATMQ	Attitude to Mathematics Questionnaire
CAfLAi	Children's Assessment for Learning Audit Instrument
CCK	Common Content Knowledge (of mathematics)
CoP	Community of Practice
CPD	Continuing Professional Development
DCYA	Department of Children and Youth Affairs
DEIS	Delivering Equality of Opportunity in Schools
DES	Department of Education and Skills, Ireland
DPMT-R	Drumcondra Primary Mathematics Test-Revised
EAL	English as an Additional Language
EFA	Embedded Formative Assessment
FA	Formative Assessment
FB	Feedback
FG	Focus Group
GERM	Global Education Reform Movement
INTO	Irish National Teachers' Organisation
IRM	Intervention Reflection Meeting
ITE	Initial Teacher Education
KCS	Knowledge of Content and Students
KCT	Knowledge of Content and Teaching
KLT	Keeping Learning on Track
KMOFAP	King's Medway Oxfordshire Formative Assessment Project
KO	Knowledgeable Other
KQ	Knowledge Quartet
LHTL	Learning How to Learn
LISC	Learning Intentions and Success Criteria

LL	Learning Log
LS	Lesson Study
LSRM	Lesson Study Reflection Meeting
MKT	Mathematics Knowledge for Teaching
MCK	Mathematical Content Knowledge
MMR	Mixed Methods Research
NCCA	National Council for Curriculum and Assessment
NCTM	National Council of Teachers of Mathematics
n.d.	No Date
NQT	Newly Qualified Teacher
OECD	Organisation of Economic Co-operation and Development
PA	Peer-assessment
PCK	Pedagogical Content Knowledge
PD	Professional Development
PISA	Programme for International Student Assessment
PL	Professional Learning
PLC	Professional Learning Community
PSA	Peer- and Self-assessment
PSC	Primary School Curriculum
PSMC	Primary School Mathematics Curriculum
PST	Pre-service Teacher
QCD	Questioning and Classroom Discussion
RJ	Researcher's Journal
SA	Self-assessment
SCK	Specialised Content Knowledge
SD	Standard Deviation
SEN	Special Educational Needs
SIGMA-T	Standardised Irish Graded Mathematics Attainment Tests
SMK	Subject Matter Knowledge
SPSS	Statistical Package for Social Sciences
SRL	Self-regulated Learning
SSE	School Self Evaluation
TC	Teaching Council

TIMSS	Trends in International Mathematics and Science
TLC	Teacher Learning Community
TLL	Teacher’s Learning Log
TQ	Technique
UK	United Kingdom
USA	United States of America
WALT	We are Learning To
WILF	What I’m Looking For
WSE	Whole School Evaluation

CHAPTER ONE: OVERVIEW OF THE STUDY

Introduction

Assessment, as argued by Gardner (2012a), is a "hot topic" (p.103) across the entire education spectrum and rarely out of the limelight. Discourse frequently emphasises process and outcome issues such as workload in schools, the results of international and national tests, and the uses to which assessment information is put, be it for summative, formative, evaluative or diagnostic purposes (National Council for Curriculum and Assessment [NCCA], 2007; Gardner, 2012b). There is ongoing discussion regarding assessment's relationship to curriculum and pedagogy (e.g., Black, 2016; Wiliam, 2014) and debate regarding the alignment of assessment practices with learning theories, for example socio-cultural learning theory (James & Lewis, 2012). Notwithstanding, scholars (e.g., Baird, Hopfenbeck, Newton, Stobart & Steen-Utheim, 2014; Cumming, Maxwell & Wyatt-Smith, 2016) highlight two key developments in the literature on assessment this century: international testing and assessment for learning. Transnational assessment systems such as Trends in International Mathematics and Science Study (TIMSS) have facilitated cross-country comparisons of national education systems and put increased pressure on countries to raise standards of learning and achievement. Some countries have responded by steering schooling systems using accountability regimes and testing, with significant implications for teachers, students and schools (Walsh, 2016). The high-stakes accountability associated with what Sahlberg (2011) identifies as the global education reform movement (GERM)[1] contrasts

---

[1] Sahlberg (2011) posits that GERM is part of an unofficial educational agenda and an offspring of globalisation. He identifies some of the following as common features of education policies and reforms which he labels GERM: standardisation of education; focus on core subjects; use of corporate management models in education; test-based accountability; increased control of schools; search for low-risk way to reach learning goals.

with the low to moderate stakes accountability consistent with reforms in Finland and other countries such as Germany, France and Italy (Conway & Murphy, 2013).

My research focused on assessment for learning, the second key development highlighted by scholars in contemporary literature on assessment, which has become ubiquitous in educational systems worldwide, with a broad take-up by policy makers in many countries. For the purpose of this research, assessment for learning (AfL) is conceptualised using the following *second generation* definition generated by the Third International Conference on assessment for learning in New Zealand in 2009 which states:

> Assessment for learning is part of everyday practice by students, teachers and peers that seeks, reflects upon and responds to information from dialogue, demonstration and observation in ways that enhance ongoing learning. (Klenowski, p.264)

This definition, explored further in Chapter Two, clearly captures the key tenets of AfL, foregrounds classroom practices, highlights the notion of AfL as a bridge between teaching and learning (Wiliam, 2011b) and, as argued by Lysaght and O'Leary (2013), views teachers and students as the primary agents of educational change. Throughout this thesis, similar to other researchers (e.g., Warwick, Shaw & Johnson, 2015), the term AfL is used interchangeably with the term formative assessment (FA). While acknowledging that some scholars distinguish between both terms (e.g., Willis, 2011), like Gardner (2012b), I believe that the terms AfL and FA are analogous since they encompass the same assessment principles and practices used to support and enhance student learning.

This research also explored the use of lesson study, sometimes known as Japanese lesson study since it originated in Japan, as a school-based vehicle of continuing professional development (CPD) in AfL. As employed in this research, lesson study (LS) is characterised as a teacher-led, peer-to-peer, research-oriented, practice-based, sustained, systematic and collaborative model of practice development and continuing professional learning (e.g., Corcoran, 2008; Dudley, 2013); the principal purpose of which is to improve the quality of teaching and learning through a collaborative, reflective and recursive process (e.g., Cajkler, Wood, Norton & Pedder, 2014). Some scholars (e.g., Cajkler & Wood, 2016; Corcoran, 2008) view the group of teachers participating in LS as a community of practice (CoP) in that it accords with Wenger's (1998) description of mutual engagement towards a joint enterprise grounded in a sociocultural view of learning. Others (e.g., Chichibu & Kihara, 2013; Gutierez, 2015) consider the LS group as a professional learning community (PLC), which Hord (2008) defines as professionals coming together in a group or community to learn. Meanwhile, Lewis and Hurd (2011) use the term Teacher Learning Community (TLC) as well as the term PLC in their discussions regarding LS. For the purpose of this research, the perspective adopted is that each of these three terms can be used to describe the LS process and so, while acknowledging subtle differences, the terms are used interchangeably.

One final key term pertaining to this research needs clarification here and that is mathematics since it is the curricular area of choice for this study. Terms such as mathematics, numeracy and quantitative or mathematical literacy have different meanings in various contexts, resulting in "difficulties in the debate about critical aspects of mathematical education" (Turner, 2012, p.1). Often there is ambiguity between the way people commonly use these terms and their intended meaning. Some view mathematics as part of numeracy, or as part of mathematical or quantitative literacy in general (Turner, 2012), while others consider numeracy as more practically oriented and a part of mathematics (Dunphy, Dooley, & Shiel, 2014). According to the Irish National Teachers' Organisation (INTO) (2013), discourse regarding terminological nuances is ongoing and precise meanings continue to be debated. Of late, the Irish government seems to favour the term numeracy in different publications, stating that "numeracy is not limited to the ability to use numbers, to add, subtract, multiply and divide" but "encompasses the ability to use mathematical understanding and skills to solve problems and meet the demands of day-to-day living in complex social settings” (Department of Education and Skills [DES], 2011a, p.8). Throughout this thesis, therefore, while recognising definitional nuances and ongoing debate regarding terms such as mathematics, numeracy, and mathematical or quantitative literacy, the terms mathematics and numeracy are preferred and are used interchangeably.

Background to the Research Problem

This section provides the background to the research. Specifically, it explores salient factors that precipitated this particular study such as research evidence attesting to the potential of AfL, and indeed LS, to positively impact teaching and learning. It also outlines the problem this research sought to address by exploring the context in which the study was conceived, and highlights apparent weaknesses in Irish education therein.

Efficacy of Assessment for Learning

Assessment for learning, with its emphasis on learning as opposed to measurement has in recent years, according to Chappuis (2014), “garnered the lion’s share of assessment attention and established a pretty good name for itself” (p.21). It has variously been described as “an international phenomenon” (Hayward, Higgins, Livingston, Wyse & Spencer, 2014, p.465), “a teaching strategy of very high leverage” (Hargreaves, 2004, p.24) and “ a powerful catalyst for learning” (Earl & Timperley, 2014, p.325). Ever since the seminal review of FA by Black and Wiliam in 1998, evidence has been steadily accumulating which suggests that when AfL practices are effectively integrated into the minute-to-minute and day-by-day classroom practices of teachers, then substantial increases in student achievement are possible, even when outcomes are measured on externally-mandated standardised tests (e.g., Leahy & Wiliam, 2012; Moss & Brookhart, 2009). Time and time again, research studies have demonstrated that AfL, when used effectively, “is perhaps one of the most important interventions for promoting high-performance ever studied" (Organisation for Economic Co-operation and Development [OECD], 2013, p.2), for example Popham (2011) states:

> Recent reviews of more than 4,000 research investigations show clearly that when [formative assessment] is well implemented in the classroom, it can essentially double the speed of student learning…it is clear that the process works, it can produce whopping gains in students' achievement, and it is sufficiently robust so that different teachers can use it in diverse ways, yet still get great results with their students. (p.2)

Various studies have also linked AfL to increased student motivation and self-esteem (e.g., Clarke, 2008), enhanced self-regulated learning and metacognitive abilities (e.g., Andrade, 2013) and better student-teacher relationships (e.g., Clarke, 2014). Furthermore, some scholars (e.g., Wiliam, 2010a) argue that AfL is a defining feature of effective teaching and have linked effective use of FA to improvements in teacher quality (e.g. Wiliam, 2011b). This is an important finding since research (e.g., Moss & Brookhart, 2009; Thompson & Wiliam, 2008) also suggests a correlation between teacher quality and student progress. The vast and ever-expanding literature that attests to the benefits of AfL when implemented effectively will be investigated and analysed in full in the literature review in Chapter Two.

In recent years, some scholars have queried the size of quantitative learning gains attributed to AfL (e.g., Bennett, 2011) and the fact that there are fewer than 20 large-scale AfL studies to review (e.g. Baird et al., 2014). There has also been criticism of the definitional discord and terminological nuances regarding AfL (e.g., Dunn & Mulvernon, 2009; Kingston & Nash, 2011). Nevertheless, as argued by Willis (2011), the assumption that AfL results in learning gains is not disputed. This ensures that AfL is an area of national and international interest and helps explain why it is worthy of further investigation in this research. However, various scholars (e.g., Earl & Timperley, 2014) argue that there is a problem with the way some teachers implement AfL in their classrooms, if they implement it at all.

Problem Regarding Effective Implementation of AfL

While AfL theory and practices may be widely disseminated and incorporated into policy in various countries, some researchers (e.g., Brookhart, 2009; Shepard, 2011) concur that implementation of AfL by teachers is often superficial, reflecting the *letter* of AfL by focusing on surface techniques, as opposed to the *spirit* of AfL based on a deep understanding of the principles underlying AfL practices (e.g., Earl & Timperley, 2014). In short, as Keeley (2016) recently remarked, “FA is research rich, yet practice poor” (p.x). Scholars (e.g., Harrison, 2013; Wiliam, 2011b) acknowledge that implementing AfL effectively is complex and often challenging for many teachers since it involves radical change in classroom practices and in how teachers perceive their role, leading to the democratisation of learning and change in pupil-teacher relationships (e.g., Black, 2015; Lysaght & O’Leary, 2013). Consequently teachers need sustained support if they are to implement AfL in the *spirit* in which it is meant. Shepard (2013), however, posits we don’t need further research documenting the limitations of current assessment practices in classrooms but instead suggests the focus should shift to supporting teacher learning in this area. Regarding teacher learning in AfL and in keeping with current thinking in relation to teacher CPD, various researchers (e.g., Wiliam, 2011b; Wiliam & Leahy, 2015) suggest teachers should engage in sustained, collaborative, school-based learning, for example TLCs, if they are to successfully embed FA practices in their classrooms. Wiliam (2006) summarised this argument as follows:

> If you’re serious about helping teachers implement AfL in their own practice, you have to help them do that for themselves…the only way to do that at scale is through school-based Teacher Learning Communities…you do not need experts to come in and tell you what to do. What you need is for you, as groups of teachers, to hold yourselves accountable for making changes in your practice. Implementing AfL requires changing teachers’ habits. Teachers know most of what I’ve talked about today already so the problem is not a lack of knowledge; it’s a lack of understanding of what it means to do AfL in practice. (p.12)

Efficacy of Lesson Study

Following the decision to focus the research on AfL practices, the most appropriate model of peer-to-peer learning that offered the greatest potential as a vehicle of teacher CPD in AfL in a school-based intervention was deemed to be Japanese lesson study. An extensive literature supports and recommends the use of LS as a school-based model of professional development (e.g., Hogan, 2015) whereby teachers can improve, for example, their "pedagogic literacy"[2] (Cajkler & Wood, 2016), thereby inducing an enquiry-oriented, reflective and holistic understanding of teaching; and can develop *cognitive empathy*, enabling them to gain deeper insights into how their students learn, thus becoming more learner responsive in their teaching. LS also facilitates development of teacher practice, knowledge and professionalism (Cajkler & Wood, 2015), improves student learning (e.g., Sibbald, 2009), and expedites the spread of best practice throughout the school (Barber & Mourshed, 2007).

Researchers proffer various other reasons why LS should be the CPD model of choice, and these are discussed in detail in Chapter Two. Although in its infancy in the Irish context and lacking an empirical base here, LS was ideally suited to this research since it was practice-based, enacted by teachers in their own setting, and could be conducted in any academic or non-academic area (Doig & Groves, 2011; Lewis & Hurd, 2011). Indeed, LS offered the possibility of not only enhancing teachers' knowledge and skills regarding AfL, but also suggested it would act as a catalyst to enhance the overall quality of teaching and learning in mathematics through a collaborative, reflective and sustained approach. It is

---

[2] Pedagogic literacy is the complex of skills, knowledge, attitudes and values that enable teachers to use their reading of the classroom to reflect-in-action and to make learner-responsive decisions that support learning in all its complexity (cognitive, social and emotional) (Cajkler & Wood, 2016).

worth noting, too, that for many years in the Irish context, Corcoran (e.g., 2011b) has advocated that LS should be the CPD strategy of choice for schools, particularly in mathematics, and this also influenced my decision. Another factor in deciding to use LS in this study was the significant parallels that became apparent between AfL and LS following the review of literature; for example, both focus on learning and the learner, emphasise the need for teachers to anticipate and adapt to pupils' responses and are premised on sociocultural/social constructivist theory, thus suggesting the possibility of an AfL/LS nexus. This is discussed in depth later.

Scholars (e.g., Walsh, 2016) posit that what happens internationally impacts national education policies and so the next section investigates what was happening in the Irish context at the time this study was conceived, with Conway and Murphy (2013) suggesting that "a suite of policy measures" was introduced here in response to "perceived deficiencies highlighted by the results" from Programme for International Student Assessment [PISA] 2009 (p.2).

Irish Context

For over a decade, Irish schools have had to cater for an increasingly diverse student population (Travers, Balfe, Butler, Day, Dupont, McDaid, O'Donnell & Prunty, 2010) and are expected to meet the needs of the 21st century learner, for example by enhancing skills such as metacognition, self-regulated learning, collaborative learning and problem solving. Coupled with this, scholars in the Irish context (e.g., Dolan, 2016; Sugrue, 2011) highlight that the education system here has undergone unprecedented change, particularly since the Education Act (Government of Ireland [GoI], 1998), resulting in increased accountability and reporting measures. Conway and Murphy (2013) suggest that Whole School Evaluation (WSE) Reports, School Self-Evaluation (SSE) and mandatory reporting of standardised tests to the Department of Education and Skills (DES) and parents are just some examples of an increased emphasis on accountability in Irish education. Indeed, they argue there has been "a move towards GERM in terms of the scope, intent and intensity of accountability mechanisms" (p.29), in contrast to Looney (2016) who believes Ireland has been *resistant* to policies generally associated with GERM. However, Looney (2016), does agree with Conway and Murphy (2013) that the impact of PISA 2009 "presented a new departure in the interaction between international tests and education policy making in Ireland" (p.76). She highlights that the PISA 2009 results, combined with the acute impact of the global financial crisis, and the appointment of a new minister for education provided the contextual factors that acted as an ignition point for educational reform in Ireland at the beginning of the second decade of this century.

A number of reports, both national (e.g., DES, 2010b) and international (e.g., TIMSS, 2011), raised concerns about various issues or inadequacies in Irish education. In particular, contemporaneous reports such as *Assessment in the Primary School* (INTO, 2010), *Incidental Inspection Findings 2010* (DES, 2010a), the *National Assessments of Mathematics and English* (DES, 2010b), and *Whole School Evaluation Reports,* highlighted various shortcomings regarding teachers' use of AfL practices in the Irish context, indicating a lack of assessment literacy among teachers. While a copy of *Assessment in the Primary School: Guidelines for Schools (*NCCA, 2007) was given to all primary teachers, few teachers received the promised CPD in assessment due to cutbacks in education. Nevertheless, teachers' self-reports from *Assessment in the Primary School* (INTO, 2010) suggested they were willing to embrace assessment in their classrooms. However, the *Incidental Findings Report* (DES, 2010a) details that, in many of the English and mathematics lessons observed by the Inspectorate, formative assessment practices were not in evidence, and also reported "particularly serious problems" (p.6) regarding assessment in certain instances. Not only did the aforementioned reports draw attention to the need for improvement in teachers' assessment practices, some (e.g., INTO, 2010) also served to highlight the immediate need for sustained, high quality, CPD in assessment for all Irish teachers to enable them to implement AfL practices effectively in day-to-day teaching and learning. Furthermore, when one considers that despite ongoing guidance and intensive support for teachers in the UK for over a decade, renowned assessment expert Paul Black still argued in 2010 that AfL "isn't happening", coupled with the acknowledgement by many AfL scholars (e.g., Willis, 2011) of how complex adopting and implementing AfL practices effectively in classroom situations actually is, the need for immediate action in the provision of CPD is evident.

Regarding primary mathematics in the Irish context at the end of the first decade of the 21st century, national (DES, 2010b) and international (TIMSS, 2011) reports highlighted similar strengths and weaknesses, with Irish students performing strongly in number but demonstrating relative weaknesses in data, measures and problem solving. In general, high levels of student engagement with mathematics were reported at primary level (McCoy, Smyth & Banks, 2012) but traditional methods of instruction, such as whole-class teaching, still predominated, with the use of textbooks very much in evidence (DES, 2010a; DES, 2010b). In TIMSS 2011, at fourth class level, Ireland was ranked 17th of 63 participating countries with a mean score of 527, significantly above the TIMSS mathematics mean of 500 but significantly lower than the mean scores achieved by pupils in 13 other countries, including Northern Ireland and England (Eivers & Clerkin, 2013). While some (e.g., INTO, 2011) questioned whether the decline in mathematics standards was real, or attributable to the actual tests and their interpretation, the end result was that numeracy and literacy came under the spotlight in Irish education.

While the reports discussed above highlight weaknesses in teachers' assessment practices, and the teaching of mathematics, as argued by Harold Hislop (2013), chief inspector with the DES, it was Ireland's poor performance in PISA 2009 that precipitated publication by the Irish government of a strategy aimed at improving standards of literacy and numeracy in Ireland: *Literacy and Numeracy for Learning and Life; The National Strategy to Improve Literacy and Numeracy for Children and Young People 2011-2020* (DES, 2011a). This strategy is one of the most significant documents pertaining to education in the Irish context in recent years, and it is especially pertinent to this research since it has particular implications for numeracy, assessment and CPD. The *Literacy and Numeracy Strategy* increased the amount of time allocated to literacy and numeracy at all class levels, and set out *ambitious* improvement targets in English and mathematics as measured on standardised tests to be achieved by 2020. Compulsory standardised testing in English and mathematics changed from two to three points in the primary cycle (second, fourth and sixth classes), with mandatory annual reporting of aggregated results to the DES to facilitate collation of a national picture of achievement. Additionally, schools must use these results as part of "robust self-evaluation" (p.40) and to prepare three-year improvement plans for the promotion and improvement of numeracy and literacy. Results also have to be given to Boards of Management and parents.

Hislop (2013) recently described the *Literacy and Numeracy Strategy* (DES, 2011a) as "groundbreaking" since it is "perhaps the only statement of public policy on how evaluation and assessment arrangements are intended to work together in the Irish school system" (p.8). The strategy (DES, 2011a) makes clear that all teachers are expected to be assessment literate. It highlights the need "to use a continuum of well-considered assessment approaches to determine the next steps in learning and in planning approaches to teaching" (p.32) and "to combine good *assessment for learning* practice with appropriate *assessment of learning* approaches" (p.74). However, while it states, "AfL should be used to inform all teaching", it also emphasises "it is not used sufficiently widely in our schools and we need to enable teachers to improve this practice" (p.74). With regard to standardised tests, while acknowledging their limitations, the document states that primary teachers are not using the information from these tests to best effect and acknowledges there are considerable deficiencies in how results are reported to parents. The strategy suggests shortcomings in assessment practices are due to "lack of focus on assessment in teachers' continuing professional development" stating "we can improve the ways in which we collect and use assessment information" (p.76). In a similar vein, Hislop (2013) acknowledged that "Ireland's educational system has paid relatively less attention to the development of teachers' expertise in assessment generally and relatively more attention to the development of their teaching skills" (p. 10). He also stated that the lack of adequate professional development concerning assessment has hampered assessment practices, particularly regarding standardised tests. A key objective of the strategy (DES, 2011a) is to:

> Provide access to approved, high-quality professional development courses of at least twenty hours duration in literacy, numeracy and assessment (as discrete or integrated themes, provided incrementally or in block) every five years for primary teachers (as an element of the continuing professional development that teachers require to maintain their professional skills. (p.36)

Some of the key targets of the *Literacy and Numeracy Strategy* that are particularly relevant to this research are as follows:

- To improve the way we use assessment;
- To increase the percentage of students performing at the highest levels and decrease the percentage of students performing at the lowest levels in National Assessments of Mathematics and English Reading by at least five percentage points;
- To promote better attitudes to mathematics among young people;
- To enable students to understand, appreciate and enjoy mathematics.

It is worth noting that the *Literacy and Numeracy Strategy* is not without its critics. Conway and Murphy (2013) suggest it demonstrates a move towards the adoption of GERM since it encompasses "a systemic move towards attainment of results-type accountability for schools in relation to literacy and numeracy" (p.28), while Ó'Breacháin and O'Toole (2013) believe that one of the main drivers behind the strategy was political and proffer that the increased focus on literacy and numeracy reduces the role of the Arts in education and threatens "the holistic nature of the Irish curriculum" (p.404).

Research Problem and Rationale

The preceding section highlighted various problems that were present in Irish education when this study was conceptualised. At a time of increased international competition there was an apparent drop in Irish students' standards in literacy and numeracy. Additionally iterative weaknesses in teacher assessment practices and capacity were identified and the lack of high-quality CPD highlighted. Various salient reports were discussed, particularly the *Literacy and Numeracy Strategy* (2011a), which signified attempts by the DES to address these shortcomings. All of these factors suggested that research into assessment, CPD/LS, and mathematics would be timely and worthwhile. Moreover, the significant research base indicating that effective implementation of AfL, and LS, offered the potential to positively impact student learning, particularly in mathematics, similarly confirmed that these areas were worthy of further investigation. Indeed, scholars (e.g. Kingston & Nash, 2015; Wiliam, 2016; Willis, 2007) have called for more research into AfL in varying contexts. Reporting on their recent cross-national project, Warwick et al. (2015) discovered that despite the ubiquitous language of AfL, FA practices were open to different interpretations in different global contexts, suggesting national differences. They concluded that more detailed investigations by both researchers and teachers are needed in specific national contexts in order to explore the import of cultural understandings of assessment practices and reveal the nuances of national assessment practices in different contexts. Furthermore, in the Irish context, scholars (e.g., Lysaght, 2009) have also highlighted the need for further research into AfL. Regarding LS, scholars, nationally (e.g., Hogan, 2015) and internationally (Cajkler et al., 2014), have similarly highlighted the need for a wider empirical base.

Research Purpose, Questions, Delimitations and Design

The purpose of this intervention was twofold: to investigate the effects of using AfL practices on the teaching and learning of mathematics at fourth class level in one girls-only primary school in the Republic of Ireland during the academic year 2012-2013; and to explore the potential of peer-to-peer learning or teacher-led school-based CPD, in this case LS, to impact teachers' knowledge and skills using AfL principles, strategies and techniques, and their beliefs towards AfL as a form of assessment. Specifically, the study investigated the following three research questions, presented as hypotheses:

H1. A nine-month school-based intervention employing assessment for learning principles, strategies and techniques will improve the standardised mathematics results of participating students in comparison to a similar cohort not involved in the intervention.

H2. The use of AfL strategies and techniques, and the adoption of AfL principles, will enhance children's mathematical confidence, and improve their engagement with, and attitudes to, mathematics.

H3. Peer-to-peer professional learning is a feasible, worthwhile, efficient and effective model of CPD in AfL and will improve teachers' skills, knowledge and use of AfL, and their attitudes and beliefs towards AfL as a form of assessment.

The primary focus of this research was therefore on one aspect of assessment only, AfL; and the potential of LS as a vehicle of CPD in AfL. The study investigated the interplay between Japanese lesson study, CPD in assessment for learning and mathematics teaching and learning. While the study's findings and recommendations may have implications for policy, or applications for post-primary and tertiary levels, the focus here was at primary level.

This is a practitioner action research case study that operated within the pragmatic paradigm[3]. A convergent parallel, mixed methods design was adopted as the best strategy, whereby quantitative and qualitative strands were given equal priority, but treated independently, before being mixed during analysis and interpretation (Creswell, 2013; Creswell & Plano Clarke, 2011). The combined use of qualitative and quantitative data aided triangulation, enhanced the study's findings and enabled better understanding of the research problem. The research design is discussed in detail in Chapter Three.

Context of the Research Study

This research project took place in the school where I teach, Scoil na nAingeal (pseudonym), from September 2012 to June 2013. It is a vertical, urban, all-girls Primary School in Ireland with an enrolment of 438 students at the time of the intervention and an all-female staff. The school is not part of the Department of Education's action plan to address educational disadvantage known as Delivering Equality of Opportunity in Schools (DEIS), although the nearby boys' senior school attended by brothers of our students has DEIS status. Previous research in the school (Gurhy, 2008), indicated that school attendance for the preceding ten years was closest to the mean annual percentage of schools designated

---

[3] The pragmatic paradigm is ideally suited to mixed methods research (Feilzer, 2010), offers a useful middle position philosophically and facilitates the selection of methodological mixes that help researchers better answer their research questions (Johnson & Onwuegbuzie, 2004).

disadvantaged and the number of students who missed twenty days or over correlated most closely with statistics for students attending DEIS schools. Regarding the student body in Scoil na nAingeal, a recent whole school evaluation report (WSE) stated: "The pupil cohort is mixed, in terms of socio-economic status and ethnicity. A significant number of pupils speak English as a second language" (DES, 2012b, p.1). These factors indicate that the level of need within the school is significant and is probably closer to that of DEIS schools who have access to a range of supports and services not open to other schools including lower pupil-teacher ratios, additional funding, extra professional development supports and access to further literacy and numeracy support such as Maths Recovery and Ready Set Go Maths (http://www.education.ie/en/Schools-Colleges/Services/DEIS-Delivering-Equality-of-Opportunity-in-Schools-/DEIS-Supporting-Information/Supports-to-DEIS-Schools.html accessed August 21st, 2016). Fifty-one students in fourth class for the academic year 2012-2013 took part in the study, along with three teachers (two class teachers and one member of the school's Special Educational Needs [SEN] team). Further details regarding the research site and participants are discussed in depth in Chapter Three.

Significance

This study is important since, to date, little empirical research has been done into the effects of AfL practices on students' mathematics learning in the Irish context at primary level. Additionally, no comparable research in the Republic of Ireland has explored the use of LS to impact teachers' skills, knowledge, and use of AfL, and their attitudes and beliefs towards AfL as a form of assessment. Indeed, to my knowledge, no intervention here at primary level has utilised a combination of AfL, mathematics and LS, and so this study will contribute to the research field. Additionally, the research can supplement research done by academics regarding AfL and LS since it provides a practitioner researcher's perspective of the field, thus *inside-outside* (Cochran-Smith & Lytle, 1993). Furthermore, since the reports above have highlighted the need for improved assessment practices in Irish schools, the demand for high-quality teacher CPD in assessment and the potential for improvement in students' mathematics achievement, this research would appear timely and topical. Notwithstanding, perhaps the most significant contribution made by this research is that it provides a unique opportunity to listen to, and contemplate, the voice of young learners as they discuss their experiences of using AfL practices in their mathematics learning and offer their opinions regarding being part of three LS cycles.

Organising Framework of the Thesis

This thesis is organised into five chapters. The first chapter introduces the research topic, defines relevant terms, provides the background and rationale for the study, outlines the research questions as hypotheses and explains the significance of the research. The next chapter reviews selected literature pertaining to the two broad areas that inform the research: assessment for learning and Japanese lesson study. It also includes mathematics education literature relevant to both. Due to the voluminous extant literature, the review is by necessity selective. It includes a critical analysis of the pertinent AfL and LS literature and identifies key themes, theories, concepts and issues from which the research evolved. Chapter Three describes and justifies the choice of research methodology, details all phases of the research process, and clarifies how ethical considerations and issues of quality control were addressed. The findings pertaining to each of the three research hypotheses are presented sequentially in Chapter Four, along with analysis and discussion of each with reference to relevant literature. Finally, Chapter Five summarises, and draws conclusions from, the main findings of this intervention, addresses the limitations and implications of the study and makes recommendations for future research.

CHAPTER TWO: LITERATURE REVIEW

Introduction

This review is structured around two broad areas that informed the research: assessment for learning (AfL) and Japanese lesson study (LS) as a model of continuing professional development (CPD). Since this research investigates the impact of AfL and LS on mathematics teaching and learning the review also includes relevant mathematics education literature. The chapter identifies key themes, concepts, theories and issues pertaining to AfL and LS but, due to the voluminous literature in both areas, the review is, by necessity, selective.

The first section begins by briefly outlining the historical development of AfL, discussing definitional issues, examining the principles of AfL and presenting a research-based rationale for the study. Next, the theoretical underpinnings of AfL are investigated followed by an exploration of various AfL strategies and techniques. Then, some key AfL projects are looked at, especially *Embedded Formative Assessment* (EFA), which particularly influenced this research. This is followed by a review of AfL in the Irish context and a brief look at AfL and mathematics. Finally, a general critique of the research on AfL is presented, attempting to identify any lacunae in previous research, looking specifically at issues such as scalability, the learner's voice and AfL and affect. The section concludes with a synopsis of the reviewed AfL literature and a brief outline of section two.

Assessment for Learning (AfL)

Etymological and Historical Underpinnings

Before embarking on a review of key AfL literature, the etymology of the word assessment is worthy of consideration. It derives from the Latin verb *assidere* which means "to sit beside", evoking notions of the teacher sitting beside the learner, discussing his/her work, sharing his/her experiences and supporting his/her learning, rather than testing his/her performance. This image and connotation of assessment as "sitting beside" the learner, whether literally or figuratively, is integral to good teaching and learning, and resonates closely with the notion of AfL or formative assessment (FA) (O'Leary, 2006; Swaffield, 2011; Wiliam, 2007).

The phrase "formative assessment" can be traced back to Scriven's (1967) concepts of summative and formative evaluation in the context of programme evaluation (Bennett, 2011; Gardner, 2012b; Popham, 2008). Later, Bloom (1969) suggested applying the same distinctions when evaluating classroom tests, stating that "evaluation which is directly related to the teaching and learning process as it unfolds can have highly beneficial effects on the learning of students, the instructional process of teachers and the use of instructional materials by teachers and learners" (p.50). Wiliam (2011b) informs us that the term "formative" was rarely used over the following two decades, although some research investigating the integration of assessment with instruction did take place. The term AfL seems to have first been used by Harry Black (1986). However, as Gardner (2012b) argues, it wasn't until the late 1980s and early 1990s that the phrase AfL came into vogue, particularly after the effect sizes of learning gains through AfL were highlighted as being "between 0.4 and 0.7 … among the largest ever reported for sustained educational interventions" (Black, Harrison, Lee, Marshall & Wiliam, 2003, p.9).

Some assessment experts consider the terms AfL and FA interchangeable, while others emphasis nuances between them and favour one term over the other. Gardner (2012b) states that there is little between the terms FA and AfL, but considers that the phrase AFL is more accessible and:

> Less likely to be used to describe the summative use of multiple assessments. The words focus squarely on the essence of our pursuit: the promotion of assessment to support learning and this is neatly contra-distinct from assessment of learning. (pp.2-3)

Similarly, the NCCA (2007) seems to priortise the term AfL over FA stating that "the concept of AfL extends the potential of formative assessment" by emphasising "the child's active role in their own learning" (p.9). Like Gardner (2012b), the stance adopted in this study is that there is little to distinguish the terms since they both encompass the same assessment practices and principles used to support and enhance student learning. Therefore, both terms are considered interchangeable in this thesis.

Definitional Discourse

Understandings of AfL and FA have developed over time (Brookhart, 2011a) and even a perfunctory review of the assessment literature reveals myriad definitions of these terms. As yet, no single universally accepted or officially agreed definition of AfL or FA exists. While Filsecker and Kerres (2012) argue that there are no right and wrong definitions of AfL, others such as Shepard (2011) and Brookhart (2009) assert that ambiguity in earlier definitions has diverted energy away from actual research, leading to misunderstandings of the principles of AfL/FA and distortion of the practices, resulting in superficial implementation of AfL/FA, i.e. conforming to the *letter* of AfL rather than embodying it in *spirit* or, as characterised by Marshall and Drummond (2006), the simple application of technique as opposed to "high organisation based on ideas" (p.137) that promote student autonomy. Bennett (2011) and Dunn and Mulvenon (2009) concur. Exploring the definitional issue, Bennett (2011) argues that, as yet, the term “formative assessment (aka, ‘assessment for learning’)” does not represent a well-defined set of practices or artefacts. He highlights the importance of a definition stating, “if we can’t clearly define an innovation, we can’t meaningfully document its effectiveness” (p.8). Similarly, the difficulties caused by definitional nuances are highlighted and explored by Swaffield (2011) but, in reality, the various definitions have much in common (Filsecker & Kerres, 2012) and could be treated as “variations on a theme” (Stobart & Hopfenbeck, 2014).

Despite various definitions and conceptualisations, this study adopts the following *second generation* definition of AfL, generated by the Third International Conference on Assessment for Learning in New Zealand in 2009, since it succinctly captures the key ideas of AfL, while also emphasising the centrality of students in the learning process.

> AfL is part of everyday practice by students, teachers and peers that seeks, reflects upon and responds to information from dialogue, demonstration and observation in ways that enhance ongoing learning. (Klenowski, 2009, p.264)

Unpacking this definition, it is clear that AfL is a process that is integral to good teaching and learning and involves all parties (teachers, learners and peers). The emphasis is on classroom practices and it highlights the notion of AfL as a bridge between teaching and learning, as advocated by Wiliam (2013) and the tripartite relationship that exists between all three. Moreover, it alludes to learner agency and manages to convey the concept of *assidere*, thus returning assessment to its roots.

Principles and Rationale

AfL principles, when effectively implemented, can help guide the development of *good* assessment practices which optimise student performance while maintaining the quality of the learning experience (Gardner, Harlen, Hayward & Stobart, 2010; Harlen, 2014; Heritage, 2013). These principles can be used as standards or underpinning values and they provide "anchor points" and a common language to guide assessment or plan improvement or CPD (Gardner et al., 2010; Harlen & Johnson, 2014; Heritage, 2013). In 2002, the Assessment Reform Group (ARG) used ten principles to delineate their definition of AfL as "Assessment for Learning is the process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how best to get there" (pp.2-3). These principles were based on their theoretical discussions and insights about AfL and were published to guide classroom practice. Their document elaborates on each principle but the headlines state that AfL should:

1. Be part of effective planning of teaching and learning;
2. Focus on how students learn;
3. Be recognised as central to classroom practice;
4. Be regarded as a key professional skill for teachers;
5. Be sensitive and constructive because any assessment has an emotional impact;
6. Take account of the importance of learner motivation;
7. Promote commitment to learning goals and assessment criteria;
8. Guide learners about how to improve;
9. Develop learners' capacity for self-assessment so they can become reflective and self-managing;
10. Recognise the full range of achievements for all learners.
    (pp. 2-3).

These research-based principles are the ones adopted in this study and, according to Greenstein (2010), it is through the weaving together of these principles that high-quality FA arises.

There are various reasons why AfL should be the focus of this research. In the AfL literature, myriad experts mention the positive effects of using AfL on both students and teachers (e.g., Florez & Sammons, 2013; Hodgson & Pyle, 2010). In the past two decades or so, numerous reviews synthesising thousands of research studies have provided quantitative evidence of the positive impact AfL practices can have on students' learning and achievement (e.g., Black & Wiliam, 1998b; Crooks, 1988; Kluger & DeNisi, 1996; Natriello, 1987; Nyquist, 2003). Additionally, major research projects developing AfL practice and based on some of the reviews above, to be discussed later in the chapter, have found that when teachers truly embrace AfL practices not only is student learning enhanced but professional and organisational learning is too (Swaffield, 2011). Furthermore, FA-related data extracted by Clarke (2014) from Hattie's (2009) synthesis of over 900 meta-analyses also suggest AfL significantly impacts learning (Table 1).

Table 1

Data from Visible Learning (Clarke, 2014)

Influences on Learning	No. of Studies	Effect Size
Assessment literate students (students who know what they are learning, have success criteria, can self-assess, etc.	209	1.44
Providing formative evaluation	30	0.90
Lesson Study	402	0.88
Classroom Discussion	42	0.82
Feedback	1310	0.75
Teacher-student relationships	229	0.72
Meta-cognitive strategies	63	0.69

Note. Clarke devised this table using data from *Visible Learning* (Hattie, 2009).

Self-regulation is widely recognised as an important skill that students need to develop in order to meet the demands of 21st century learning and, in the past decade, new advances in thinking about AfL mean that current conceptualisations include contemporary theories about student learning and increasingly link self-regulated learning (SRL) and AfL practices (Andrade, 2010; Baas, Castelijns, Vermeulen, Martens & Segers, 2015; Black &

Wiliam, 2009; Brookhart, 2013; Clark, 2012; Heritage, 2013; Lysaght, 2015; Nicol & Macfarlane-Dick, 2006; Wiliam, 2014). These scholars believe that through engagement in effective AfL principles, strategies and techniques, students become increasingly autonomous in their learning and ultimately become equipped with a wide range of cognitive and metacognitive strategies to enable them to self-regulate their learning. SRL, a key characteristic of effective learning (Clark, 2014; Pintrinch, 2000; Zimmerman & Schunk, 2001), has been defined by Pintrinch (2000) as:

> An active, constructive process whereby learners set goals for their learning and then attempt to monitor, regulate, and control their cognition, motivation, and behavior, guided and constrained by their goals and the contextual factors in the environment. (p.45)

In general, researchers suggest that SRL includes goal setting, metacognition and the use of metacognitive strategies (Andrade, 2013; Earl, 2013; Vrugt & Oort, 2008; Zimmerman, 2000).

Andrade (2013) highlights that AfL emphasises many regulatory processes and goals similar to SRL i.e., the process of collecting, evaluating and using evidence of student learning to support, monitor and improve learning. In concert with this view, Wiliam (2014) observes how SRL overlaps considerably with FA, and quotes from earlier work by Black and Wiliam (2009) stating that FA is concerned with “the creation of, and capitalisation upon, ‘moments of contingency’ in instruction for the purpose of the regulation of learning processes” (p. 6). He also suggests that SRL is particularly linked with the “unpacking” of FA as five key strategies, proposed by Leahy, Lyon, Thompson and Wiliam (2005). By way of explanation, Wiliam (2014) points out that the first strategy (clarifying, sharing, and understanding learning intentions and success criteria) helps learners become clear about the goals they wish to pursue. Once learners have embraced a specific goal, then both SRL and the second AfL strategy (activating students as owners of their own learning) emphasises the means by which learners use a variety of strategies to pursue their goal (Wiliam, 2014). Regarding feedback, the third AfL strategy, this is considered very important in both FA and SRL with Clark (2011) contending that formative FB reinforces and actualises SRL strategies and promotes lifelong learning among students. Similarly, in a summary of the literature on FB and SRL, Nicol and Macfarlane-Dick (2006) broadly define good FB as “anything that might strengthen the student’s capacity to self-regulate their own performance” (p.205). However, they also provide a note of caution stating:

> If formative assessment is exclusively in the hands of teachers, then it is difficult to see how students can become empowered and develop the self-regulation skills needed to prepare them for learning outside university and throughout life. (p.200)

This highlights the need for students’ involvement in AfL and SRL and emphasises that the agency for learning now resides with the learner, with students being considered the *definitive* source of FA (Andrade, 2010; Brookhart, 2013).

Figure 1 is an adaptation by Andrade (2013) of an earlier model of SRL and FB by Nicol and Macfarlane-Dick (2006). Here, Andrade’s emphasis is on the central and active role that students occupy in all FB processes, particularly in monitoring and regulating their progress towards desired goals and the evaluation of the efficacy of various strategies to attain these goals. Processes that are internal to the students, such as domain knowledge and activating motivation, are depicted inside the shaded area (B-F). Andrade’s main modifications put increased emphasis on other-regulation via FB from teachers, peers and technology (H), the incorporation of the processes of interpreting FB (I) and the closure of the FB loop for teachers (J).

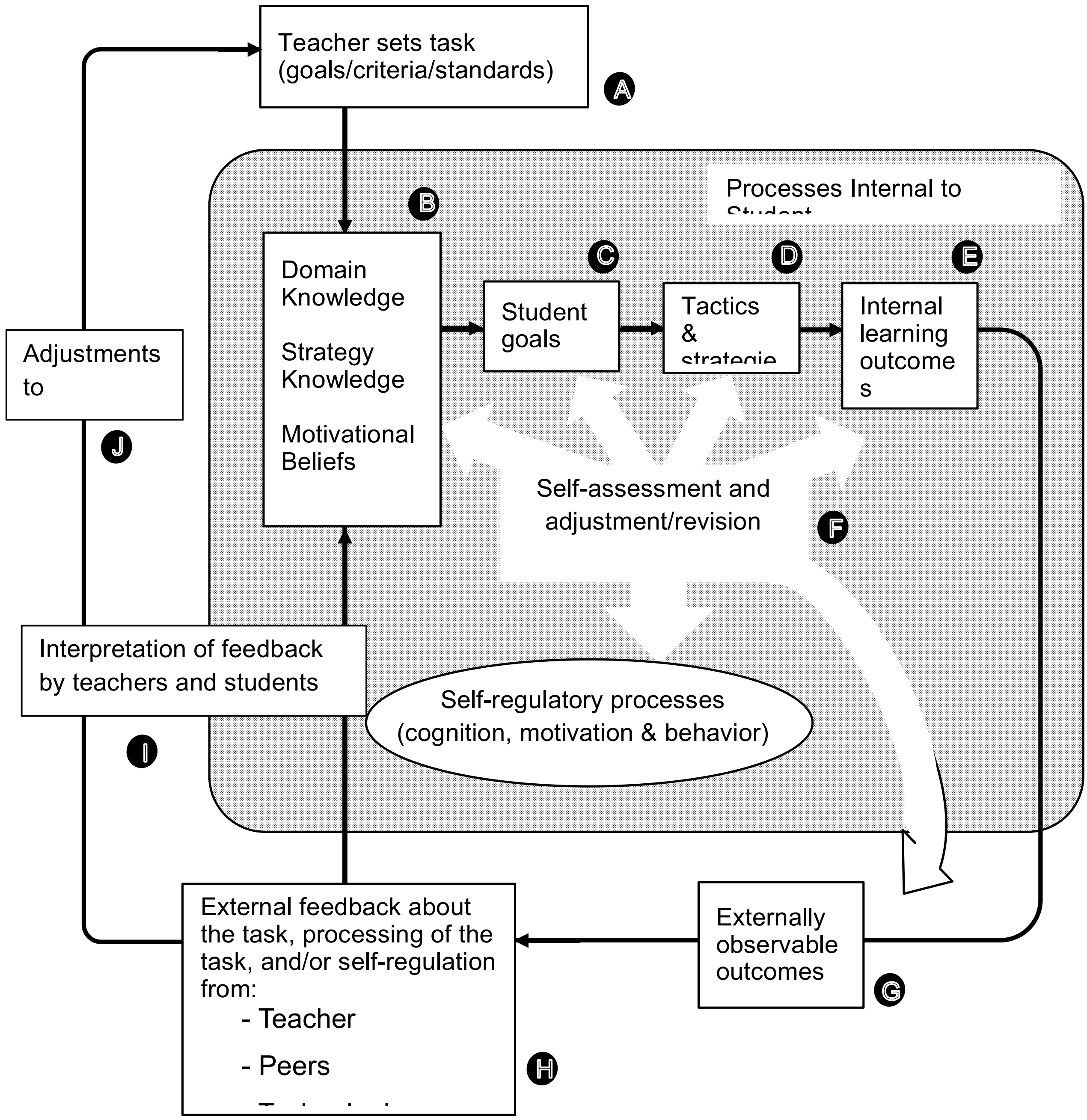


Figure 1. Model of assessment as the regulation of learning by oneself and others (Andrade, 2013)

Research shows that students can be useful sources of FB information via self-assessment (e.g., Andrade, Du & Wang, 2008). Andrade (2010) informs us that theories of self-regulation and self-assessment have much in common with recent scholarship on FA. Drawing on research from self-assessment and SRL, both of which involve students thinking about the quality of their own work and processes rather than relying on the teacher as the only source of evaluative judgment, Andrade makes the case that students are key producers and consumers of FA information. She views them as complementary processes that can result in significant improvements in academic achievement and autonomy. Figure 2 shows Andrade's conceptualisation of formative assessment and self-assessment as two aspects of self-regulation. She draws on Zimmerman's three phases of self-regulation: *Forethought*, when learners set goals and plan how to reach them; *Performance and Control*, which occurs during learning and consists of self-monitoring and the use of learning management strategies; and *Reflection*, when learners evaluate and reflect on their work. Andrade's model also incorporates Hattie and Timperley's (2007) conception of FB in learning as closing the gap between one's goal and current understanding by addressing the following three key questions: "Where am I going?" "How am I going?" and "Where to next?"

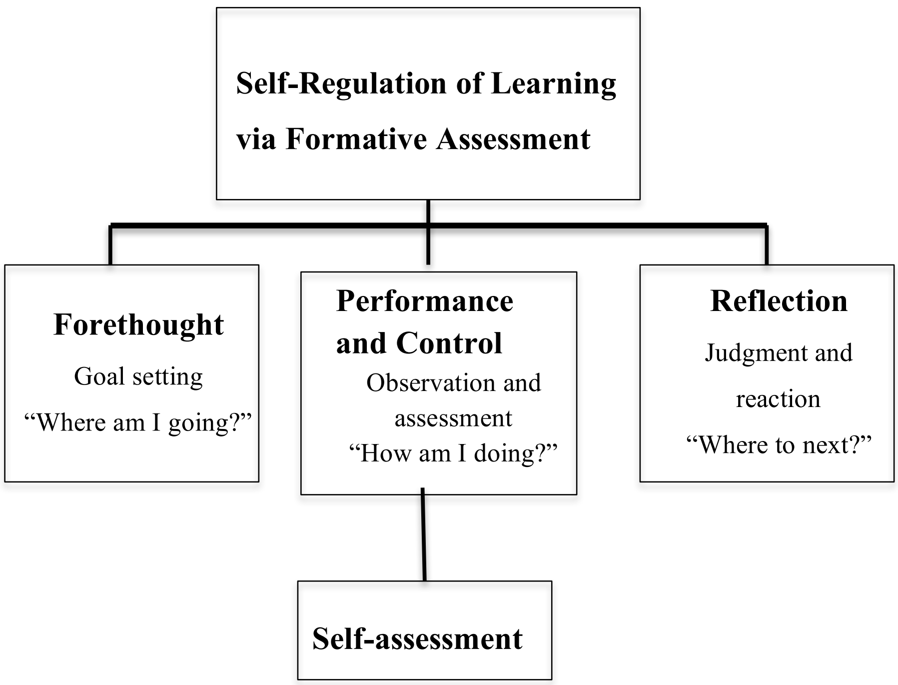


Figure 2. Self-regulated learning and formative assessment (Andrade, 2010, p.96)

Finally, it is noteworthy that a growing body of evidence acknowledges that SRL includes an important motivational component and thus addresses the roles of motivational processes such as goals, task interest, intrinsic motivation, self-efficacy beliefs, attribution, self-esteem, emotions, values, self-concept, outcome expectations and self-evaluations, in addition to strategies to regulate affect and motivation (Baas et al, 2015; Vandervelde, Keer & Rosseel, 2013; Schunk & Zimmerman, 2012; Zimmerman, 2008).

To conclude this section, it is important to note, as Brookhart (2013) does, that FB has the potential to help all children and that:

> When students are assisted into the self-regulation process with FA methods, such as deliberately teaching what students are to be learning and what constitutes quality in that learning, the provision of FB and opportunities to use it, even unsuccessful students learn. (p.44)

In other words, self-regulation is learnable (Andrade, 2010; Pintrinch, 1995; Zimmerman & Schunk, 2001), and can be developed through the use of AfL practices.

Although much has been written about AfL since the early 1990s, Gardner (2012b) still argues that "the extent of existing knowledge and understanding of such a complex process and set of techniques is still in its early stages" (p.284) and is a work in progress. This may, in part, explain the ongoing definitional and conceptual debate regarding AfL and provides one of the key reasons for the research at the heart of this thesis. Others (e.g., Kingston & Nash, 2015; Wiliam, 2016; Willis, 2007) have also called for further research on AfL in varying contexts. Reporting on their cross-national project, *Assessment for Learning in International Contexts* (ALIC), Warwick et al. (2015) discovered that despite the ubiquitous language of AfL, FA practices were open to different interpretations in different global contexts, suggesting national differences and leading the researchers to conclude that more detailed investigations, by both researchers and teachers, are needed in specific national contexts in order to explore the import of cultural understandings of assessment practices and reveal the nuances of national assessment practices in different contexts.

Theoretical Underpinnings

The import of exploring the theoretical underpinnings of AfL research were highlighted by Shepard (2005) when she observed that:

> Although teachers and teacher education students often have little patience with theory, big-picture understandings are especially important when we are trying to change our teaching practices. Theory helps us think about what to do when we can't rely on past experience. (p.10)

It has been argued that despite the extensive influence of AfL empirical research (e.g., Shepard, 2009; Taras, 2007), theory in this area remains rudimentary with much of the published research having implicit rather than explicit theoretical bases (Baird et al., 2014). Indeed, Black and Wiliam (2009), when trying to develop a theory of FA, acknowledged that their earliest work on FA (1998a), which many credit with promoting international interest in this area, did not evolve from any pre-defined theoretical base. For over a decade, various assessment scholars have endeavoured to provide a theoretical grounding for FA, to ensure the separate elements of effective practice work together and make sense (e.g., Black & Wiliam, 2006, 2009; James, 2006; Shepard 2005, 2009). Some have attempted to align FA practices with various theories of learning (Elwood, 2006; James, 2006: Shepard, 2000), although Baird et al. (2014) emphasise the difficulty of doing this since different assessment practices and learning theories have "co-existed chronologically" (p.5). Nevertheless, most would agree with Stobart and Hopfenbeck (2014) who state that current understandings of AfL "can be aligned with a range of learning theories and that each will be reflected in differing formulations and practices" (p.30).

AfL-related learning theories

Learning theories differ in various ways, such as their guiding principles, general assumptions and emphases, but all address key questions such as what learning is, how it happens, the role motivation plays and what influences student development and achievement (James & Lewis, 2012; Schunk, 2012). According to Watkins (2003), in general, learning theories tend to fall into three broad categories or philosophical frameworks, each taking a different stance on who is doing what to (or with) whom. In America, learning theories are broadly delineated as behaviorist, cognitivist and situated, while in Europe, the terms behaviourist, constructivist and sociocultural or activist are favoured (James & Lewis, 2012). Both sets of labels are basically analogous and have been simplified by Watkins (2003) as: learning is being taught, learning is individual sense-making, and learning is building knowledge as part of doing things with others. Table 2 provides a summary of how Stobart and Hopfenbeck (2014) have aligned learning theories with AfL and presents some of the implications and typical practices associated with each.

Table 2

Summary of Theories of, and Implications for, FA (Stobart & Hopfenbeck, 2014, p.40)

Theoretical Orientation	Associated with	Formative assessment emphasis	Typical practices
Behaviourist/ neo-behaviourist	Thorndike Gagne Bloom Popham (early) Test publishers	Atomised/step-by-step mastery; Regular testing for error detection and correction; Tests as formative assessments (product not process).	Learning objectives; Tests to establish what is not known; Test feedback to teacher to modify instructions; Feedback to student corrective.
Cognitive/ constructivist	Piaget Bruner (early) Simon Chomsky Bransford Pellegrino Ramaprasad Sadler Roos and Hamilton	Need for learners to "make sense" of information and developmental schemas; Importance of learner understanding; Learner objectives and success criteria ('standard'); Feedback as dynamic process (cybernetics).	Negotiated learning; Intentions and success criteria; Feedback information for learner to close gap; Self-regulated and self-monitoring learning.
Social constructivist	Crooks Shepard Cobb Sfard ARG Black and Wiliam (1998a)	Importance of school and classroom ethos; Dialogue and negotiated learning; Self- and peer-assessment; Motivation through engagement.	Classroom expectations; Encouraging learner engagement; Active learning – dialogue, group work, self- and peer assessment.
Sociocultural	Vygotsky Lave and Wenger Torrance and Pryor Pryor and Crossouard Black and Wiliam (2006) Ecclestone et al. (2012) Allal Perrenoud	Learner identity and changed teacher role and identity; Negotiating understandings of task and quality criteria; Apprenticeship model of learning; Social context central to learning – classroom ethos (regulation).	Renegotiated learner identities; Collaborative classrooms; Learning through active social processes and interactions; Changed classroom 'contract' around learning.

Some scholars believe that assessment and learning are frequently “out of sync” and suggest the need to pursue a more inclusive and complete theory of learning to inform the practice of teaching and assessment (James, 2006, 2008; James & Lewis, 2012; McCormick & Murphy, 2008). Shepard (2000, 2005, 2009) has argued for social constructivist or sociocultural theories of learning as a framework for clarifying and making coherent insights from the AfL literature, which would help explain how FA works. Others (Hayward, 2012; Willis, 2011) posit similar arguments. Exploring the links between assessment and learning thus far, James and Lewis (2012) outline three generations of assessment they believe are underpinned by behaviourist, cognitivist/constructivist and sociocultural/‘situated’ views of learning respectively. However, they find little evidence that third generation assessment exists in schools. Baird et al., (2014) apparently concur with James and Lewis (2012) and contend that although the attempts to draw sociocultural theory into assessment practices is encouraging, it is unclear that this has been successful and the state of play is, they believe, that assessment is primarily based upon cognitive constructivist approaches, possibly because it can take decades for the implications of theoretical advances to work into practice.

The debate regarding the alignment of sociocultural theories of learning with assessment continues and is still emerging (Cowie & Moreland, 2015; Elwood & Murphy, 2015; Hickey, 2015). Notwithstanding, Shepard (2000) and James and Lewis (2012) agree that possibilities exist for the emergence of a more complete theory of learning and assessment, possibly formed from a synthesis of key elements from previous theories. It is, however, worthwhile remembering Edwards’ (2012) argument that “while assessment must be driven by theory it is inherently practical” (p.208) and so the next section explores how AfL practices might be actualised in classroom settings.

One "Big Idea", Five Strategies and Multiple Techniques

Seeking to understand what AfL might look like on the ground, Thompson and Wiliam's (2008) formulated their "Big Idea" which envisions "students and teachers using evidence of learning to adapt teaching and learning to meet immediate learning needs minute to minute and day by day" (p.6). In other words, evidence about learning is used to adjust instruction ensuring teaching is adaptive to students' learning needs (Wiliam, 2011b). Unpacking this idea, and attempting to provide an improved theoretical basis for FA, Thompson and Wiliam utilised the following three questions from Ramaprasad (1983):

- Where the learner is going?
- Where the learner is right now?
- How to get there?

Explicating the implications of the processes from these questions and crossing them with the people involved in teaching and learning (teacher, peer and learner), Thompson and Wiliam (2008) were then able to conceptualise and unpack the "big idea" of AfL into five key strategies (Figure 3). These elements:

> Can be integrated within a more general theoretical framework of the regulation of learning processes … within such a framework, the actions of the teacher, the learners, and the context of the classroom are all evaluated with respect to the extent to which they contribute to guiding the learning towards the intended goal. (Wiliam, 2005, pp.31-32)

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

Figure 3. Aspects of formative assessment (Wiliam, 2014)

The five key strategies of FA are:

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

Wylie and Lyon (2015) inform us that these five key strategies support teachers' understanding of AfL and target the identification and use of evidence in different ways. They are intended to provide contingent information that teachers and students can use to progress student learning (Wiliam, 2011b; Wylie & Lyon, 2015). Warwick et al. (2015) view these five strategies as the "underlying pedagogic foundations of AfL" and believe that "a trajectory towards self-regulated learning through the use of the strategies, with appropriate scaffolding related to the contingent position of the learner, is both implicit and explicit" (p.42). They contend the strategies have "an underlying connection to instructional practices designed to foster metacognitive awareness in students" (p.42), which they argue helps develop important characteristics such as persistence, effort and critical judgment. It should be noted that researchers differ in the numbers of strategies they suggest are involved in AfL. Some scholars proffer that there are six strategies (Moss & Brookhart, 2009), others seven (Clarke, 2008; Stiggins, Arter, Chappuis & Chappuis, 2004) and some four (Lysaght & O'Leary, 2013). Notwithstanding, all seem to address the same conceptual territory and promote the same message, albeit in different ways (Wiliam, 2016). Thompson and Wiliam's (2008) "Big Idea" and five key AfL strategies formed the backbone of their 'Tight but Loose' framework and provided a solid base to ensure fidelity yet facilitate flexibility as teachers implemented AfL in local contexts.

Each AfL strategy can be instantiated through myriad techniques. Thompson and Wiliam (2008) view techniques as specific, concrete, ways that teachers can choose to implement the various AfL strategies. Wiliam (2010) distinguishes between strategies and techniques, considering that strategies define the territory of FA whereas teachers have responsibility for the choice of techniques, thereby facilitating customisation and implying democratisation and contingency in learning and teaching. Furthermore, these techniques are low-cost, low-tech and feasible for each teacher to implement (Leahy et al., 2005; Thompson & Wiliam, 2008). Their implementation involves only minor changes in teaching practice but can leverage significant improvements in student learning (Leahy et al., 2005; Wiliam, 2011b). A wide and ever-expanding range of techniques is available (Appendix A). As argued, with careful planning and the thoughtful use of these techniques, teachers can make the classroom a much more engaging place for their students, one in which teachers can make quick and effective instructional adjustments to meet the learning needs of all their students (Wiliam, 2011b). The five AfL strategies and sample techniques will now be explored in turn.

Clarifying and sharing learning intentions and criteria for success

Research evidence shows that, for AfL to be effective, students need to know what they are learning and why, where they are going and what counts as good quality work (Brookhart, 2009; Wiliam, 2011b; Wylie & Lyon, 2015). This first strategy, sharing clear learning intentions and success criteria (LISC) with students, helps them monitor their own progress and become more accountable and autonomous in their learning (Wiliam, 2011b; Wylie & Lyon, 2015). Wiliam (2011b) argues that there is no simple formula for communicating LISC to students and suggests it is up to teachers to exercise their professional judgement (Wiliam, 2011b). However, recent findings by Lysaght and O'Leary (2013) in the Irish context show that this strategy is at best *emerging* and have highlighted that teachers frequently find sharing LISC challenging. This suggests it may take time for teachers to acquire the necessary expertise in formulating appropriate LISC. Notwithstanding, researchers (e.g., Clarke, 2014; Wiliam 2011b; Wylie & Lyon, 2015) seem to agree that learning intentions should focus explicitly on the learning rather than the activity, with the timing of when they are shared depending on the lesson being taught. These scholars also concur it can be useful to develop or co-construct learning intentions and success criteria with students. Nonetheless, Wiliam (2011b) proffers that this should not be considered a democratic process since, in his opinion, the teacher knows more about the subject being taught and therefore retains overall responsibility. Regarding sharing LISC, teachers frequently use the acronymns WALT (We are learning to) and WILF (What I'm looking for). These can be useful when working with children, although Wiliam (2011b) cautions that students need to develop the habits of mind that define FA and so should come to terms with the *official* terminology of that process.

Engineering effective classroom discussions, questions and learning tasks that elicit evidence of learning

Various scholars (e.g., Brookhart, 2009; Clarke, 2014; Wylie & Lyon, 2015) agree that, in order to achieve the learning gains promised by AfL and plan better instruction, teachers need to develop and utilise effective questioning and engage in classroom dialogue that solicits evidence from each of their students that establishes the current state of their thinking and learning/understanding, including their prior knowledge and misconceptions. Indeed, Shepard alludes to how she believes this strategy (questioning and classroom discussion [QCD]) and the previous one work best when she states:

> Formative assessement practices enhance learning when students are positioned as thoughtful contributors to classroom discourse and have a sense of ownership in criteria used by a community of practice. (Shepard, 2013, p.xx)

Notwithstanding, it is worth noting Lysaght and O'Leary's (2013) premise regarding findings from their recent research, specifically regarding this strategy, that suggest a tension between pupil-led approaches to assessment and the more traditional, teacher-led approaches; and they also highlight their fear that getting teachers to implement AfL in ways that democratise learning and change the pupil-teacher relationship may prove challenging.

Providing feedback that moves learners forward

Hattie (2009) defines feedback (FB) as "information provided by an agent (e.g., teacher, peer, book, parent, or one's own experience) about one's performance or understanding" (p.174) and, after many years of research, places FB among the top ten influences on student achievement. Few scholars would argue with the assertion that providing good quality, effective, FB is central to the success of the AfL process, although most would agree that providing effective feedback is challenging (e.g., Hattie, 2012; Wiliam, 2011b). To help with this task, Shute (2008) provides nine guidelines for using FB to enhance learning, which include suggestions that FB should be simple, clear, specific, unbiased, objective, provided in manageable units and elaborate enough to describe the *what*, *how* and *why*. Regarding FB, and reminiscent of earlier work by Sadler (1989; 1998), Wylie and Lyon (2015) state that for FB to be formative "it must identify gaps between the desired learning goal and the student's present status, provide actionable suggestions for how to close the gap and students must then act upon the provided suggestions" (p.142). This definition implies two things: first that FB should provide details that help students improve their work and, second, it presupposes the requirement of student action or revision. Other AfL researchers concur with this (Black & Wiliam, 1998a; Brookhart, 2009; Shute, 2008), although most acknowledge that not all FB is useful (Brookhart, 2009; Wylie & Lyon, 2015).

Various studies have investigated the impact of FB on student learning (e.g., Butler, 1988; Kluger & De Nisi, 1996). Scholars inform us that FB should be focused and related to the LISC, produce a cognitive rather than an emotional response and involve more work for the recipient than the donor (Brookhart, 2009; Wiliam, 2011b). Additionally, research has shown that FB with comments and without grades encourages a "growth mindset" (Butler, 1987, 1988; Dweck, 2000), the belief one can always improve by one's own effort, which helps students cope better with challenging tasks and become more effective learners (Black, 2015). Specifically, investigating FB in mathematics with students from 5th to 12th grades, Deevers (2006) found when students received constructive FB they focused on learning rather than performance. Shute (2008) argues that written or computer-based FB is preferable, but most scholars believe it does not matter what form the FB takes, whether written or verbal, and agree that the impact of FB is not always positive but can sometimes be ineffective or, worse still, counterproductive (Clarke, 2014; Kluger & DeNisi, 1996; Wiliam, 2011b). Students can receive different FB from various sources and their responses to FB can vary. Using findings from one of the most systematic research studies on the effects of various types of feedback (Kluger & DeNisi, 1996), Wiliam (2011b) summarised possible student responses to FB with only the two italicised responses likely to improve performance (Table 3).

Table 3

Possible Responses to Feedback (Wiliam, 2011b)

Response type	Feedback indicates performance exceeds goal	Feedback indicates performance falls short of goal
Change behaviour	Exert less effort	*Increase effort*
Change goal	*Increase aspiration*	Reduce aspiration
Abandon goal	Decide goal is too easy	Decide goal is too hard
Reject feedback	Ignore feedback	Ignore feedback

Building on the research of others (e.g., Kluger & DeNisi, 1996; Sadler 1989; 1998), Hattie's work has added much to our understanding of FB and his research is worthy of further consideration here. He (2012) believes feedback typically comes second, after instruction. He stresses that to be effective, the key factor is that students must act upon the FB that is received. Similar to Shute (2008), he argues that errors should be welcomed since they are the key levers for enhancing teaching and learning. Hattie (2009) emphasises that "it is the feedback to the teachers about what students can and cannot do that is more powerful than feedback to the student and it necessitates a different way of interacting and respecting students" (p.4). Thus, similar to what Lysaght and O'Leary (2013) highlighted regarding QCD, Hattie believes providing effective FB also necessitates change in pupil-teacher relationships and leads to increased democratisation of classroom practice. Over the past decade or so, Hattie (2007; 2009; 2012) has been working on a model of FB to enhance learning (Figure 4) and explains that the purpose of FB is to reduce the "gap" between where the student "is" and where they are meant to be, i.e. between current achievement and the success criteria (Hattie, 2012).

	Levels	Major questions	Three feedback questions
1	Task	How well has the task been completed; is it correct or incorrect?	Where am I going? What are my goals?
2	Process	What are the strategies needed to perform the task; are there alternative strategies that can be used?	How am I going? What progress is being made towards the goal?
3	Self-regulation	What is the conditional knowledge and understanding needed to know what you are doing? Self-monitoring, directing the processes and tasks.	Where to next? What activities need to be undertaken next to make better progress?
4	Self	Personal evaluation and affect about the learning.	

Figure 4. Feedback levels and questions (Hattie, 2012)

Hattie (2012) posits that FB addresses three major FB questions ("Where am I going?", "How am I going?" and "Where to next?"). Respectively, the three questions relate to the learning intentions/goals/success criteria, the progress FB and the progression towards new goals. These FB questions work at four levels (task, process, self-regulation and self) that correspond to phases of learning (novice, proficient, competent). FB has different effects across these four levels but, ideally, the first three levels are integrated and form a progression from the task, to the process, to self-regulation and should be distinguished from the fourth (self) level, which can be subsumed under the notion of praise and can detract from the other three levels and dilute the power of feedback. In sum, Hattie (2012) states, "the aim is to provide feedback that is 'just in time', 'just for me', 'just for where I am in my learning process' and 'just what I need to help me move forward'" (p.122).

To conclude, most scholars believe that the primary purpose of FB should be to increase the extent to which pupils are owners of their own learning, leading ultimately to self-regulation (e.g., Shute, 2008; Wiliam, 2011b). Indeed, Nicol and Macfarlane-Dick (2006) specifically propose seven principles of good FB practice in relation to the development of self-regulation (Appendix B). Nevertheless, it is worth noting that Hattie and Timperely (2007) caution that, while FB combined with effective instruction can greatly enhance learning, it has its limitations and suggest "with inefficient learners, it is better for a teacher to provide elaborations through instruction than to provide FB on poorly understood concepts" (p.104). In short, FB "is a consequence of performance" (Hattie & Timperley, 2007, p.81) and can only build on something; it is of little value without initial learning.

Activating students as the owners of their own learning

Within AfL, a strong emphasis is placed on students' active involvement in the assessment process, particularly through the use of peer- and self-assessment (PSA), both of which can engage students and promote learning (Harris, Brown & Harnett, 2015; Sebba, Crick, Yu, Lawson, Harlen & Durant, 2008; Willey & Gardner, 2010). However, certain conditions, such as the commitment of teachers to allow learners greater control of their own learning and the development of a more interdependent relationship, where teachers and students discuss learning, affect the success of PSA (Hayward, 2013; Sebba et al., 2008). As stated earlier, self-assessment (SA) has been viewed as a key aspect of self-regulation (Andrade, 2010). In SA, which Andrade (2010) advises should be near-term, students must regulate their own learning by assessing their progress towards the LISC, reflecting on that progress and engaging in metacognition (Wiliam, 2011b; Wylie & Lyon, 2015). Research has shown that when students take ownership of their own learning, then student learning improves since students are more intrinsically motivated and engaged (Fontana & Fernandez, 1994; McMillan & Hearne, 2008; White & Frederiksen, 1998). Student self-assessment is a critical skill that not only enhances achievement but also increases student motivation (McMillan & Hearn, 2008). While the student obviously plays a central role in this strategy, the teacher's role is to scaffold the self-assessment process (where necessary) by informing students of the intended learning, providing time for students to internalise the success criteria (SC), and providing guidance to support accurate and informative self-reflection and assessment (Wylie & Lyon, 2015). Ultimately, self-assessment must be used to highlight the next steps in the students' learning, but it takes time and perseverance to achieve the aforementioned benefits (Wiliam, 2011b; Wylie & Lyon, 2015).

Activating students as instructional resources for one another

This final strategy focuses on the students' role in the AfL process. Peer-assessment (PA) had been defined by Lui and Andrade (2014) as "a process during which students consider the quality of a peer's work or performance, judge the extent to which it reflects targeted goals or criteria, and make suggestions for revision" (p.1). Research indicates that students can engage successfully in PA and play a role in improving the learning of their peers (Black, 2015; Topping, 2009, 2010). However, researchers (Black, 2015; Lui & Andrade, 2014; Wiliam, 2011b; Wylie & Lyon, 2015) emphasise that the crucial aspect of successful PA is that students must receive explicit guidance and act under specific structures when engaging in this strategy. It has been suggested that PA promotes reflection and can help develop transferable skills for life, e.g., social skills (Lui & Andrade, 2014; Topping, 2010). It can lead to increased student engagement and ownership of the teaching-learning-assessment process as students experience a greater sense of accountability and responsibility, be more immediate than other forms of assessment, and save teachers' time (Topping, 2009; 2010). Topping (2010) considers PA an effective but underutilised type of FA. He argues that PA has a powerful affective component since "a trusting relationship with a peer who holds no position of authority might facilitate self-disclosure of ignorance and misconception, enabling subsequent diagnosis and correction that could not occur otherwise (p.65). Topping (2009; 2010) also posits that students act differently to FB from peers and adults, the former often being viewed as giving richer FB open to negotiation. At first, assessors and assessees may experience initial anxiety about PA, particularly if it involves appearing unpleasant to their friends; but the process can be beneficial to both and so should be introduced over time, with scaffolding by the teacher, if necessary, in the initial stages (Lui & Andrade, 2014; Topping, 2009, 2010).

The previous sections have defined the domain of FA and identified critical aspects of each strategy. All these aspects are crucial to AfL, but it is also necessary to recognise and comprehend the interdependencies between the strategies as, failing this, the quality of AfL implementation would be weakened (Wylie & Lyon, 2015). The next section explores key AfL research that influenced this study.

Key Projects

Significant research into AfL with teachers and students has been undertaken since the publication of Black and Wiliam's (1998a) seminal review, particularly in the United Kingdom (UK) and the United States of America (USA). Much has been learned from large-scale AfL projects such as the *King's Medway Oxfordshire Formative Assessment Project* (KMOFAP; Black, Harrison, Lee Marshall & Wiliam, 2003; 2004), the *Learning How to Learn Project* (LHTL; James, Black, McCormick, Pedder & Wiliam, 2006) and the *Assessment is for Learning Project* (AifL; Hayward, Priestley & Young, 2004) in the UK, and the *Keeping Learning on Track Programme* (KLT; Wiliam & Thompson, 2007) in the USA. A more recent large-scale AfL study, *Embedded Formative Assessment* (EFA, Wiliam, 2011b; Wiliam & Leahy, 2015), also influenced this research. It is worth noting that, bar AifL, Wiliam has been involved in all the aforementioned projects, and from his experience he (2016) has identified six leading indicators that he believes provides evidence of progress with respect to developing FA:

1. Teachers are given time to meet, and do so;
2. Teachers increasingly act as "critical friends" to others;
3. The prevalence of FA practices is increasing;
4. Students are more engaged in classrooms;
5. Teachers modify the techniques in appropriate ways, indicating an understanding of the underlying theory;
6. There is a shift in the ownership of the reform (pp. 219-227).

All these major research projects have found that when teachers truly embrace AfL practices, not only is student learning enhanced but professional and organisational learning is too (Swaffield, 2011). These key AfL projects have made a significant contribution to this field of research and so Table 4 offers a distillation of each project, providing a synopsis of the foci, findings and areas needing further research. Subsequently, the state of AfL in the Irish context will be explored.

Table 4

Key AfL Research Projects

Studies	Key Foci	Findings	Lacunae/Future Research
***The King's Medway Oxfordshire Formative Assessment Project* (KMOFAP).** **England.** **(2000-2002)**	Attempted to bridge research-practice divide. Focus on content and process of CPD. Exploration of four key areas: questioning, feedback through marking, peer- and self-assessment, formative use of summative tests.	Effect size of intervention was 0.32 SD. Collection of FA practices emerged which improved pupil performance and enriched teaching and learning. Demonstrated teachers and students openness to change.	Involved only a small number of teachers, who were supported by researchers throughout. Lacked sufficient detail to enable other teachers to put FA into practice in their classrooms.
***Learning How to Learn* (LHTL).** **England.** **(2001-2005)**	Large-scale research project. Combined research on AfL and CPD to explore kinds of professional learning most suited to development of AfL practices in classrooms and promotion of optimal learning conditions for students.	Only one fifth of lessons embodied *spirit* of AfL and promoted learner autonomy. Teachers' beliefs about learning thought to impact way they applied AfL in classroom. Recognition of need for input on learning theory and for need to modify four original headings under which AfL practice had been conceived (questioning, feedback, sharing criteria and self-assessment).	Principles of AfL insufficient to leverage teacher change. Teachers encouraged to engage in collaborative inquiry into own practice and to use relevant research. Future research to be placed on conditions and practices that hold potential to promote autonomy in learning.
***Assessment is for Learning* (AiFL).** **Scotland.** **(2002-2008)**	Attempt to integrate research, policy and practice into single framework. Emphasis on role of assessment in supporting each pupil's learning.	Patchy implementation at times but significant change in schools which embraced AifL Programme. Impact greatest in primary schools and those where principals involved. Reported deeper learning by students, especially with SEN.	Challenges remain regarding: • Scalability • Sustainability • Replicability • Generalisability
***Keeping Learning on Track* (KLT).** **USA.** **(2005 to date)**	Key principle = "One Big Idea". Teacher CPD programme focused on minute-to-minute, day-by-day instruction. Key components: content, process, empirical /theoretical. Addressed scalability issue.	Successful adoption of AfL strategies. Sharing experiences with other teachers was the most helpful aspect of KLT. Influenced teacher practices and student engagement but little evidence KLT impacted student achievement over course of study (NWEA, '15).	Evaluation of KLT is ongoing.
Embedded Formative Assessment (EFA). Singapore, UK, Australia, US. (2011 to date)	One "Big Idea" (Five strategies, myriad techniques). TLC as vehicle of change ('Tight but Loose'). Two-year programme and resource pack. Whole school. Addresses scalability.	Teachers found EFA helpful. They believed EFA impacted students' engagement and learning. When implemented as designed EFA programme *effective* and *manageable* (Wiliam, 2016). Time identified as biggest obstacle to implementation.	Evaluation of EFA is ongoing.

AfL in the Irish Context (Primary Level)

Section 22 of the Education Act (GoI, 1998) imposes a statutory obligation on all schools and teachers to assess their students and report results to parents regularly. The Irish primary school curriculum (PSC) emphasises the centrality of assessment to successful teaching and learning, stating that "an awareness of available assessment tools and the use of appropriate assessment procedures and practices are essential in providing children with an effective learning experience" (p.19). Sugrue (2004), however, criticises the cursory manner in which assessment was dealt with in the PSC and the lacuna regarding AfL that was evident in official curriculum documentation. Two phases of review of the primary curriculum (NCCA, 2005; 2008) revealed that teachers were unclear of the purpose, function and role of assessment in supporting teaching and learning, highlighting the need for teachers to receive advice and support regarding assessment practices. In this context, the DES published Circular 0138/2006, which detailed plans for ongoing CPD in assessment from 2007/8 onwards, stating that AfL would be "an important national priority in that context" (p.2). However, only small numbers of teachers received CPD, in standardised testing only, and no further DES-initiated CPD in assessment has materialised. In 2007, the NCCA published *Assessment in the Primary School Curriculum: Guidelines for Schools*, which built on the functions of assessment described in the PSC and focused on the uses and differences of two principal approaches to assessment:

- Assessment for learning (AfL) where "the teacher uses evidence on an ongoing basis to inform teaching and learning" and;
- Assessment of learning (AoL) where "the teacher periodically records children's progress and achievement for the purpose of reporting to parents, teachers and other relevant persons" (p.8).

The guidelines focused on assessment at both the classroom and school level and presented a continuum of assessment methods with multiple exemplars of good assessment practice. Since very few teachers have received CPD regarding these guidelines, the INTO (2010) argue they may remain underused and more recently contend they "contributed to a sense of overload among teachers" (INTO, 2015, p.7). Insights into recent assessment practices in Irish primary schools garnered from whole school evaluation (WSE) and Incidental Inspection findings reveal scope for 'significant' improvement in teachers' assessment practices, particularly in literacy and mathematics, to ensure they become embedded in teaching and learning (DES, 2010a; Ó Donnchadha & Keating, 2013). Interestingly, a recent INTO (2015) publication recommends that "teachers should be supported in developing their knowledge and skills in assessment" (p.44).

AfL research in the Irish context

The increasing national and international prominence of AfL in both policy and practice is reflected by the ever-expanding body of research studies into AfL undertaken in the Irish context, especially at postgraduate level (e.g., Collins, 2010; Darcy, 2011; Doyle, 2012). Most studies, to date, have been small-scale, single-case, action research projects, leaving Ireland at a disadvantage in comparison to jurisdictions such as the UK and the US, where large-scale studies have taken place. The doctoral study by Lysaght (2009), which explored the potential of a TLC as a vehicle of professional development (PD) to enhance teachers FA practices and student achievement in reading in one DEIS school in Co. Louth, warrants special mention at this point since it particularly influenced the design of the current study. Lysaght, in collaboration with O'Leary, subsequently developed the AfL audit instrument (AfLAi), which seeks to enable teachers and schools to accurately gauge the extent to which they implement AfL with fidelity, and provide them with information to guide future staff CPD in assessment. To date, results from an in-depth snapshot of the FA practices of over 450 primary teachers in Ireland indicate that teachers view questioning and classroom discussion (QCD), sharing learning intentions and success criteria (LISC) and feedback (FB) (rank order) as *emerging,* rather than *embedded* or *established*, while peer- and self-assessment (PSA) is reported as being more *sporadic* (Lysaght & O'Leary, 2013). Nevertheless, while this information provides a valuable insight into current Irish teachers self-reported AfL practices, research into AfL in the Irish context remains sparse, especially in the area of mathematics.

AfL and Mathematics

It has been argued that AfL has both generic features that apply to all school subjects and across all stages while also having features that are specific to particular subjects (Black et al., 2003; Hodgen & Wiliam, 2006). However, while the principles of effective AfL can be applied equally to all subjects, much of the research internationally has been carried out in the areas of mathematics and science (e.g., Harlen & Winter, 2004). While, the paucity of large-scale research into AfL has already been highlighted, it is worth noting that all the key large-scale projects previously discussed had mathematics as the primary or secondary focus (Table 4). For example, regarding their choice of mathematics in KMOFAP Black and Wiliam (2003) stated that they started with mathematics and science since "these were subjects where we felt there were clear messages from the research and also where we had expertise" (p.630). Consequently, mathematics is a subject area with a strong research base in AfL, thereby making it a suitable subject for this research. Balan (2012) posits that in mathematics education, research into FA can be divided into two main categories: studies that focus on teachers, for example attempting to improve their assessment literacy (e.g., Koh, 2011) or studies directed at students' mathematical learning (e.g., Andersson, 2014).

Different disciplines have different approaches to inquiry. Regarding mathematics, research evidence suggests that focusing on the use of day-to-day AfL is one of the most powerful ways of improving learning in mathematics classrooms and can result in large learning gains (Wiliam, 2007). It has been argued that FA is an essential process in mathematics education that helps ensure students master the crucial contextual and conceptual knowledge they need to use mathematical procedures appropriately (National Council of Teachers of Mathematics [NCTM], 2013; Noyce & Hickey, 2011). The NCTM (2013) in the USA recently clarified their position on the role of FA in mathematics education stating:

> Through formative assessment, students develop a clear understanding of learning targets and receive feedback that helps them to improve. In addition, by applying formative strategies such as asking strategic questions, providing students with immediate feedback, and engaging students in self-reflection, teachers receive evidence of students' reasoning and misconceptions to use in adjusting instruction. By receiving formative feedback, students learn how to assess themselves and how to improve their own learning. At the core of formative assessment is an understanding of the influence that assessment has on student motivation and the need for students to actively monitor and engage in their learning. The use of formative assessment has been shown to result in higher achievement. The National Council of Teachers of Mathematics strongly endorses the integration of formative assessment strategies into daily instruction. (n.p.)

In the Irish context, the importance of regularly using AfL to enhance the teaching and learning in mathematics is also recognised by the DES (2011a) and the NCCA (2007; 2016), as well as some post-graduate researchers of mathematics (e.g., McDonnell, 2013). Notwithstanding, teachers need further support than they have received thus far if they are to develop and enhance their skills, knowledge and effective use of AfL.

General Critique of AfL

While AfL has many advocates, it is notable that within the education literature there is an increase in the number of reviews criticising claims made regarding its efficacy, resulting in some scholars questioning whether the effects of FA on learning have been over-sold (e.g., Baird et al., 2014). Although lauded for over a decade, and widely credited with being the catalyst for the increased interest in AfL, in recent years the publication by Black and Wiliam (1998a) has come in for some criticism, particularly regarding effect size claims (Bennett, 2011; Briggs, Ruiz-Primo, Furtak & Shepard, 2012; Dunn & Mulvernon, 2009). Bennett (2011) concludes that Black and Wiliam's review has been treated as a meta-analysis, which it is not, leading to a "mischaracterisation that has essentially become the educational equivalent of urban legend" (p.12).

Critics of AfL research primarily highlight issues such as definitional discord and terminological nuances, emphasising the need to clarify the nomenclature and develop a clear and shared lexicon for FA (Dunn & Mulvernon, 2009; Kingston & Nash, 2011), with Bennett (2011) concluding that AfL and FA are under-defined and currently represent a range of approaches. Dunn and Mulvernon (2009) and Kingston and Nash (2011) highlight what they consider are exaggerated claims regarding AfL's effectiveness, criticising the validity of effect size claims and the paucity of sound empirical evidence regarding the impact of FA in education. Others have in turn criticised these reviews, while Kingston and Nash (2015) recently issued an erratum regarding some of their previous claims (2011).

Bennett's (2011) critical review of AfL is the most extensive and widely cited and, in it, he discusses six interrelated issues. He stresses the need for prudence with regard to any claims made about AfL and highlights the need to tackle measurement issues, since many of the studies claiming effect sizes of 0.4 to 0.7 come from "untraceable, flawed, dated or unpublished sources" (p.5). Bennett suggests that AfL be instantiated and conceptualised within specific domains. He concedes much time and professional support would be needed if most teachers are to become proficient users of AfL and recognises that FA exists within the larger educational context. Wiliam (2011a) has acknowledged the importance of

Bennett's review while Stobart and Hopfenbeck (2014) have used it as a framework in their recent critique of FA.

It appears from the arguments and counter-arguments that little consensus exists among experts in the ongoing debate about FA's efficacy and that conceptions of AfL are still evolving. Notwithstanding, the bottom line, upon which most researchers agree, is that better quality research covering all aspects of AfL is needed (Briggs et al., 2012; Kingston & Nash, 2012; McMillan, Venable & Varier, 2013). There are, however, some specific issues that are regularly highlighted in the AfL literature as needing further research and these are discussed forthwith.

Scalability

A common criticism leveled against AfL research is the lack of large-scale research projects that affirm claims made about its warrant. Indeed, while a recent review of the state of the field (Baird et al., 2014) reviewed 481 conference proceedings and 907 peer-reviewed articles on AfL/FA, less that 10 of these studies could be described as large-scale and most involved just one or two schools and relatively few pupils. Furthermore, Baird et al. highlighted that few randomised control trials have been carried out. While both the KLT and EFA projects try to address the issue of scalability, the challenges of implementing effective CPD at scale are manifest in both programmes (Leahy & Wiliam, 2012). Problems such as replicability, sustainability, and dissemination have also been highlighted in the literature. Furthermore, the need to accept and embrace the diversity of schools means that *one size fits all* interventions seldom succeed. Some recent AfL projects, nationally (e.g., Lysaght, 2009; Lysaght & O'Leary, 2013) and internationally (e.g., DeLuca, Klinger, Pyper & Woods, 2015; Hopfenbeck, Petour & Tolo, 2015; Jonsson, Lundahl & Holgrem, 2015), have tried to explore how the successful support conditions of small-scale AfL projects can be scaled up to national level. Hopfenbeck et al., (2015) found that AfL can be implemented successfully at scale provided that factors such as trust, dialogue and teacher agency are present in schools. They identified the following factors that should be considered when implementing AfL on a large scale:

> External testing, teacher resistance to peer- and self-assessment, teacher resistance to change teacher and student roles, lack of commitment from senior staff, shortcomings in teachers' disciplinary knowledge and assessment skills, superficial understanding of the approach, busy classrooms and lack of knowledge on how to put AfL into practice. (p.47)

Regarding scalability, Thompson and Goe (2009) believe that increasing scalability is possible, but that an iterative research and development process will be needed. Nevertheless, Clark (2011) is of the opinion that large-scale implementation of AfL remains a distant prospect.

The Learner's Voice

New advances in thinking about AfL mean that current conceptualisations recognise the importance of student involvement in the assessment process (Brookhart, 2011b; Shepard, 2006; Shute & Kim, 2013). As argued:

> What's new in formative assessment is the importance of *students* as formative decision-makers who need information of a certain type (descriptive) at a certain time (in time to act) in order to make productive decisions about their own learning. (Brookhart, 2011a, p.4)

Underpinning the potential of AfL to impact learning is the belief that learners know where they are in their learning, where they are going and how they can close the gap (Chappuis, 2005; Willis & Cowie, 2014). Therefore, understanding participant thoughts and actions with regard to AfL interactions is essential, and their perspectives are at least as important as those of adults (Hayward, 2013; Stiggins, 2007: Willis and Cowie, 2014). Furthermore, Andrade (2010) argues that since students "have exclusive access to their own thoughts and actions, they can and should be considered as the definitive source of FA information (p.12). These views echo Sadler (1998) who argues that:

> We need to let students into the secret, allowing them to become insiders of the assessment process. We need to make provision for them to become members of the guild of people who can make consistently sound judgments and know why those judgments are justifiable. (n.p.)

Traditionally, however, learners have had a quiet voice (Hayward, 2012). To date, research into students' perspectives on FA practices is somewhat limited. Most studies have taken place outside of Ireland, in tertiary settings, and predominantly investigated students' perceptions of FB (e.g. White, 2009).

A study by Hargreaves (2004) identifies six commonalities to student voice and AfL (increased student engagement, responsibility and meta-cognitive skills, open collaborative relationships with staff, more capable and confident social skills and enhanced active participation) and suggests that if schools work on one, either AfL or student voice, it acts as a bridge to fostering the development of the other. One of the earliest studies exploring the student's voice suggests that even children in infant classrooms are able to understand and interpret FB (Tunstall & Gipps, 1996). Another study, examining second-level students' and teachers' perceptions of existing in-class FA practices in New Zealand (Rawlins, 2010), found disparity between teachers' and students' FB preferences, with teachers perceiving oral FB as effective but students preferring to engage with written FB through peer-oriented interactions. In contrast, a further study, also in New Zealand (Harris, Gavin & Hartnett, 2012), found that "the source and content of FB was strongly associated with the teacher giving written comments or grades of various sorts" (p.12) and found that "there is a long journey ahead to persuade students that FB can come legitimately from themselves and their peers" (p.13).

In her research, Williams (2010) particularly emphasises the importance of the learners' viewpoint and notes that, although children have a right to speak for themselves, their voices regarding FB are rarely heard. She posits students could offer valuable insights and "round out the picture being developed by other researchers" (p.304). Studying participants in New Zealand, aged 11-13 years, Williams found that "the ability of the children to comment on their own learning was established to a remarkable degree" (p.312). They understood the concept of FA and what makes it useful to them. Specifically regarding FB, they could articulate the elements most helpful to them to improve their learning, the type of FB they preferred and the most appropriate timing of that FB. Indeed, Williams reported that many of their ideas regarding FB were found to closely match Sadler's (1989) FB principles.

Referring to the AiFL study (2008) with primary students in Scotland, Hayward (2012) also gives a clear picture of the student perspective with regard to AfL. She argues that where teachers and students engage in FA practices in such a way that the *spirit* of AfL is in evidence, this precipitates a change in relationships between both and necessitates listening to learners' perceptions of their AfL experiences, not just those of teachers. She found that students had a good understanding of AfL, were enthusiastic about their learning and were positively disposed to their new role. The students understood what they were to learn, recognised the central role of the teacher as the "arbiter of success" but also linked peer explanation to effective learning. They stressed the importance of listening and being listened to, while further evidence suggested, "consultation and choice were key features of changed teacher-pupil relationship and that such changes were necessary conditions of motivation and engagement" (p.135).

Various authors have emphasised the centrality of the learner in the learning process (e.g., Wiliam, 2014). If we accept, as Earl (2013) argues, that "learning is an active process of constructing thoughts and making sense of the world" (p.109), then we need to provide opportunities for learners to be reflective and to give them opportunities to engage in dialogue. We need to hear their voice, to let them articulate what is going on in their hearts and minds as they engage in AfL.

AfL and Affect

While many accept that effective use of AfL leads to gains in student achievement, some also believe that AfL can positively impact affective factors such as students' motivation (Cauley & McMillan, 2010; Clarke, 2008; Gardner, 2012b; Heritage, 2013; Stiggins, 2006), self-confidence (Chappuis et al., 2004; Clarke, 2008; Stiggins, 2006) and self-esteem (Clarke, 2008; Heritage, 2013; Miller & Lavin, 2007). Miller and Lavin (2007), however, caution that the empirical evidence to support such claims is somewhat limited. One study by Yin, Shavelson, Ayala, Ruiz-Primo, Brandon, Furtak, Tomita and Young (2008) reported that FA did not have any statistically significant effect on student motivation whereas findings from a study by Miller and Lavin (2007), with upper primary students in Scotland, found gains in students' beliefs about competence and in self-esteem, although differences were found in relation to ability, gender and confidence levels. Additionally, the data suggested that benefits in terms of student self-perception might only be evidenced in the long term. Furthermore, certain caveats from the literature regarding AfL and student affect should also be noted: first, the negative impact of summative assessment on self-esteem and self-confidence (Harlen, 2012; Miller & Lavin, 2007) and second, Hattie and Timperley's (2007) findings which reveal that the impact of FB on student motivation can be either positive or negative depending on the type of FB the learner receives.

Conclusion

This section of the literature review has outlined the development of AfL, discussed definitional nuances and AfL principles, and presented a research-based rationale for the study. It has examined the theoretical underpinnings of AfL and its conceptualisation as various strategies and techniques. Additionally, it has explored key AfL projects, reviewed AfL in the Irish context and AfL and mathematics. The section concluded with a general critique of current literature that highlighted the lack of consensus in the field with regard to definitions, theories, conceptualisations and the actual impact of AfL. Furthermore, attention was drawn to the need for more and better quality research, including further research investigating the student voice, AfL and affect and how AfL might be scaled up with fidelity yet flexibility. Black (2015), entitled a recent article *FA - an optimistic but incomplete vision*, drawing attention to the need for further research. It seems what Black and Wiliam (1998b) highlighted all those years ago remains the case: “The improvement of formative assessment cannot be a simple matter” and if the “substantial” rewards promised by FA are to be realised it will take time and sustained CPD “since lasting and fundamental improvements in teaching and learning can only happen in this way” (p.15).

In sum, the involvement of reflective, engaged professional teachers is central to the development of effective classroom-based FA (e.g., Black, McCormick, James & Pedder, 2006; Warwick et al., 2015) and each teacher has to incorporate the AfL strategies and techniques “into his or her own practice in his or her own way” (Black & Wiliam, 1998a, p.62). Consequently, the next section looks at how teachers can improve their skills, knowledge and implementation of AfL through one particular vehicle of professional learning, Japanese lesson study (LS). Starting with a brief look at CPD in general, the remainder of the section investigates LS as a school-based model of CPD through which teachers can learn collaboratively about AfL and how it can best be implemented in their classrooms.

Continuing Professional Development

Few scholars would argue with the assumption that providing high-quality and relevant CPD, at all stages along the teacher continuum, is essential in efforts to achieve a high-quality teaching force and improve instructional practices and student learning (Borko, 2004; Bubb & Earley, 2007; Opfer & Pedder, 2011), making CPD a major policy priority for most education systems worldwide (Banks & Smyth, 2011; OECD, 2005). Various terms pervade the relevant literature, sometimes reciprocally, to describe teacher professional development or learning: CPD, in-service education, staff development, lifelong learning, professional development (PD), professional learning (PL), and in-career development. In recent government (DES, 2011a), and Teaching Council (TC) documentation (TC, 2011), CPD appears to be the phrase of choice and, more recently, PL (TC, 2016). For the purpose of this research, the terms CPD, PL and PD are used interchangeably.

While acknowledging CPD may be a common phenomenon, myriad definitions exist and current conceptualisations and meanings are varied and debated, although in reality they have much in common. One definition by the TC (2011), the body charged with oversight of teacher CPD in Ireland, states:

> Continuing professional development (CPD) refers to life-long teacher learning and comprises the full range of educational experiences designed to enrich teachers' professional knowledge, understanding and capabilities throughout their careers. (p.19)

Inherent in this definition, and congruent with interpretations by others (Bubb & Earley, 2007; Day, Sammons, Stobart, Kington & Gu, 2007; Fullan, 1991), is the notion that CPD includes not only formal learning activities but also comprises all informal learning experiences across the career span. Some scholars expand their definition of CPD to include using technology to enhance teaching (Grant, 1996; Mohan, 2011) while others suggest that CPD should not be limited to rejuvenating practice or expanding one's professional repertoire but, rather, should also promote personal growth and increase self-confidence, self-esteem, resilience, and enthusiasm for teaching, thereby enhancing job satisfaction (Bubb & Earley, 2007; Pachler & Field, 2004). This study adopts the TC definition.

While there is general consensus among teachers and educationalists about the import of teachers' PL as one way of improving education, there is little agreement about how this process occurs, what form it should take or what it should focus on (e.g., Van Driel, 2014; Wiliam, 2011b). There is, however, acceptance that traditional approaches, such as "one-shot" workshops and conferences, where teachers listen passively, are limited in their effectiveness (Bubb & Earley, 2007; Wiliam, 2011b). Additionally, there is increasing concurrence that effective CPD attends to context, process and content elements (Lysaght, 2009; Wylie, Lyon & Mavronikolas, 2008). The context should be conducive to teacher learning and concomitant changes leveraged by participation in the CPD. The content should be research-based, relevant to participants' needs and likely to impact positively on student learning, while the process concerns ways of leveraging teacher change.

An extensive search of the voluminous research literature on teacher CPD revealed that effective CPD for teachers comprises many common elements as identified by researchers in the field. Table 5 summarises these key elements, in descending order of frequency and reveals consensus about what scholars consider the core features of effective CPD to be. As outlined, the most effective CPD incorporates many of the elements listed in the Table in that it is sustained, fosters teacher professional collaboration and is closely connected to teachers' work in the classroom, to name but a few.

Table 5
Key Elements of Effective CPD

Characteristic	Guskey (2000)	Hirsh (2009)	Desmione (2011)	King (2012)	O'Sullivan (2011)	Darling-Hammond & Richardson (2009)	Wei et al. (2009)	Bubb & Earley (2007)
Job-embedded	✓	✓	✓		✓	✓		✓
Sustained		✓	✓	✓		✓	✓	✓
Collaborative		✓	✓	✓		✓	✓	
Active learning/linked to practice			✓	✓	✓	✓	✓	✓
Builds Relationships			✓		✓		✓	
Intentional	✓	✓						
Systemic	✓	✓						
Intensive						✓	✓	
Focus on content			✓		✓	✓		✓
Evidence-based				✓	✓			
Supported by leadership				✓	✓			

Three primary influences are identified in the literature as impacting the effectiveness of CPD and include the individual teacher, the supports and structures offered by the school for teacher learning and the actual CPD experience (King, 2012; Opfer & Pedder, 2011; Smith & Gillespie, 2007). Individual factors mediating the influence of CPD include teachers' motivation, concerns, self-efficacy and learning styles (e.g., WeiBenrieder, Roesken-Winter, Schueler, Binner & Blömeke, 2015). School factors emphasise collegiality and the crucial role played by school leadership, particularly the principal, in supporting successful CPD, especially if it is school-based (Murchan, Loxley & Johnston, 2009;

Robinson, Hohlepa & Lloyd, 2009; Timperley, 2011). With regard to the actual CPD experience, scholars highlight the importance of teachers adopting a professional inquiry stance and engaging in their own research and the research of others as part of their CPD journey (e.g., Bell, Cordingley, Isham & Davis, 2010; Cordingly, 2015b; Borko, 2004; Cochran-Smith & Lytle, 2009).

Continuing Professional Development: The Irish Context

In the Irish context, *The Literacy and Numeracy Strategy* (DES, 2011a) reveals the government's belief in the critical importance of providing high-quality CPD for teachers, throughout their careers, to enable them to enhance their pedagogical skills and understanding, improve literacy and numeracy and make a substantial difference to the quality of students' learning. Interestingly, reminiscent of findings by Delaney (2005), a recent report using data from Trends in International Mathematics and Science Study (TIMSS) and Progress in International Reading Literacy Study (PIRLS) 2011, *National Schools, International Contexts* (Eivers & Clerkin, 2013), found that "Irish teachers were far less likely to engage in regular CPD than were teachers in most countries" (p.88), were less likely to have participated in mathematics CPD in the previous two years, and also spent less time on reading-related CPD. Additionally, in comparison to the TIMSS (2011) study average, Irish teachers had "particularly low rates of participation for CPD related to assessment" (p.89). Regarding collaborative practices, recommended by the TC's code of professional conduct and the DES guidelines for School Self-Evaluation (SSE), the report states "Ireland is unusual, in international terms, for the very low level of collaboration and sharing of professional expertise among teachers of fourth class pupils" (Clerkin, 2013, p.100). More specifically, in the two years prior to TIMSS 2011, just 16% of pupils were taught by teachers classified as being *very collaborative*, less than half the TIMSS average.

Furthermore, 25% of pupils were taught by teachers who say they *never/almost never* discuss teaching, 27% by teachers who say they *never/almost never* collaborate in preparing materials with another teacher, while 82% of Irish pupils were taught by teachers who say they *never/almost never* visit another classroom to learn about teaching, compared to 53% in all the TIMSS countries (Clerkin, 2013). Perhaps the low uptake of CPD may in part be explained by the fact that, in Ireland, participation in CPD is not linked to teacher certification. Nonetheless, it is worth remembering that twenty hours of CPD related to numeracy, literacy or assessment is now mandatory over every five-year period (Circular 0056, DES, 2011b).

On the topic of teacher CPD in the Irish context, the authors of a recently published NCCA report (2014b) have recommended that "school-based lesson study should be given due attention". Although speaking specifically about mathematics teacher development, they believe that LS is a "particularly effective vehicle" of CPD since it is "enacted by a community of educators working in their own setting" and involves colleagues being "mutually engaged in the shared enterprise of developing mathematical proficiency in their learners" (p.123). Many other scholars also assert that LS is an ideal strategy for building, developing and sustaining communities of educators, while also addressing students' learning needs and improving instruction (Chichibu & Kihara, 2013; Lewis & Hurd, 2011; Lewis, Perry & Hurd, 2009; Lewis, Perry & Murata, 2006; Lieberman, 2009; Roberts & Pruitt, 2009). Interestingly, Desforges (2015) recently argued that traditional models of teachers' CPD, which took place outside the classroom, assumed that learning would be transformed and applied in the classroom context, which often was not the case, whereas "one of the great, in-principle attractions of LS is that such transfer is unnecessary" (p.xvii). Consequently, Desforges (2015) and other LS scholars, nationally (e.g. Corcoran, 2011b) and internationally (e.g., Dudley, 2015), believe that LS should be the CPD strategy of choice for schools or, as Dudley (2015) puts it, "professional learning for our time" (p.1). The remainder of this chapter investigates LS in detail.

Lesson Study

Background to Lesson Study

Lesson study (LS) is described as "a practice that is currently foregrounded in the literature as a significant development on school-based professional development" (NCCA, 2014b, p.123). The words "lesson study" are derived from two Japanese words, *jugyou* meaning lesson or instruction and *kenkyuu* meaning study or research. These words can equally well be translated in reverse order as research lesson (Lewis, 2000). LS can typically be characterised as a sustained, teacher-led, peer-to-peer, practice-based, research-oriented, systematic and collaborative model of continuing PL and practice development (e.g., Cajkler et al., 2014; Corcoran, 2008; Dudley, 2013; Murata, 2011). The principal purpose of LS is to improve the quality of teaching and learning through a collaborative, reflexive and recursive process (Cajkler, Wood, Norton, Pedder & Xu, 2015; Dudley, 2011), which Cajkler et al. (2014) claim can help build teachers' "professional capital" (Hargreaves & Fullan, 2012). LS originated in Japan where it is widely considered the principal method of professional development for teachers, particularly in elementary (1st to 6th grade) and middle schools (7th to 9th grade) (Lewis 2000; Stigler & Hiebert, 2009; Yoshida, 1999). In Japan, there are manifold versions and sizes of LS, ranging from small-scale, in-school initiatives to large-scale, nationwide ones, with most Japanese teachers belonging to at least one LS group (Kieran, Krainer & Shaughnessy, 2013; Sloane, 2005). For many decades, LS has been viewed as the *linchpin* of the school improvement process in Japan and is widely credited with facilitating instructional improvement in schools there and contributing to their high ranking in international assessments such as TIMSS and PISA (Stigler & Hiebert, 2009; Fernandez & Yoshida, 2004). LS is deeply integrated within Japanese educational culture and teachers' participation is voluntary (Baba, 2007; Verhoef, Tall, Coenders & Van Smaalen, 2013). Fernandez and Yoshida (2004) emphasise, however, that while most elementary and middle schools in Japan engage in LS, the quality of LS "varies widely depending on the quality of school leadership, the quality of the teachers in the building, the bonds that exist between them, and their inherent interest in *konaikenshu*" (in-service education within the school) (p.17). Additionally, Sloane (2005) cautions that Japanese performance in international comparisons is not only based on high-quality LS but also on the proliferation of "cram schools" where students focus on drill and memorisation. Lewis (1995), in contrast, has argued, that "cram schools" are only a factor in junior high, the last stage of compulsory Japanese schooling.

As reported, LS can be conducted in any 'academic' or 'nonacademic' area, (Doig, Groves & Fujii, 2011; Lewis & Hurd, 2011) although, outside of Japan, most LS research has focused on mathematics (e.g., Murata, 2011; Williams, Ryan & Morgan, 2014). LS may also vary according to its focus, which can be on the development of teaching and learning methods, the quality of students' learning experience or the development of a particular lesson (Elliot & Lo Mun Ling, 2011). Interest in, and use of, Japanese LS has grown exponentially since the 1990s, especially following the publication of *The Teaching Gap* (Stigler & Hiebert, 1999). For almost two decades, LS research has been reported extensively in English language journals and has migrated across Asia, the US and Canada and is now used in Europe, the Middle East and Africa (Dudley, 2013). Its use has increased to such an extent that the World Association of Lesson Studies (WALS) was formed in 2007 to advance LS research and promote LS practices and there is now a dedicated peer-reviewed journal (*International Journal for Lesson and Learning Studies*), published since early 2012.

The Lesson Study Process

According to Desforges (2015), "creativity is the *sine qua non*[4] of LS" (p.xix), resulting in a great deal of autonomy in how LS is understood, interpreted and implemented. Conversely, Dudley (2015) cautions that because of its adaptability and apparent simplicity there is a danger that LS could be:

> Corrupted by a teaching profession that too often has been encouraged to simply innovate for innovation's sake …and is unused to adopting professional levels of clinical discipline when applying or honing classroom innovations. (p.5)

Notwithstanding, among LS experts, opinions vary with regard to the optimal number of participants in a LS group with Stigler and Hiebert (2009) suggesting three to five members as ideal, Morris and Hiebert (2011) three to six; and Dudley (2011) of the opinion that three works well or even two. Regarding the LS cycle, LS scholars typically outline from four to six steps (Corcoran, 2011b; Murata, 2011) although Stigler and Hiebert (2009) and Cerbin and Kopp, (2011), propose seven. Murata (2011) asserts, however, that most LS cycles typically include the steps outlined in Figure 5 at a minimum.

---

[4] An essential condition; a thing that is absolutely necessary. (Retrieved from https://en.oxforddictionaries.com/definition/sine_qua_non).

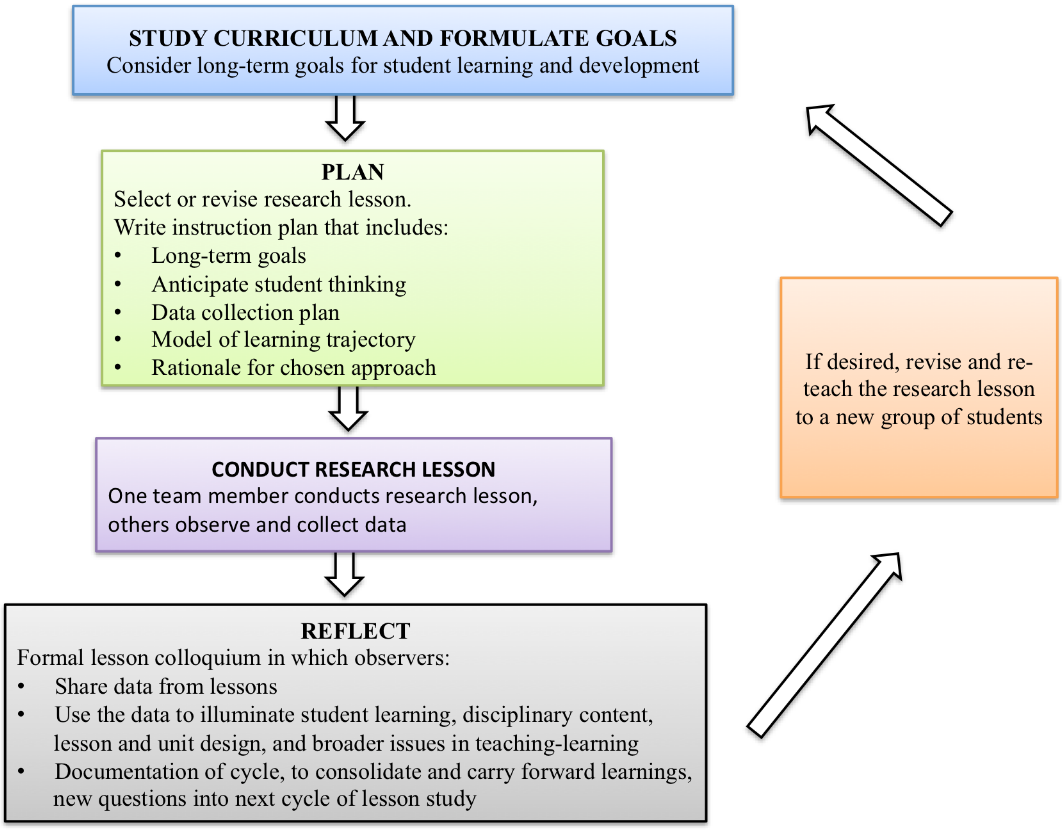


Figure 5. Lesson study cycle (Adapted from Murata, 2011)

As outlined, the LS group meet to study the curriculum and formulate goals, they plan and teach the research or live lesson, debrief and reflect on this lesson and then revise and reteach it, if necessary. These steps are discussed and explained in detail in Appendix C.

LS groups can elect to focus on particular case pupils (Dudley, 2011; Gardner, Galanouli, Devlin, Magee, McSweeney, McHenry, McVeigh & Mitchell, 2012) or not, to see how students' thinking evolves, although in a recent study Gardner et al. (2012) considered that these pupils were probably self-conscious and acted differently when they knew they were being observed. Other researchers prefer to focus on whole class groups and this has been the norm in the Irish context (Corcoran, 2011a; Leavy, 2010; McLoone, 2011; Ní Shúileabháin, 2015).

Knowledgeable other

Much of the extant LS literature mentions the role of "Knowledgeable Other" (KO), coach, external expert or adviser from outside the LS group who is invited to help support and guide the LS process (Doig & Groves, 2011; Dudley, 2011; Fernandez, 2005; Kieran et al., 2013; Watanabe, 2002). This person is usually a university lecturer, school principal or experienced teacher who has expertise in LS, specialises in the particular subject area or is especially interested in the topic (Murata, 2011; Wang-Iverson & Yoshida 2005; Yoshida & Jackson, 2011). Their role is to provide information about subject matter or its teaching, to add new perspectives and ideas, to raise questions, provide comments and be a co-researcher and to share the work of other LS groups, all without being authoritative in their approach (Fernandez, 2005; Murata, 2011; Wang-Iverson & Yoshida, 2005). Nevertheless, since LS is relatively new to the UK and especially Ireland, this can present challenges since few researchers are experienced in LS and/or have developed the expertise as KOs, even if they are experts in their own field. Outside of Japan, therefore, the role of the KO is an area that requires further attention and development (Fernandez, 2005; Lewis et al., 2006; Murata, 2011). There isn't general consensus as to the necessity for a KO, with some researchers advising that the inclusion of an outside scholar is not required (Lewis & Hurd, 2011; Watanabe, 2011), while feedback from LS groups and most LS researchers suggests they can greatly enhance the LS process (Hart & Carriere, 2011; Lewis & Hurd, 2011; Wang-Iverson & Yoshida, 2005). Without the contribution of a KO, scholars (Cajkler at al., 2014; Fernandez, Cannon & Chokshi 2003; Hart & Carriere, 2011) suggest that it is of paramount import that teachers have good pedagogical content knowledge (PCK) so that they can engage in substantive discussions about content and pedagogy and question each other's assumptions, thus maintaining the quality of the LS process.

Theoretical Underpinnings

LS can be viewed as a type of research (Dudley, 2011; Fernandez et al., 2003) with even the etymology of the words and the centrality of the research lesson in the LS process suggesting this. Japanese teachers also consider LS to be research, but Lewis (2002a), one of the most prolific LS writers in Western contexts, emphasises that there are two important distinctions between LS and most educational research. First, she posits, the goal of LS is to improve instruction and knowledge in one's own purview as opposed to generating knowledge that others will apply. Second, LS explores an active improvement effort rather than just any idea or question. In short, and reminiscent of Cochran-Smith and Lytle's (2009) *knowledge-of-practice*, Lewis (2002a) states that "in a traditional research model research is applied to practice. In LS practice is research" (p.86). In a more recent article, Lewis (2012b) elaborates further on this idea of LS as research and suggests that four features of LS, in particular, give it an edge over other educational research methods common in the West and the power to transform practice and inform policy. She states that LS: "welcomes and learns from variation"; "provides a natural mechanism for scale-up of innovations"; provides "a window for policy makers to conduct formative research on policy" and "creates demand for research" (pp.1-2).

Regarding LS as research, Fernandez (2002) considers that one of the main challenges facing teachers is "trying to adopt the research focus that is inherent in lesson study" (p.393). In an exploration of the strong claims made about the potential of LS, Fernandez, et al., (2003) asked a group of US teachers to engage in LS (in mathematics) with the support of some Japanese teachers from a nearby international school. They concluded that, in order to benefit from LS, teachers would first need to learn how to apply the following three critical lenses to their examination of lessons or practice: the *researcher lens*, the *curriculum developer lens* and the *student lens*, something Japanese teachers appear to do automatically in the LS process. In the study, the Japanese teachers encouraged their US counterparts to focus on the research process rather than the research lesson, and to view themselves as researchers by generating and testing meaningful hypotheses, gathering evidence of their findings and articulating what they had learned (*researcher lens*). For the research lessons, the Japanese teachers encouraged the Americans to choose lessons on core topics and make curricular connections across the grades, thus connecting study lessons to previous and future learning (*curriculum developer lens*). Adopting this perspective allows LS participants to deepen their knowledge of the curriculum, take cognisance of students' prior learning, better understand the development of content within a lesson and how to structure it and, ultimately, improve their teaching. Finally, the Japanese teachers encouraged the Americans to explore all aspects of the lesson through the eyes of their pupils by anticipating students' thinking, learning, understanding and behaviour, using this to plan and evaluate their lessons (*student lens*). In sum, this study suggests that, in order for teachers to fully engage in the LS process, they need to become critical reflectors on their practice and make use of the three critical lenses discussed above.

Corcoran (2007) informs us that various researchers have proposed different theoretical frameworks for LS. Dudley (2015) proffers that the LS process promotes "the conditions necessary for learning as set out in socio-cultural learning theory" including "the building of community and the suppression of concerns about one's self within the group" (p.16). This is similar to Bocala (2015) who uses a sociocultural perspective to explore the development of teachers' learning in LS. Wood (2015), on the other hand, recently argued that although there appears to be an implicit theory of learning underpinning Japanese LS, it is not clear how this comes to guide teachers in the design of learning situations, but he believes that constructivism has a role to play. Many researchers view LS as being underpinned by situated learning theories (e.g., Corcoran, 2008; Lenski & Caskey, 2009;

Lewis et al., 2009); others have drawn on Cultural-Historical Activity Theory (CHAT) to provide theoretical underpinnings for their research (e.g., Dudley, 2013; Tsui & Law, 2007; Wake, Swan & Foster, 2015), while some have adopted an exploratory, open-ended approach (e.g., Cajkler et al., 2015). Many researchers internationally (e.g., Cajkler & Wood, 2016; Gómez-Blancarte & Miranda, 2014) and in the Irish context (Corcoran, 2008; McLoone, 2011; Ní Shúilleabháin, 2015) have conceptualised LS groups as communities of practice (Wenger, 1998). Regarding the various attempts to describe LS from different theoretical perspectives, Krainer (2011) acknowledges that, while this can be fruitful he believes one should not impose a theoretical frame on the *culturally situated approach* but should rather take an open approach when reflecting on LS. In reality, the theoretical base for LS is in the early stages of development.

Perhaps more than any other LS researcher in the West, Lewis, in collaboration with colleagues, has tried to elucidate why LS improves instruction. For over a decade she has been working on a theoretical model that demonstrates how she thinks LS produces instructional improvement (e.g., 2009; 2011; 2015). Drawing on Japanese sources, she includes the visible features of LS (e.g., planning, teaching, observing and revising) and identifies certain pathways for teacher learning through which LS, she believes, improves instruction and ultimately student learning (Figure 6). The practice-based cycles of LS (featured in the left column of Figure 6) are hypothesised to improve instruction by simultaneously improving the basic inputs to instruction (centre column). Lewis' model is informed by cognitive theories of teacher learning and situated learning theories and she posits that LS makes teachers' ideas about pedagogy and student learning visible, strengthens professional community and improves instruction by developing teachers' knowledge (of content, pedagogy and student thinking) and by improving learning resources. Lewis (2009) does not claim that her theoretical model is definitive but calls for researchers to build on it.

Pathways for Teacher Learning

Visible Features of Lesson Study:

Investigation

Consider students' current characteristics

Consider long-term goals for student learning and development

Study the content area: key concepts, existing curricula, standards, learning trajectory, research

Planning

Select or develop research lesson

Try task in order to anticipate student solutions

Write up instructional plan, including goals for student learning and development anticipated student thinking, data collection points, rationale for lesson design, connection to long-term goals

Research Lesson

Conduct research lesson

Team members observe and collect data during live research lesson

Reflection

Share and discuss data from research lesson in post-lesson colloquium

Team members (and other observers) draw out implications for research lesson redesign, for teaching-learning more broadly, and for understanding of students and subject matter

Summarise in writing what was learned from cycle, to consolidate the learning

Revise and reteach lesson (optional)

Becomes Thinking Visible

Community Norms, tools, identity, and participation develop

Intervening Changes in:

Teachers' Knowledge, e.g.:

Subject matter knowledge, pedagogy, student thinking and how to capture it

Curriculum materials, long-term goals of discipline

How to connect daily practice to long-term goals

How to integrate diverse, sometimes fragmented information sources (e.g. textbooks, assessment results, standards, prior practice, professional development)

Teachers' Skills, Beliefs and Dispositions, e.g.:

Sense of efficacy (by changing my practice, I can improve student learning

Belief that change can make a difference

Motivation to learn and improve

Attention to student thinking

Inquiry stance toward practice

Teachers' Professional Community

Norms and relationships that support daily learning from colleagues and continuing improvement, e.g.:

Sense of accountability to colleagues' to provide high-quality instruction

Commitment to improvement of practice, and to instruction that is coherent across classrooms

Sense of mutual responsibility: 'our' students, not 'my' and 'your' students

Learning Resources and Tools, e.g.:

Lessons that reveal and promote student thinking

Tools that support planning, discussion, and observation

Assessment strategies that reveal student thinking

System Features, e.g.:

Time for collaborative planning, observation and analysis

Structures that use lesson study to vet new curricula, socialize new teachers, meet policy mandates, etc.

Improvement Of Instruction (Research Lesson and Daily Practice)

Improvement of Student Learning

Figure 6. Lesson study theoretical model: How lesson study produces instructional improvement (Adapted from Lewis, 2009; 2011)

Impact of Lesson Study

An extensive literature supports and recommends the use of LS and evidence attesting its benefits continue to grow (e.g., Cajkler et al., 2015; Hogan, 2015). According to research, LS encourages greater teacher collaboration, enabling teachers to participate in more focused and in-depth discussion from multiple perspectives (Dudley, 2013; Sibbald, 2009), share knowledge, ideas and resources (Lewis et al., 2009; Sibbald, 2009) and accept joint responsibility for decision making (Lawrence & Chong, 2010). As reported, through engagement in the process of LS teachers can become increasingly knowledgeable and skilful, with gains in their subject matter knowledge (Dudley, 2013; Fernandez 2005), PCK (Dudley, 2015; Fernandez, 2005; Lewis et al., 2009; Sibbald, 2009) and knowledge about students (Fernandez 2005; Lewis 2009). Through their participation in LS researchers suggest that teachers develop greater awareness of their students and their students' needs (Perry & Lewis, 2009; Ylonen & Norwich, 2012). As argued, they become more attentive to students' prior knowledge (Dotger, 2011), unmask the "hidden characteristics of pupils" (Dudley, 2013) and develop what Cerbin (2011) terms "cognitive empathy", leading to "more contextualised insights into their students' learning needs and … more learner-responsive teaching" (Cajkler et al., 2015, p.194). According to Dudley (2013), this not only sharpens teachers' understanding of proximal development needs but the focus on student learning fuels teachers' dispositions to learn. Indeed, Cajkler et al., (2014) argue that participation in LS can also lead to "greater willingness for pedagogic risk-taking, leading to lessons that include frequent hands-on opportunities and independent opportunities for learning" (p.525). Another recent argument worth noting is Dudley's (2015) assertion that teachers' use of LS should impact student assessment:

> Because LS can provide deeply formative and diagnostic insights into pupils' learning which are strengthened by the fact that they are made in the learning context it should be developed as an assessment tool. (p. 25)

Scholars have also made claims that participating in the LS process impacts teachers' attitudes and beliefs about teaching, increases their confidence to work with new concepts (Lewis et al., 2006; Sibbald, 2009), improves their sense of self- (Meng & Sam, 2011; Sibbald, 2009) and collective-efficacy (Cajkler et al, 2014; Lewis et al., 2006) and develops more positive attitudes towards teaching mathematics (Corcoran, 2008; McLoone, 2011). Additionally, some believe in the potential of LS to develop in teachers an inquiry stance and disposition for reflective practice (Fernandez, 2005; Ricks, 2011) and in its potential to contribute to and support teachers' lifelong learning (e.g., Suratno, 2013; Wood, 2015; Yoshida, 2012). Others argue that LS fosters teachers' intrinsic motivation to improve practice throughout their teaching careers (Chong & Kong, 2012; Lewis, 2005; Lewis & Hurd, 2011; Stigler & Hiebert, 2009). Indeed, Cerbin (2011) posits that LS promotes the scholarship of teaching and learning (SoTL) bringing to mind Shulman's (2004) statement on the matter that:

> It is only when we step back and reflect systematically on the teaching we have done, in a form that can be publicly reviewed and built upon by our peers, that we have moved from scholarly teaching to the scholarship of teaching. (p.166)

Similarly, Cajkler and Wood (2016) posit that LS leads to growth in what they characterise as "pedagogic literacy", inducing, they claim, an enquiry-oriented, holistic and reflective understanding of teaching (Wood & Cajkler, 2013).

Regarding the impact of LS, it is worth noting too that experts highlight that it is through the LS process that teachers' tacit knowledge and learning is made visible when otherwise it might be lost forever (e.g., Dudley, 2015; Pedder, 2015). Teacher knowledge of practice and the expertise underpinning how teachers think and support their pupils' learning effectively remains hidden and finds expression in what teachers do rather than in what they say (Pedder, 2015). Since teachers generally teach in isolation, their interventions with students in the classroom are invisible, unconscious and automatic and teachers often don't remember them in their conscious memory (e.g., Dudley, 2011; Pedder, 2015). Consequently, according to Dudley (2015), unless a way is found to access this tacit knowledge and to share it with colleagues, then “we will take most of our professional knowledge to our graves” (p.16). This is where LS comes in, i.e. the knowledge and insights gained from participation in the LS process can be shared and others can use it too (Dudley, 2015). These ideas are reminiscent of Chenoweth’s (2000) comments on LS:

> When a brilliant American teacher retires, almost all the lesson plans and practices that he or she developed also retire. When a brilliant Japanese teacher retires, he or she has left a legacy to be enhanced by future teachers. (n.p.)

Ultimately, LS is believed to improve the quality of classroom instruction (Corcoran, 2011a; Lewis et al., 2006), expedite the spread of best practice throughout the school (Barber & Mourshed, 2007) and is seen as an effective way of enhancing student learning (Cajkler et al., 2015; Lewis & Hurd, 2011; Sibbald, 2009; Stigler & Hiebert, 2009), including students with mild learning difficulties (Dudley, 2012a; Ylonen & Norwich, 2015). Figure 7 provides a theoretical model that illustrates the impact of LS as it shapes individual teachers, the curriculum and the teaching culture. It demonstrates, according to Lewis (2011) that, over time, LS not only improves planning and instruction but also supports the growth of individual teachers and the teaching community- for example, development of norms that expect de-privatisation of practice, inquiry and self-improvement effort. As teachers engage more in LS the quality of LS improves, possibly explaining why scholars such as Bocala (2015) and Lenski, Caskey and Anfara (2009) posit that teachers new to LS tend to focus primarily on learning the LS routine while more experienced practitioners concentrate on how they elicit student thinking.

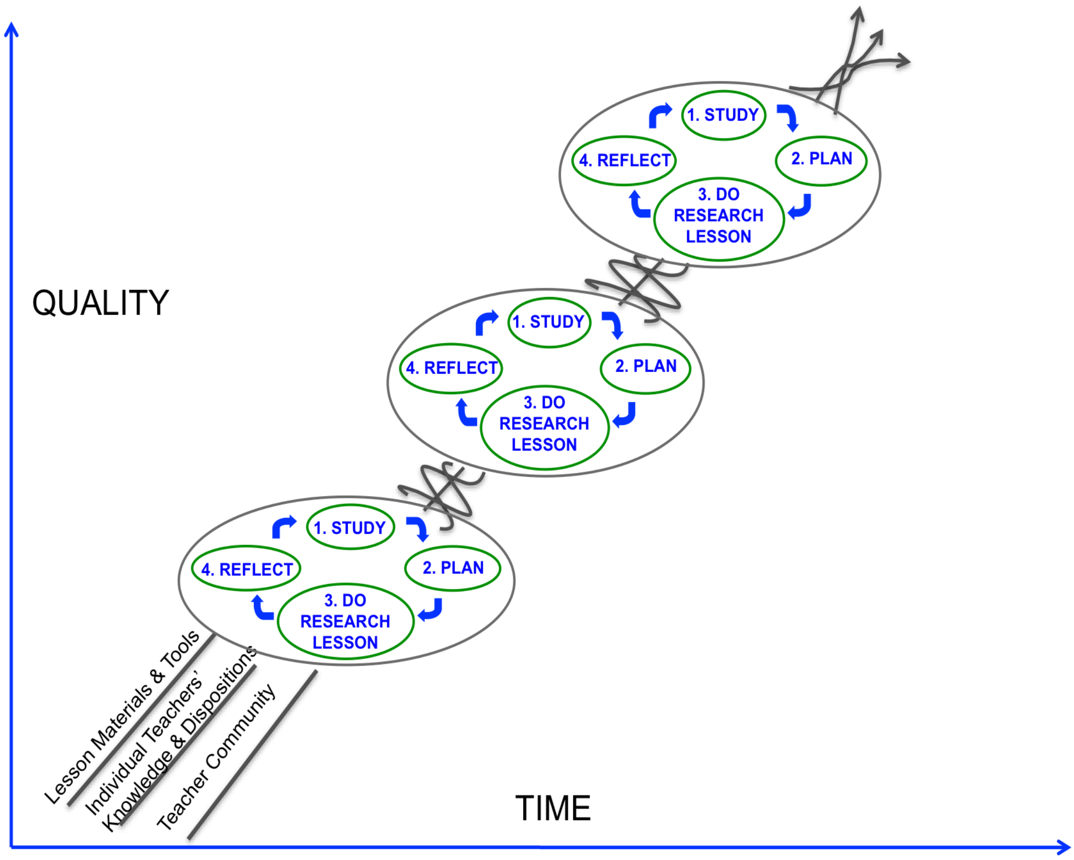


Figure 7. Impact of lesson study on lesson materials, teachers, and teacher community over time (Lewis, 2011, p.238)

In sum, LS establishes learning, the preeminent professional focus of teachers, as a visible and explicit focus of teachers' talk (Pedder, 2015) and facilitates "development of teacher knowledge, practice and professionalism" (Cajkler et al., 2015, p.194). Not only this, as argued by Stigler and Hiebert (2009), it enables participating teachers to see themselves as contributing to their own professional development.

Lesson Study as CPD

Lesson study seems to embody many of the key characteristics of effective, high-quality CPD described in the literature and discussed earlier, e.g., active teacher participation and sustained teacher learning connected to practice and built on strong collegial relationships (e.g, Darling-Hammond & Richardson, 2009; Desmione, 2009). In addition to those characteristics, scholars such as Bird and Little (1986) have also emphasised the benefits of peer observation to support teachers' learning, although this practice remains uncommon nationally (e.g., Clerkin, 2013) and internationally (e.g., Little, 2007). As reported, LS offers the potential to overcome the traditional norms of privacy and professional isolation prevalent in teaching, in a structured and non-threatening way, across the career continuum. In fact, Perry and Lewis (2009) argue that the potential of LS has already been demonstrated in Japan where LS is the most common form of CPD and where there is a strong record of student achievement, suggesting a relationship between student learning and LS, albeit not tested in controlled studies.

In addition to embodying many of the core features of effective CPD, LS seems to "fit" satisfactorily into Cosán (Figure 8), the recently published framework for teachers' learning (TC, 2016), as it fulfills many of the criteria described in the framework as essential to fostering a culture of "powerful professional learning" for teachers. The framework stresses the importance of teachers taking ownership of their CPD and being actively involved in their own learning for their own benefit and that of their students, thus enhancing professional autonomy. Additionally, the document emphasises that CPD should be accessible, relevant and flexible, with the potential to suit the context, culture and career-stage of all teachers. As reported, this is the type of CPD that can make an impact on teaching and learning - a process that allows for sustained commitment to quality teaching and learning and continued professional growth. LS, therefore, appears to "fit" this standard for teacher CPD.

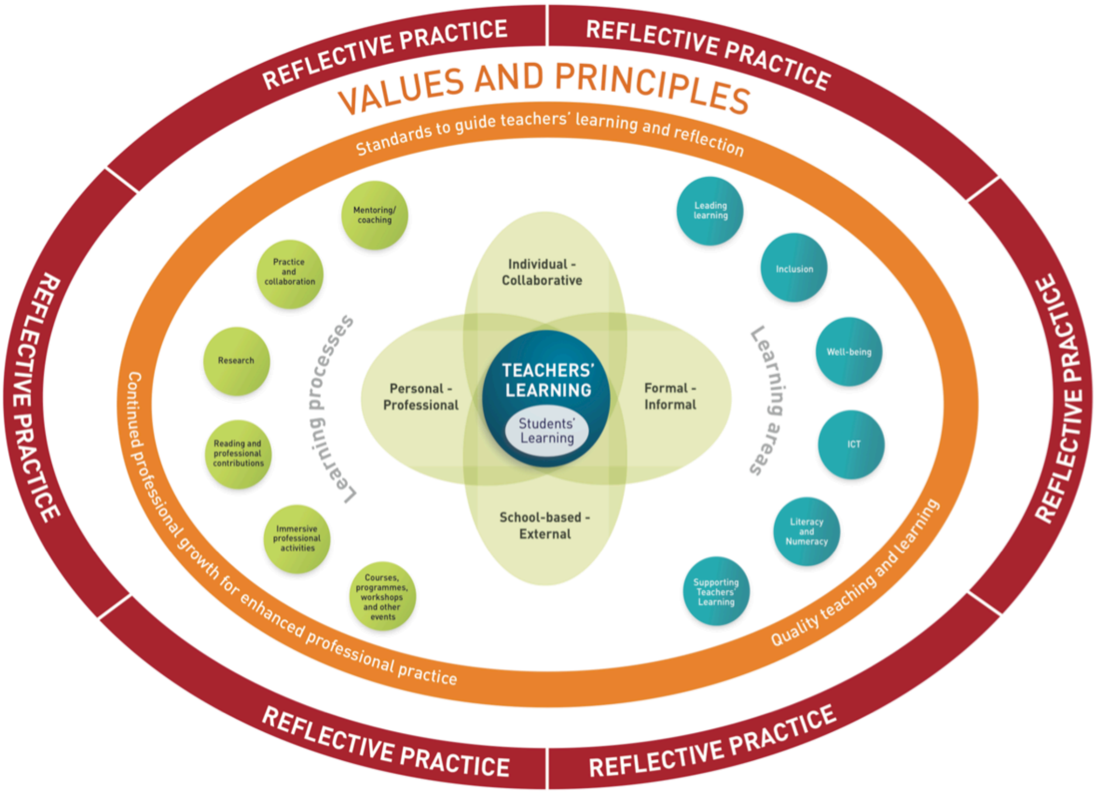


Figure 8. Key elements of Cosán (Teaching Council, 2016)

Scholars (e.g., Lewis & Hurd, 2011; Stigler & Hiebert, 2009) posit that the idea behind LS is relatively simple: If you want to improve instruction, what better place to do so than in the context of a real classroom lesson which teachers have prepared and reflected on collectively. After all, as Lewis et al. (2012) remark, "We would not expect surgeons or tennis pros to learn just from books or videos, and we should not expect teachers to learn without actual practice and feedback from colleagues" (p.373). Similar to Cochran-Smith and Lytle (1999), Lewis (2008) believes teachers learn from, and in, practice. Exploring ways of improving teaching, specifically through LS, Lewis et al. (2012) identified three supports for teacher learning *within practice* that would improve teaching: high-quality instructional resources, practice-based opportunities to learn and collegial learning that facilitates the development of shared knowledge and commitment among teachers. These supports are closely intertwined so teachers can draw on one another's knowledge, collaborate with colleagues and build professional norms and motivation. Well-designed LS would, they argue, already include all three components. Figure 9 from Lewis, Perry, Friedkin and Roth (2012) illustrates the status quo side-by-side with the ideal.

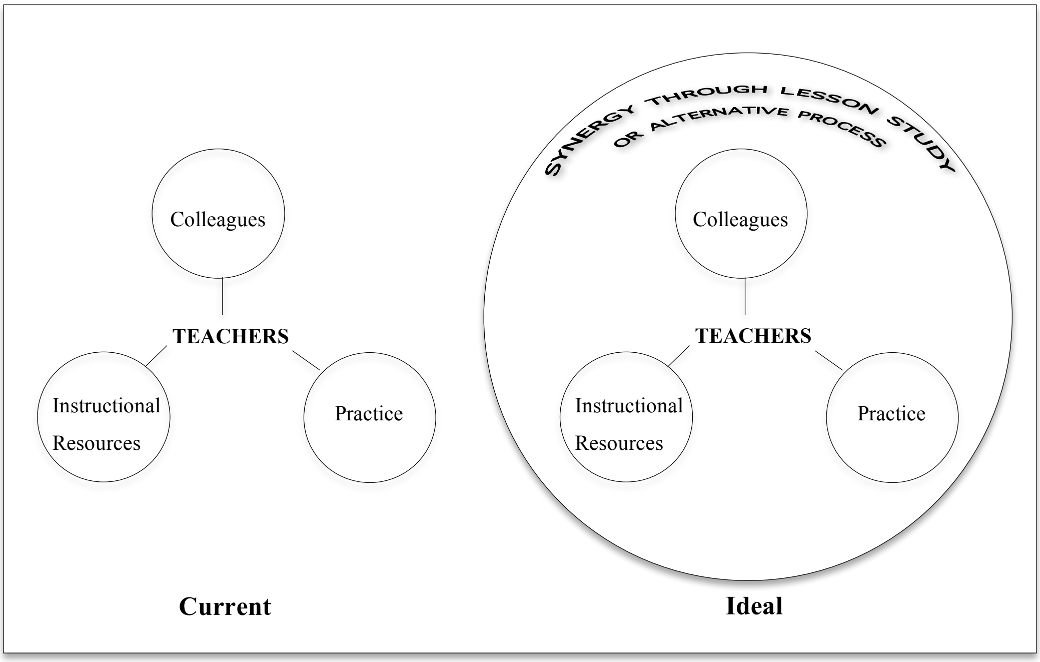


Figure 9. Resources for improving teaching: From disconnection to synergy (Lewis et al., 2012, p.369)

Currently, teachers' opportunities to learn from colleagues, practice and resources are limited, occurring in isolation from each other (left of Figure). This is in contrast to LS (right) where "the three elements are closely intertwined, so that teachers draw on one another's knowledge to bring to life high-quality instructional materials in the classroom" - moving from isolation to synergy (p.369).

Versus other CPD

In contrast to coaching and mentor-mentee contexts, peer-to-peer learning is at the core of the LS process (Gardner et al., 2012) where experienced teachers and newly qualified teachers (NQTs) learn from each other (Dudley, 2011). Through the LS process:

> Novice teachers who experience the lesson with experienced teachers are apprenticed into the profession through participation. This community aspect goes beyond the idea of collaboration among individual persons. It is about a way of further developing a profession (by engaging novices into serious academic activity and thus fostering identity building). (Krainer, 2011, p.10)

This vision of LS is in concert with Lave and Wenger's (1991) concept of communities of practice, where people "engage in a process of collective learning in a shared domain of human endeavor " and "learn how to do it better as they interact regularly" (Wenger, 2015, Retrieved from: http://wenger-trayner.com/introduction-to-communities-of-practice/).

Various LS scholars emphasise that LS is a "bottom-up" rather than "top-down" model of teacher development (Corcoran & Pepperell, 2011; Day & Sachs, 2004; Galanouli, 2010). Consequently, in contrast with the more traditional models of CPD, teachers are actively involved in the decisions regarding their own learning, rather than being passive recipients. Table 6 by Lewis and Hurd (2011) succinctly contrasts traditional CPD with LS and reveals the differences between both. Essentially, then, LS as a form of CPD is the joint enterprise of a group of teachers but can also involve the participation of schools and even governments, without them imposing or dictating from above (Galanouli, 2010).

Table 6

Comparing Views of Professional Development (Lewis & Hurd, 2011, p.7)

Traditional	Lesson Study
Begins with answer	Begins with question
Driven by outside 'expert'	Driven by participants
Communication flow: Trainer to teacher	Communication flow: Among teachers
Hierarchical relations between trainer and learners	Reciprocal relations among learners
Research informs practice	Practice is research

Finally, in a recent comprehensive review of literature pertaining to studies of professional development interventions related to mathematics in grades 1-8 in the U.S. undertaken by Gersten, Keys, Rolfes & Gonchar (2014), out of 643 reviewed, only five met the researchers' standards of rigour regarding effect, i.e. whether the PD approach caused improvements in students' mathematics proficiency. Of these, only two studies had statistically significant positive effects and one of these used LS as its approach. This study by Perry and Lewis (2011) is discussed later as it is particularly pertinent to this research.

Lesson Study Research

The growth of LS as a particularly promising model of PD for teachers has resulted in a concomitant rise in research interest in many European countries, adding to an already established research base in the Far East and the U.S. (Hogan, 2015). While much has been written in the extant literature regarding the positive benefits of engaging in LS, some researchers (Cajkler et al., 2015; Hogan, 2015) have recently highlighted the lack of high-quality, well-controlled LS studies and have emphasised that a stronger, wider, empirical evidence base is needed. Indeed, a review of the extant LS literature from 2000-2010 by Cheung and Wong (2014) could only identify nine studies in this field that they deemed sufficiently rigorous. Furthermore, in the UK, Xu and Pedder (2015) have recently criticised LS researchers for their lack of attention to important issues such as:

> The micropolitical dimensions of teachers' collaborative work in LS contexts, such as the building of trust, establishing norms of collegiality characterised by the sharing and exchange of resources and ideas, and the resolution of conflict. (p.49)

Additionally, it is argued that future research needs to explore the relationship between LS and students learning in a controlled fashion (Perry & Lewis, 2009), scrutinise the impact of LS on student outcomes, expand the descriptive knowledge base (Krainer, 2011; Saito, 2012), identify the scale of its impact and sustainability (Desforges, 2015), as well as explore how LS can be integrated into prevailing school cultures (Cajkler et al., 2015).

In sum, despite some concerns, the overwhelming message from the extant literature is that LS is a powerful dynamic approach to teachers' professional learning (Cajkler et al., 2014; Hogan, 2015; Lewis et al. 2012). In the Irish context, Corcoran (2011) identifies LS as "a form of continuing professional development which appears particularly suitable to mathematics teaching" (p.134) and suggests "its potential for use by practising teachers to develop and enhance the teaching and learning of mathematics at all levels appears worthy of further investigation" (2010, p.5).

Lesson Study and Mathematics

While mathematics is the most common area for LS research, and many have advocated using LS to improve mathematics teaching and learning, until the recent study by Perry and Lewis (2011) there was little evidence from controlled trials to support this argument. Their research involved LS teams of four-nine educators across 66 schools (87%

at elementary level) and took place over a five-month period. The LS group used research-based resource materials for teaching fractions, an area identified as being problematic for U.S. students. While teachers took turns to lead the group, experts were available to answer teachers' questions. Findings from this randomised, controlled trial of LS supported by mathematical resource kits indicated a positive and statistically significant impact on teachers' and students' fractions knowledge and students' mathematics proficiency. They also found that LS supported by mathematical resources increased teachers' mathematical knowledge for teaching, improved teachers' efficacy beliefs and student learning itself (Lewis, Perry, Friedkin & Roth, 2012). This would seem to affirm Corcoran's (2008) thesis that "to engage in LS is to learn mathematics for teaching" (p.xvii). Indeed, LS is increasingly being used in Initial Teacher Education (ITE) to explore the impact of different instructional approaches on the development of children's mathematical understanding (Leavy 2010; Leavy & Hourigan, 2014; Murata 2011). Additionally, scholars (e.g., NCCA, 2014b) argue that the focus on students' thinking, inherent in LS, helps teachers to promote the processes of mathematization (i.e. children interpreting and expressing their everyday experiences in mathematical form and analysing real problems in a mathematical way) by knowing how and when to increase the task challenge level, thus improving the quality of their teaching.

The import of teachers having good PCK when engaging in LS, especially without a KO, has already been highlighted. Refining Shulman's (1986; 1987) concept of PCK, Ball, Thames and Phelps (2008), developed a theory regarding the mathematics primary teachers need to know, which they called mathematical knowledge for teaching (MKT). They defined MKT as "…the mathematical knowledge needed to carry out the work of teaching mathematics" which includes absolutely "everything that teachers must do to support the learning of their students" (p.395), including planning, assessment, parent-teacher meetings, homework and much more. MKT, they argue, includes not only PCK but also subject matter knowledge (SMK). Figure 10 shows the relationship between PCK and SMK and illustrates how PCK consists of the subdomains of knowledge of content and students (KCS), knowledge of content and teaching (KCT) and additionally includes Shulman’s idea of curricular knowledge. This knowledge of the curriculum, teaching, content and students helps teachers make mathematical ideas understandable for students.

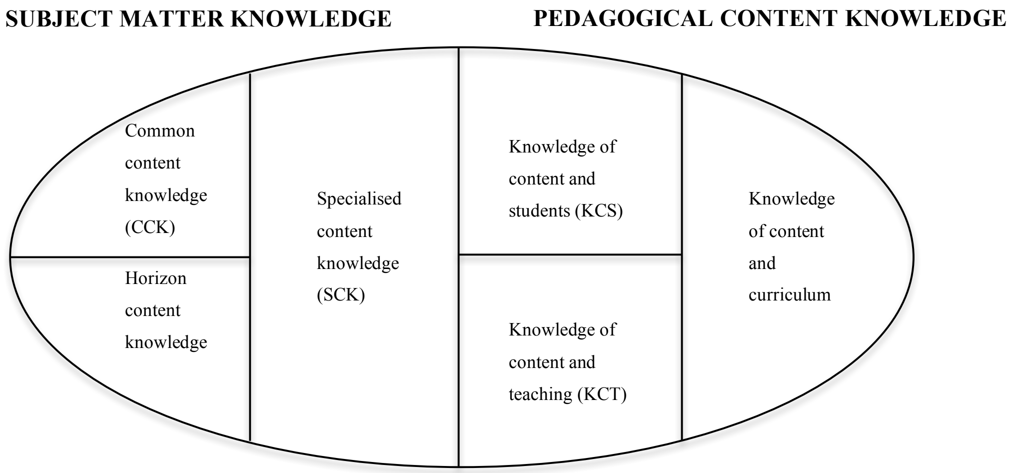


Figure 10. Mathematical knowledge for teaching (Ball et al., 2008)

Ball et al., also subdivide Shulman’s domain of subject matter knowledge into common content knowledge (CCK) and specialised content knowledge (SCK) as well as horizon content knowledge. Teachers and nonteachers alike need CCK, while SCK is unique to the work of teaching. Horizon content knowledge refers to “an awareness of how mathematical topics are related over the span of mathematics included in the curriculum” (Ball et al., 2008, p.403). In the Irish context, scholars (e.g., Delaney, 2008) highlight that MKT is an essential component of teaching which greatly influences student mathematical achievement, while Corcoran and Pepperell (2011) claim that LS enhances MKT.

The Knowledge Quartet

Several LS researchers have used the Knowledge Quartet (KQ) (Rowland, Huckstep & Thwaites 2005) in the Irish context as a framework or lens for discussing research lessons (Carroll, 2013; Corcoran, 2008). Scholars argue that reflective frameworks such as the KQ "facilitate critical inquiry and the use of a common language for talking about learning and teaching mathematics" (NCCA, 2014b, p.124). Corcoran (2008; 2010; 2011a), as part of her LS research, used the KQ as a framework to analyse the different dimensions of the mathematical content of research lessons, not only to identify teachers' mathematical knowledge but also "to build community by becoming part of the shared repertoire of ways of doing things" (2011a, p.256). Originally developed for use with pre-service and beginning teachers, Rowland, Turner, Thwaites and Huckstep (2009), now consider that the KQ can be used by all primary teachers as "a way of building up professional knowledge for mathematics teaching" since it allows teachers "to engage critically with actual lessons and teaching episodes – their own and others' – with the aim of learning from teaching-in-action" (p.xv). As reported, the KQ has been extensively "road tested" and, by using it, teachers can develop and improve their teaching of mathematics through structured observation and reflection as part of their ongoing development, even without expert support (Appendix D).

Background to Lesson Study in the Irish Context

In the Irish context, LS lacks an empirical base but is an emerging and growing field, although still in its infancy. LS was first mentioned in an Educational Studies Association of

Ireland (ESAI) article when Kelly and Sloane (2003) offered insights from LS as a possible solution to the problem of aligning educational research and practice. LS was again highlighted by Conway and Sloane (2005) in their report to the NCCA on international trends in post-primary mathematics education, and by Sloane (2005) in his presentation to the first national conference on research in mathematics education. However, Corcoran's (2006-2007) *Dublin Study* was the first research project to investigate and theorise LS here. As part of her doctoral research, and working with six pre-service teachers (PSTs), she investigated LS as a process of mathematics teacher development across the four dimensions of Rowland's Knowledge Quartet. Adopting a *situated* perspective and locating the research in a theory of social practice, the LS group was conceived as a community of practice dedicated to learning how to teach primary mathematics well (Corcoran, 2008). Findings suggest that engaging in LS enhanced teachers' mathematical content knowledge (MCK) and the development of a mathematics teacher identity.

Since Corcoran's study, a small number of research projects have explored LS as a vehicle to help pre-service teachers (PSTs) bridge the theory-practice gap and improve their mathematics instruction (e.g., Carroll, 2013; Leavy, 2010; Leavy & Hourigan, 2014) while others examined the use of LS as a model of CPD for teachers at primary (McLoone, 2011), post-primary (Brosnan, 2014; Ní Shúilleabháin, 2015) and tertiary level (Corcoran, Reilly, Breen, Dooley & Ryan, 2011; Leavy, Hourigan & Carroll, 2014 [ongoing]). While Leavy and her colleagues at Mary Immaculate College, Limerick have amassed data from seven years of LS research with 140 PSTs in 28 primary schools, the biggest Irish LS project thus far has been linked to the introduction of the Project Maths curriculum at second level (Brosnan, 2014). With support from the DES and NCCA, LS was promoted as a central approach in advancing the aims of *Project Maths* and a pilot project involving 250 mathematics teachers in 24 pilot schools was undertaken during the academic year 2008-

2009. Brosnan, the national co-ordinator of *Project Maths*, claims the project provided significant insights regarding the introduction of LS in an Irish context, including insights into dealing with teachers' apprehensions, misconceptions and resistance to LS. Similar to findings by Xu and Pedder (2015) in the UK, it appeared that teachers and KOs found LS a highly complex learning process which, to be successful, requires a shared professional culture where participants feel comfortable with each other and are supported by school leadership. While teachers revealed they had gained a great deal from their involvement in LS, they showed considerable initial resistance and found LS to be onerous and time consuming. Emergent findings indicated that research lesson plans, where produced, were inadequate, there was little live observation, post-lesson discussions were largely lacking and teachers did not have the required skill to reflect critically on lessons. The LS initiative was stopped after one year but was reintroduced for the 2012-2013 academic year and extended for the academic year 2014-2015 when 40 schools participated voluntarily. Teachers have presented their work and shared their experiences at National Conferences in 2013 and 2015, with resulting lessons available on the Project Maths website (A. Brosnan, personal communication, August 11, 2015). This documentation and dissemination of the LS process while viewed as an integral part of LS in Japan and the U.S, where public lessons at the end of the LS process are common, is relatively new to the Irish context. However, it is reflective of a growing movement towards teacher scholarship, which recognises the importance of teachers recording, sharing and passing on the knowledge they have gained (Cerbin, 2013; Dudley, 2012b).

Irish LS research

For the most part in Irish research to date, LS has been conceptualised as a community of practice (Corcoran, 2008; Corcoran & Pepperrell, 2011; Gardner et al., 2012; Ní

Shuilleabháin, 2015) with mathematics educators acting as KOs. Similar to international trends, LS research in Ireland has been predominantly into mathematics (Brosnan, 2014; Corcoran, 2008; Leavy & Hourigan, 2014; McLoone, 2011; Ní Shúilleabháin, 2015). Most LS research projects have been small scale and have taken place in primary school settings. Findings have been generally positive and, for the most part, researchers reveal that teachers enjoy participating in the LS process (Corcoran, 2008; Ní Shúilleabháin, 2015) and find it professionally rewarding and worthwhile (Hogan, 2015; Leavy et al., 2014), although there seems to be conflicting evidence of this from the initial stage of the Project Maths intervention (Brosnan, 2014). Most of the benefits of LS highlighted in other contexts have been emphasised by Irish researchers too. Similar to teachers in a UK study by Cajkler et al. (2015), Irish participants emphasised the importance of shared decision-making, collective agency and the importance of insights gained about students. There was general agreement that the research lessons had been successful although evidence supporting teacher learning is mainly qualitative (Ní Shúilleabháin, 2015). To recap, claims made by Irish LS researchers reveal that participating in the LS process enhanced participants' knowledge of students, pedagogical capacities, and provided deep insights into student thinking and learning (Hogan, 2015; Ní Shúilleabháin, 2015). It also helps teachers develop new perspectives on their teaching (Corcoran, 2011b; Ní Shúilleabháin, 2015) and increases teacher reflection (Ní Shúilleabháin, 2015).

To date, in the Irish context, however, only one recent small-scale cross-border research project has examined AfL as part of their LS research, and this at second level in one Belfast school. There, a team of three teachers worked with students from years eight to eleven, across the subjects, drama, history and sociology, from August 2010 through to September 2011 (Gardner et al., 2012). Findings confirm the effectiveness of LS as an approach to staff development which generates "collegial engagement on improving aspects of classroom practice" (p.7).

Some additional observations regarding LS in the Irish context are worth considering. First, little LS research has been carried out with, or by, experienced teachers. Second, to my knowledge, research utilising a combination of AfL, LS and mathematics has never been carried out here. Third, little or no LS research has been carried out in Ireland without using outside experts. Fourth, one of the main differences between LS research carried out in the Irish context to date and that carried out in Japan, and even in the US, is the lack of public lessons at the end of the LS process. Although speaking specifically about LS research in the UK, Hogan (2015), a well-known Irish educator, recently stated about LS research "Much done, and an abundance to do" (p.181). His thoughts are, I believe, even more applicable to the Irish context since LS is less established here.

Challenges to Lesson Study Implementation

Researchers have identified substantial challenges that need to be overcome if LS is to be successful outside of Japan. It has been argued that if LS is to be effective teachers need to clearly understand the LS process and its primary focus on student learning (Cajkler et al., 2014; Fernandez et al., 2003) and they need to have good content and pedagogical knowledge (Hart & Carriere, 2011; Murata, 2011; Yoshida, 2012). Others raise concerns that LS might be treated as a passing fad (Cajkler et al., 2014; Murata, 2011) or question the sustainability of LS outside of Japan (e.g., Murata, 2011; Perry & Lewis, 2009). Interestingly, Murata (2011) mentions the cost of engaging in LS as a potential obstacle to participating in LS but this contrasts with Dudley's (2015) opinion that LS is a relatively cheap form of CPD.

In many countries, teachers are not accustomed to engaging in research, studying the curriculum, collecting data and drawing conclusions, and so multiple LS studies have highlighted the difficulties teachers have in maintaining the *researcher lens* (Akiba & Wilkinson, 2016; Fernandez et al., 2003; Yoshida, 2012) discussed earlier. Additionally, many teachers do not have the resources or opportunities to develop the content and PCK necessary for facilitating LS by themselves (Akiba & Wilkinson, 2016; Fernandez, 2005; Hart & Carriere, 2011, Murata, 2011). Previous studies have, however, highlighted the benefits of involving KOs (Fernandez et al., 2003; Hart & Carriere, 2011; Perry & Lewis, 2009) and of using high-quality instructional resource materials (Akiba & Wilkinson, 2016; Lewis & Perry, 2014).

By its very nature LS throws up surprises and unintended outcomes so this can also be challenging (O'Shea, Teague, Jordan, Lang & Dudley, 2015). Nevertheless, a lack of time is the overriding reason identified in the literature as an obstacle to adopting LS (e.g., Cajkler et al., 2015; Fernandez et al., 2003; Gardner et al., 2012). LS is a slow process yielding small, incremental improvements (Stigler & Hiebert, 2009) and it requires time and sustained commitment to take root and for teachers to reap the full benefits to their thinking and practice (Akiba & Wilkinson, 2016; Cajkler et al., 2015; Perry & Lewis, 2009). It is also clear that there are organisational, logistical and workload challenges when implementing LS. It can be difficult to use LS on a sustained basis, especially for second-level schools where emphasis can be on performance (Cajkler et al., 2015). Organising staffing schedules and the school timetable to enable teachers to observe the same research lesson three times could be problematic, although these challenges are not insurmountable (Cajkler et al 2014; O'Shea et al., 2015). In sum, therefore, scholars (e.g., Cajkler et al., 2014) suggest that LS requires time, management support, and freedom from quick-fix expectations linked to what Hargreaves and Fullan (2012) term "business capital thinking" (p.526).

Considering cultural nuances

As research evidence regarding the positive benefits of LS increases, the question of how LS can be transferred to the practice of schools in different parts of the world becomes more important. Some have suggested that to be effective, each country needs to adapt the LS process to suit its own context so that it is rooted in the new culture as an intrinsic part of teachers' professional lives (Lewis et al., 2006; Murata, 2011; Watanabe, 2002). In this regard, Doig and Groves (2012) suggest that we need to *adapt* rather than *adopt* LS when working in another culture. However, Lewis and Hurd (2011) caution that the adoption of LS would still represent "a paradigm shift for most educators outside of Japan" (p.v). They acknowledge that, over time, educators will want to adapt LS to their local circumstances but argue the importance of understanding how LS differs from other practices, identifying the following common misconceptions:

- LS is lesson planning;
- LS means writing original lessons;
- LS produces a library of perfect lessons;
- The research lesson follows a rigid script and is a demonstration/expert lesson;
- LS is basic research and;
- It's all about the lesson plan. (Lewis and Hurd, 2011, pp.90-96).

These misconceptions, while a natural part of the efforts to understand LS and bring it to life around the world, need, they argue, to be worked through in order to build a robust version of LS outside Japan.

According to Murata (2011), teaching is a highly localised practice and so with regard to the LS process, modifications are essential and expected if this new PD approach is to be adopted and used effectively in various countries. Adaptation and implementation of LS can, however, be challenging (Wang-Iverson & Yoshida, 2005) and so, to maintain the integrity of the LS process, Murata (2011) identifies the following key characteristics of LS that must be maintained as a minimum, although she admits this list is not exhaustive and will be modified in time as more knowledge is gained through local contexts. LS must be: centered on teachers' interests, student focused, have a research lesson and be reflective and collaborative. According to Lewis (2006), an adaptation of LS has been successful if, on completion of the LS cycle, teachers can look back and say that they learned something about their teaching, their students, the subject matter and course materials. Additionally, they should be able to identify an increased capacity to work with colleagues, feel an improved sense of self-efficacy and an increased knowledge of subject matter in their teaching. Nevertheless, teachers can still benefit from participation in LS even from adapted or less time-rich versions than those in Japan (Cajkler et al., 2014), especially in mathematics.

The Primary School Mathematics Curriculum and its Interpretation

To reiterate, this study investigates the impact of AfL and LS on mathematics teaching and learning. Mathematics has always been an area of great interest to me, both as a student and as a teacher and has been prioritised by government in the *Literacy and Numeracy Strategy* (2011). However, first-hand experience as a primary teacher has shown me that mathematics is an area where students often lack confidence and/or understanding, and one that teachers, even at first-level, can find difficult to teach effectively. In the Irish context, the Primary School Mathematics Curriculum (PSMC) (GOI, 1999a) guides the teaching and learning of mathematics at primary level. Based on constructivist principles, it comprises five strands (Number, Algebra, Shape and Space, Measures and Data) that are considered interrelated and are subdivided into various strand units. The content of the PSMC is divided into four stages or levels (infants, 1st/2nd classes, 3rd/4th classes and 5th/6th classes), delineated by year and accompanied by separate Teacher Guidelines (GOI, 1999b).

The curriculum identifies six mathematical skills which children need to develop (Applying and Problem-Solving, Communicating and Expressing, Integrating and Connecting, Reasoning, Implementing, and Understanding and Recall) and encourages each child "to be confident and to communicate effectively through the medium of mathematics" (p.2). The PSMC promotes a wide range of teaching approaches and methodologies such as guided discussion and a hands-on approach. It recommends the use of collaborative and active learning in a mathematics-rich environment, along with the use of concrete learning resources and digital technology for all classes. Discussion and the development of mathematical language are highlighted as central to children's mathematics learning as is the development of estimation skills. Real-life problem solving is considered a key element of the curriculum, since it helps develop higher-order thinking skills, and highlights how mathematics can be used in everyday life. While the PSMC outlines what should be assessed and offers a range of assessment practices to elicit information regarding pupils' progress, it has been criticised by some scholars (e.g., Sugrue, 2004; 2011) regarding the apparent disconnect between curriculum and assessment, with assessment ostensibly treated as an "add-on" activity (Sugrue, 2004; 2011). Other criticisms of the PSMC (1999) include the failure to mention the need to develop teachers' mathematical knowledge for teaching to ensure successful curriculum implementation (Delaney, 2008) and the argument that the use of strands and strand units "offers a narrow and consequently limiting list of teaching objectives" (Corcoran, 2008, p.17). Indeed, Corcoran (2008) also suggests that the PSMC "is open to contradictory interpretations" (p.17), as well as highlighting two significant lacunae concerning substantive and syntactic mathematics. In a review of the PSMC by the NCCA (2005), strengths identified included children's enjoyment of mathematics, being child-centred, involving practical work and children's success in specific content areas, e.g., number. However, the curriculum fails to reference international research into mathematics education (Corcoran, 2008), and the range of curriculum supports is limited, as is the provision of exemplars, which lags behind other countries (Burke, 2014). Mindful of these criticisms and the fact that the PSMC is now one of the oldest Mathematics Curricula in Europe (Burke, 2014), depending on how it is used and interpreted, the PSMC can still provide a useful framework for planning the teaching and learning of mathematics at primary level.

Adopting the constructivist model of teaching mathematics inherent in the PSMC allows scope for us (teachers) to make use of learning in mathematics that is situated, collaborative and self-regulated, enabling our students to develop a mathematical disposition and to become mathematically proficient (De Corte, 2004). For example, in the Teachers Guidelines (1999) it states: "all number work should be based as much as possible on the children's own experiences and real-life examples" (p.9), thus highlighting the need, as suggested by research (e.g., Cai & Lester, 2010), for children to relate what they are learning in mathematics lessons to their everyday lives and to the real world. Additionally, the view of mathematics as a problem solving activity permeates the PSMC. Therefore, the use of real-life problems as a context to learn mathematics, deepen children's understanding and develop higher-order mathematical skills is already encouraged, although such problems currently need to be sourced or developed by teachers since they are lacking in curriculum resources and in most Irish mathematics textbooks. Furthermore, in the PSMC students are encouraged to explain and express "mathematical ideas, processes and results in oral and written form" (p.12) thus promoting maths talk which is also part of good mathematics pedagogy. Taking another example, the PSMC (1999) states that children should "use their knowledge of one area of mathematics to explore another" (p.15) which is similar to "connecting" considered part of the process of mathematization (NCCA, 2014a), a process that is key to the development of mathematical proficiency and mathematical reasoning. In sum, we (teachers) can choose to use the PSMC as a starting point, and can supplement this with knowledge of national and international research and best practice in mathematics and with collaborative and reflective practice with colleagues within our own school setting to become the best teachers of mathematics we can be.

The AfL/LS Nexus

An analysis of the extant literature on AfL and LS reveals significant parallels between both, such as their focus on learning and the learner, their emphasis on collaborative practice and the need for teachers to anticipate and adapt to students' responses. Both are premised on sociocultural/social constructivist theory and are relatively low-cost, low-tech and imminently scalable, albeit reasonably complex and time-consuming to implement correctly, especially in the initial stages. Experts from both traditions have used the analogy of the *black box* when explaining why AfL/LS might improve student learning, whilst the necessity and importance of teachers possessing good PCK is central to the success of both. Teachers would not only require good content knowledge and pedagogical skills but would also need to develop adaptive expertise if their teaching is to meet the minute-to-minute learning needs of their students inherent in AfL practices or to anticipate their students' responses in LS. Developing adaptive expertise would enhance teachers' metacognition and self-regulation and guide the lifelong learning needed to help their students achieve (Hammerness, Darling-Hammond & Bransford, 2005). Hattie (2012) emphasises that teachers as adaptive experts would know and model multiple ways of teaching and learning and could adapt strategies and resources to help students attain worthwhile learning intentions. They could detect if students are learning or not, where they are on the continuum of learning and where to go next, while creating a classroom environment conducive to attaining learning goals. In a similar vein, Lysaght (2012) highlights the similarities in the skills needed to successfully implement AfL practices with those linked to adaptive experts. Finally, proponents of both AfL (Heritage, 2013; Wiliam, 2011a) and LS (Conway, 2013; Murata, 2013; Shulman, 2005a) have suggested that their respective fields warrant investigation as potential signature pedagogies for the teaching profession. According to Shulman (2005):

> A major challenge for the education of teachers and the professional development of veteran teachers for this next generation will be to recognize that we desperately need a suite of signature pedagogies that are routine, that teach people to think like, act like, and be like an educator. (p.15)

While acknowledging that there are differences between both, it is nevertheless possible to identify an AfL/LS nexus that contains many elements common to both (Figure 11). The central column of Figure 11 itemises the features common to both AfL and LS, ergo the AfL/LS nexus. Meanwhile the left and right columns summarise the key elements of AfL and LS respectively.

Figure 11. AfL/LS nexus

Conclusion

This section of the literature review began by briefly exploring CPD in general and in the Irish context. Then, the history, process, theoretical underpinnings and potential impact of LS were investigated. Next, the benefits of LS as a model of CPD were highlighted, followed by discussion of the concept of LS as research and a review of the literature regarding LS and mathematics, including use of the KQ. Subsequently, there was an exploration of LS in the Irish context, followed by consideration of possible challenges to LS implementation. The section concluded with a discussion regarding the possibility of an AfL/LS nexus. Chapter Three follows with a description of the research design, instrumentation and methodology.

CHAPTER THREE: RESEARCH DESIGN, INSTRUMENTATION AND METHODOLOGY

Introduction

In the preceding chapter the salient literature pertaining to assessment for learning (AfL) and lesson study (LS) was reviewed, including their theoretical and historical underpinnings, their value and impact on teaching and learning and the status quo regarding both in the Irish context. Particular attention was paid to how AfL and LS can be made real in classrooms by exploring previous research and practice. Additionally, the review highlighted significant parallels between AfL and LS and suggested the possibility of an AfL/LS nexus. Furthermore, it emphasised there is little or no research combining AfL, LS and mathematics and suggested the need for further research into both AfL and LS in the Irish context. This study emerged as a response to the literature. It investigated the effects of using AfL practices on students' learning of mathematics in fourth class, in an all-girls primary school, and explored the potential of peer-to-peer learning (LS) to impact teachers' knowledge and skills using AfL principles, strategies and techniques.

This chapter discusses and justifies the research design and details the instruments and methodology used to test each of three research hypotheses. As argued (Pallant, 2013), good research depends on meticulous planning and the careful execution of the study, which in turn has implications for data quality and analysis. Although many models have been posited, no single blueprint exists for planning research (Robson, 2011). Instead, research design is based on the notion of *fitness for purpose*. The purposes of the research determine its methodology and design (Cohen, Mannion & Morrison, 2010). This thesis utilises Maxwell's (2013) interactive model of research design (Figure 12), with minor adaptations in structure and content, and his model is also used to provide a structure for this chapter.

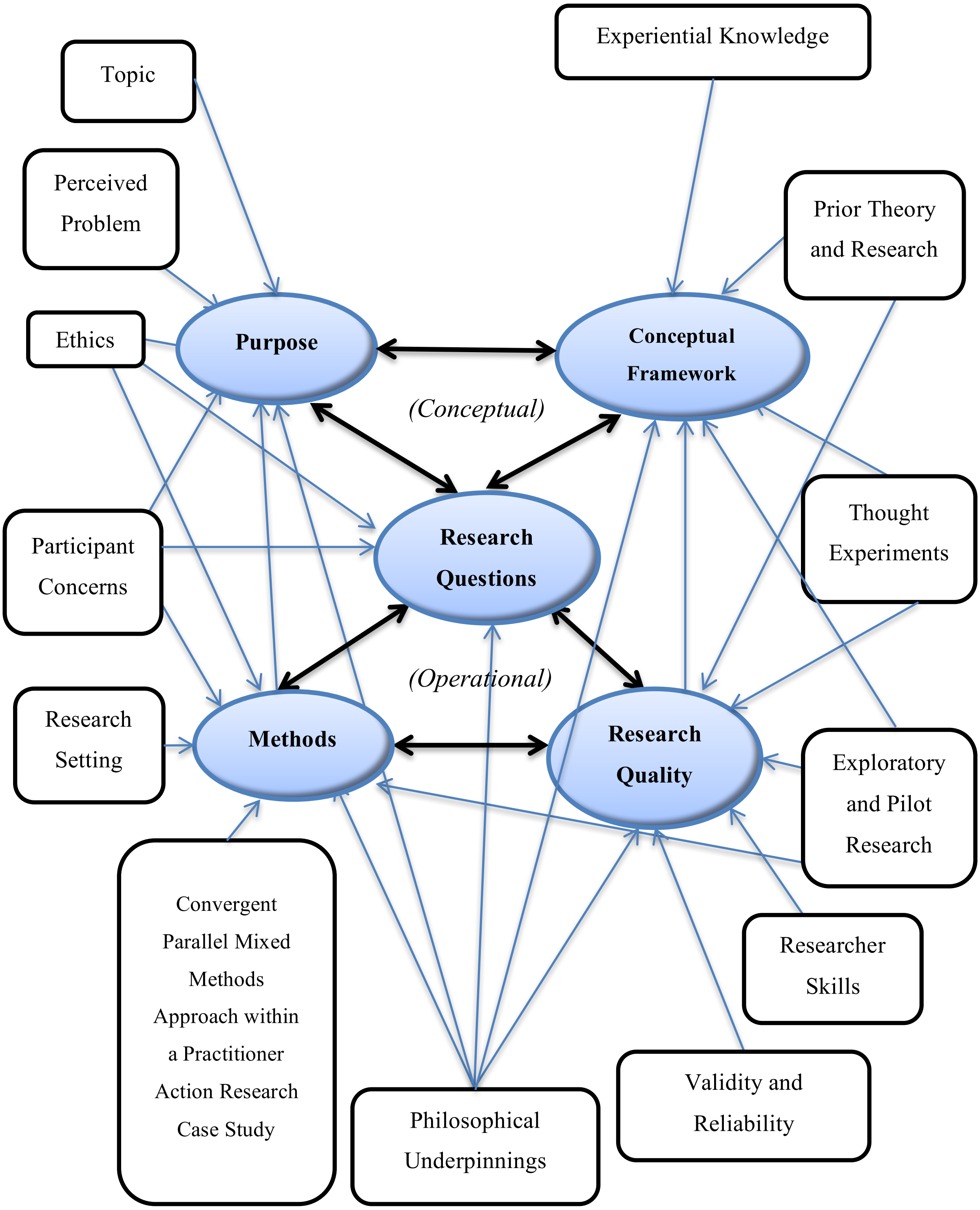


Figure 12. Key elements of this project's research design (Adapted from Maxwell, 2013, p.6)

Maxwell's model consists of five main components (coloured blue), each of which addresses specific concerns: purpose, research questions, conceptual framework, methods and research quality. As detailed in Figure 12, these components form an integrated and interactive whole, with the arrows highlighting the relationships between them, although these relationships are not rigid. The peripheral components of Figure 12 present some of the key factors in the environment that may influence the design and conduct of the study. The research questions, presented as hypotheses in this study, are in the centre of the figure since they are at the heart of the design and, according to Maxwell (2013), directly link, influence, and respond to every other part of the project, especially the study's methods and conceptual framework. The upper triangle formed by the arrows deals with the more conceptual elements of the study, while the lower triangle is more operational. Maxwell (2013) posits that other factors, for example perceived problem, ethics and the research setting, etc., are not part of the research design per se, although they do influence it and are taken into account.

The next section re-presents the purpose or goals of the study and includes the research problem and hypotheses. Then, the conceptual framework is presented, thus completing the discussion regarding the conceptual triangle at the top of Figure 12. After this, the philosophical worldview underpinning the research is considered, before exploring the project's strategies of inquiry, including justification for a case study approach. This leads on to an in-depth exploration of the research methods, which includes details regarding the case, site and participant selection; the instruments and methods used to collect data; and the data handling. The chapter concludes by detailing how issues of quality control and ethical considerations were addressed.

Research Purpose, Problem and Hypotheses

As stated earlier, the problem this study addressed had multiple dimensions. First, data from national and international reports suggested that Irish students were underperforming in mathematics (e.g., DES, 2010b). Second, there appeared to be a lack of assessment literacy among teachers and reports, cited in the literature review, highlighting insufficient use of AfL in their teaching (e.g., DES, 2010a). Third, although the Department of Education and Skills (2011a) acknowledged the need to enable teachers to improve their practice of AfL, no relevant CPD has been provided (INTO, 2015). In addition, with regard to AfL, there has been little research in the Irish context into AfL practices in schools, especially with regard to numeracy. Furthermore, issues have been raised about the validity of some of the effect sizes quoted when promoting AfL, thus raising questions about its warrant and suggesting the need for further research (Bennett, 2011; Dunn & Mulvenon, 2009).

To reiterate, the purpose of this study was to investigate the impact of AfL practices on the teaching and learning of mathematics in fourth class in a vertical, all-girls primary school in a large provincial town in the Republic of Ireland, over a nine-month period. To fulfill this aim, the three participating teachers formed a lesson study (LS) group, which met twenty four times during the intervention. Using LS as a model of peer-to-peer professional learning, the teachers learned about AfL principles, strategies and techniques, before implementing them in their mathematics classes. To assess whether this project achieved its goals, specific research questions were posited and these are now presented as three research hypotheses:

H1. A nine-month school-based intervention employing assessment for learning principles, strategies and techniques will improve the standardised mathematics results of participating students in comparison to a similar cohort not involved in the intervention.

H2. The use of AfL strategies and techniques, and the adoption of AfL principles, will enhance children's mathematical confidence, and improve their engagement with, and attitudes to, mathematics.

H3. Peer-to-peer professional learning is a feasible, worthwhile, efficient and effective model of CPD in AfL and will improve teachers' skills, knowledge and use of AfL and their attitudes and beliefs towards AfL as a form of assessment.

Conceptual Framework

The role, significance and use of a conceptual framework in doctoral research is now widely accepted (Berman, 2013; Leshem & Trafford, 2007) since it has heuristic value (Teddlie & Tashakkori, 2009) and enhances conceptual thinking. Various authors present the notion of conceptual frameworks differently, but it is generally agreed that research is of better quality when they are made explicit (Miles, Huberman & Saldana, 2014). As argued by Maxwell (2013), conceptual frameworks are a key part of the research design, and help clarify "the system of concepts, assumptions, expectations, beliefs, and theories that supports and informs your research" (p.39), thus providing a lens for making sense of things. Figure 12, at the beginning of this chapter, illustrates how conceptual frameworks are closely linked with the research purpose and research questions, completing a conceptual triangle at the top of Maxwell's (2013) model. According to Ravitch and Riggan (2012), literature reviews and theoretical frameworks should be viewed as aspects of conceptual frameworks. They posit that conceptual frameworks are comprised of three main elements: personal interests, topical research and theoretical frameworks, all with different functions within the conceptual framework itself. They explain that: "topical research describes the *what* of the study, while theoretical frameworks clarify the *why* and the *how*" (p.13). In this study, I adopt a similar stance and utilise the conceptual framework in the broadest sense, thus viewing it as inclusive of the theoretical one. Reflecting on the findings from an initial trawl of pertinent literature, and drawing on my own professional experience and the research of others, especially Lysaght's doctoral study (2009), I developed a conceptual framework for the study (Figure 13) that evolved over time.

As advocated by Miles et al., (2014), this conceptual framework details graphically the main things to be studied in the research project: the key factors and constructs and their interrelatedness. The conceptual framework reads from the bottom up and is underpinned by the pragmatic paradigm, which is discussed in the following section. Starting at the bottom of the Figure, the problems this study sought to address are highlighted, as are the main topics and the proposed solution. The various factors that influenced the research are detailed on the left side of the conceptual framework, all of which impacted on the model of CPD, in this case LS, which is depicted on the right. In keeping with Berman's (2013) view that conceptual frameworks help the researcher to anticipate various conceptual and practical outcomes the research might bring, I found the process of developing it invaluable when trying to resolve the conundrum of how to combine research into AfL, LS and mathematics, and when trying to predict outcomes for teachers, students and classroom practice following the study, all of which are detailed at the top. Developing this framework challenged me to explore my thinking and beliefs about AfL, CPD, LS and mathematics, and to consider factors that might impact the research such as relevant theories, previous research, experiential knowledge, participants and context.

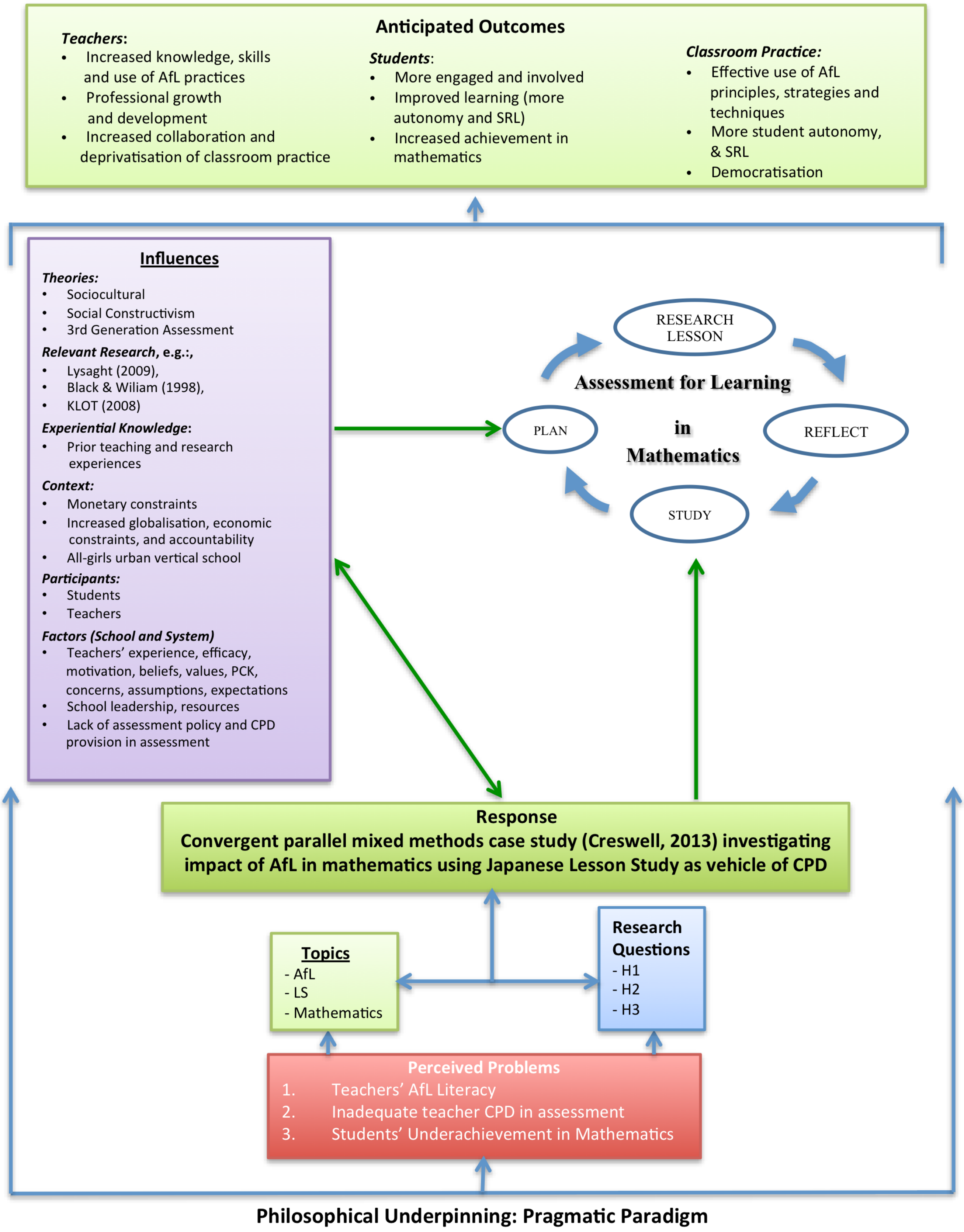


Figure 13. Conceptual framework for this study

This enabled me to make my ideas explicit, thus improving my understanding of the research problem and the various issues involved. In particular, using a conceptual framework helped me formulate the three research questions, presented as hypotheses, and plan the research methods and design.

Philosophical Underpinnings

Brannen (2005) argues that methodological choice is not made in a "philosophical void" but instead is "chiefly driven by philosophical assumptions – ontological and epistemological- which frame the research or the researcher's frame of reference" (p.7). Although these philosophical ideas often remain implicit, they still influence the practice of research. Consequently, Creswell (2009) suggests that researchers "make explicit the larger philosophical ideas they espouse" (p.5), thus clarifying their choice in research design. Scholars writing on these issues often refer to paradigms (Teddlie & Tashakkori, 2009), or worldviews (Creswell, 2009), when attempting to explain the beliefs and ideas underpinning their research. Conceptualisations of the various paradigms continue to evolve but, according to Cameron (2011), inconsistency is evident across the literature as to how paradigms are labelled, dichotomised, and polarised.

Focusing on the current study, alternative paradigm suggestions for mixed methods include pragmatism, critical realism and the transformative paradigm (Teddlie & Tashakkori, 2010). Notwithstanding, various scholars have proposed that pragmatism is the best paradigm for justifying the use of mixed methods research (Feilzer, 2010; Teddlie & Tashakkori, 2010), with some scholars (e.g., Denscombe, 2008) viewing it as its "philosophical partner". As argued (Feilzer, 2010), pragmatism offers "an alternative worldview to those of positivism/postpositivism and constructivism" since it "accepts philosophically, that there are singular and multiple realities that are open to empirical enquiry and orients itself towards solving practical problems in the real world" (p.8). Teddlie and Tashakkori (2009), however, suggest that a "…continua of philosophical orientations rather than dichotomous distinctions, more accurately represent the positions of most investigators" (p.94). Placing pragmatism in the middle of their proposed continuum, they consider it "particularly appealing when engaging in mixed methods research" (p.103). With that in mind, this thesis adopts a pragmatic stance and construes the pragmatic paradigm as:

> A set of interlocking philosophical assumptions and stances about knowledge, our social world, our ability to know that world, and our reasons for knowing it - assumptions that collectively warrant certain methods, certain knowledge claims, and certain actions on those claims. (Greene & Caracelli, 1997, p.6)

Choosing the pragmatic paradigm allowed me freedom to use the methods, procedures and techniques that best met my needs and to utilise both qualitative and quantitative strategies for data collection and analysis. Pragmatism provides the best fit epistemologically, ontologically, axiologically and methodologically. In short, it was the most suitable underlying philosophy for this study because:

> It offers an immediate and useful middle position philosophically and methodologically; it offers a practical and outcome-orientated method of inquiry that is based on action and leads, iteratively, to further action and the elimination of doubt; and it offers a method for selecting methodological mixes that can help researchers better answer many of their research questions. (Johnson & Onwuegbuzie, 2004, p.7)

Strategies of Inquiry

As argued, strategies of inquiry (Creswell, 2009), or research methodologies (Mertens, 1998), are types of quantitative, qualitative and mixed methods models or designs that provide specific guidance for procedures in a research design. A mixed methods strategy was considered the most appropriate for this study since the combined use of quantitative and qualitative data aids triangulation, enables better understanding of the research problems/hypotheses and enhances the study's findings. The selected strategy of inquiry is a convergent parallel mixed methods design (Figure 14) and, as advocated by various scholars (e.g., Creswell, 2013; Creswell & Plano Clarke, 2011), the quantitative and qualitative strands are given equal priority, but treated independently, before being mixed during analysis and interpretation. Researchers (e.g., Creswell, 2009; Robson, 2011) consider case studies a particular strategy of inquiry and justification for its use in this study is discussed forthwith.

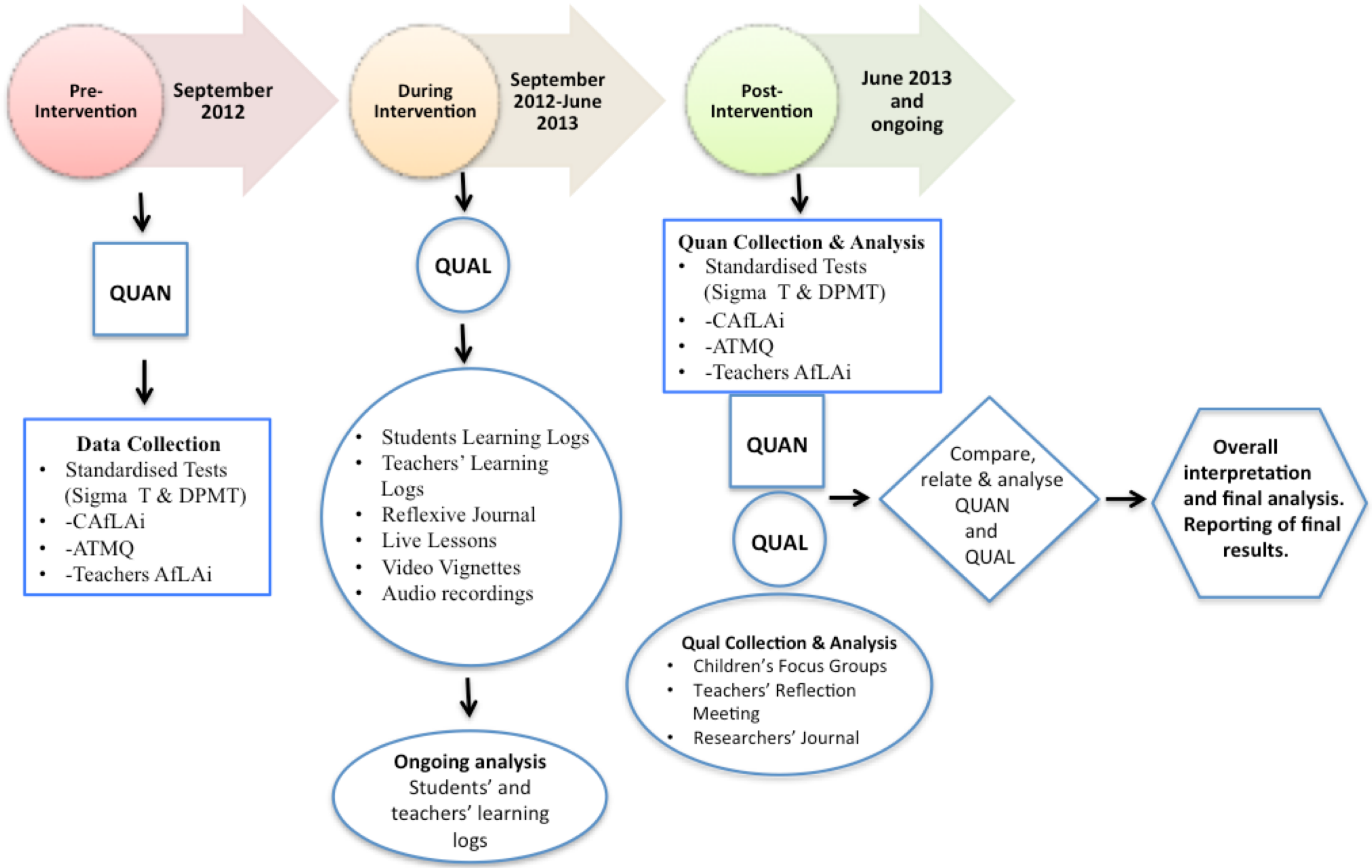


Figure 14. This study's convergent parallel mixed methods design

Justification for Case Study Strategy

An explorative case study approach was considered appropriate for this research since, as defined by Robson (2011), it is "a strategy for doing research which involves an empirical investigation of a particular contemporary phenomenon within its real life context using multiple sources of evidence" (p.136). While some authors (e.g., Stake, 2005) consider that case studies are primarily qualitative, it is now widely accepted that they can make use of both quantitative and quantitative data collection methods (Gerring, 2007; Yin, 2014). In essence, according to Yin (2014), case study research "comprises an all-encompassing method" which "can embrace different epistemological orientations" (p.17) and which covers the research design, data collection techniques, and particular approaches to data analysis. A case study approach allows the researcher to explore a process, programme, event or activity, in some depth (Creswell, 2009), which may lead to discoveries that might not have come to light through more superficial research (Denscombe, 2010). Pring (2004) argues that case study research emphasises "the uniqueness of events or actions, arising from their being shaped by the meanings of those who are participants in the situation" (p.40). Using a case study strategy is therefore ideally suited to the needs of individual or small-scale researchers (Bell, 2010; Blaxter, Hughes & Tight, 2006) and is an appropriate approach for this research.

It is important to acknowledge that case studies have both strengths and limitations (Yin, 2014). Many of the advantages have been signalled above. However, a major concern about case studies is that generalisation is not always possible (Bell, 2010). Nonetheless, Bassey (1981) states that "the relatability of a case study is more important that its generalisability" (p.85). Additionally, as recommended by Gall,

Gall and Borg (2003), by providing "thick description" readers should be able to more easily compare a study such as this one with their own situations. Overall, the decision to use a case study approach in the current research was a strategic one since it allowed for the "development of detailed, intensive knowledge of a single case" (Robson, 2002, p.89) in its natural setting.

Research Methods

As stated in Chapter One, this research utilises a convergent parallel mixed methods approach within a practitioner action research case study strategy. The research design has many similarities with that adopted by Lysaght (2009) and makes use of both quasi-experimental and qualitative strategies. The following sections explore this study's research methods beginning with site and participant selection. Then, starting with the quantitative data, details of the research instruments, data collection techniques and data handling are investigated, followed by a similar discussion pertaining to the qualitative data set. Issues of validity and reliability are dealt with under quality assurance later in the chapter, along with any ethical concerns.

Site Selection

This research project was carried out in Scoil na nAingeal (pseudonym), the school in which I work, between September 2012 and June 2013. Undertaking research in one's workplace has advantages and disadvantages. Some scholars (e.g., Bogdan & Biklen, 2003) advise the novice researcher to avoid researching in one's place of work since it may be difficult to distance oneself from personal concerns. Others, such as Blaxter et al. (2006), highlight further disadvantages such as difficulties in maintaining anonymity, having one's conclusions ignored or rejected, thinking that one knows the answers already and overlooking "the significance of things that seem obvious" (p.47). However, Blaxter et al. (2006) also emphasise things such as insider knowledge, ease of access and knowing some of the answers already, as being advantageous. On balance, given the nature of the project, the unlimited access, and full-time work commitments, my own school was deemed the most appropriate as a research site.

Research Site

Scoil na nAingeal is a vertical, urban, single-sex school with an all-female staff and an enrolment of 438 girls at the time of the research. Although not falling within the Delivering Equality of Opportunity in Schools (DEIS)[5] scheme, the brothers of our students attending the nearby boys' senior school have DEIS status. A recent Whole School Evaluation (WSE) Report in 2012 described the pupil cohort in Scoil na nAingeal as "mixed, in terms of socio-economic status and ethnicity" and commented that "a significant number of pupils speak English as a second language" (DES, 2012b, p.1). The following section provides specific details about the students and teachers from Scoil na nAingeal who participated in the intervention.

---

[5] Delivering Equality of Opportunity in Schools (DEIS) the Action Plan for Educational Inclusion, was launched in May 2005 and remains the Department of Education and Skills policy instrument to address educational disadvantage. The action plan focuses on addressing and prioritising the educational needs of children and young people from disadvantaged communities, from pre-school through second-level education (3 to 18 years). - See more at: http://www.education.ie/en/Schools-Colleges/Services/DEIS-Delivering-Equality-of-Opportunity-in-Schools-/#sthash.7w7DnHPP.dpuf

Research Participants

The research participants included all fifty-one students enrolled in fourth class for the academic year 2012-2013, the average age of whom was ten years in September 2012. The comparison group consisted of all students enrolled in fourth class for the previous academic year (2011-2012), 52 girls, the average age of whom was also ten in September 2011. Both groups were pre-formed, intact groups and so random assignment was not considered appropriate. While the standardised test results of the comparison group were utilised as part of the research analysis, these students were not actively involved in the intervention *per se* and consequently did not receive the treatment. Nevertheless, the belief is that following the project, all teachers and students in our school will engage in AfL practices as part of normal teaching and learning, thus the comparison group should ultimately benefit.

Three teachers participated in this project, two other teachers and myself as insider-researcher. According to Dwyer and Buckle (2009), being an insider-researcher involves conducting research with populations of which you are a member and share experiences, in my case undertaking research with colleagues and students in the school where I teach. This has both advantages and disadvantages with scholars such as Gray (2014) and Robson (2011) suggesting it may be beneficial since it can provide ease of access, and in-depth knowledge and experience of the research context and participants, thereby offering valuable insights that might not otherwise be possible. In contrast, Kanuha (2000) cautions that although being an insider-researcher might enhance understanding of the population you are studying, questions regarding objectivity, bias, reflexivity and the authenticity of the research are raised because you are too close, or know too much, or are too similar to those being studied. Additionally, as highlighted by Brannick and Coughlan (2007), you can also struggle with “loyalty tugs” (p.70) since you do not want to paint your school in a bad light. Consequently, adopting the insider position presented the challenge of balancing my role as teacher participant with that of researcher, something I struggled with particularly in the initial stages of the reseach process. Nevertheless, I think that being an insider improved rather than impeded my research and so this approach was justified as it automatically gave me access to data and provided a level of openness and trust that may have been lacking otherwise. As a result, I believe that my colleagues and the students were more willing to share their experiences because they felt I was one of them and understood where they were coming from. I realise it could be argued that this might also impede the research process but I was careful not to allow my personal experiences and beliefs to cloud my perceptions of what the participants were saying or doing. Furthermore, at all times I tried to remain aware of the potential impact my insider status might have on interviews or other data collection instruments such as the learning logs. Since I was deeply immersed and invested in the research process I also endeavoured, as Asselin (2003) suggests, to gather data with my “eyes open”, guarding against any assumptions I might have.

Scholars (e.g., Burns, Fenwick, Schmied & Sheehan, 2012) argue that insider-researchers can experience role ambiguity to varying degrees, especially during data collection, and so I maintained an ongoing reflexive dialogue throughout the research process, utilised respondent validation as recommended by Maxwell (2013), remained vigilant about my subjectivity and reported all evidence fairly. I made use of techniques such as writing in my researcher’s journal and using diagrams or concept maps as suggested by Buckley and Waring (2013) to aid reflexivity. Such retrospection became an integral part of the research process and helped me understand and undertake my dual role. In particular, using the researcher journal enabled me to reflect on issues and experiences that happened on the research journey. This helped lessen researcher bias and increase the *trustworthiness* of the research process, helping ensure that the voice of the other participants was heard in the narratives that follow and that this researcher shares. Additionally, I used triangulation of data, sources and methods to reduce responder and researcher bias and I was fully honest and authentic in my interactions with participants. Particular ethical issues can be associated with being an insider-researcher and these are discussed in further detail later in this chapter under the heading 'Ethical Considerations'. As part of these considerations, I have given pseudonyms to the other two teachers and the children in an attempt to preserve their anonymity. However, my own contribution to the data should be clearly evident.

I had been allocated fourth class for that year and so asked the other fourth class teacher and the SEN teacher working at this level to participate in the project. The teachers each had a minimum of twenty years teaching experience at primary level, had taught most class levels and also said that they felt confident teaching mathematics and enjoyed doing so. It should be noted that in Scoil na nAingeal, classes from 4^{th} to 6^{th} are streamed for mathematics. As recommended by Cohen et al., (2010), in order to exercise control over extraneous variables, the same teacher was assigned to the same ability group for both the comparison and intervention groups, while mathematics was taught from 11am for one hour, Monday to Thursday, for both academic years. The following section focuses on the quantitative data, specifically how it was collected and handled.

Quantitative Data Collection and Handling

Beginning with the children's quantitative data, and in keeping with this study's quasi-experimental design, pre- and post-test data were collected from the comparison and intervention groups to test research hypotheses one and two. The following instruments were utilised to collect this quantitative data:

1. The Standardised Irish Graded Mathematics Attainment Tests (SIGMA-T) and the Drumcondra Primary Mathematics Test Revised (DPMT-R), both norm referenced standardised tests used in the Republic of Ireland to ascertain students' mathematics achievement. The test's authors, Wall and Burke (2007), state that the SIGMA-T "is a highly reliable test" (p.43) having obtained good scores on three types of reliability tests (test-retest evidence, parallel form evidence and internal consistency evidence). The authors also provide content-related and criterion-related evidence to support the validity of the SIGMA-T. The reliability measures provided for the DPMT-R include standard error of measurement and internal consistency evidence (Educational Research Centre, 2007), while evidence provided from content and construct validity testing suggest that the test at Level 4 is valid;
2. The Attitudes to Mathematics Questionnaire (ATMQ)[6], developed by the researcher using an amalgam of one scale of an instrument used in Trends in International Mathematics and Science Study (TIMSS) 2007 and items from Tapia and Marsh's (2004) Attitude to Mathematics Inventory (ATMI);
3. The Children's Assessment for Learning Audit Instrument (CAfLAi) developed by the researcher but based on the same four scales as Lysaght and

[6] Further details regarding the ATMQ, the CAfLAi and TIMSS 2007 are provided later.

O'Leary's (2013) audit for teachers (AfLAi), with one additional scale measuring AfL techniques.

One further quantitative instrument, the Assessment for Learning Audit Instrument (AfLAi), was utilised in the study. Developed by Lysaght and O'Leary (2013), the AfLAi (Appendix E) is an instrument that helps gauge teachers' individual and collective levels of understanding and use of AfL. It consists of 58 item statements distributed across four scales based on the following four key AfL strategies: sharing learning intentions and success criteria (LISC), questioning and classroom discussion (QCD), feedback (FB) and peer- and self-assessment (PSA). The AfLAi can be used as a tool to identify gaps in teachers' AfL practices and guide school-based CPD in AfL for teachers. Psychometric analyses indicate that outcomes for the reliability of the four scales are very satisfactory, with Alpha Reliabilities ranging form 0.83 to 0.92. The AfLAi, used to test hypothesis three, pertains to the teachers and will be discussed in due course.

The next section considers the use of the SIGMA-T and DPMT-R in this study and is followed by a discussion on the development of the ATMQ and the CAfLAi. Then, details regarding test administration are provided. Finally, data handling for all the quantitative instruments is considered together, since the raw data from each instrument were treated in a similar fashion.

Standardised norm referenced mathematics tests

Mindful of Leahy and Wiliam's (2012) belief that the positive effects of AfL on student learning "are indeed achievable in real classrooms even where the outcomes are measured using externally-mandated standardised tests" (p.54), it was considered appropriate, in the current study, to measure students' mathematical achievement by using standardised tests. As required by school policy, students in Scoil na nAingeal complete the Sigma-T at the end of each academic year from first to sixth class. To strengthen the findings of this study, the DPMT-R was also administered to the intervention group at the end of third class, and to both the intervention and comparison groups at the end of fourth. This is what Teddlie and Tashakori (2009) refer to as the nonequivalent control group design and may be represented as follows:

O	X	O
O		O

Figure 15. Nonequivalent control group design

While some quasi-experimental designs such as pre-test post-test single designs have at times come in for criticism, Robson (2002) sees no particular problem when it is used “simply to determine whether there is an increase of performance after a treatment or even to assess its statistical significance” (p. 137).

Attitude Towards Mathematics Questionnaire (ATMQ)

Chapter Two highlighted how effective use of AfL can positively impact students’ self-confidence (Clarke, 2014; Stiggins, 2006), self-esteem (Clarke, 2014; Heritage, 2013), motivation (Clarke, 2014; Gardner, 2012) and attitudes towards learning (Hayward, 2012). It also suggested that student engagement is integral to academic achievement (e.g. Appleton, Christenson & Furlong, 2008; Perdue, Manzeske & Estell, 2009) and that students’ attitudes towards mathematics, their motivation to do mathematics and their self-confidence regarding mathematics can directly influence their mathematical achievement (e.g. Duerr & Harlow, 2013; Lim & Chapman, 2013; Mullis, Martin, Foy & Arora, 2012; Stankov, Lee, Luo & Hogan, 2012). Attempting to capture whether the use of AfL practices impacted children's engagement with mathematics and, in particular, if it affected children's attitudes towards mathematics, their self-confidence regarding mathematics and their motivation to do mathematics and, unable to locate a suitable instrument, I developed the Attitude Towards Mathematics Questionnaire (ATMQ). This instrument evolved from an amalgam of items from one scale of an instrument used in the Trends in International Mathematics and Science Study (TIMSS, 2007) and many items from Tapia and Marsh's (2004) Attitude To Mathematics Inventory (ATMI). Scale one of the ATMQ (ATMQ-TIMSS) uses statements 8a-8h of the TIMSS 2007 Grade 4 Student questionnaire verbatim (p.8). These statements examine "students' general attitudes towards mathematics" and "their self-confidence in learning mathematics" (Mullis, Martin & Foy, 2008, p.173). Leaving the TIMSS questionnaire unchanged facilitates comparative analysis with national and international data and ensures the reliability of this scale (median reliability coefficients across all TIMSS countries at fourth grade was 0.83, Mullis et al., 2008, p.401). The second scale of the ATMQ (ATMQ-SCLM) measures students' self-confidence regarding mathematics, with all statements taken from the ATMI (Tapia & Marsh, 2004) except for minor changes in wording to render them age and culturally appropriate, e.g., replacing the word 'mathematics' with 'maths' or 'advanced' with 'difficult'. Tapia and Marsh (2004) provided evidence that the ATMI had good content validity while the reliability coefficient alpha of the ATMI was 0.97. The first five statements of the final scale (ATMQ-MOT) measuring motivation are also taken from the ATMI, again with minor changes in wording, while the researcher composed the remaining statements of this scale to capture other aspects of students' motivation, for example intrinsic motivation. The ATMQ was piloted twice in Scoil na nAingeal with mixed ability groups from third and fifth classes respectively and any necessary adjustments were made. Scoring was done using a four-point Likert scale, with response options ranging from *Agree a lot* to *Disagree a lot* (Appendix F).

Children's AfL Audit Instrument (CAfLAi)

I was unable to find an instrument that gauged children's assessment literacy in mathematics and the extent to which they used AfL in their learning. Although conscious of Fraenkel and Wallen's (2006) admonition that developing an instrument is not easy to do and their recommendation to choose "an already developed instrument when appropriate" (p.114), I was, nevertheless, anxious to hear the voice of the child with regard to AfL. Through a process of iterative improvement I developed an evaluation instrument that attempted to measure the children's baseline levels of understanding with regard to AfL practices and the extent to which AfL strategies were embedded in their learning of mathematics. This instrument has two sections. Section one uses the same four independent scales used by Lysaght and O'Leary (2013) in their AfLAi but in the children's instrument each scale contains five items. Section two consists of twenty items and attempts to measure the children's use of some of the most commonly used AfL techniques presented in the AfL literature (Wiliam, 2011a). The instrument contains examples for the children to do prior to completing the instrument and the language is age appropriate for fourth class students (Appendix G). For quantitative purposes, each of the scale points was given the following numeric values:

5 = *Always*

4 = *Often*

3 = *Sometimes*

2 = *Never*

1 = *I don't understand what this means.*

Test administration

Sigma-T and DPMT-R guidelines for test administration were strictly adhered to while administering both standardised tests. In an attempt to limit inconsistency in the administration of the CAfLAi and the ATMQ, the researcher administered both instruments, pre- and post-intervention, with test conditions matched as closely as possible on each occasion. The three teachers participating in this study completed the AfLAi independently, before and after the intervention. This gauged the teachers' baseline understanding of AfL practices and the extent to which AfL was embedded in their teaching and subsequently measured whether participation in the intervention had affected their knowledge, skills and practices with regard to AfL.

Quantitative data handling

To help organise, quantify and analyse the quantitative data, raw data from the children's standardised tests, the ATMQ, the CAfLAi and the teachers' AfLAi, were coded or categorised, recorded and prepared in Microsoft Excel (Appendix H). Variable names and numerical labels were assigned to every response and a codebook was compiled for each data set. Taking the ATMQ as an example, Appendix I gives details of how the raw data were coded, collated and cleaned. Once this preparation was complete, the Excel files were imported into a software package called the Statistical Package for Social Sciences (SPSS 21), where various statistical tests and analyses were executed, for example descriptive statistics, histograms, t-tests, etc. These will be discussed in detail in Chapter Four.

Qualitative Data Collection and Handling

In conjunction with the quantitative data previously discussed, multiple qualitative data sources were utilised to test research hypotheses two and three and to provide in-depth accounts of participants' thoughts and opinions on AfL, LS and mathematics, and about the intervention in general. Not only did these data illustrate and supplement the quantitative data by providing *thick descriptions* (Geertz, 1973) of the intervention, they also allowed for triangulation of evidence and, as recommended by Lincoln and Guba (1985), helped achieve external validity. The importance of listening to children when engaging in research has been highlighted already in the literature review. By choosing the methodologies discussed in the next two sections, I attempted to empower children as central voices in this project, and ultimately in their own learning. Following this, the teachers' qualitative data collection and data handling is investigated in detail.

Student learning logs

In previous large-scale AfL projects (e.g., Thompson & Wiliam, 2008), student learning logs have been used successfully as a strategy to activate students as owners of their own learning. Over the course of this project, students were asked to reflect on their learning by completing a learning log at the end of various mathematics lessons throughout the year. The learning logs contained a series of prompts, which the children responded to, and were a useful way to encourage them to take ownership of their learning (Appendix J). The teachers read and analysed the learning logs regularly, although they did not amend or mark them. This provided feedback to each teacher as to where her students were in their learning and, where appropriate, led to modification of instruction. Additionally, when reading back over their learning logs at the end of the year, students could see what they had learned over the course of the project. Using learning logs was one way of collecting qualitative data from the children and, in particular, listening to their voice.

Focus groups

In an effort to triangulate data from the children's questionnaires and mindful of my desire to hear the voice of the child, I decided to use focus groups as the best method to engage the children in meaningful dialogue, enabling them to critically reflect on their experiences with regard to AfL and to give voice to their ideas, perceptions, feelings and attitudes in an age-appropriate way (Wilson, 1997). Experts (e.g., Cohen et al., 2010; Robson, 2002) suggest that focus groups are not without their drawbacks. Nevertheless, in this case, the advantages seemed to outweigh the disadvantages. As argued (Bogdan & Biklen, 2003), focus groups are a data-gathering method that "encourage self-disclosure" and because of the interaction within the group participants, "can stimulate each other to articulate their views or even to realise what their own views are" (p.101), views, according to Gall et al., (2003), they may not express if interviewed individually. This can provide the researcher with insights into not only what people think but why (Denscombe, 2010). Focus groups provide an opportunity to interview several people at the same time and "consist of people you have specifically selected for their experience in relation to whatever you are studying" (Kane & de Brun, 2005, p.273). In addition, Teddlie and

Tashakkori (2009) inform us that they allow access to content that you are interested in, such as the attitudes and experiences of your informants.

Following an initial analysis of the children's learning logs in June 2013, I identified various themes and, from that, possible candidates for the focus groups. Then, following discussion with the collaborating teachers, children were chosen to participate in the interviews based on these themes and discussions. I endeavoured to include a reasonable cross-section of all students so that students from the SEN, middle and top mathematics groups were chosen for each focus group interview, while some students with English as an additional language (EAL) were also selected. The focus group was piloted with five children, the necessary adaptations were made and then two focus groups, each with six children, were formed. The focus groups were held during school time, in a small room where the children felt at ease and were comfortable with self-disclosure. Each focus group interview started with a preordained routine that stated the time and date and emphasised the voluntary nature of the children's participation and their ability to withdraw at any time. The purpose of the interview was explained and further assurances about confidentiality and anonymity were given (Appendix K). The sessions were video-recorded, with permission, and lasted approximately one hour. As researcher, I remained conscious of Hill, Laybourn and Borland's (1996) advice that:

> The influences that affect interviews with adults are also relevant to children, such as the need to establish rapport, ensure confidentiality or pose questions clearly and concisely, but there are also additional factors. Communication with children needs to be adapted to their level of cognitive and linguistic development; to make use of materials, techniques and settings with which they are familiar; and to convey instructions in a manner that makes sense from a child's perspective. It is very important to be aware of extensive evidence that children in interview situations are very affected by the perceived power and status of adults and by the presumptions about what answers are expected. (p.133)

Teacher learning logs

Following every meeting of the PLC or LS group, the three participating teachers completed a learning log (Appendix L), consisting of two parts. In section one, using prompts from Wiliam, (2011b, pp.157-158), we reflected on our learning and then responded to any three prompts of our choice. Similar to Lysaght (2009), section two was developed to identify our needs and concerns with regard to AfL and LS, with the responses being used by the researcher to guide preparation for the next meeting of the PLC or LS group.

Research journal

Some scholars (e.g., Gray, 2014; Robson, 2011) consider it good practice to keep a research journal from day one of a research project. Consequently, I used a large-size diary from beginning to end of the research process. In it, I recorded a variety of information about my reflections on the research and, as recommended by Lincoln and Guba (1985), included the weekly schedule and "logistics of the study" in addition to "methodological decisions and accompanying rationales" (p.327). These entries provided a detailed portrait of events that helped facilitate *thick description* for subsequent data analysis and interpretation, while also acting as an aide memoir for reflection purposes (Gray, 2014).

Lesson study observations and reflections

As recommended by LS scholars (e.g., Lewis & Hurd, 2011), the teachers observing the research lesson recorded their observations and shared them subsequently at the LS reflection meeting. These observations were generally written onto the live lesson plans and were guided by prompts developed by the researcher based on similar guides in the LS literature (Appendix M). Additionally, the teacher who had taught the live lesson also wrote her reflections on that lesson. Corcoran (2009) explains the importance of such observation thus:

> The collation of multiple accounts of observers, noticing different elements of the teaching/learning episode helps build a more complex, possibly more challenging and certainly more useful picture of the learning ecosystem, which is the mathematics classroom. (p.39)

Video taping

The use of video in social research has increased in recent times (Jewitt, 2012). Video recordings are a "powerful resource" (p.7) as they are capable of capturing large amounts of information in comparison to what a human can observe in real time (Barron, Pea & Engle, 2013). In this project, video was used for all research lessons in cycles two and three of the LS process and for the children's focus groups. While acknowledging the potential of video recordings, proponents of LS, such as Corcoran (2009), argue that the value of first hand observation by fellow teachers in the live lessons is superior to video footage since it is not limited by the eye of the person holding the camera. Therefore, in the *spirit* of lesson study, the teacher observations were deemed more important than video footage and, consequently, provided the main data source for the live lessons. Nevertheless, using video provided me with multiple viewing opportunities of the live lessons and the focus groups. This stimulated recall and reflection and helped define and develop data for the project, thus helping to strengthen my research findings. Furthermore, combining video recordings with other forms of data provided "opportunities for triangulation across multiple sources of evidence" (Barron et al., 2013, p.205).

Teachers' meetings/ Professional development

A key part of this study was the use of peer-to-peer learning as a vehicle of continuing professional development (CPD). Meeting on average every two weeks, over the course of the intervention the teachers learned about AfL strategies and techniques before implementing them in their mathematics lessons on a phased basis prior to the next meeting. Reflective practice and collaborative learning were vital components of this process. Input was by the researcher at the start of each session. What the teachers learned about AfL consisted of an amalgam of ideas taken from various sources but was especially focused on the five key AfL strategies, "One Big Idea" and multiple techniques (Appendix N) identified by the literature (Thompson & Wiliam, 2008). The core content of the programme was taken from the following sources:

- A commercially available Scottish programme called *The Learning Set* (Learning Unlimited, 2004);
- Lectures from the online Masters in Education Programme (2011), St. Patrick's College of Education (Assessment Module);
- *Sine Qua Non of Assessment and Learning* (Lysaght, 2012);
- *Embedding Formative Assessment* (Wiliam, 2011) and various other works by Wiliam (e.g., 2007; 2009; 2013);
- Selected research by other experts in AfL such as Clarke (2008), Hayward (2012), Heritage (2013), James (2006), Shepard (2006; 2008), Stiggins (2006; 2007) and Swaffield (2011);
- The work of John Hattie (2009; 2012), especially his work on feedback.

In addition to learning about AfL we also learned about Japanese lesson study and how to plan and implement it, as well as studying other relevant content and theories as necessary. Appendix O outlines the CPD timeline, while Appendix P provides a synthesis of the CPD content, compiled by the researcher for use at the intervention reflection meeting (27/06/2013) at the end of the study.

Audiotapes and transcripts from these meetings comprise part of the qualitative data set. Meetings were site-based, usually lasted from one to two hours, took place after school and were not part of the Croke Park Agreement[7] hours. Leahy and Wiliam (2012) highlight the benefit of adopting a standard structure for such meetings so that the learning is foregrounded. Therefore, a similar structure was adopted for each session to include:

- Introduction – sharing of learning intentions;
- New learning in AfL, e.g., a particular AfL strategy and relevant AfL techniques; or new learning/focus in LS, e.g., LS as research;
- Review of how things were going;
- Planning for next two weeks;
- Reflection.

The meetings were, however, “tight but loose” in that the exact content of the CPD and the meeting schedule were emergent. In addition to the planned meetings, many informal chats happened on an almost daily basis throughout the year and these were really important. It should be noted that there is a spirit of collaboration in Scoil na nAingeal. This is important since the three participating teachers already had a close working relationship, thus making the formation of a LS group a natural progression. Additionally, since the LS process necessitated supervision of the classes not involved

[7] Under the Croke Park Agreement (2011) teachers at primary level are required to work an additional 36 hours non-contact time per annum (DES, Circular 0008/2011).

in the live lesson, this was kindly facilitated by the principal or members of the SEN team.

Qualitative data handling

Given the research topics, hypotheses, context and my relative inexperience as a researcher, I chose thematic analysis as the best analytic approach for analysis of the qualitative data set. As discussed above, there were multiple sources of qualitative data, for example, children's learning logs, transcripts of the focus group (FG) interviews, teachers' learning logs, teachers' observations from the live lessons and transcripts of various teacher meetings regarding AfL and LS. As recommended by Miles et al. (2014), data collection and analysis were interwoven from the outset and during data collection I consistently searched for "patterns of meaning and issues of potential interest in the data" (Braun & Clarke, 2006, p.15). To familiarise myself with the data and gain an overall sense of any emerging themes, I repeatedly listened to the audiotapes, looked at the different video-recordings and read and re-read the teacher and student learning logs and various transcripts. I used Braun and Clarke's (2006) six-step approach to thematic analysis as my guide when analysing each data item individually and subsequently the complete qualitative data set (Table 7). Appendix Q provides a worked example of this thematic analysis and the extract is taken from one of the focus group interviews with the children at the end of the intervention.

Table 7

Phases of Thematic Analysis (Synthesis from Braun & Clarke, 2006)

Phase	Description of the process
1. Familiarisation with data:	Immersion in data. Transcription of audio files, reading and re-reading the data, noting down initial ideas. Marking ideas for subsequent phases.
2. Generation of initial codes:	Identifying interesting aspects and patterns. Systematically coding interesting features of the data across the entire data set. Collating data relevant to each code.
3. Searching for themes:	Collating codes into appropriate themes. Gathering all data relevant to each potential theme, sub-theme and overarching theme.
4. Reviewing themes:	Checking if the themes work in relation to the coded extracts (Level 1), and the entire data set (Level 2). Visual representation of themes.
5. Defining and naming themes:	Ongoing analysis to refine the specifics of each theme, and the overall story the analysis tells. Generating clear definitions and names for each theme. What's interesting and why?
6. Producing the report:	Selection of vivid, compelling, extract examples, final analysis of selected extracts, relating back of the analysis to the research hypotheses and literature. Embedding abstracts within analytic narrative and producing a scholarly report of the analysis.

Additionally, as advocated by Braun and Clarke (2006), I maintained an ongoing reflexive dialogue throughout the analytic process, I was thorough and consistent in my analysis, and reported truthfully about the data. I used an inductive approach to identify themes strongly linked to the data themselves and am confident the resulting themes represent an accurate reflection of the entire qualitative data set. As recommended by various scholars (e.g., Maxwell, 2013; Robson, 2011) I utilised respondent validation or member checks with the children and teachers to ensure that I was reporting accurately what they had said/written and to guard against researcher bias. Throughout the analysis process, which is discussed in more detail in Chapter Four, I remained vigilant about my subjectivity and using the strategies described above was conscious of Yin's (2014) admonition to report all evidence fairly.

Quality of Research

It is generally accepted that rigour is important in conducting high-quality research and that sound research, regardless of paradigm, should, as Brown (2016) suggests, involve systematic and principled inquiry. Broadly speaking, scholars (e.g., Brown, 2016; Teddlie & Tashakkori, 2009) posit that the quality of quantitative research can be defended by the researcher, and evaluated by the reader, in terms of validity, reliability, replicability and generalisability, while the qualitative analogues for these terms are credibility, dependability, confirmability and transferability or, in one word, trustworthiness (Lincoln & Guba, 1985). Consequently, since mixed methods research (MMR) systematically combines qualitative and quantitative methods, all these issues must also be addressed. Not only that, in the past decade, scholars (e.g., Brown, 2014; Onwuegbuzie & Johnson, 2006) have recommended that the quality of MMR can be enhanced, defended and evaluated by using a concept called *legitimation*, defined by Brown (2014) as "the degree to which MMR integration of qualitative and quantitative research strengthens and provides legitimacy, fidelity, authority, weight, soundness, credibility, trustworthiness, and even standing to the results and interpretations in MMR" (p.128). Thus, in order to enhance the research and the resulting inferences at the MMR or integration level of the study, I remained mindful of *legitimation* throughout the project (Brown, 2014). The next section discusses particular ways in which the validity and reliability of the quantitative data were enhanced. After that, attempts made to ensure the trustworthiness of the qualitative data are explored, followed by other data quality assurance measures.

Validity and Reliability

Yin (2014) judges the quality of case study design by applying three types of validity criteria and one reliability criterion: construct validity, internal validity, external validity and reliability. Robson, however, cautions "there is no easy, single way of determining construct validity" (2002, p.102), i.e. "the degree to which the constructs under investigation are captured/measured; the degree to which inferences may be made about specific theoretical constructs, on the basis of the measured outcomes" (Teddlie & Tashakkori, 2009, p.298). Notwithstanding, scholars (e.g. Robson, 2011) agree that striving for construct validity is essential and so, as advised by Yin (2009), I used multiple sources of evidence to increase construct validity. Additionally, terms such as AfL, CPD, PLC, and attitudes to mathematics (Chapter Two), were clearly defined while the measurement instruments used to evaluate the impact of the CPD programme on AfL literacy, attitudes and mathematics achievement were clearly described and tested for reliability (to be discussed). Since this research involved the use of a quasi-experimental design, numerous threats to internal validity, such as differences between the comparison and intervention group and the streaming of mathematics groups, were identified and their potential to affect inferences noted. With regard to enhancing external validity, multiple data sources were used to ensure triangulation and this will be discussed in the next section. Furthermore, since this is an action research project, I also adhered to the five validity criteria posited by Anderson and Herr (1999) for practitioner research (Appendix R). Yin (2003) suggests that the best way of ensuring reliability in case study research is to "make as many steps as operational as possible and to conduct research as if someone were always looking over your shoulder" (p.38). I took heed of his advice. Finally, the reliability and validity of the SIGMA-T and the DPMT-R have been discussed previously and so the next section discusses the statistical analyses undertaken to assess the reliability of the ATMQ and CAfLAi.

Instrument reliability for the ATMQ and the CAfLAi

Cronbach alpha coefficient values were sought to ascertain the internal consistency of the instruments as a whole and for each of their subscales separately. Cronbach's alpha reliability scores for the three scales of the ATMQ suggest that the reliability of each scale is acceptable since, as argued (Cohen et al., 2010), they are above the usually accepted alpha reliability threshold of 0.70 (Table8).

Table 8

ATMQ Scale Alpha Reliabilities

Scale	Acronym	No of Items	Alpha Reliabilities
Trends in International Mathematics and Science Study	ATMQ-TIMSS	8	0.89
Self-Confidence in Learning Mathematics	ATMQ-SCLM	12	0.88
Motivation to do Mathematics	ATMQ-MOT	10	0.78

Regarding the CAfLAi, the composite score for the first four scales measuring the strategies is reliable (Cronbach's alpha of .76). However, reliability is low for each of the following scales individually: LISC, QCD, and PSA scales (Table 9). While, as Gray (2014) suggests, the low level of these Cronbach's coefficients may have been affected by the fact there are fewer than 10 items in each scale, and by the fact there was not a large number of participants, given that the scores are significantly lower than the usually accepted alpha reliability threshold of 0.70 (Cohen et al., 2010), findings from the CAfLAi, and in particular the LISC, QCD and PSA scales, are interpreted with caution in Chapter Four.

Table 9

CAfLAi Scale Alpha Reliabilities

Scale	Acronym	No of Items	Alpha Reliabilities
Sharing Learning Intentions and Success Criteria	LISC	5	0.31
Questioning and Classroom Discussion	QCD	5	0.25
Feedback	FB	5	0.74
Peer- and Self-Assessment	PSA	5	0.60
Techniques	TQ	20	0.84

Trustworthiness of Qualitative Data

Referring primarily to qualitative data, Lincoln and Guba (1985) suggest that maintaining an audit trail helps determine both dependability and confirmability and state that it "may be the single most important trustworthiness technique" (p.283) available to the researcher. Consequently, I audited the research process from the beginning in accordance with the six Halpern audit trail categories as outlined by Lincoln and Guba (1985, p.319). This helped ensure best practice in relation to the following aspects of the research process:

1. Raw data collection;
2. Data reduction and analysis products;
3. Data reconstruction and synthesis products;
4. Process notes;
5. Materials relating to intentions and dispositions;
6. Instrument development information.

Lincoln and Guba (1985) further suggest that triangulation and the keeping of a reflexive journal dovetail with the audit to ensure confirmability. As argued (Sarantakos, 1997), "triangulation intends to offer a stereoscopic view of the issue in question and improve the quality of the findings" (p.186). In view of the small sample size, it was vital to document, triangulate, and validate results, using as many methods and reliability checks as possible, thus "getting a fix on it from two or more places" (Robson, 2002, p.371). Although Teddlie and Tashakkori (2009) question whether the term *triangulation* is still useful or not, others (e.g., Bell, 2010; Denscombe, 2010) recommend its use as a way of reducing responder and researcher bias, viewing it as a powerful way of demonstrating validity, especially in qualitative research (Cohen et al., 2010). In this study, triangulation of data (pre-and post-intervention tests), sources (focus groups and observations) and methods (qualitative and quantitative) was used in an attempt to improve accuracy and gain a broader understanding of the issues under investigation. It is, however, worth noting that Maxwell (2005) states, "in the final analysis, validity threats are ruled out by *evidence*, not methods" (p.112).

Additionally, as discussed earlier, I also kept a reflexive journal for the duration of the project. This technique, suggests Teddlie and Tashakkori (2009) provides information for all four of Lincoln and Guba's (1985) trustworthiness criteria: credibility, transferability, dependability and confirmability.

Other Data Quality Assurance Measures

According to Blaxter et al., (2006) "the value of pilot research cannot be overestimated" (p.137) and so each method of data collection was trialled beforehand. The focus group was piloted with five students from the intervention group and any necessary adaptations were made. Additionally, as discussed previously, the CAfLAi and the ATMQ were both piloted with a group of children from third and fifth classes respectively.

In order to minimise errors and bias in the study, I avoided using the case study on its own to substantiate a preconceived position. Bell (2010) cautions that it can be easy to fall into the 'bias trap', particularly for individual researchers and so I remained open to "competing explanations and discrepant data", thus ensuring that the research was not "simply a self-fulfilling prophecy" (Maxwell, 2005, p.126). However, Bogdan and Biklen (2003) advise that biases are difficult to eliminate since we cannot separate our beliefs, values and past experiences from our research. In order to minimise errors and reduce bias in the study, I employed a range of strategies to guard against my own biases by being vigilant, triangulating, regularly questioning my practice and being critical of my analysis of the data. As argued by Teddlie & Tashakkori (2009), member checking "is perhaps the most important strategy for determining the credibility of the researcher's interpretation of the participants' perceptions" (p.213). Therefore, participating teachers were given the opportunity to review and discuss the results of their AfL audits (pre- and post-intervention), to verify and respond to what they had written in their learning logs and, finally, to review the penultimate draft of the findings chapter. Additionally, on several occasions, the children were also given the opportunity to explain, discuss and clarify what they had written in their learning logs, particularly during the FG interviews.

In sum, throughout the research process, I made efforts to be as precise, objective, and consistent as possible. Furthermore, to ensure the quality of this research, I tried at all times to ensure that the research was systematic, principled and of the highest possible quality. I believe these measures, taken in tandem with the other measures discussed above, have helped ensure the reliability, credibility, trustworthiness and authenticity of this study.

Ethical Considerations

Ethical issues are inherent in all research designs and apply to each research paradigm (Creswell, 2009). Basic ethical principles include:

> A commitment to the well-being, protection and safety of participants; a duty to respect the rights and wishes of those involved; a responsibility to conduct high-quality scientific research; and a commitment to disseminate and communicate the results to stakeholders. (Department of Children and Youth Affairs, [DCYA], 2012, p. 8)

In addition, as argued by Morrow and Richards (1996), in research involving children, other ethical considerations, such as children's vulnerabilities and competencies, also need to be addressed. Consequently, this research was carried out within the framework of *Children First: National Guidelines for the Protection and Welfare of Children* (DCYA, 2011), while the core ethical principles and concepts outlined in *Guidelines for Developing Ethical Research Projects Involving Children* (DCYA, 2012) regarding child-related research were also followed with care. I was particularly mindful of the power differential that exists between students and teachers, which Morrow and Richards (1996) have identified as the greatest ethical challenge when doing research with children, since it has the potential to affect children's ability to decide freely whether to participate in the research or not. When reflecting on what Waldron (2006) describes as a "complex matrix of power relationships: that between adult and child, between researcher and researched" (p.91), it is worth noting her premise that this imposes an even greater onus on researchers to ensure the research is necessary, meaningful, purposeful and of positive benefit to the children involved.

Bogdan and Biklen (2003) suggest that the two issues which dominate the "traditional official guidelines of ethics in research with human subjects are: informed consent and the protection of subjects from harm" (p. 43). Since this study was being conducted as part of the day-to-day practice of normal teaching and learning, the research was deemed to present no ethical risks to participants. Nevertheless, in an effort to mitigate any potential risks the following key ethical issues were also identified and specifically addressed: gaining access; informed consent and assent; and confidentiality and anonymity.

The approval and support of the principal is central to any research in schools. The opinion of the principal "carries a great deal of weight" and s/he is the "key gatekeeper" in each school (Bogdan & Biklen, 2003, p.76). This project was therefore first discussed with the school principal, who was very interested in it. Subsequently, a letter was submitted to the school's Board of Management and formal permission was obtained from the Board in June 2012 (Appendix S). The nature of the project was then fully explained to the two participating teachers who subsequently signed consent forms (Appendix T). Next, information sheets and consent forms were distributed to the parents of all potential participants in the intervention (Appendix U) and comparison groups (Appendix V). Once parental approval was obtained from the parents of children in the intervention group, assent forms were distributed to the children (Appendix W). These forms were first explained and discussed orally, and the children were advised that they could agree to take part or not, without repercussion. All forms were age- and language-appropriate and clearly explained the purpose, procedures, and possible uses of the research. As advocated by Miller and Bell (2002), all participants, both children and teachers, were advised that not only was their participation in the research voluntary but, they were free to withdraw at any stage.

Additionally, I took heed of the advice of Teddlie and Tashakkori (2009) by addressing the privacy rights of all participants. Confidentiality was guaranteed. In addition, every effort was made to ensure that the identity of the participants in the research study was protected. Pseudonyms were used for the names of the participants and the school. Nevertheless, participants were told that it was not possible to guarantee their anonymity since the research took place solely in the school where I teach and the participants are students and teachers in that school.

Creswell's (2009) warning that "ethical practices involve much more than merely following a set of static guidelines" (p.88) was also taken into consideration. This is reminiscent of Hesse-Biber and Leavy (2011) who argue that since ethics exist in a social context rather than a vacuum, ethical guidelines alone cannot cover all possible ethical dilemmas. They (Hesse-Biber & Leavy, 2011) also highlight the difference between ethical codes and ethical values and advocate that research be guided by "ethical principles beyond informed consent" (p.82). Throughout the project, I therefore engaged, as they suggested, in self-reflexivity, which helped clarify my own ethical principles as a researcher. I was also conscious of other ethical researcher dilemmas particular to research within one's own workplace such as "divided loyalties" (Bell & Nutt, 2002) and "double agency" (Ferguson, Yonge & Myrick, 2004). Additionally, from the project's inception to completion, I conscientiously and rigorously adhered to the guidelines and procedures outlined in the St. Patrick's College Research Ethics Protocol and it was also signed off by my supervisors. Furthermore, I did not engage in any deceptive practices such as suppressing, falsifying or inventing findings (Creswell, 2009) and, as recommended by Denscombe (2010), ethics approval was obtained prior to beginning data collection.

Conclusion

This chapter discussed and justified the research design, instrumentation and mixed methodology. It reiterated the research purpose, problem and hypotheses and provided the conceptual framework and philosophical underpinnings for the research. Data collection methods and techniques, and data handling were also discussed. Additionally, the chapter outlined how issues of quality control were addressed and how ethical issues, such as informed consent and privacy, were attended to, especially when dealing with children. In sum, this chapter described how the research data was collected in response to the research questions. A complete and systematic analysis of this data is presented in detail in the next chapter.

CHAPTER FOUR: FINDINGS AND ANALYSIS

Introduction

This chapter presents the findings of the research and systematically describes, evaluates, discusses and analyses the data corpus. It comprises quantitative and qualitative data that are organised into three main analytical sections, each pertaining to one of three research hypotheses. The focus of both the first and second hypothesis is on participating students and the effects using AfL practices had on their attitudes, learning and achievement regarding mathematics, while the third hypothesis concentrates on the intervention's impact on participating teachers. In the first section, quantitative findings from the standardised mathematics scores of all participants are investigated to ascertain if the intervention impacted students' mathematics achievement in a way that can be measured by these instruments. Next, data from both quantitative and qualitative methods are used to determine if, over the course of the intervention, children's mathematical confidence, their engagement with, and attitudes to, mathematics were improved by their use of AfL practices. Finally, in the last section, qualitative and quantitative data from teachers' learning logs, reflection meetings, the researcher's journal and teachers' AfL audits are analysed and discussed in a bid to establish if teachers' participation in this model of CPD, i.e. LS, positively impacted their knowledge of AfL, their use of AfL practices and their attitudes and beliefs towards AfL as a form of assessment. The utilisation of multiple sources, for example surveys, learning logs and focus groups, and the combination of quantitative and qualitative approaches, enabled both data and methodological triangulation and enhanced the rigour of this research (Gray, 2014; Robson, 2011).

Table 10

Quantitative and Qualitative Data Sources

Hypothesis	Data Source	Data Type	Timing	Date	Participants	n
1	SIGMA-T	Quan	Pre-	19/05/2011	Comparison Group	50
			Post-	03/06/2012	Comparison Group	52
	SIGMA-T	Quan	Pre-	30/05/2012	Intervention Group	51
			Post-	30/05/2013	Intervention Group	51
	DPMT-R	Quan	Pre-	n/a	n/a	n/a
			Post-	21/06/2012	Comparison Group	52
	DPMT-R	Quan	Pre-	21/06/2012	Intervention Group	51
			Post-	06/06/2013	Intervention Group	51
2	ATMQ	Quan	Pre-	13/09/2102	Intervention Group	51
			Post-	07/06/2013		
	CAfLAi	Quan	Pre-	14/09/2012	Intervention Group	51
			Post-	12/06/2013		
	Childrens' Learning Logs	Qual	During inter-vention	Throughout	Intervention Group	51
	Childrens' Focus Groups	Qual	Post-	(FG1)20/06/2013 (FG2)21/06/2013	Intervention Group	6 6
	Teachers' Learning Logs (TLL)	Qual	During inter-vention	Throughout	Teachers	3
3	AfLAi	Quan	Pre-	31/08/2012	Teachers	3
			Post-	12/06/2013	Teachers	3
	Teachers' Learning Logs (TLL)	Qual	During inter-vention	Throughout	Teachers	3
	Audit Reflection Meeting (ARM)	Qual	Post-	26/06/2013	Teachers	3
	Intervention Reflection Meeting (IRM)	Qual	Post-	27/06/2013	Teachers	3
1,2,3	Researcher Journal (RJ)	Qual	Through-out	Throughout	Researcher	1
1,2,3	AfL Meetings	Qual		31/08/2012 01/10/2012 15/10/2012 22/10/2012 07/01/2013 21/01/2013 12/02/2013 26/02/2013	Teachers	3

Note. For LS data see Table 27

Table 10 provides an overview of primary data sources for each hypothesis, recognising that some overlap of sources occurred during the analysis stage when attempting to substantiate findings.

Quantitative Findings and Analysis

In this study, both parametric and non-parametric tests were used in the analysis of the quantitative data. Since parametric tests, for example t-tests, are potentially more powerful than non-parametric ones, they were used once certain assumptions such as normal distribution and sufficient sample size were met (Pallant, 2013). Otherwise, non-parametric tests were used. When exploring hypothesised relationships between variables, one must determine what the probability is that the relationship exists and, if it exists, how strong that relationship is through strength of association or effect size. Significance tests were used here to determine whether or not the null hypothesis was rejected in favour of the alternative or research hypothesis (Hinton, 2008; Levine & Hullett, 2002). This involved making a decision about whether the distributions were the same or different by using a decision criterion, the significance level or *p*-value, typically set at $\alpha = 0.1$, 0.05 or 0.01. This limited the risk of Type 1 errors, falsely rejecting the null hypothesis (Hinton, 2008; Salkind, 2011). However, since significance tests and p-values only provide an outcome where one rejects or fails to reject the null hypothesis, effect size statistics measuring the magnitude of the effect, are also provided (Hinton, 2008; Levine & Hullett, 2002; Pallant, 2013). Eta squared is one of the most commonly used effect size statistics and is used in this study to indicate the proportion of variance of the dependent variable (students' achievement in mathematics) that is explained by the independent (group) variable (Pallant, 2013). Values range from 0 to 1 and the guidelines proposed by Cohen (1988, pp.284-287) for interpreting eta squared are:

.01 or 1% = *small effect*

.06 or 6% = *moderate effect*

.14 0r 14% = *large effect*

Research Hypothesis One

A nine-month, school-based intervention employing Assessment for Learning principles, strategies and techniques will improve the standardised mathematics results of participating students in comparison to a similar cohort not involved in the intervention.

This section provides a summary and analysis of the results that test the first research hypothesis. It has been suggested that in the social sciences knowing the demographic characteristics of the individuals you are studying can help interpret a study's findings and enhance the quality of the research (e.g. Connolly, 2013). Therefore, this section begins with a synopsis of the demographic characteristics of the children in both the intervention and comparison groups (Table 11), previously discussed in detail in Chapter Three.

Table 11

Demographic Characteristics of Comparison and Intervention Groups

		Age (Beginning 4th Class)		Nationality		Stream		
	n	Mean	Range	Irish	New-comers	Top	Middle	SEN
Comparison	52	9yrs 11 mths	9yrs 2mths-12yrs 8mths	41	11	23	20	9
Intervention	51	9 yrs 11 mths	9yrs 5mths-10yrs 8mths	43	8	23	20	8

The data set for this first research hypothesis consists of students' scores from two Irish standardised graded mathematics attainment tests, Standardised Irish Graded Mathematics Attainment Tests (SIGMA-T) and Drumcondra Primary Mathematics Test-Revised (DPMT-R). It should be noted that since the comparison group did not complete the DPMT-R at the end of third class, analyses of these results are more limited. Data were prepared for analysis (see Chapter Three), and SPSS (Version 21) was used to conduct statistical tests. Preliminary analysis involved exploring the data visually and conducting descriptive analyses. Descriptive statistics are useful since they not only provide the data in a manageable and meaningful format that facilitates data interpretation (Gray, 2014; Punch, 2014) but also check that the variables are not violating any of the assumptions made by the statistical tests used to address the research question/s (Pallant, 2013). The descriptive statistics displayed in Tables 12 and 13 reveal the main features of the quantitative data for the SIGMA-T and DPMT-R, comparison and intervention groups, pre- and post-intervention, and provide the means and standard deviations (SD) for the dependent variable in the current study.

Table 12

Descriptive Statistics for the Sigma-T

SIGMA-T (All pupils)					
	Administered	Group	n	Mean	SD
Pre-	End of 3rd	Comparison	50	98.60	14.25
Pre-	End of 3rd	Intervention	51	103.55	14.56
Post-	End of 4th	Comparison	52	105.71	16.62
Post-	End of 4th	Intervention	51	108.67	15.98

Table 13

Descriptive Statistics for the DPMT-R

DPMT-R (All pupils)					
	Administered	Group	n	Mean	SD
Pre-	End of 3rd	Comparison	n/a	n/a	n/a
Pre-	End of 3rd	Intervention	51	103.55	14.876
Post-	End of 4th	Comparison	52	101.27	17.960
Post-	End of 4th	Intervention	51	101.96	15.290

As discussed earlier, many parametric statistical tests make the assumption that the distributions they are investigating are normally distributed. A normal or Gaussian distribution is represented by a bell-shaped curve that has the greatest frequency of scores in the centre and the smaller frequencies towards the extremes (Pallant, 2013). Normality assumes that the dependent variable is normally distributed and can be assessed graphically with, for example, histograms or numerically with tests such as Shapiro-Wilk. Tabachnick and Fidell (2013) recommend using histograms to inspect the shape of the distributions, especially with large samples. It is common, however, in the social sciences that variables are not normally distributed (Pallant, 2013), although Gray (2014) advises that t-tests can still be used even when the distribution is not perfectly normal. Other descriptive statistics reported here include frequency distributions that provide graphical representations or histograms of the scores obtained by students from the comparison and intervention groups in both tests. The histograms, Figures 16 and 17, depict the actual shape of the distribution for each group in the SIGMA-T and DPMT-R respectively. Visual assessment of the histograms for the standard scores on both standardised tests revealed that the distributions appeared to be reasonably normally distributed and so further normality tests such as Shapiro-Wilk or Kolmogorov were not deemed necessary.

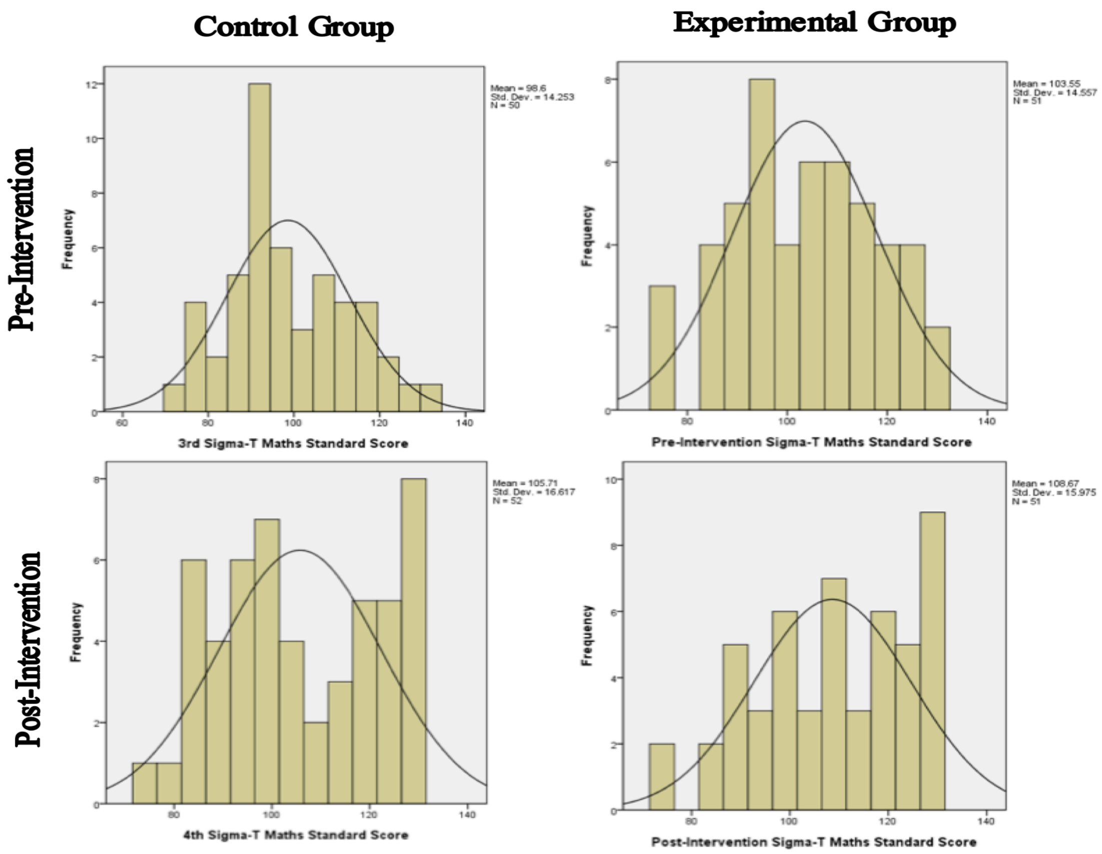


Figure 16. Histograms showing SIGMA-T standard scores

.

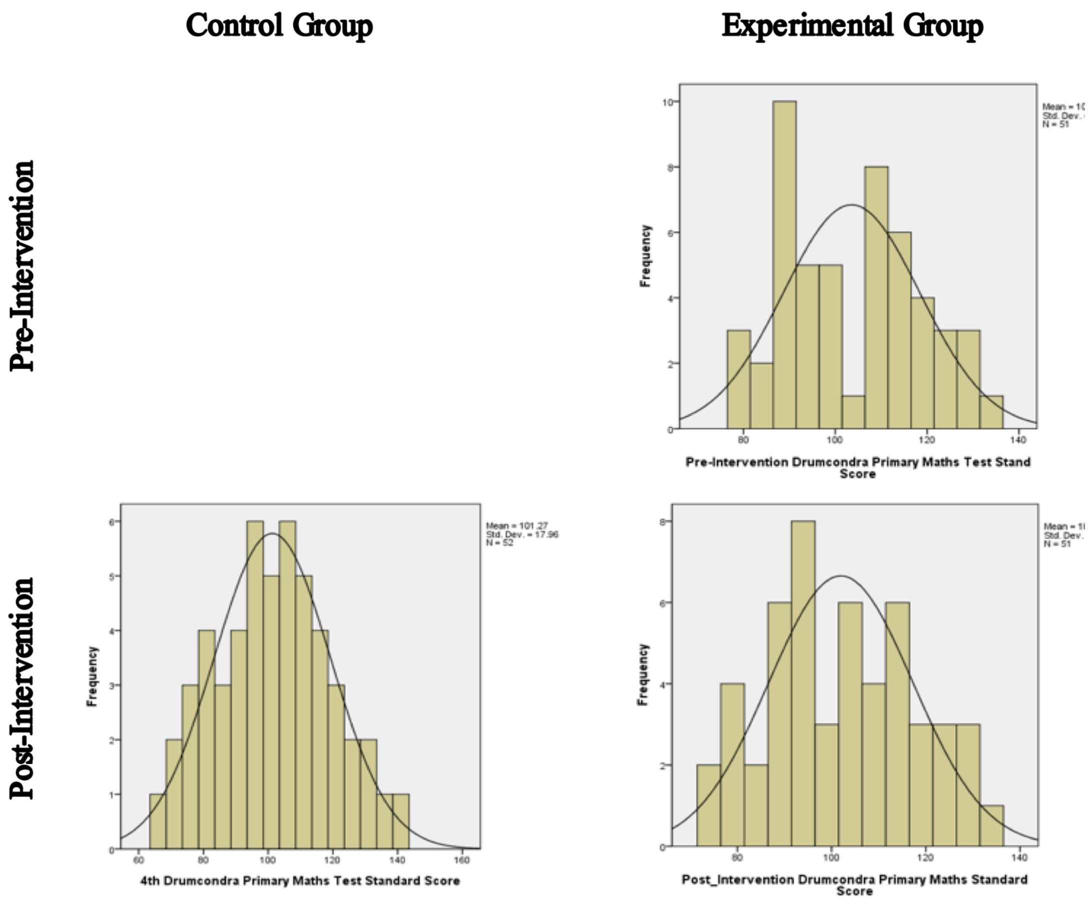


Figure 17. Histograms showing DPMT-R standard scores

Analysis of Data from the SIGMA-T

Since the descriptive statistics had revealed the distributions to be normally distributed, it was possible to proceed with the use of inferential statistics using parametric tests. An independent samples t-test was conducted to compare the SIGMA-T standard scores for the comparison and intervention groups at the end of Third Class, i.e. prior to the research intervention. Results indicated that there was no statistically significant difference in scores for the comparison group (M = 98.6, SD = 14.3) and the intervention group (M = 103.6, SD = 14.6; t (101) = -1.73, p = .09). The magnitude of the differences in the means was small (eta squared = .029), indicating that the achievement of both cohorts as measured by these tests was similar prior to the intervention. A second independent samples t-test was conducted which compared SIGMA-T standard scores for the comparison and intervention groups at the end of Fourth Class, i.e. post-intervention. These results indicated that there was no statistically significant difference in scores for the comparison group (M = 104.6, SD = 17.0) and the intervention group at the end of the intervention (M = 108.7, SD = 16.0; t (101) = -1.23, p = .495). The magnitude of the differences in the means was small (eta squared = .015). The results from these independent sample t-tests indicate that there was no statistically significant difference between the comparison and intervention groups in mathematics achievement as measured by the SIGMA-T, either before or after the intervention.

Results from dependent t-tests comparing SIGMA-T standard scores indicated that there was a statistical difference in scores for the comparison group pre-intervention (M = 98.6, SD = 14.3) and post-intervention (M = 104.6, SD = 17.0; t (49) = -6.62, p = .000) and the magnitude of the difference in the means was large (eta squared = .470). Similarly, results for the intervention group indicated that there was a statistically significant difference in pre-intervention (M = 103.6, SD = 14.6) and post-intervention scores (M = 108.7, SD = 16.0; t (50) = -6.5, p = .000) and that the magnitude of the differences in the means was also large (eta squared = .458). However, while the dependent t-tests revealed a large difference in the pre- and post-test scores for the intervention group, this was not statistically significant since the difference in mean standard scores for the comparison group was also large (Figure 18). Therefore, the improvement cannot be attributed directly to the intervention.

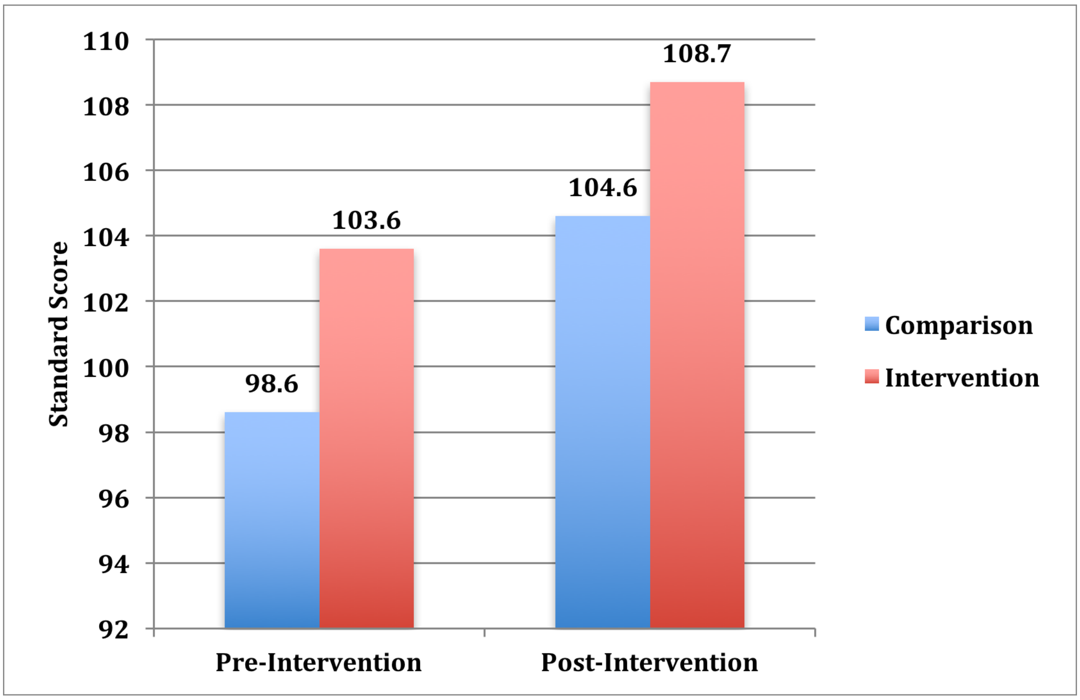


Figure 18. Comparison of standard scores for the SIGMA-T

Analysis of Data from DPMT-R

As stated earlier, the DPMT-R was not administered to the comparison group at the end of third class and so an independent samples t-test was conducted to compare standard scores for the comparison and intervention groups at the end of Fourth Class only. Results indicated that there was no statistically significant difference in scores for the comparison group (M = 101.3, SD = 18.0) and the intervention group (M = 102.0, SD = 15.3; t (101) = -0.2, p = .83). The magnitude of the differences in the means was small (eta squared = .0004), indicating once again that the performance of both cohorts on the day did not signal any differences that were statistically significant.

Results from dependent t-tests comparing DPMT-R standard scores for the intervention group, pre- and post-intervention, indicated that there was a statistical difference in scores from the end of third class (M = 103.6, SD = 14.9) to the end of fourth class, i.e. after the intervention (M = 102.0, SD = 15.3; t (50) = 2.0, p = .052). The magnitude of the difference in the means was small (eta squared = .038). However, while the dependent t-tests revealed a small difference in the pre- and post-test scores for the intervention group, this was not statistically significant. Therefore the improvement cannot be directly attributed to the intervention.

Analysis by Stream: Sigma-T

Since both independent and dependent t-tests failed to find any statistically significant difference between the mathematical achievement of the comparison and intervention groups in either standardised test, and since the literature suggests that low achievers may benefit most from using AfL practices (Black & Wiliam, 1998a; Stiggins, 2009; Stiggins & Chappuis, 2008), it was decided to further analyse the results by stream. As mentioned in Chapter Three, students from 4^{th} to 6^{th} in Scoil na nAingeal are streamed into three ability groups for mathematics, determined by cumulative standardised test results, teacher-assessments and teacher professional judgement. Consequently, using these criteria, approximately ten children were selected for a SEN group, usually comprising children from both classes. Since a full data set is not available for the DPMT-R, the focus here is on the SIGMA-T only. Descriptive statistics are presented in tabular form according to stream and detail measures of central tendency and measures of spread for each group (Table 14).

Table 14

Descriptive Statistics for the SIGMA-T by Stream

		Top Stream			
	Administered	Group	n	Mean	SD
Pre-	End of 3rd	Comparison	21	112.76	8.46
Pre-	End of 3rd	Intervention	23	116.74	7.67
Post-	End of 4th	Comparison	23	120.43	9.11
Post-	End of 4th	Intervention	23	122.43	7.24
		Middle Stream			
	Administered	Group	n	Mean	SD
Pre-	End of 3rd	Comparison	20	91.75	3.75
Pre-	End of 3rd	Intervention	20	96.9	5.57
Post-	End of 4th	Comparison	20	96.85	6.05
Post-	End of 4th	Intervention	20	102.6	8.69
		SEN Stream			
	Administered	Group	n	Mean	SD
Pre-	End of 3rd	Comparison	9	80.67	5.74
Pre-	End of 3rd	Intervention	8	82.25	5.29
Post-	End of 4th	Comparison	9	81.11	5.53
Post-	End of 4th	Intervention	8	84.25	6.61

Given that the sample size in each stream was less than 30, a series of non-parametric statistical tests were conducted (Independent Samples Mann Whitney U and Related Samples Wilcoxon Signed Rank) and adjusted for multiple comparisons (six for each) using the Bonferroni adjustment, giving an alpha of 0.008. When performing both these tests, no assumption was made regarding the data (except that they were ordinal) or about the populations' underlying distribution (Hinton, 2008).

An examination of results from the Mann Whitney U test (non-parametric equivalent of the independent t-test) for the pre-test SIGMA-T scores of the students in the top comparison and intervention groups (Table 15) did not show any statistically significant difference (z=-1.567; p=0.117>0.008). Similarly, the post-test results (z=-.50; p-0.617>0.008) failed to indicate any significant statistical difference between these two groups. Analysis therefore indicates that the scores achieved on the SIGMA-T by the top intervention group post-intervention revealed no statistically significant difference between scores obtained by their peers in the comparison group. Wilcoxon Signed Rank Tests (non-parametric equivalent of dependent t-tests) for the top stream revealed differences in the pre- and post-test scores of both the comparison (z=-3.923, p=<0.00005) and the intervention groups (z=-3.886, p=<0.00005) but the difference in scores for both groups was not statistically significantly different.

Table 15

Mann Whitney U Results Comparing Pre-test Scores of the Comparison and Intervention Groups on the SIGMA-T (Top Stream)

	Administered	Group	n	Mean	SD	Z	p
Pre-	End of 3rd	Comparison	21	112.76	8.461	-1.567	0.117
Pre-	End of 3rd	Intervention	23	116.74	7.671	-1.567	0.117
Post-	End of 4th	Comparison	23	120.43	9.11	-0.5	0.617
Post-	End of 4th	Intervention	23	122.43	7.241	-0.5	0.617

As evidenced by the data in Table 16, the Mann Whitney U Test on the pre-test scores of the middle comparison and intervention groups revealed a statistically significant difference (z=-2.879; p=0.004>0.008). Mean pre-test scores for the intervention group were approximately five points higher than for the comparison group, while post-test the difference was approximately six points. Thus, although the data reveals that the intervention group improved slightly more, this was not statistically significant. Wilcoxon Signed Rank Tests for the middle stream revealed differences in the pre- and post-test scores of both the comparison (z=-2.982, p=<0.003) and the intervention groups (z=-3.179, p=<0.001) but these are not statistically significant. In sum, data from the middle group do not indicate any difference between the comparison and intervention groups that can be attributed directly to the intervention.

Table 16

Mann Whitney U Test Results Comparing Pre-test Scores of the Comparison and Intervention Groups on the SIGMA-T (Middle Stream)

	Administered	Group	n	Mean	SD	Z	p
Pre-	End of 3rd	Comparison	20	91.75	3.754	-2.879	0.004
Pre-	End of 3rd	Intervention	20	96.9	5.572	-2.879	0.004
Post-	End of 4th	Comparison	20	96.85	6.046	-1.993	0.046
Post-	End of 4th	Intervention	20	102.6	8.69	-1.993	0.046

An examination of results from the Mann Whitney U test (Table 17) for the pre-test SIGMA-T scores of the students in the SEN comparison and intervention groups did not show any statistically significant difference (z=-.436; p=0.663>0.008). Similarly, the post-test results (z=-1.11; p-0.267>0.008) failed to indicate any statistically significant difference between these two groups. Analysis indicates that the scores achieved by students in the SEN intervention group in the SIGMA-T after the intervention revealed no statistically significant difference when compared to their peers in the comparison group. Wilcoxon Signed Rank Tests for the SEN stream revealed differences in the pre- and post-test scores of both the comparison (z=-.773, p=<0.439) and the intervention groups (z=-.563, p=<0.574) but again these are not statistically significant. Therefore, data from the SEN group do not indicate any difference between the comparison and the intervention groups that can be attributed directly to the intervention.

Table 17

Mann Whitney U Results Comparing Pre-test Scores of the Comparison and Intervention Groups on the SIGMA-T (SEN Stream)

	Administered	Group	n	Mean	SD	Z	p
Pre-	End of 3rd	Comparison	9	80.67	5.74	-0.436	0.663
Pre-	End of 3rd	Intervention	8	82.25	5.285	-0.436	0.663
Post-	End of 4th	Comparison	9	81.11	5.53	-1.11	0.267
Post-	End of 4th	Intervention	8	84.25	6.606	-1.11	0.267

To summarise, Figure 18 illustrates the mean standard scores in the SIGMA-T for each of the ability groups, pre- and post-intervention, comparison and intervention. It is clear from this figure that while the achievement trend was upwards for all groups, there was no statistically significant difference between the comparison and intervention groups, pre- and post-intervention, regardless of stream. In other words, findings were consistent across tests and across streams.

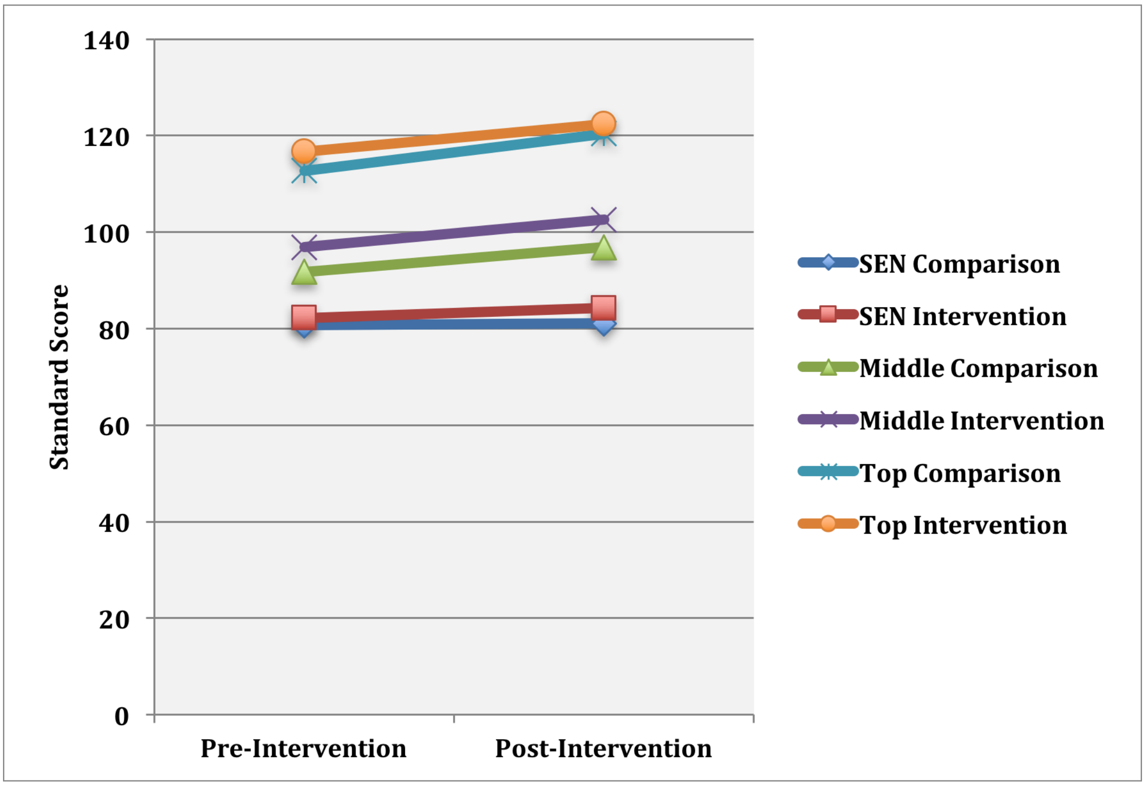


Figure 19. Mean standard scores of the SIGMA-T: Comparison by stream

In conclusion, analysis of the data set for research hypothesis one indicates that the intervention programme did not result in any statistically significant difference to the standardised mathematics performance of the intervention group as measured by the SIGMA-T and DPMT-R when compared to that of the comparison group and consequently the research hypothesis is rejected. These findings corroborate those of Lysaght (2009) who, in a similar study in the Irish context, investigated the impact of AfL practices on children's reading achievement and found that the intervention did not lead to an increase in children's reading scores on standardised tests, although a small difference was found in relation to children with SEN. Gardner, Harlen, Hayward and Stobart (2008) believe that many assessment initiatives have failed to provide significant data on the impact of changing assessment practices on student achievement since it is only in the long term that such changes can be expected to have an effect on student outcome. Consequently, the duration of this intervention may just have been too short to make a quantitative difference on standardised test scores. Notwithstanding, as has been suggested elsewhere (Lysaght, 2009; Wiliam, 2011b), the effects of AfL cannot necessarily be measured by standardised test scores and so attention now turns to investigate how AfL impacts students' mathematical dispositions.

Research Hypothesis Two

The use of AfL strategies and techniques, and the adoption of AfL principles, will enhance children's mathematical confidence, and improve their engagement with, and attitudes to, mathematics.

While investigation of the first hypothesis could be seen as an attempt to validate claims made in the literature regarding the positive impact using AfL practices can have on student achievement (e.g., Black & Wiliam, 1998b; Nyquist 2003), exploration of this Second Research Hypothesis concentrates on an empirical investigation of other claims made in the literature regarding the impact of using AfL strategies and techniques, namely the affective impact of AfL. In this regard, and to synopsise what was discussed in Chapter Two, it has been claimed that effective use of AfL can positively impact students' motivation (Clarke, 2014; Gardner, 2012b), self-confidence (Clarke, 2014; Stiggins, 2006), self-esteem (Clarke, 2014; Heritage, 2013) and attitudes towards learning (Hayward, 2012). Additionally, while acknowledging a paucity of empirical research in this area, it is also contended that AfL can help students develop skills such as self-regulated learning (Earl, 2013; Nicol

& MacFarlane-Dick, 2006; Wiliam, 2014), and metacognition (Clarke, 2014; Earl, 2013) which, in turn, can improve students' ability to reflect critically on their own learning, identify the next steps in that learning (Hayward, 2012; Williams, 2010), and ultimately take increased responsibility for it (Hayward, Priestly & Young, 2004). Student engagement is integral to academic achievement (e.g., Appleton et al., 2008; Perdue et al., 2009) and students' attitudes towards mathematics, their motivation to do mathematics and their self-confidence regarding mathematics can directly influence their mathematical achievement (e.g. Duerr & Harlow, 2013; Lim & Chapman, 2013; Mullis et al., 2012; Stankovet al., 2012).

Consequently, the question here is whether, in the current study, the use of AfL strategies and techniques enhanced children's mathematical confidence or improved their engagement with, and attitudes to, mathematics. In short, did the use of AfL practices leverage change, and, if so, can the data that follow provide empirical evidence to substantiate any claims made? Both qualitative and quantitative data were used to decide whether to accept or reject the second research hypothesis. The quantitative data were collected by using two instruments: the Attitudes to Mathematics Questionnaire (ATMQ) and the Children's Assessment for Learning Audit Instrument (CAfLAi), both developed by the researcher specifically for the current study. The qualitative data included transcripts and video from focus group interviews, teachers' learning logs (TLL), students' learning logs (LL) and the researcher's journal. In addition to aiding triangulation, these data supplemented the quantitative data by facilitating a more in-depth analysis of students' views about using AfL strategies and techniques in their learning of mathematics. The next two sections detail the results, analysis and discussion of the quantitative data pertaining to the second research hypothesis, specifically data from the ATMQ and subsequently the CAfLAi. When reading the sections that follow, it is important to bear in mind that some researchers have raised concerns regarding the reliability of self-reported data (e.g. Lysaght & O'Leary, 2013), while others have highlighted that school children may comply with the wishes of authority figures in order to please (e.g., Coyne, 2010) and so, both these factors must also be taken into consideration when interpreting the findings from the analysis. Additionally, the possibility of the "Hawthorne effect", where participants react positively to the "experiment" since they realise they are being studied, must also be borne in mind (Cohen et al., 2010; Robson, 2011), as should issues such as the reliability and validity of both instruments.

Results and analysis of the Attitudes To Mathematics Questionnaire (ATMQ)

As detailed in Chapter Three, I was unable to locate an appropriate instrument that measured primary school children's attitudes to mathematics in the Irish context and so developed the ATMQ. This instrument was designed specifically to capture whether the use of AfL practices impacted children's engagement with mathematics, and, in particular, if it affected children's attitudes towards mathematics, their self-confidence regarding mathematics and their motivation to do mathematics. Data from each of the three scales of the ATMQ (pre- and post-), ATMQ-TIMSS, Self-Confidence in Learning Mathematics (ATMQ-SCLM) and Motivation to do Mathematics (ATMQ-MOT), are summarised in Tables 18, 20 and 21 and presented and discussed forthwith.

ATMQ-TIMSS

Beginning with the TIMSS scale of the ATMQ, a dependent (paired) samples t-test was considered appropriate because of this study's pre-test/post-test experimental design. These t-tests were conducted in order to compare the scores of the same participants at Time 1 and Time 2, to ascertain whether or not the intervention had an impact (Pallant, 2013). Results indicated that there was a statistically significant difference between pre-test ($M = 2.03$, $SD = 0.73$) and post-test scores ($M = 1.56$, $SD = 0.42$; $t(49) = 5.09$, $p < .001$) for the ATMQ-TIMSS scale, i.e. post-test scores indicated more positive attitudes towards mathematics (mean values in post-test scores were closer to 1 = *Agree a lot*). The magnitude of the difference between the pre- and post-test means can be interpreted as being large (eta squared = .35). This indicates that 35% of the variance in students' self-confidence and positive affect towards mathematics (measured by this scale) can be explained by the variables used in the model. Accepting the fact that this is not a true experimental design and that there are other confounding variables at play, for example time, it is encouraging to note that in all likelihood the use of AfL practices in mathematics positively impacted students' attitudes towards mathematics. Table 18 presents combined *agree* percentages for pre- and post-results for each statement in the ATMQ-TIMSS scale.

Table 18

ATMQ-TIMSS Scale

	Statement	Subscale	n	Combined % Agreeing PRE-TEST	Combined % Agreeing POST-TEST
a	I usually do well in Maths	SCM	50	76	96
b	I would like to do more Maths in school	*	50	60	70
c	*Maths is harder for me than for most other students in my class*	SCM	50	28	14
d	I enjoy learning Maths	PATM	50	80	98
e	*I am not good at Maths*	SCM	50	26	6
f	I learn things quickly in Maths	SCM	50	58	70
g	*Maths is boring*	PATM	50	26	6
h	I like Maths	PATM	50	80	100

Note. % Agreeing = Agree a lot +Agree a little

Italicised text highlights statements that were recoded (c, e, g).

For discussion purposes, percentage scores for the three PATM statements and the four SCM statements were amalgamated to provide a composite percentage score for these scales which could then be compared with TIMSS data.

One child did not complete this questionnaire, hence n=50 for ATMQ.

*Regarding statement b, Clerkin (personal communciation, April 15, 2015) suggests it was probably originally intended to be part of the PATM scale but following factor analysis was found not to represent positive affect in the same way as the other scale items and so was excluded.

A more detailed exploration of the eight statements from the ATMQ-TIMSS scale revealed that following the intervention the combined *agree* percentages scores for the five positively-worded statements (a, b, d, f, h) had increased by between 12 and 20 percentage points while the percentage of students agreeing with the three negatively-worded statements (c, e, g) had decreased by between 14 and 20 percentage points. This suggests that children believed the intervention had positively impacted their general attitudes towards mathematics and their self-confidence in learning mathematics. Specifically, regarding the three statements which measured students' general affect towards mathematics (d, g, h), results indicated that following the intervention almost 100% of the participants *agreed a little* or *a lot* with these three statements (Table 19).

Table 19

PATM Scale Data Comparisons (Intervention and National Data from TIMSS, 2011)

	Statement	**Subscale**	**TIMSS 2011 Irish Girls non-DEIS**	**Pre-**	**Post-**
d	I enjoy learning maths	PATM	80	80	98
g	*Maths is boring*	PATM	32	26	6
h	I like maths	PATM	79	80	100

Note. % Agreeing = Agree a lot +Agree a little

Italicised text highlights statements that were recoded

This result is significantly higher than the 72% of all 4th grade students who agreed a little or a lot with these same statements in TIMSS 2007 (Mullis, Martin & Foy, 2008). Ireland did not participate in that study but the three statements were included again in the 4th grade student questionnaire in TIMSS 2011, when Ireland did participate. In that study, 79% of girls surveyed from fourth class, in non-DEIS schools similar to Scoil na nAingeal, liked or somewhat liked mathematics (Clerkin & Creaven, 2013). Therefore, pre-intervention, scores for participants in this study (i.e. 80%) were similar to Irish findings from TIMSS (2011) for a similar cohort for the statement ‘I like mathematics’, while post-intervention 100% of the participants said they liked or somewhat liked mathematics, a significant increase. As can be seen from Table 19, regarding the other two statements from the PATM scale (d, g), pre-intervention percentages are also similar to those from TIMSS 2011 for a similar cohort nationally (Clerkin, personal correspondence, April 15th, 2015), while post-intervention percentages have significantly increased, indicating that, following the intervention, students’ attitudes towards mathematics were more positive.

ATMQ-SCLM

In TIMSS 2007, at fourth grade, students expressed reasonable self-confidence in their mathematical ability, with 57% agreeing a little or a lot with all four statements (a, c, e, f) in the Index measuring Students’ Self-Confidence in Mathematics (Mullis et al.,2008). This is slightly below the figure of 70% for students in this study who, prior to the intervention, agreed a little or a lot with the same four statements (percentage score were obtained by averaging scores for the four statements) and is significantly lower than the 87% who agreed with these statements afterwards. No comparable statements were utilised in TIMSS 2011.

In an analysis of data from TIMSS 2007, Olsen, Martin and Mullis (2008) discovered that the statement “I like mathematics” had the highest factor loading for the latent factor PATM and therefore the strongest relationship with it, while the statement “I usually do well in mathematics” had the highest factor loading and strongest relationship for the latent factor SCM. These two statements also received the highest combined *agree* percentage scores in this study for these two factors post-intervention, at 100% and 96% respectively, and were two of the areas of greatest change in this scale. In sum, analysis of data from the ATMQ-TIMSS scale would seem to indicate that, by the end of the intervention, children believed their self-confidence regarding mathematics had increased and their general attitudes towards mathematics were more positive.

A dependent (paired) samples t-test was conducted to compare pre- and post-test scores of the 51 students in the intervention group for the ATMQ-SCLM scale. Results indicated that there was a statistically significant increase in scores between the pre-test ($M = 2.27$, $SD = 0.63$) and the post-test ($M = 1.59$, $SD = 0.40$; $t(49) = 8.36$, $p = <.001$) i.e. mean values in post-test scores for the ATMQ-SCLM were closer to 1= *Agree a lot*. It should also be noted that the magnitude of the difference could be considered large (eta squared = .60). Table 20 presents frequencies for the pre- and post-test data from this scale.

Table 20

ATMQ Self-Confidence in Learning Mathematics Scale (ATMQ-SCLM))

	Statement	**n**	**Combined % Agreeing PRE-TEST**	**Combined % Agreeing POST-TEST**
a	*Maths is my least favourite subject*	50	40	40
b	*My mind goes blank when doing maths*	50	34	18
c	*Doing maths makes me feel nervous*	50	42	10
d	*When I hear the word maths I start to daydream*	50	30	4
e	*It makes me nervous to even have to think about doing a maths problem*	50	40	10
f	Maths does not scare me	50	66	94
g	I am very confident when it comes to maths	50	62	90
h	Solving maths problems is easy for me	50	52	68
i	I expect to do fairly well in most of my maths classes	50	78	94
j	*I get confused in my maths class*	50	64	14
k	I learn maths easily	50	46	84
l	I think I am good at solving maths problems	50	58	78

Note. % Agreeing = Agree a lot +Agree a little

Italicised text highlights statements that were recoded

Analysis of the results displayed in Table 20 (ATMQ-SCLM) revealed that the biggest combined 'agree' percentage increase/decrease was in statement j: "I get confused in my maths class". Before the intervention, 64% of students felt that they got confused in their mathematics classes, in contrast to just 14% after the intervention. This would appear to confirm findings in other studies (e.g. Collins & O'Leary, 2010; Wiliam 2011b), that when teachers use AfL strategies and techniques effectively, particularly when they share learning intentions and success criteria, then students are clearer regarding what it is they are supposed to be learning and therefore less confused. A significant difference in the percentage of students who agreed a lot or a little with the statement "I learn maths easily" was also evident, increasing from 46% pre-intervention to 84% afterwards. Answers to statements c, e and f, relating to feeling nervous about doing mathematics, indicated that after the intervention students felt less anxious when doing mathematics, resulting in a combined *agree* difference of between 28-32% for these statements. Interestingly, following the intervention, there was no change in *agree* percentage scores for the first statement of this scale, "Maths is my least favourite subject", which appears at odds with scores from the final statement of the ATMQ 'TIMSS' scale "I like mathematics", which had increased significantly. Notwithstanding, by the end of the intervention, 90% of students agreed a little or a lot that they were very confident when it comes to mathematics (Statement g), which seems to confirm findings from the rest of this ATMQ-SCLM scale, and indeed from the SCM index of the ATMQ-TIMSS scale, previously discussed. Furthermore, these findings, suggesting students are less confused, less nervous and like mathematics more, are also confirmed by qualitative data, which will be discussed later in this chapter.

ATMQ-MOT

Finally, a dependent (paired) samples t-test was conducted to compare the pre- and post-test scores for the Attitudes to mathematics motivation (ATMQ-MOT) scale. Results indicated that there was a statistically significant increase in motivation scores between pre-test ($M = 2.04$, $SD = 0.53$) and post-test scores ($M = 1.68$, $SD = 0.34$; $t(49)= 4.91$, $p = <.001$). The magnitude of the difference can be considered large (eta squared = .34). Table 21 summarises combined ‘agree’ statistics for pre- and post-test results for each item in the ATMQ-MOT scale.

Table 21

ATMQ Motivation Scale (ATMQ-MOT)

	Statement	**n**	**Combined % Agreeing PRE-TEST**	**Combined % Agreeing POST-TEST**
a	I am confident that I could learn difficult maths	50	22	80
b	*I would like to avoid using maths in secondary school*	50	30	14
c	I would be willing to do extra maths	49	59	54
d	When I get into secondary school I would love to do extra maths	50	32	62
e	I think maths is fun because you have to figure things out	50	68	96
f	I think I can do even the hardest maths if I keep trying different ways to find the answer	50	64	92
g	*If I'm not one of the best in my maths class then I don't try at all*	49	6	0
h	I try my best at maths because I want to learn new things	50	92	100
i	I try my best at maths when there's a reward	50	82	66
j	*If I find maths difficult I give up straight away*	50	14	2

Note. % Agreeing = Agree a lot +Agree a little

Italicised text highlights statements that were recoded

Statement (a) of the ATMQ-MOT scale "I am confident that I could learn difficult maths" resulted in the biggest percentage increase of any item in the ATMQ instrument. Only 22% of students agreed with this statement prior to the intervention, whereas 80% agreed with it afterwards. Furthermore, by the end of the intervention 100% of students indicated that they would try their best at mathematics because they want to learn new things (statement h), while the percentage of students who revealed that they try their best when there is a reward decreased from 82% to 66% (statement i). This appears to suggest that, following the intervention students felt motivated more by intrinsic rather than extrinsic factors. Similarly, statement j "If I find maths difficult I give up straight away" indicates that there was an increase of 12% in participants who were more motivated to persevere, even if the mathematics proved difficult. Statement c 'I would be willing to do extra maths' is the only item of the AMTQ instrument where there was a decrease in *agree* percentage scores and that by only 5%. This could be attributed to the fact that due to the *Literacy and Numeracy Strategy* (DES, 2011a), students were timetabled to spend more time doing mathematics than they had in third class.

In conclusion, while once again acknowledging the fact that this is not a true experimental design and that other confounding factors may have been at play, nevertheless, given how carefully the intervention was designed and implemented, in all likelihood the use of AfL practices did positively impact students' attitudes towards mathematics. Furthermore, the quantitative data from the three scales of the AMTQ seem to indicate that following the intervention, children believed that their attitudes towards mathematics were substantially more positive; they felt more confident about doing mathematics, and they thought their engagement with, and motivation to do mathematics had improved.

Results and analysis of Children's Assessment for Learning Audit Instrument (CAfLAi)

As mentioned in the previous chapter, the CAfLAi was developed by the researcher to provide baseline data regarding children's use and understanding of AfL strategies and techniques and to measure changes, if any, that had occurred following their participation in the intervention. The CAfLAi contains the same four scales as the teachers' assessment for learning audit Instrument (AfLAi) created by Lysaght and O'Leary (2013): sharing learning intentions and success criteria (LISC), questioning and classroom discussion (QCD), feedback (FB) and peer- and self-assessment (PSA), plus an extra scale labelled techniques (TQ). The first four scales measuring AfL strategies contain five items each, while the final scale measuring techniques contains twenty. Although the composite score for the four scales measuring the strategies is reliable (Cronbach's alpha of .76), reliability is low for each of the following individual scales: LISC, QCD, and PSA scales (Table 9). While, as Gray (2014) suggests, the low level of these Cronbach's coefficients may have been affected by the fact there are fewer than 10 items in each scale, given that the scores are significantly lower than the usually accepted alpha reliability threshold of 0.70 (Cohen et al., 2010), findings from the CAfLAi and, in particular, the LISC, QCD and PSA scales, must be interpreted with caution. All children from the intervention group (n=51) completed the CAfLAi, pre- and post-intervention, the data were coded, collated into an Excel file and subsequently inputted into SPSS 21 for analysis.

Dependent (paired) samples t-tests were conducted to compare the pre- and post-test scores of all CAfLAi scales. It should be noted that the closer the mean ratings are to 5, the more likely it is that the AfL strategy *always* happens, while ratings closer to 2 suggest it *never* happens. Beginning with LISC, results indicated that there was a statistically significant increase in scores between the pre-test ($M=3.32$, $SD=0.41$) and post-test ($M=4.06$, $SD=0.49$; $t(50)= -10.03$, $p=< .001$) i.e. the mean value in post-test scores was closer to 4=*Often*. Thus, the magnitude of the difference in the means could be considered large (eta squared = .67). With regard to QCD, there was also a statistically significant increase between the pre-test ($M=3.05$, $SD=0.36$) and post-test scores ($M=3.25$, $SD=0.42$; $t(50)= -2.87$, $p=<.006$) i.e. the mean value in post-test scores was closer to 3=*sometimes*. It should also be noted that the magnitude of the difference in the means for QCD could be considered large (eta squared = .14). The results of the dependent (paired) samples t-test for the pre- and post-test scores for FB indicated that there was a statistically significant increase in scores from pre-test ($M=3.55$, $SD=0.59$) to post-test ($M=4.29$, $SD=0.47$; $t(50)=-8.41$, $p=<.001$) i.e. the mean value in post-test scores was closer to 4=*often*. Once again, the magnitude of the difference between the means could be considered large (eta squared =.59). Results for the PSA scores indicated that there was a statistically significant increase in scores between the pre-test ($M=3.11$, $SD=0.45$) and post-test ($M=3.89$, $SD=0.47$; $t(50)=-7.942$, $p=<.001$) i.e. the mean value in post-test scores was closer to 4=*often*. The magnitude of the difference between pre- and post-test scores for PSA was also large (eta squared =.56). Finally, results from the dependent (paired) sample t-test comparing the pre- and post-test scores of TQs show that there was a statistically significant increase in scores between the pre-test ($M=1.87$, $SD=0.37$) and post-test ($M=3.47$, $SD=0.36$; $t(44)=-22.188$, $p=<.001$) i.e. the mean value in post-test scores was closer to 3=*Sometimes*. The magnitude of the difference between pre- and post-test scores for TQs could again be considered large (eta squared = .91).

With regard to the four AfL strategies, the data suggest that over the course of the intervention, the children believed that the most statistically significant change in their practice of AfL took place in the LISC strategy, changing from *sometimes* to *often* (eta squared= .67). In contrast, the data indicate the children considered that the least statistically significant difference occurred in QCD. Mean ratings here were also lowest both before (M= 3.05=*Sometimes*) and after the intervention (M=3.25=*Sometimes*). This finding is interesting as many teachers, those in this study included, assert that QCD is an integral part of their classroom practice. Nevertheless, the data here appear to confirm findings by Lysaght and O'Leary (2013), albeit from the teachers' perspective, that *democratised* QCD, with pupil-led as well as teacher-led approaches, is, as yet, not well established in Irish classrooms. According to the data, students also believed that over the course of the project, a statistically significant change had taken place in their use of PSA, changing from *sometimes* and getting closer to *often*. Reliabilities for the FB scale of the CAfLAi are acceptable so it seems feasible to explore data from this scale a little further (Table 22).

Table 22

Average Ratings for the CAfLAi Feedback Scale (Pre- and Post-Intervention)

	Statement	n	Pre-test X̄(sd)	Post-test X̄(sd)	Diff. in mean score	Eta Squared
1	I discuss with my teacher how to improve my work as I go along	51	2.96 (.66)	3.57 (.75)	0.61	.30
2	I think that the teacher really listens to what I have to say about my work because of the comments she makes and the questions she asks me about it	51	3.75 (.96)	4.53 (.88)	0.78	.33
3	When the teacher paises my work she explains what was good about it	51	3.65 (.80)	4.69 (.62)	1.04	.51
4	When my teacher looks at my work she uses different ways to show me what I have done well and where I might improve	51	3.65 (.84)	4.39 (.75)	0.74	.31
5	I make changes to my work after the teacher corrects it so that I learn from my mistakes and my work improves	51	3.75 (.91)	4.29 (.94)	0.54	.18

The post-test data in Table 22 suggest that children in the intervention group believed FB was used regularly in their classrooms to guide teaching and learning. Following the intervention, children thought they discussed more regularly how they might improve their work with their teacher (statement 1). Moreover, students also felt that teachers listened more carefully to what they had to say about their work (statement 2). Additionally, there is consistency between the data from statements 3 and 4 where students are responding about teacher-feedback, especially regarding what students did well; and this is the area where children reported the intervention had greatest impact (Cohen's d=0.59). Statement 5, which explored students' use of teacher-FB to guide the next steps in their learning, showed least change following the intervention and revealed greatest variability in how children responded (SD=0.94). Making effective use of feedback is central to the process of successful FA and pivotal in the development of SRL (e.g. Sadler, 1989; Wiliam, 2011b). Children reported that this practice happens *often*. It is interesting to note that, while the magnitude of the difference between pre- and post-test mean scores for the statements regarding teacher-FB to students was medium, there wasn't a corresponding increase in students' application of these suggestions to their work (statement 5). Overall, following the intervention, FB practices were reported by the children to be happening on a more regular basis. Notwithstanding, in retrospect, it may have been an oversight not to have included a statement in this scale that questioned children regarding student-led FB practices, an area highlighted by Lysaght and O'Leary (2013) as needing attention in the Irish context.

Table 23 presents the overall mean rating for each of the five scales of the CAfLAi in rank order (post-intervention), beginning with the most embedded. The average ratings suggest that following the intervention the children viewed the first three strategies, FB, LISC and PSA, as happening *often* in their classrooms (average close to 4), while the TQs and QCD happened *sometimes*. Students' responses to statements for the various scales were reasonably consistent, with all scores within 0.5 standard deviations of the mean. Prior to the intervention, average ratings for each scale were lower but the first three strategies had still been placed in the same rank order. Children considered that FB was the most embedded strategy both pre- and post-intervention while, following the intervention, children's responses to the TQ scale showed the biggest mean increase, suggesting that many of the techniques were either new to the children or they were coming from a low base. It has been argued that techniques should be at the heart of changes in classroom assessment practice since it is by using AfL techniques that the AfL strategies become embedded (Lysaght & O'Leary, 2013; Wiliam, 2011b). Therefore, the apparent increase in student use of AfL techniques in this study could be indicative of changing assessment practices in the classrooms being studied.

Table 23

Post-Intervention CAfLAi Scale Comparisons

	Scale	n	Mean	SD	
1	FB	51	4.29	0.47	*Often*
2	LISC	51	4.07	0.49	*Often*
3	PSA	51	3.89	0.47	*Often*
4	TQ	51	3.47	0.36	*Sometimes*
5	QCD	51	3.25	0.42	*Sometimes*

Although results from the CAfLAi t-tests indicated there was a statistically significant difference in pre- and post-test scores for all scales, with large eta squared values, it is necessary to be mindful of the limitations of this instrument and the aforementioned caveats, and therefore prudent to draw only tentative conclusions here. While these data appear positive and reflect opinions expressed by the children in their learning logs and the focus group interviews, to be discussed in the next section, the data provide little more than a yardstick or approximate measure of the impact of the intervention on children's AfL practices.

In sum, with regard to the quantitative data pertaining to the second hypothesis, the data appear to indicate that the children believed their use of AfL strategies and techniques had increased by the end of the intervention. Additionally, post-test quantitative data also suggested that following the intervention children considered that their engagement with, and attitudes to, mathematics had improved, as had their mathematical confidence. Attention now turns to an exploration of the qualitative data relating to the second research hypothesis.

Qualitative Findings and Analysis

The qualitative data pertaining to the second hypothesis were transcribed and preliminary analysis was undertaken through multiple close readings of the transcripts. Additionally, video recordings of the focus groups (FG) were watched multiple times, in an attempt to capture things that could not be gleaned solely from the written word, e.g. facial expressions or group dynamics. It should be noted that teachers did not correct children's entries in their learning logs (LL) and consequently there are times when dates have been omitted and when children wrote comments which sometimes appear unclear or are without elaboration. Through meticulous and repeated scrutiny of the transcripts, children's constructs, ideas and opinions were analysed, interpreted and grouped into thematic units. Most of these themes can be seen as directly related to the central issues of this thesis. The following five main themes were identified:

1. Enjoying the AfL journey;
2. Growing positivity and self-confidence in mathematics;
3. A changed classroom dynamic;
4. Peer- and self-assessment: - a highlight for children;
5. Unexpected insights.

The findings in relation to each of these themes are now presented, analysed and discussed sequentially.

Enjoying the AfL journey

Quantitative findings from the CAfLAi, discussed earlier, suggested that following the intervention there was a statistically significant increase in children's use of AfL strategies and techniques. Analysis of the qualitative data set for research hypothesis two confirmed this and indicated that by the end of the intervention students had also developed quite a good understanding of these AfL practices. Not only had the children become familiar with different strategies and myriad techniques, they had also obviously reflected on these practices, internalised them and appropriated them for their own purposes. The children were able to use the language of AfL, interchanging child-friendly acronyms such as WALT and WILF for more *adult* terms such as learning intentions and success criteria although, at times, some were confused between the terms "strategy" and "technique".

There was considerable evidence from the qualitative data set that the children enjoyed their participation in the intervention. The following comment from Amelia, a student with EAL, is typical of the views they expressed:

> Maths this year was fun for me because we had to do fist to five and *thumbs* up and *traffic lights*. I most liked the traffic lights because you can tell about your work, and the teacher would know how you did, and then she can tell more. (FG2)

When describing their experience of the intervention, the children regularly used words such as "fun" and "enjoyment". It was evident from early on that not only did they like learning about, and using, the various AfL practices during mathematics lessons but they understood their value and effects, e.g. "I was surprised by how

WILFs and WALTs make maths so much more useful and clearer. After this session I feel a lot more confident and smart" (Mia, LL, 08/11/2012). One child even wrote the following in her LL: "I think parents would like these strategies, because they will know that their children can do maths well and like doing maths" (Emily, LL, n.d.). Throughout the intervention, the children's self-reported positivity regarding their AfL journey was also apparent to the three participating teachers and was described by me in the following excerpt from my teacher learning log (TLL):

> It is really evident from the children's learning logs and from talking to them that they are enjoying using AfL. They love every time I introduce a new technique as they think it's just a fun way of learning. They even use them in other subjects too. (13/02/2013)

In sum, the children's enthusiasm regarding using AfL practices in mathematics can be succinctly summarised by Maria's comment: "I love doing AfL and I would like to continue doing it" (LL, 12/06/2013).

In addition to enjoying using AfL practices the children also appeared to have developed a good understanding of the various AfL strategies and techniques, which this section now discusses. The children did not question the importance of the teacher's role in the learning and teaching process but, similar to other studies (e.g., Hayward, 2012), welcomed the use of AfL strategies to enhance their learning. The following comment from Mia's LL exemplifies the children's understanding of LISC: "WALT helps me realise what we're really learning today. WILF is what makes me successful and helps me to break things down" (12/06/2013). She later explained during the FG interview, "with WALT and WILF up on the board, you'll be able to check while you're going along". Other children similarly mentioned the benefits of having the LISC on display throughout the lesson: "Maths this year seems more clearer to me. I know exactly what to do" (Kate, LL, 31/01/2013). This notion that using LISC lessened children's confusion in mathematics lessons was prevalent throughout the qualitative data, possibly explaining why participating teachers, as with Collins and O'Leary (2010), were asked fewer procedural questions, observed increased focus on tasks and witnessed more independent learning.

Similar to the teachers participating in the intervention, questioning and classroom discussion (QCD) was the strategy least written or spoken about by the children, perhaps because their experience of QCD did not differ substantially from previous years and so was not particularly noteworthy. In contrast, the children frequently wrote about feedback in their learning logs and discussed it in some depth in both focus groups, e.g.,:

> With FB the teacher would tell you, would show you where you went wrong, and then you could learn from that and say "ok, for the next time we're doing it, I know now what to do. (Sarah, FG2)

The children acknowledged that feedback could legitimately come from different sources, as Amelia remarked: "not only the teachers can like give you feedback, but like when they [classmates] give me feedback, I like more improve as well, not only when the teachers help me" (FG2). They seemed to enjoy receiving FB as these comments exemplify: "I just like getting FB 'cause…it shows you where you are in your learning" (Hollie, FG1) or, as Sophie explained, "you want to know what's going on and where you are getting it wrong" (FG1). They also liked giving FB to each other. Emily's comment was typical:

> I like when you are giving feedback because the people mightn't even notice it themselves…but when you do it, you kind of just look through it all over again, and you see the mistakes they haven't seen. (FG2)

Lily even suggested that getting FB from your friend might be advantageous:

> Well maybe if you get feedback from your friend, like say if you were doing peer-assessment, they might give you, they might write something that you wouldn't write, like you might just want to keep telling yourself that it's ok and it's good, but they can give you feedback and you can improve your work the next time, and make sure that it's how it should be, and not just be saying

> it to make yourself feel that you think it is good, but you know that there are a few things that aren't really good. (FG2)

She also highlighted the utility of giving herself FB:

> I liked using two stars and a wish, because when you're doing your wish, it's kind of like giving yourself feedback and, like you're judging your own work so the next time you can look back at what your wish was the last time, and you can see if you have that done. (FG2)

Nevertheless, most of the children expressed the view that they found it easier to give FB if they were given specific guidelines on how to do so. The following comment was typical: "I like giving feedback, especially the rubric, because it actually shows you" (Kate, FG2). Maria concurred stating:

> I like giving FB because of the rubric, and it's much easier, and when you look at somebody else's work and then you look at your work you can compare them and see which one is better, and see what you can improve on, and then you could see what they could improve on and just tell them. (FG2)

Some children also explained how they used the FB they received to improve their learning: "Feedback has helped me because I learn from my mistakes and I'll remember it the next time I'm doing a problem etc." (Sarah, LL, 12/06/2013) or

> If you were doing a test, and say you were doing time and didn't carry sixty over to the minutes, and then the teacher or peers or a person asked, told you in your feedback to remember to do that next time you do a test, then you'd remember that and that's the next step in your learning. (Ruby, FG1)

The children's understanding of PSA was also clear from the data and appeared to be the highlight of the intervention for many of them. Consequently, PSA will be discussed in due course as a theme in its own right.

Over twenty new AfL techniques had been introduced to the children on a phased basis during the intervention and some examples of these are presented in Appendix X. The children loved learning new techniques and mentioned them frequently in their learning logs and in the focus group interviews, e.g.

> I think they're really good, because they really help you and it makes maths more fun, because before I didn't like maths and then we started using all these techniques and then I said "oh, maybe this will be fun". (Hollie, FG1)

Use of these techniques proved helpful in embedding AfL practices and greatly contributed to the children's enjoyment of the AfL journey. Some typical responses which demonstrate the children's positivity towards a few sample AfL techniques are quoted here:

> I liked the *two stars and a wish* in the assessment for learning because you see the wish is how you'd make your work better and that would be the main point in the assessment for learning and the *two stars* are just things that you think you did good. (Kate, FG2)

> I rather maths this year because I think it's easier because there's more strategies and the teacher can tell when you know what you're doing and when you don't with *thumbs* and *fist to five*, because otherwise you mightn't want to tell the teacher if you didn't do well, so I think it's easier this year. (Maria, FG2)

The following comment makes explicit the views of some children who believed that the techniques also helped teachers ascertain how children were doing:

> Well, if you're doing *fist to five*, the teacher just says "show me *fist to five"* instead of going around and asking everybody or checking their work. Like if you put up five it means that you're doing well. Say if there was like a lot of fours, it would mean that the teacher would probably have to go over it a bit better, and if there were threes, she'd have to do the lesson again, and if there was twos she'd have to like really explain it so that they'd understand it, like she'd have to try and explain it a way that they'd understand it a bit more. (Lily, FG2)

Another child suggested that using AfL techniques helped motivate fellow students to take greater interest in their learning:

> With the *lollipop sticks*, I think they're good, because usually in maths it could be the same hands up all the time, and with the *lollipop sticks*, it's whoever teacher picks out, so everyone has to know the answer, because there could be a person at the back of the room who never puts her hand up. (Sarah, FG2)

However, Chloe cautioned:

> With the *think-pair-share*, you have to be careful when you're doing it because your friend might be relaxing and then she'll let you do all the work, but you have to kind of sit back yourself and let her do a few sums as well, you have to make sure that she's working. (FG1)

By the end of the intervention many of the children recommended that more teachers and students should use AfL techniques e.g.

> I liked it all, how we used the techniques and that it actually helped people and I think that teachers should actually use them in older classes and in younger classes because it really helps in maths class. (Emily, FG2)

Although the intervention with the children focused specifically on mathematics lessons, it became clear that the children had appropriated the AfL strategies and techniques in various ways: "Yeah, I use it in all my other work as well. I don't just use it in maths, so it helps me in every work, all the pieces of work that I do in school" (Ruby, FG1). Additionally, during the FG interviews many of the children revealed they regularly remembered the LISC from earlier in the day as an aid to helping them to do better homework. Emma's comment was typical: "I think this year was way better because if you got maths homework, then by using all them techniques, that could help you do your homework well and help you learn in your work" (FG1).

All in all, it was clear that the children had enjoyed their AfL journey and had readily embraced AfL practices in their learning of mathematics. In particular, they linked their use of AfL with an increased enjoyment of mathematics, the development of more positive attitudes towards mathematics, and enhanced confidence regarding mathematics.

Growing positivity and increased self-confidence in mathematics

Quantitative findings from the ATMQ, discussed earlier, suggested that children's engagement with and attitudes towards mathematics had changed substantially since the start of the intervention and in a positive way. These findings are corroborated by much of what was revealed from the qualitative data discussed forthwith. The children's attitudes towards mathematics seemed underpinned by their enjoyment of mathematics and many of them equated mathematics with *fun* e.g. "Each day I look forward to maths because it's fun and you learn new things" (Lily, LL, 31/01/2013). For many students, this *fun* and their increased enjoyment of mathematics was linked to their use of AfL strategies and techniques. Amelia's comment was typical: "This year in maths I actually enjoyed it because you had a lot of strategies to do and they actually helped me" (FG2). Some children also believed that using AfL practices made mathematics easier and increased their liking of mathematics: "It makes maths so much more fun, 'cause like in third class I used to hate maths and then like now that there's all these different strategies, it just makes maths so much easier" (Sophie, LL, n.d.). Others, like Chloe, believed that, in addition to making mathematics easier, using AfL lessened their anxiety about mathematics:

> In the other classes before…I didn't want to do maths 'cause I didn't like it really, then this year I can't wait to do maths when I go into the classroom. Maybe it's easier…it's much easier with the WALT and the WILF that you're not afraid to do it…like I used to be afraid of doing maths because I'd think "oh I suppose I'm going to get something wrong". (FG1)

Hollie, a pupil with SEN, also spoke about this issue during the FG interview and, echoing research by Stiggins (2006), linked using AfL with her increased confidence regarding mathematics:

> Well all the years when I was doing maths, I used to get really nervous when I was doing it, because I didn't think that I was able to do well enough and now I think I'm able to do…now I think I'm more confident and all the strategies and techniques really helped me…like WALT and WILF helped me…by saying what we're learning to do and what's the next step in my learning. (FG1)

Interestingly, at the end of this interview when the I asked the children if there was anything they would like to add or ask, Hollie stated the following:

> When I go into 5th class, if we don't really do all these things [AfL strategies and techniques] maybe I won't be as confident, because I think that all of them just helped me up my game, but if I don't really do them all the time, then maybe I'll be bumped back down, but I think that I'll still carry them with me, even though we might not do them. (FG1)

Mia, a student from the middle stream (whose parents had brought up the issue of her lack of confidence regarding mathematics at a PT-meeting earlier in the year) then added: "I agree with Hollie, that even if we don't do them next year, I'll still bring them, but if we don't I'll probably lose a bit of my confidence because they helped me a lot." Indeed, participating in the intervention did appear to help both these students as they attained higher scores in the SIGMA-T and DPMT-R than previous years. Notwithstanding, because of the small sample size, one cannot draw a definitive conclusion here that links positive attitude towards mathematics with mathematical achievement. Furthermore, while most students standardised test scores had increased post-intervention, some had not, including students who professed a more positive attitudes towards mathematics and who believed their mathematics had improved. Nevertheless, it is clear that students themselves reported a growing positivity and increased self-confidence in mathematics as a result of using AfL practices. Various studies (Hattie, 2009; Ma, 1999; Mullis et al., 2012) have linked positive attitudes towards mathematics with increased achievement. Therefore, it is hoped that findings from this study, which indicate an increase in students' positivity towards mathematics following the intervention, will in time result in increased mathematical achievement.

Regarding mathematics, another issue that some children addressed in their LLs and FG interviews was their belief that using AfL enhanced their ability to solve problems. Emma wrote the following in her LL, “This year problems are way more clear to me. I never really liked to do problems but I think the techniques help a lot, but I still need a bit more help on them” (31/01/2013), while Sarah expressed the following opinion at the FG interview:

> Well I didn’t really like maths at the start either but now I really do like it, and I think the FB also helped me, because if say on a problem I got something wrong and I did division instead of multiplication, the next time I’d know to do multiplication, so it helped me for the next time I’d do a problem. (FG2)

In sum, children’s growing positivity and increased confidence in mathematics seemed intertwined with their use of AfL strategies and techniques.

A changed classroom dynamic

Although the intended focus of this intervention was on AfL, LS and mathematics, ultimately it was about learning and teaching and what was actually happening in each classroom. If then, as some AfL experts suggest (Marshall & Drummond, 2006; Hayward, 2012), engaging in AfL practices in the *spirit* in which it is meant impacts learning and teaching and acts as a leverage for change, then *ipso facto* changes in classroom practices and in the roles and relationships therein should have been evident to participating students and teachers alike. Some children identified various instances where the teachers used AfL strategies and techniques to support their teaching. Emily gave the following example:

> I like using all the different techniques and everything, because it actually makes the work easier, but, with the *ABCD cards* I thought they really made

> the teacher see how people know what to do, and if they got the right answer, because some people, they mightn't put up their hands to answer questions, and when the teacher said "put up the ABCD cards", you have to put up one, and then teacher will know the answer, but if you get it wrong, she'll have to know where you got it wrong and how you thought it was the right way. (FG2)

Another student, Ruby, explained the benefits of using *fist to five*:

> Let's just say you're doing *fist to five* or something, and the teacher would say to you "okay, do fist to five" and then, say we had some fives, some threes and some twos, the teacher would say "okay, were not really ready to move on yet", so that she'd explain and then everybody would know it more and be able to do it in a test, or they'd be able to do it in their copies in the future. (FG1)

Several students also referred to teachers as learners:

> I'd say the teachers learned a lot this year as well, because like we're doing something and we really like it. That means you [researcher] learned from it, that you should do it again, like to the other classes, that you'll have next year. (Amelia, FG2)

In particular, children linked the idea of teacher as learner with the LS process e.g.

> It's just like we're playing the piano and they're the examiners, just the same as we're doing maths and they're looking at us. I think they might have been checking what we've learned, and how you've [researcher] learned us, because say if they were just looking at one pair, they'd be figuring out how do they do multiplication or division, when you're working it out, to see if you do it right or wrong. (Kate, FG2)

Nevertheless, echoing Hayward (2012), the children also recognised the teacher as overall guide and arbiter in the classroom as the following statement illustrates:

> With feedback the teacher would tell you, would show you where you went wrong, and then you could learn from that and say "okay for the next time we're doing it I know now what to do". (Sarah, FG2)

This student also explained that the teacher would sort any disagreements regarding peer-assessment:

> Then you have to show it to the teacher, and you [pointing to me] have to say which one, who you agreed with, whether it was the person who was peer-assessing it, or the girl who was arguing, saying she didn't get that wrong or anything.

This idea of teacher as arbiter is similar to Hayward's findings as part of the AiFL project (2008), reported in Gardner (2012b). The children agreed that the teacher was guiding the overall learning process, but recognised that they, too, had an important role to play. Several students mentioned feeling like a teacher when engaging in peer-assessment. Mia explained it thus:

> I feel like a teacher when I do it because you've the page in front of you and you get to correct them or mark them wrong and give them points, and you show them what's wrong, and with the rubric it sort of helps you like you're a teacher as well, because it's sort of feedback. (FG1)

Additionally, some children also showed increased awareness of teachers' work e.g. "Today I learned it is very hard to be a teacher correcting a maths test" (Amelia, LL, 08/11/2012) while others mentioned how "exhausting" they found it.

In this era of lifelong learning, learning how to learn or SRL is an important skill. By the end of this intervention, it was obvious that students were becoming conscious of themselves as learners. It was evident that they were, as Sophie put it "thinking way more", that they wanted to be more actively engaged in their own learning and that they enjoyed it. During the FG discussions, the children talked a lot about learning and revealed an awareness of the processes involved. Their understanding of particular AfL strategies and techniques has already been discussed, and so, this section takes a broader view and looks at their understanding of learning in general after the intervention. Ruby demonstrated that she had reflected critically on her own learning and use of AfL practices, stating:

> I think I've developed from last year, and all the new techniques that we've done, it's made it a bit easier to do maths this year, because we're doing say WALT and WILF and feedback I'm getting more things to help me to see the next step in my learning, so it's helping me this year. (FG1)

Lily also appeared to connect using AfL practices with her own learning:

> If you write down like a wish, you can use that, say if you weren't happy with what you did with one piece of work you can make sure that you do it in the next piece of work. (FG2)

For others like Emily, confirming what was discussed earlier in the quantitative section, using AfL practices lessened confusion and enhanced learning: "I thought with the WILF and WALT, it showed you what you are learning at the start of maths so that you weren't confused" (LL, n.d.). There was evidence too that for some children engaging in AfL practices was helping them become more independent learners, and that they accepted their new role and the responsibility it brought. For example, Chloe commented, "I think it's helped me, that it's not letting the teacher correct all your work, that you, you kind of have to check it, and you have to, because there's some silly mistakes that you could make" (FG1). Others displayed an increased confidence in their ability to assess, both themselves and others: "I think that when you correct it yourself, or for your friend, you know what you're correcting, you know why you're correcting, what you did wrong, so you know what you're doing" (Emma, FG1). Meanwhile Sarah pointed out that using AfL strategies meant that everyone had to engage in classroom learning and that there was no place to hide:

> The *lollipop sticks*, I think they're good because usually in maths it could be the same hands up all the time, but with the *lollipop sticks* it's whoever she picks out so everyone has to know the answer because there could be a person at the back of the room who never puts her hand up and she's sitting there not even listening and then her name comes out. (FG2)

Interestingly, perhaps due to media coverage regarding Junior Cycle reform, she subsequently commented:

> When we are in third year, we won't be doing our Junior Cert, we'll be doing the assessment for learning, so I think that it's very good that we already know what it is and we're a bit ahead of every other, all the other people. (FG2)

Additionally, Lily seemed to realise the potential impact learning about AfL might have on her future learning:

> I feel more confident about maths and say even if we were in 3rd or 6th class if we don't use the techniques anymore, I'll still remember them and they'll still help me to learn, and I can still use them like for the rest of my life. (FG2)

She also highlighted the value of collaborative learning:

> When we were doing the LS you get to work with your partner, and say if you know something that she doesn't know, you can talk about it, and if she knows something that you don't know, you can like compare, the other person can tell you about it…so what this person knows and what you know, you can put them together to try and get the right answer. (FG1)

Overall, many of the ideas expressed by the children in their LLs and the FGs are reminiscent of findings by Williams (2010) and Hayward (2012), in that the children seemed to understand the concept of AfL and could articulate their ideas about using AfL practices, as some of the previous statements illustrate. Additionally, the children were beginning to monitor their own learning and to evaluate their progress. This resulted in a growing awareness of their own learning and suggests that through the process of engaging in AfL strategies and techniques over the course of one academic year, these students were moving towards self-regulated learning:

> If we were doing a piece of work that we haven't seen before, and say you're doing a worksheet on it, and then you get feedback, I'd like that because it shows that that's the kind of unravelling you have to do the next time, so that at the end of the year you'll have it perfect. (Ruby, FG1)

Peer- and Self-Assessment - a highlight for children

Although self- and peer-assessment are distinct processes, they are discussed together here as they were combined in the teachers' and students' AfL audits and both are examples of student-generated FB. The literature review has emphasised the importance of getting students actively involved in assessment, especially through using PSA, and has highlighted the potential of using PSA to increase student self-regulation and achievement (Harris et al., 2015; Warick et al, 2015; Wiliam, 2011b). As mentioned earlier, in this study, using PSA was undoubtedly a highlight for participating children as this comment illustrates: "I love self-assessment and peer-assessment" (Hollie, LL, n.d.). The following comment from Ruby demonstrates that by the end of the intervention, the children had developed a good understanding and appreciation of PSA:

> I thought the self-assessment was excellent because we were judging ourselves and could learn from our mistakes. Peer-assessment was brilliant for your partner or pair could judge your work and spot mistakes that you might not have spotted yourself. (LL, 12/06/2014)

While most children said they liked both peer- and self-assessment, some expressed a preference for one over the other. Most favoured peer-assessment and this comment from Maria was typical:

> I like doing peer- assessment more than self-assessment because with somebody else's work you get to like compare them sort of in a way, and you get to see what somebody else is doing, and like see if you could do better, or see if you could do the same…or if you could improve it. (FG2)

Chloe, meanwhile, viewed self-assessment as inclusive of, or almost a prerequisite for peer-assessment, commenting: "You're going to do the self-assessment before you're going to do the peer-assessment anyways, so you have to kind of check it before you give it to them" and added "If you didn't do self-assessment and you just handed it up to the teacher, then you're actually just taking the lazy way out of it" (FG1). Some children believed that peer-assessment had advantages over self-assessment. Lily highlighted that when you assess your own work you can miss things: "Well, sometimes if you're checking over your work you wouldn't really notice a mistake, and they might notice it more than you". However, most children readily admitted they found it difficult to do peer-assessment unless it was scaffolded in some way.

They particularly found rubrics and two stars and a wish, useful in this regard. Maria explained:

> I really like peer-assessment, but the first time we were doing it I said "Oh no, what will I write down?" but then we got the *rubrics* and I kind of found it really easy, because the *rubric* sort of like guided you along and told you what to do, sort of what you're supposed to learn and stuff, and now I think peer-assessment is really good. (FG2)

Meanwhile, Mia emphasised that peer-assessment can be used to help motivate fellow students:

> If you're doing a peer-assessment with the *rubric* and you colour in some points, and the person that you're giving it to says "oh, that's not enough points for me, I want to get more", they'll work harder at it. (FG1)

Some children made reference to the fact that using peer-assessment could help improve student learning, e.g.

> Well, if you are doing two stars and a wish on someone else's work, it would be easy to show them like what they did very well, like they were neat and everything, but, then you could show them something that they'd need to improve on, so that would really help them along for the next time they do something like that. (Sarah, FG2)

Additionally, Emma felt that using peer-assessment improved her own work too:

> Like if you're correcting your friend's work, and she has a lot of silly mistakes, it just shows you how careful you need to be in your next work, you know, so you don't lose any points or fall behind in maths because you didn't check. (FG1)

Others held the view that peer-assessment not only enhanced learning but was also fun:

> Well if you're doing peer-assessment, it's kind of fun as well because you're judging the other person's work, and you'd like check it and then you'd write down like two good things, and then you'd write like something that they would need to improve on…they might just write something like "I could do less rubbing out" but you could write something like that they need to make sure that they carry their one or something like that, something a bit more important. (Lily, FG2)

Similar to research by Topping (2010), trusting your peers was mentioned by a number of participants as integral to good peer-assessment practice. For example, Hollie remarked: "I think peer-assessment is the best because you get to like trust your friends more, so they'll be more honest with you in the future" (FG1). Sophie expressed a similar view:

> I prefer peer-assessment because I find it fun, like looking at your friend's work, but you wouldn't say anything mean about them, like you won't go off and tell other people if she got this wrong or that, you'd keep it to yourself, because you wouldn't want her going off to her friends. (FG1)

In sum, for the children participating in this study, PSA was undoubtedly a highlight. They enjoyed using it, especially once they had a rubric or specific guidelines to follow. In contrast, the three teachers participating in the intervention found PSA challenging, especially at the beginning.

While the potential benefits of PSA have been highlighted in Chapter Two, Harris et al. (2015) have raised concerns about its utility in helping pupils make progress in their learning. Teachers in this study also had reservations regarding the value of using PSA, particularly at the outset. In the initial stages, all three teachers found that many students were inclined to either overestimate or underestimate the quality of their work. The teachers were also conscious that, in situations where students were assessing their friends' work, their judgements appeared to be based more on friendship or peer pressure than on the quality of the work being assessed. This is in keeping with findings by Harris et al., (2015) who stated that the "potential of interpersonal relationships as a mediator when providing FB to social peers warrants further investigation" (p.276). The three teachers questioned the value and accuracy of student-generated FB, especially when compared to their own professional judgement. From about February 2013 onwards, the teachers realised that if their students were to provide better quality self- and peer-assessment, then they (the teachers) needed to play a greater role in this process by providing students with more concrete examples of the type of FB they expected, and by scaffolding that FB with rubrics, or more specific guidelines. The crucial role played by the teacher in guiding PSA has been highlighted in the literature (Black, 2016; Warwick et al, 2015) and findings from this study also reveal that it was only when children were given detailed guidance and support that they became comfortable using PSA. Consequently, both teachers and students felt happier using PSA once this process was scaffolded. The quality of students' FB to themselves and to each other improved considerably as a result and as the children became more experienced in using PSA. Towards the end of the intervention I was able to write the following in my LL:

> I've been trying out peer-assessment for some months now and at first I had doubts about how successful it would be. I felt that perhaps the children would be afraid to be honest with each other or that they would be incapable of actually assessing the work that was done in an in-depth fashion…It was only when the children used rubrics over a number of lessons that they really began to understand how they should be assessing the work of their peers. Subsequently the standard of what they were saying in the 'two stars and a wish' format also improved. I now see how important it is that we help children more when they are assessing their own work, and that of others. Above all, I think it will take time for students to develop the skills necessary for good self- and peer-assessment practices. (12/06/2013)

Unexpected insights

Discussions during the FG interviews also resulted in some unexpected insights that are discussed briefly here. During the discussion on FB, reviewed earlier, the children in FG1 used what Black (2015) has termed reasoning words such as "think", "because", "would" or "should", e.g. Hollie, a pupil with SEN, stated the following: "I think that FB is really good because, I know all teachers give you FB but I think this year the FB they gave me I understood more. I just like getting FB because… it just shows you where you are in your learning" or as Chloe suggested:

> I think FB, that it can be good FB or bad FB…like good FB would be like if you were doing a sum, a really hard sum… and the teacher would go over and say “keep on doing that, that’s brilliant”…and then just normal FB would be like to remember to check your answers or remember to carry over your one or stuff like that. (FG1)

Participants in FG1 also considered the merits, or otherwise, of what they termed “positive” and “negative” FB. Mia began the debate stating: “I think negative FB helps you move on to the next stage when you’re going to do something else, but when you get positive, you won’t really know if you’ve done something wrong or not”. Sophie then argued: “I think positive is good as well because if you get positive…you go on and you think you’re great and you keep on going and you get better answers”. Each of the group in turn then contributed to this debate with most preferring what they termed “negative” FB. Here are some of their comments:

> I think that the negative FB is actually the best because it says like what you're going to do next. The positive FB is good but…and it tells you to keep going the way it is…but with the negative FB, it’s not exactly positive yet, so you still have some few things to correct and a few things to look over and a few things to work on, but the next time then you’ll get more positive FB and less negative. (Ruby, FG1)

> I think that negative is actually better. Positive is kind of bad at the same time because if you think that you just say “yeah I’m great at this and I know how to do this” but like you could just go off with yourself and think that you're great and then accidently do something wrong. (Hollie, FG1)

> I think that without negative FB, I don't think that you’d be very good at whatever you're doing that you got the FB on, because if there was no negative FB, people might just go on and think they’re brilliant and then they get a shock when their test comes…but maybe if you want you shouldn't put all just negative, maybe just a little comment to say “keep going on like this”. (Emma, FG1)

Ultimately, the comments above suggest that students in this study valued feedback, regardless of whether they perceived it as “positive” or “negative”. This mirrors findings by Voerman, Meyer, Korthagen & Simmons, (2012) who highlighted the need for students to receive both “positive” and “negative” feedback.

Comment-only-marking

At the beginning of the year, in line with recommendations from the AfL literature (Wiliam, 2011a; Butler, 1988), participating teachers had agreed to give only comments, written and/or oral, as their feedback to students. Consequently, regarding mathematics, students never received grades throughout the intervention. This decision had not been discussed with students and they had never commented on it until during one of the FG interviews, when the matter came up incidentally. Ruby, a high achieving student, immediately commented:

> I'm glad that you didn't tell us that we weren't going to get marks because then everyone would kind of be 'oh we're not going to get marks, we won't really bother', so it's better that you didn't tell us so we were working hard because we thought that we were. (FG1)

When I asked her to clarify her stance she went on to explain that if she had to choose, she'd prefer feedback "because we'd know that for the next time in our learning that's what we have to do, for the next test or the next worksheet or whatever we do". Chloe, from the middle group, then added:

> I would have liked marks and the feedback. I thought the marks would have taught you how you have done in the year from each test to another and then the feedback would be saying that this is what you have to do next time and to learn from your mistakes. (FG1)

Each of the remaining four students then all agreed with Chloe, revealing that they would have preferred both marks and feedback. Two students even went on to explain that receiving marks as well as feedback would have enabled them to see if they had done better in their SIGMA-T in comparison to previous years. This was an interesting discussion, and contrary to what experts such as Black (2015) and Wiliam (2011b) suggest is best for students, with Black (2015) recently suggesting that grades can even damage students' development as learners, encouraging what Dweck (2000) terms a "fixed" rather than a "growth" mindset.

Rubrics

One other item of interest that many of the children wrote about in their learning logs and discussed in the focus groups was their opinion about using rubrics. Although it was briefly discussed elsewhere, it is worth highlighting here in it's own right. The rubric was probably the AfL technique that the children believed had greatest impact on their learning and assessment. In particular, it seemed to give them confidence when doing self- and peer-assessment as it scaffolded the assessment process so they did not feel that the full responsibility was on them. Maria wrote: "My favourite thing about the AfL was using the rubric" (LL, 12/06/2013) and later explained why:

> I like using the rubric because when we were first going to do peer-assessment I was like 'oh God, what will I say was wrong?' and 'I don't know what to do here', but then you showed us the rubric, and I was like "oh, it's ok", because it said four for the really good, three for the good but it could be a little bit better, and two for the good enough and then one for the you'll have to try better. Otherwise I wouldn't have known what to do. Otherwise, I'd just be like "oh, gosh, what'll I say was wrong?" because I wouldn't really know what to assess. But, with the rubric, it tells you what you're supposed to assess, and like otherwise you could've just been like telling them something that they did wrong, like something that they didn't really need to know, but there could have been something that they did wrong that was important, that they should have known. (FG2)

The children explained how using rubrics helped pinpoint where they needed to focus on to improve their work e.g.

> Well if you're doing the rubric, and like say you get points, and I think, say if you only got like 10, there's obviously somewhere you need to be working on very hard and then somewhere you're really good at. (Sophie, FG1)

Rubrics made it easier for children to assess work and also helped them when they were giving feedback. Sophie also highlighted another advantage of using rubrics: "It is very good on telling you how you did, without hurting any feelings" (LL, n.d.). Finally, Mia's comment is typical of what the children thought about rubrics: "The rubric is so good because you feel like a teacher and you can tell people exactly what is good about the work or bad" (LL, 12/06/2013).

Conclusion

In sum, quantitative findings relating to the second hypothesis suggest that, firstly, there was a statistically significant increase in students' use of AfL practices and that, secondly, students' mathematical confidence, engagement with, and attitudes towards mathematics had also statistically significantly increased. The qualitative analyses confirm these findings, elaborate on them and provide richer insights, particularly through the student voice, a perspective that some (e.g. Florez & Sammons, 2013) argue is often missing from assessment research and dialogue. Consequently, both quantitative and qualitative findings provide evidence that suggests the second hypothesis can be accepted.

Research Hypothesis Three

Peer-to-peer professional learning is a feasible, worthwhile, efficient and effective model of CPD in AfL and will improve teachers' skills, knowledge and use of AfL and their attitudes and beliefs towards AfL as a form of assessment.

This final section explores and analyses the data set relating to the third research hypothesis and again includes both quantitative and qualitative data. The quantitative findings pertain to the teachers' AfL audit, while the qualitative data set, comprises teachers' reflection tools, the researcher's diary, LS observations and reflections, as well as transcripts from review meetings regarding the audit, AfL, LS, and the intervention in general. Chapter Three clarified how this research adopts a pragmatic stance since it provides a good fit for mixed methods research and is particularly appropriate for “solving practical problems in the real world” (Feilzer, 2010, p.8). Since pragmatism “offers a practical and outcome-orientated method of inquiry that is based on action” that leads to further action (Johnson & Onwuegbuzie, 2004, p.7), it became clear, as the research progressed, that my two colleagues were becoming increasingly engaged in the study and this allowed me to work with them as participant/researcher, so that ‘I’ and ‘they’ became ‘we’. Notwithstanding, to preserve their anonymity, the other two teachers have been assigned the pseudonyms Niamh and Eimear. Similar to hypothesis two, thematic analysis (Braun & Clarke, 2006) was employed to analyse these data. This section begins by describing, analysing and discussing pre- and post-test data from the AfLAi designed by Lysaght and O’Leary (2013) and also includes pertinent qualitative data. The remainder of the chapter explores the qualitative data pertaining to Hypothesis Three, investigating whether LS is an effective vehicle for CPD in AfL and whether the intervention has impacted teachers’ skills, knowledge and use of AfL and their attitudes and beliefs towards AfL as a form of assessment.

Analysis of AfLAi

As discussed in Chapter Three, the AfLAi instrument was designed to elicit data to help gauge teachers’ baseline understanding and use of AfL practices in teaching and learning. Table 24 provides details regarding the AfLAi’s rating scales, with scores ranging from 6 downwards, signifying how *embedded*, or otherwise, each strategy is.

Table 24

AfLAi Rating Scale (Adapted from Lysaght & O'Leary, 2013)

Mean Score	Rating	Meaning
6	*Embedded*	Happens 90% of the time
5	*Established*	Happens 75% of the time
4	*Emerging*	Happens 50% of the time
3	*Sporadic*	Happens 25% of the time
2	*Never*	Never happens
1		I don't understand what this means

In an attempt to track changes that may have occurred in our assessment practices due to the intervention, the three participating teachers completed the AfLAi in early September 2012 (pre-intervention), and again in June 2013 (post-intervention). As highlighted by Lysaght and O'Leary (2013), concerns have been expressed regarding the reliability of self-reported data, and the ability of respondents to assess and/or report accurately what they do in class, and so one must remain mindful of these factors when exploring the audit's findings. The mean scores and ratings for each teacher on the four AfLAi scales, pre- and post-intervention, are now presented (Table 25) and discussed in turn. Additionally, relevant quotations from qualitative sources are also included to enrich data analysis.

Table 25

Mean Scores for the AfLAi Strategies for Participating Teachers

		Pre-Intervention		**Post-Intervention**	
Strategy	Teacher	Mean	Rating	Mean	Rating
LISC	Ann Marie	2.75	*Sporadic*	5.19	*Established*
	Niamh	3.06	*Sporadic*	4.94	*Established*
	Eimear	2.69	*Sporadic*	4.94	*Established*
	Group Mean	2.83	*Sporadic*	5.02	*Established*
QCD	Ann Marie	2.88	*Sporadic*	4.94	*Established*
	Niamh	2.13	*Sporadic*	4.69	*Established*
	Eimear	2.88	*Sporadic*	5.25	*Established*
	Group Mean	2.63	*Sporadic*	4.96	*Established*
FB	Ann Marie	3.00	*Sporadic*	4.92	*Established*
	Niamh	3.42	*Sporadic*	4.67	*Established*
	Eimear	2.92	*Sporadic*	5.17	*Established*
	Group Mean	3.11	*Sporadic*	4.92	*Established*
PSA	Ann Marie	2.21	*Sporadic*	3.86	*Emerging*
	Niamh	3.57	*Sporadic*	3.93	*Emerging*
	Eimear	2.29	*Sporadic*	4.29	*Emerging*
	Group Mean	2.69	*Sporadic*	4.03	*Emerging*

Initial analysis comparing pre- and post-test scores for the LISC strategy revealed that all three of us believed our practice in this area had improved following the intervention, moving from *sporadic* at the start to *established* afterwards. More specifically, mean scores for Eimear and I had increased on all 16 audit statements in this scale, meaning that, following the intervention, we believed our practice of sharing LISC had increased across the board. At the audit reflection meeting (ARM, 26/06/2013), Eimear explained why she thought her practice in this area had improved so much, stating: “I was obviously starting from a low base in lots of them because I wasn’t including AfL practices some of the time, so it was easier to improve from a low base”. Mean scores for Niamh improved in 13 of the 16 statements but, regarding the September audit she remarked, “I even consider I scored myself too highly then. I just didn't truly understand the whole meaning of learning intentions”. Reminiscent of comments by teachers following their participation in Lysaght’s (2009) study, this suggests that Niamh believed she had a much deeper understanding of this strategy following the intervention, even if this did not show in results from some audit statements. Prior to the intervention, none of us had identified LISC as the strategy we considered most or least *embedded* in our teaching and mean ratings for this strategy are similar for all three of us. Nevertheless, qualitative data from the study provides evidence that, for the first few months, we all found sharing LISCs a challenge, e.g. I wrote the following in my TLL (05/11/2013): “I’m finding it really challenging to write clear and useful LISCs that are accessible to all children in my class”. Even at the end of the intervention, we still clearly remembered our struggle with this strategy, with Eimear stating: “I suppose the most frustrating thing was the time we spent when we got so bogged down in the success criteria and learning intentions”. This initial difficulty with sharing appropriate LISCs is not new and was also highlighted by Lysaght and O’Leary (2013) in the Irish context. Judging by the increase in mean scores from pre- to post-intervention, I believed my biggest improvement was in the LISC strategy whereas, for the other two teachers, the biggest improvement was in the next strategy, QCD.

Regarding QCD, Niamh remarked, “That was definitely the one that I’ve taken the biggest strides in” (ARM, 26/06/2013). Data from the QCD strategy suggest that mean scores for each of the sixteen statements improved for each of us, except for one statement of mine (No. 7), which stayed the same. Similar to findings from the LISC scale, a comparison of scores for the QCD strategy reveals that all three of us were of the opinion that our practice in this area had improved following the intervention, moving from *sporadic,* pre-intervention, to *established* afterwards. In their research Lysaght and O'Leary (2013) were surprised by findings that infer that many of the techniques linked to QCD were not more *established* since, as they conclude, most would argue that QCD is a normal part of teaching and learning in Irish primary classrooms. Similarly, in this study, the fact that, pre-intervention, we (the teachers) considered our use of QCD as only *sporadic* was also surprising but may, as Lysaght and O'Leary (2013) suggest, be due to tensions between pupil-led and the more traditional teacher-led approaches. By the end of the intervention, however, we believed that QCD was *established* i.e. happening more than 75% of the time in our mathematics lessons.

The importance and centrality of FB to the success of the AfL process has already been highlighted in the literature review. An examination of data pertaining to pre- and post-test scores for the FB strategy shows that the three of us were of the opinion we used FB *sporadically* before the intervention, in comparison to afterwards when we considered our use of FB was *established*. Prior to the intervention, two of us felt we *never* provided '*closing-the-gap-FB',* while the third teacher thought she provided it *sporadically*; in comparison to post-intervention when we believed this practice was *emerging* or *established*, respectively. Statement 7, where pupils are involved formally in providing information to their parents/guardians, was one of the few statements from the first three scales of the audit where there was little or no difference in pre- and post-intervention results for any of us. To involve students in this way would likely necessitate policy change at whole school level and was therefore beyond the scope of this study. Statement 2 ("Assessment techniques are used during lessons to help the teacher determine how well pupils understand what is being taught") is the only FB statement where all three of us believe our practice was *embedded* following the intervention. Interestingly, for statement 12 of the scale, which Lysaght and O'Leary (2013) identify as encapsulating the essence of AfL in that pupils are provided with FB on their learning on a *minute-to-minute, day-by-day*, basis, Eimear and I marked *never* in the pre-intervention audit whereas afterwards we marked *established,* while Niamh marked *sporadic* initially and subsequently *emerging*.

Experts have highlighted the import of PSA in the development of pupils' ability to self-regulate (Andrade, 2013; Lysaght & O'Leary, 2013; Lysaght, 2015), a skill required for lifelong learning and previously discussed in Chapter Two. Pre- and post- scores for PSA reveal that this was the area in which all three of us considered we had made least progress. Indeed, quite a number of the statements from the PSA strategy showed little or no difference in mean scores, pre- and post-intervention. Similar to the FB scale, there is only one statement (No. 6) where all three of us believed our practice was *embedded* following the intervention and, once again, this statement refers to assessment techniques. Pre-intervention, Eimear and I obtained our lowest mean score in the PSA scale, whereas post-intervention, the mean scores for all three of us was lowest in this scale. Furthermore, the smallest mean increase from pre- to post-intervention for each of the three teachers was also in this scale. This resonates with findings by Lysaght and O'Leary (2013) who found that PSA was the strategy that featured most infrequently in the Irish classrooms they surveyed. Ultimately, the PSA data confirmed our own perceptions and opinions regarding our practice of PSA. The following quotation of mine demonstrates how we were less comfortable using PSA than any other strategy and that it was only in the later stages of the intervention that we began to feel more comfortable with its use:

> With peer- and self-assessment, it was definitely only in the last few months that I was really getting into that. Yes, they were doing self-assessment for ages, because they were doing *thumbs* and *traffic lights* and whatever, but I'm talking about the more in-depth self-assessment…I now know more about it and so instead of applying AfL to the *letter* you're now embodying the *spirit* of AfL more, that you are actually doing it. (ARM)

Participating pupils, on the other hand, seemed to struggle less with adopting PSA practices than their teachers and, as discussed earlier, engaging in PSA practices was, for them, a highlight of the intervention. Similar to findings by Lysaght and O'Leary (2013), audit results from the PSA strategy from this study also highlight the challenge of introducing student-led approaches in the classroom.

To conclude, data from the AfLAi demonstrates that teachers in this study considered that, prior to the intervention, our use of the four AfL strategies (LISC, QCD, FB, PSA) was at best *sporadic* whereas, post-intervention, we believed our use of LISC, QCD and FB was now *established*, with PSA *emerging*. Therefore, taking the teachers as a group, pre-intervention ratings for all strategies were less *embedded* for the three of us in this study in comparison to average ratings nationally (Lysaght & O'Leary, 2013) whereas, post-intervention, ratings were more *embedded* (Table 26), probably due to our participation in the intervention. These findings mirror those in a similar study by Lysaght (2009) who found that quite radical changes in teachers' AfL knowledge and practices had taken place following their participation in her study.

Table 26

Comparison of Participant Ratings with National Ratings for the AfLAi

Strategy	Pre-	Post-	Nationally
LISC	*Sporadic*	*Established*	*Emerging*
QCD	*Sporadic*	*Established*	*Emerging*
FB	*Sporadic*	*Established*	*Emerging*
PSA	*Sporadic*	*Emerging*	*Sporadic*

The teachers concluded that results from the audit demonstrated quantitatively what we had deduced from our classroom practice of AfL, i.e. that our skills, knowledge and use of AfL had greatly improved, as well as our attitudes to AfL as a form of assessment. As Niamh remarked regarding the audit:

> Well, it tells us that we're using AfL, well it tells me anyway, that I certainly have embraced AfL, and that I know more about it and that I have some idea about what I'm doing. …I've come a long way from September. (ARM)

Nonetheless, several other issues are worth noting. Firstly, the three of us were conscious that findings from this research would be disseminated outside our school and so this may or may not have impacted on how we answered the audit or influenced the comments we made. The following comment from Niamh during the ARM was typical of how all three of us felt regarding this concern:

> You know, there is that whole thing as well of wanting to present the school, like if you felt this was going somewhere out of the school, to be read by people who weren't involved with the school…there probably was that feeling of, well you know, we're not going to let the side down.

This resonates with findings by Lysaght (2009), who highlighted that teachers can be reluctant to report on practices they consider may reflect negatively on their school.

Nevertheless, the three of us believed we had answered the audits as honestly as we could, both pre- and post-intervention. A second issue raised at the ARM was that we felt that when we had completed the initial audit our level of understanding of AfL was much less than when we answered it post-intervention, prompting Niamh to comment:

> We've a better understanding as well, so we were probably being harsher on ourselves now than we were in September because we know more…so that's probably even more of a jump to be honest.

This suggests that with hindsight, she believed some of her answers should have been less *embedded* in September than she had marked. Notwithstanding, as I then commented:

> The main thing I'll take out of it is that we have moved positively…I was delighted because it shows that our knowledge of this has really improved and you know our understanding and our application of AfL has really moved…we've all jumped hugely in each of the strategies.

Remaining Qualitative Findings and Analyses

The rest of this chapter uses data from qualitative sources (teachers' learning logs, researcher's diary, and reflection meeting transcripts) to investigate whether peer-to-peer professional learning, specifically LS, is an effective model of CPD in AfL and explores if participating teachers' skills, knowledge and use of AfL, and their attitudes and beliefs towards AfL as a form of assessment, were impacted by the intervention. Table 27 provides details regarding the three LS cycles: the strand and topic for each live lesson, the teacher teaching the lesson and when each live lesson and subsequent reflection meeting took place.

Table 27

Lesson Study Schedule

LS Cycle:	Teaching/Reflection	Date	Teacher	Topic	Strand
Cycle 1:	Planning Live Lesson	05/11/2012	All		
	Planning Live Lesson	19/11/2012	All		
	Research Lesson 1:	26/11/2012	Ann Marie	Nets of cube	Shape and Space (Problem Solving)
	Reflection Meeting	26/11/2012	All		
	Research Lesson 2:	27/11/2012	Niamh		
	Reflection Meeting	27/11/2012	All		
	Research Lesson 3:	14/12/2012	Eimear		
	Reflection Meeting	14/12/2012	All		
Cycle 2:	Planning Live Lesson	05/03/2012	Ann Marie & Eimear		
	Planning Live Lesson	12/03/2012	All		
	Research Lesson 1:	10/04/2013	Niamh	Garden Design	Measures (Problem Solving)
	Reflection Meeting	10/04/2013	All		
	Research Lesson 2:	16/04/2013	Ann Marie		
	Reflection Meeting	16/04/2013	All		
	Research Lesson 3:	25/04/2013	Eimear		
	Reflection Meeting	25/04/2013	All		
Cycle 3:	Planning Live Lesson	20/05/2013	All		
	Planning Live Lesson	27/05/2013	All		
	Research Lesson 1:	05/06/2013	Eimear	Area	Measures (Problem Solving)
	Reflection Meeting	05/06/2013	All		
	Research Lesson 2:	10/06/2013	Ann Marie		
	Reflection Meeting	10/06/2013	All		
	Research Lesson 3:	11/06/2013	Niamh		
	Reflection Meeting	11/06/2013	All		

Employing Braun and Clarke's (2006) six-phase guide to thematic analysis, discussed in Chapter Three, it was possible to identify two primary themes from the aforementioned qualitative data pertaining to this Hypothesis:

1. Embodying the *spirit* of AfL: An evolving journey for teachers;
2. LS: Effective CPD for our time.

These themes are portrayed in Figure 19 and are then delineated into various subthemes that are analysed and discussed later in the chapter. The next section begins with a discussion on why attempting to embody the *spirit* of AfL could best be perceived as an evolving journey for teachers involved in the intervention. Subsequently, both AfL subthemes will be considered.

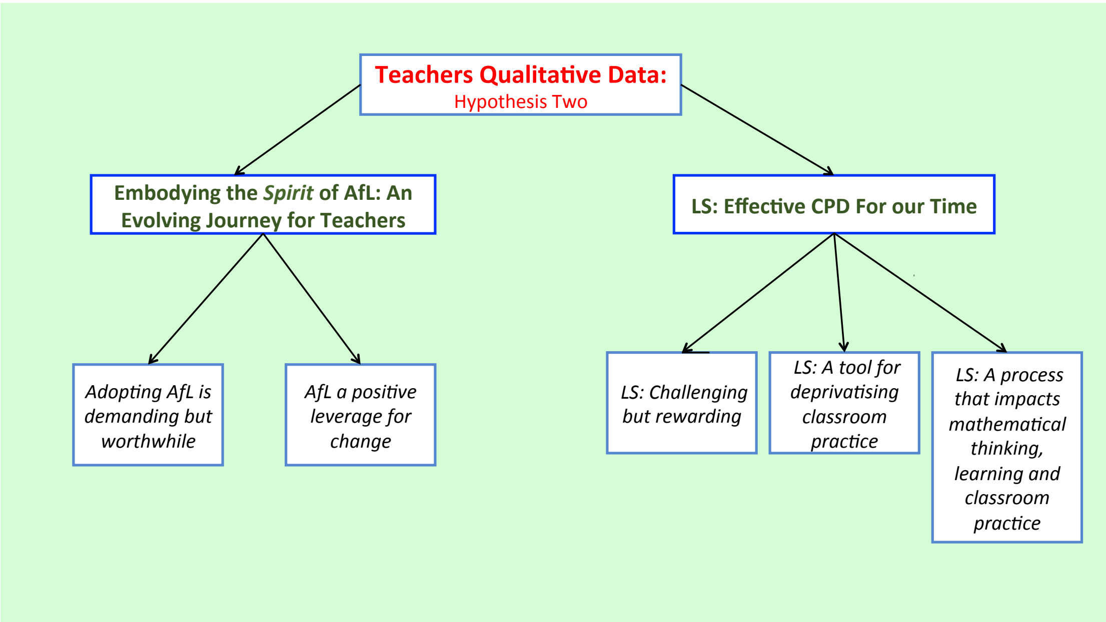


Figure 20. Hypothesis two qualitative data themes

Embodying the *spirit* of AfL: An evolving journey for teachers

A robust literature attests to the benefits of embracing AfL practices and suggests a positive relationship between effective use of AfL and students' and teachers' learning, improved classroom practices and student achievement (e.g., Black & McCormick, 2010). However, many scholars have also highlighted the significant challenges of adopting AfL practices successfully (e.g., Lysaght, 2009). Klenowoski (2009) warns of the real danger of "implementing a narrow interpretation of FA in classrooms using techniques" (p.265) that results in "procedural compliance" (p.263) whereby AfL practices are implemented mechanically or superficially, conforming to the *letter* of AfL rather than capturing what Marshall and Drummond (2006) coined the "*spirit* of AfL". Embodying the *spirit* of AfL necessitates, among other things, the promotion of learner autonomy and the democratisation of classroom practice (Lysaght & O'Leary, 2013) in the "here and now of learning" (Swaffield, 2011, p.441). The previous section discussed findings from the AfLAi and, through analysis, traced how teachers in this study started from a "low base" regarding our AfL practices but ultimately through personal experience, witnessed a sea change in our skills, knowledge and use of AfL, in addition to noting changes in our attitudes and beliefs towards AfL as a form of assessment e.g.,

> Well certainly in terms of the whole self-assessment, like I probably wouldn't have placed as much value on it as I would now…definitely my attitude to assessment has changed. I would say hugely, actually. (Eimear, IRM)

Myriad examples could be given of how we struggled with specific AfL practices, especially initially, sometimes doubting our ability to correctly implement the particular AfL strategy or technique we were learning about, or even questioning our teacher efficacy. Niamh voiced concern she was "not up to the mark in some areas" (TLL, 09/11/12) and reiterated this up to half way through the intervention: "As always, that I’m not up to the mark! That I may not be implementing the strategies already learned correctly” (TLL, 12/03/13). Similar to teachers elsewhere (e.g., DeLuca et al., 2015), it took time for teachers in this study to effectively embed AfL practices and principles in our classrooms and, as Niamh observed, to gain confidence in this regard:

> I think it actually took me a long time to understand AfL. I basically didn't understand it [AfL] at all in September, to be honest with you. It was after Christmas definitely before I really understood what we were at. To be honest like, you know, get a proper grip of what exactly this was about…and I think from then my confidence started to increase, like even in the beginning implementing *thumbs* and *traffic lights* and stuff, I actually used to wonder was I doing it right…you know it took the guts of the full year to understand it like, it wasn't something that I would have said that by October “Oh yeah, I know, I get it”. (IRM)

Nevertheless, by the end of the intervention, we believed that AfL had become embedded in our regular classroom practice. One point that teachers in this study discovered retrospectively was mentioned in the AfLAi discussion and is worth noting again here – we believed our assessment of how successfully we were implementing AfL seemed mediated by the depth of our knowledge, skills, understanding and use of the AfL process at the particular time, e.g., about halfway through the intervention I remarked:

> I think that it is only now, after spending the last few months trying out the various AfL strategies and techniques, that I realise that previously I was probably implementing AfL on a superficial level, in comparison with the way I’m implementing it now – embracing the *spirit* of AfL, rather than implementing it to the *letter* in my classroom. (TLL, 07/01/13)

Attempting to embody AfL practices in the *spirit* in which it is meant was not a straightforward journey, however, and certainly involved some anticipated and unanticipated highs and lows. All three of us, even to this day, identify a particular period in January 2013 as the abiding low point of the intervention. At that time we had become repeatedly frustrated in our attempts to write what we considered appropriate LISC and, ultimately, had recourse to professional expertise before overcoming this difficulty to our satisfaction. This obstacle had not been anticipated but, subsequently, we felt that we made greater progress in our implementation of AfL practices. In contrast to the children, the undoubted "high" of the intervention for the teachers was our involvement in the LS process, to be discussed in due course, while a close second was the joy we experienced from our observation of how engaged the children were in their mathematics learning as a result of using AfL principles, strategies and techniques on a daily basis:

> The children having such a positive attitude to maths, having enjoyed AfL so much, like it has really rejuvenated us as well. You know that you just get such a buzz out of them and you really feel passionate about it and get a great sense of achievement as well. (Eimear, IRM)

Attempts by teachers in this study to embody AfL practices and principles in the *spirit* in which they are meant had been a slow and evolving journey, albeit a positive one:

> I can't believe that I'm using so many of the strategies so naturally now, when I was so, even reading about them in the beginning I was thinking "Oh God" you know, yeah, you just get in on them and they [children] get in on them. I think it's because they sort of, they're so delighted to do them that you kind of start to feel, you lose your own sort of inhibitions about them really, because they love it. (Niamh, IRM)

Indeed, it is interesting to note that by the end of this particular journey, we teachers had indeed done what Black and Wiliam (1998a) suggested all those years ago in that each of us had incorporated the AfL strategies and techniques into our own practice in our own way:

> Because we were watching each other in the LS and we were watching the way, you know Eimear might have a different way of using *fist to five* or you might have a different way of using *thumbs*…. You actually learn from each other different ways of using the AfL strategies, so I found that really interesting as well, that by the time the end of the year was, you were using them for everything easily. (Niamh, IRM)

Adopting AfL is demanding but worthwhile

While the previous quotation suggests participants were comfortable using AfL practices by the end of the intervention, each teacher readily acknowledged that adopting AfL had been challenging and demanding, albeit beneficial. Some scholars have argued that teachers lack the competency (e.g. Bennett, 2011) or the systemic support structures to fully implement AfL practices (DeLuca et al., 2015). Additionally, challenges such as time, class size, (DeLuca et al., 2015), accountability, resources, and student diversity (Bennett, 2011) have all been highlighted in the literature as obstacles to effective AfL implementation. Similar to DeLuca et al. (2015), coming to terms with the "jargon" or "learning the terminology of AfL" was an early concern for all three of us teachers in this study, a factor we felt made adopting AfL practices more demanding. A further issue we highlighted was the demand of trying to accept the change in classroom dynamics that engaging in AfL necessitates. For us teachers, this involved relinquishing a certain amount of control, especially in the area of self- and peer- assessment. The following comment from Niamh's TLL, approximately two-thirds of the way through the intervention, reflects honestly how she felt about this issue:

> I think it will take me some time to trust peer- and self-assessment, even though I completely recognise that it must be of value. There is a huge amount of *control freakery* in teaching and I feel that this part of AfL requires a person relinquishing a certain amount of that control…this might prove a problem for me. (12/03/13)

Time was the other significant factor that the three of us identified as making AfL demanding. In the first instance, we had given of our free time to meet together and learn about the various AfL strategies, principles and techniques. Since current conceptualisations of AfL are so complex, it then took time to develop our knowledge of, and skills using AfL, as well as perseverance to ensure AfL practices were implemented regularly in our classrooms. As teachers, we clearly found it demanding to engage in AfL on a minute-to-minute, day-by-day basis. Notwithstanding, perhaps the factor we believed most impacted teacher time was the length it took us to give students "closing-the-gap" feedback:

> I have no doubt that with time 'feeding back' in maths will become easier, but I am a long way from the point where "feedback should be more work for the recipient than the donor". I found it extremely time-consuming and may not yet have mastered the art of "less is more". I think the idea of 'acting on the feedback' for it to be effective is important, and therefore I feel time should be allocated within class time for this, so that if difficulties arise the teacher is there to assist...I do see a huge problem with regard to time and feedback though. (Niamh, TLL, 22/01/2013)

During the year, the three of us had experimented with both oral and written feedback, and as recommended in the literature, no grades were given to our students. However, adjusting to this new way of doing things took time for our students and for us. Nevertheless, by the end of the intervention we were conscious of the benefits that using formative feedback accrued:

> I've definitely moved substantially where feedback is concerned. It is more time consuming though … but it's more beneficial too. The other was a waste of time, probably, so I mean, which would you prefer? (Eimear, IRM)

We were mindful of the various other benefits using AfL brings. As mentioned earlier, we were aware that using AfL had increased collaborative learning and student engagement and motivation. Writing about her use of ABCD cards, Eimear reflected that "anything that fully engages every member of the class as this did all the time must surely be of benefit" (22/01/2013). We believed that using AfL had improved our teaching and made it more focused. Niamh wrote: "I find it [AfL] not only focuses the children's learning but it makes me clearer as a teacher about what I want to achieve in a lesson" (TLL, 22/01/2013). Similarly, Eimear commented:

> AfL is proving very interesting and thought provoking. I certainly feel that I'm thinking much more critically about how the pupils are learning and I'm trying

> to plan lessons so that they are always actively involved and are clear about what they're doing and why they're learning to do it. (13/02/2013)

Additionally, we believed our use of AfL techniques throughout the intervention had also impacted student learning and reflection and, concomitantly, enhanced our teaching. Writing specifically about student learning logs, I stated:

> Having the children write into their learning logs has been a revelation. It has been quite amazing to see what they are capable of writing when given the time to reflect on their learning. It is also evident from their logs that they are really enjoying using AfL. Using learning logs is definitely something I'd consider using with any class I have from now on. It's a great way of getting student feedback and of hearing the "voice of the child", especially from students who tend to be less vocal in class. (TLL, 13/02/2013)

Furthermore, we could see that AfL arguably offered the opportunity to give students more responsibility and involve them more in decision-making and in the assessment process. Other benefits of AfL, such as more time spent on task and asking less procedural questions, were also mentioned and discussed previously.

In sum, while the three of us evidently considered AfL to be demanding we also deemed it very worthwhile. Nonetheless, during the final reflection meeting Eimear and Niamh pondered whether perceived improvements in classroom practice were solely attributable to AfL or an amalgam of things, comprising AfL, LS and the particular children and teachers involved. Eimear commented:

> The improvement in say our attitude, our energy, enthusiasm for it [maths] and the children's scores and their attitude as well, was that, would that, could you put that down to AfL, or would it have happened anyway, with us planning and collaborating, or, but it's hard, that's hard to split really, because AfL is so much part of what we were doing and the success of it.

Niamh concurred and her reply was similar:

> In my wildest dreams, when I looked at the NRIT of the group I had this year, I didn't think that we would be able to get the results that we did, so like something worked. Whether it was the buzz around maths because of what we were doing, or whether it was the actual AfL itself, I don't know, but it definitely worked, whatever we did this year. Like the results surely prove that.

AfL as leverage of change

Previous sections, and indeed the last quotation, have demonstrated that participants in this study, both students and teachers alike, considered AfL had positively affected the teaching and learning of mathematics in our classrooms. Students' positivity towards AfL was palpable throughout the intervention and it has been discussed how they believed that their engagement in AfL practices had enhanced their mathematical confidence and improved their engagement with, and attitudes to, mathematics. Furthermore, they revealed a surprising understanding regarding the impact using AfL had on classroom practices, in addition to unexpected insights regarding the benefits of different types of feedback and their preferences regarding the various AfL strategies and techniques.

The three of us teachers also welcomed the changes that AfL leveraged:

> That was the main thing that the children are more involved and they're more hands on. They got to give feedback or ask questions or explain why they did such and such a thing. I think that's important. (Eimear, IRM)

We acknowledged the project had enhanced our assessment literacy, had positively impacted our discourse on teaching, learning and assessment and, ultimately, had benefited classroom practices:

> It has changed the way we teach. I don't think there's any question about that. We continued to use the strategies and techniques because they were working well and I would feel that I am teaching maths better than I taught it last year. That's it in a nutshell. I think that the lessons are more enjoyable for the kids. It's as simple as that. They're more involved and they're more engaged…it came at a very good time for me and I can continue to implement this next year and hopefully…I definitely will implement the strategies and techniques. (Niamh, IRM)

The literature highlights that engaging in the *spirit* of AfL necessitates change in the roles and relationships within the classroom. In the initial stages, there seemed little evidence of this with Niamh writing: "I feel we are nowhere near the point where the child is in charge of her own learning. We are still doing the bulk of the work" (TLL, 22/01/2013). However, towards the end of the intervention I, as teacher/researcher, had observed a more visible shift in classroom roles and relationships:

> At this stage, it is clear that students have readily accepted a greater role in their own learning and are taking more responsibility for it. They are eager to engage in self- and peer-assessment and can articulate what they believe are the next steps in their learning. I still think they need a lot of guidance and support from us teachers as they are relatively new to the whole AfL process at this level, but a noticeable change is definitely underway. (RJ, 24/05/ 2013)

Indeed, the key to this change in classroom dynamics was possibly due to the use of PSA. In the initial stages both students and teachers shared concerns about students' ability to self- and peer-assess, with us teachers being most reticent. For both parties, the key to accepting the benefits of PSA was the use of scaffolding e.g., by using rubrics. The other teachers and I came to see value in self-assessment, as this comment from my learning log illustrates:

> I think that honest self-assessment is key to helping students to fulfill their full potential. It actively involves them in their own learning and helps them to recognise for themselves the steps that they need to take in order to move forward in their learning. This will be key in their learning as they get older, especially in secondary school and even more so in college. If they learn to honestly assess their own work then their learning will be maximised and they will become independent learners. Self-assessment also has benefits for us as teachers since it gives us a better insight into our students' thinking, and gives us some time to monitor their progress more effectively. (12/03/13)

By the final reflection meeting, even though the three of us still had some reservations regarding PSA, we could see its potential and acknowledged that the children loved using it. Eimear summed up the feelings of the group:

> The peer- and self-assessment…that's just second nature to them now, isn't it? And they really love it. I know maybe sometimes we're a bit concerned is there a bit of a mismatch between their self-esteem and their positive attitude and their actual assessment…but then I suppose you have to start with a positive attitude. (27/06/2013)

Additionally, Niamh was then able to revisit her opinion discussed earlier:

> I think I had said in one of my reviews that the peer- and self-assessment involves relinquishing control to a large extent, which as a control freak I'm not sure I'm ready for; so the rubric does actually allow for peer-assessment, but it's controlled peer-assessment…it's probably the ideal way to do it really. (27/06/2013)

The literature review emphasised the importance of developing the skill of self-regulation if students are to meet the demands of 21st century learning. Additionally, it explained how current conceptualisations of AfL increasingly link SRL and AfL practices, and also suggested that self-regulation is learnable through the use of AfL principles, strategies and techniques. Admittedly, none of us teachers participating in the study, or indeed students, mentioned SRL *per se*. However, much of the data, both quantitative and qualitative, alludes to the perception that students in this study, through their use of AfL practices, had become more actively involved in their own learning and assessment. They had been given and had accepted more responsibility for their learning and were becoming more autonomous in that learning; *id est* they were learning how to learn and were thinking about their thinking. If one accepts this premise, then students were displaying signs that they were beginning to self-regulate. Of course this move towards SRL was only in the initial stages and similar to Jonsson et al., (2015), by the intervention's conclusion, AfL practices, while changed, were still primarily teacher-led or at least scaffolded by them. Nonetheless, over the duration of the project, both students and teachers had observed real change in classroom practices and classroom dynamics, a change they attributed to their use of AfL practices. Ultimately, for participants in this project, AfL had been a positive leverage of change and all concerned recognised this.

Similar to the students, the three of us teachers in this study had plenty to write and say about our experiences regarding our use of AfL practices. Nevertheless, analysis of the qualitative data set reveals we had even more to say about LS. Indeed, each of us expressed that it was our involvement in the LS process that had been the greatest highlight of the intervention. The next section explores the qualitative data pertaining to LS. It begins with a brief introduction to the overall theme of "LS: Effective CPD for our Time". Subsequently, the three subthemes are discussed in turn, which should then provide an overall picture as to why the participants believed LS to be such an effective model of teacher CPD.

LS: Effective CPD For Our Time

The Literature Review highlighted the importance of high-quality and relevant CPD for teachers throughout their careers. Additionally, it emphasised some widely accepted characteristics of effective CPD such as being job-embedded, sustained, collaborative and linked to practice; criteria which LS as a model of CPD successfully fulfills. Indeed, following this intervention, all three of us teachers participating in the study were agreed that our involvement in LS had been far more beneficial than any other type of CPD we had previously undertaken. Eimear stated: "LS had far more impact than doing a summer course (Appendix Y), or a one-day course, or a night-time course; to do it in school and throughout the year collaboratively did make a big difference" (IRM). Niamh concurred, commenting:

> You do a summer course in July, and by the time September comes around the chances of implementing anything from it…you really have forgotten an awful lot of it. Even if you were doing courses throughout the year, at nighttime or whatever, it's just not the same. It just isn't the same as doing it on-site, with your stuff and planning on an ongoing basis. I wouldn't even compare them. Isn't it easier to be more relevant as well when you're actually in school and although you mightn't feel like it at three o'clock in the day the children are still fresh in your heads. It's easier too not to have to get into the car at 7 o'clock in the evening; it's a lot handier, you have your books, everything there at hand.

Furthermore, all three of us believed that the effectiveness of the LS process was in part due to the fact we felt accountable to each other, on an ongoing basis e.g.:

> The fact that we were meeting again in two weeks or whatever it was, you basically had to have your homework done. It was literally like that wasn't it? So you actually had to implement it straight away; so the school context was ideal. (Niamh, IRM)

Indeed, one interesting anecdote here is that I only discovered during the final reflection meeting that, outside of the planned meetings throughout the year, the other two teachers regularly inquired from each other as to how things were going in their respective mathematics classes. In sum, all three of us involved in the project believed that LS was a very effective model of CPD.

LS: Challenging but rewarding

Many scholars highlight the benefits of engaging in LS and advocate its use as the CPD model of choice for teachers, particularly in the area of mathematics, both nationally (e.g., Corcoran, 2011a) and internationally (e.g. Dudley, 2015). Post-intervention, reflecting on our engagement in three cycles of LS over the course of one academic year, we teachers saw "great value in this process" and found it "very worthwhile" and a "huge success". Each of us concurred with Eimear when she stated that she "got so much satisfaction and such a sense of achievement out of LS" (IRM). She went on to pinpoint LS as the highlight of the intervention, especially "the time we spent together planning lessons", commenting:

> There was so much to be learned, with sharing our ideas and adapting and reviewing. There really was a great pooling of knowledge and expertise, because people do know so much, and I think sometimes there are limits on your own. I think we've come a long way from when we started. I think we've learned an awful lot. There's so many things we've learned that you could say you would definitely use again and continue to use.

Multiple benefits were highlighted by all three of us, many of which confirm findings from previous research highlighted in the review of literature (Chapter Two, pp.59-67). These benefits include the apparent positive impact LS has on classroom practice, teaching and learning, teachers' PCK, collaborative practice and the deprivatisation of teaching, all of which will be discussed in due course. Not only did we appreciate the value of LS, we also evidently enjoyed engaging in the process, for example at the intervention reflection meeting I commented:

> I have really enjoyed it, even though it's been hard work, but I really enjoyed meeting…I just think that we all learned so much and I found it a very positive experience; with ups and downs. (27/06/2013)

Analysis of the previous quote suggests that engaging in LS had also posed particular challenges that had to be overcome. Indeed, as is often the case in schools we concurred that "the biggest challenge was the whole time aspect" (Niamh, IRM); and this lack of time is also identified in the literature as the primary obstacle to adopting LS successfully (e.g., Cajkler et al, 2014). By participating in this study, we gave up a significant amount of our free time (circa 40 hours) in order to meet after school to learn about AfL strategies and techniques, to study research, to plan for mathematics in general, to source resources, to plan specifically for the live lessons and, subsequently, to reflect on their implementation. However, it was the live lessons that presented particular logistical challenges since, by necessity, these lessons had to take place during school hours when students were present. In total, there were nine live lessons and students not involved in the live lesson on a particular day could not be left unattended. Without the extensive support of the principal and members of the school's SEN team, who kindly took care of these students, it would have been extremely difficult to carry out any live lessons or continue to implement the intervention in full. The LS literature highlights the importance of support from school leadership (e.g. Xu & Pedder, 2015) to the success of the LS process and the facilitation of the aforementioned supervision, if nothing else, supports this theory.

The three of us also highlighted other aspects of LS that, through our experience, we believed contributed to making LS challenging. Besides having to "get out of your comfort zone" for the live lessons, to be discussed later, teachers participating in this study particularly highlighted the importance of trust, which we deemed a prerequisite for successful LS. The following comment from Niamh exemplifies her beliefs that having trust in each other was crucial:

> There is the whole trust thing as well. I think you do need to know each other very well… I think it took us a while probably to build up trust in each other, which we did thankfully, and that was evident in the lesson study and even in what we all said after the lesson study. (IRM)

Eimear concurred that trust was important but also emphasised her belief "that people need to feel comfortable with each other". Regarding this issue of compatibility, Niamh then added:

> I don't think LS would work with just any random cohort of teachers. I think it would definitely have to be teachers with a common interest, teachers who are comfortable with each other…so therefore I'd say on a staff of 24 there might be 8 different sets of 3 that might work well together, or there might be 12 that didn't buy into it at all.

This emphasis on the need for trust and compatibility brings to mind recent remarks by Xu and Pedder (2015) who criticised LS researchers for their failure, thus far, to pay sufficient attention to what they term the "micropolitical dimensions of teachers' collaborative work in LS contexts" (p.49), which includes the building of trust and the establishment of collegiality norms.

One final issue which the other two teachers felt had contributed to their belief that LS could be demanding was their conviction that, although LS is collaborative, “you would need one person to be the driving force” (Eimear, IRM), leader, and/or KO. Niamh articulated her thoughts on this matter to me as follows:

> LS would need a leader. Whether you like it or not, you are the Knowledgeable Other. I know you’re not inclined to, you don’t want to take that mantle but I would see you as that. I think LS probably does need a leader, you know, somebody to just take charge and lead. I mean the amount of work you put in and the background knowledge that you had about the LS, that made a huge difference…I mean in fairness, you definitely were the Knowledgeable Other. (IRM)

Eimear’s views were similar:

> We did have our own in-school expert. We mightn’t have had an outside expert but you did put so much work into it. It wasn’t a three-way split. You mightn’t be considered from outside but you definitely were the KO. (IRM)

In sum, while all three of us teachers participating in the study clearly believed our engagement in LS was very worthwhile, we also acknowledged it had been demanding in terms of time, logistics, openness, trust and leadership. In short, it takes you “out of your comfort zone”.

LS: A tool for deprivatising teaching

Observing and being observed

The most demanding aspect of LS for the three of us teachers involved in the study was undoubtedly teaching the live lessons, particularly in the initial stages. This is interesting because all of us are experienced professionals. Additionally, we had all taught multiple class levels and had declared our enjoyment of, and confidence in teaching mathematics. However, until this intervention, none of us had ever invited a colleague into our classroom to specifically observe our teaching. Each of us admitted to feeling nervous, fearful or apprehensive before we taught our first live lesson in front of our two colleagues, but gradually this became less daunting. I taught the first live lesson and it is interesting to review one research journal entry regarding observation by my colleagues of that very first live lesson:

> I felt that it was difficult enough to get away from the notion that my colleagues were not observing me and were really only observing the students' thinking and learning, as suggested in the LS literature. In retrospect, I also think that I felt under pressure to teach a 'really good lesson' since I had initiated the LS process. (Audit Reflection Meeting [ARM], 26/11/2012)

While each teacher's journey was obviously unique, there were many similarities. Consequently, tracing Niamh's LS journey through the live lesson process should provide some insight into what we as teachers experienced over the course of the intervention. Beginning with a comment regarding teaching her first live lesson, she stated: "I approached the actual teaching of the lesson with some trepidation and I was unprepared for this nervousness that beset me in front of my colleagues" (27/11/2012). However, she was satisfied with how things went, commenting:

> Overall, personally, I felt a certain sense of achievement when the lesson was over. I was removed from my comfort zone in teaching in front of two colleagues for whom I have great respect. Now that I have done this once, I expect the next lesson not to be as daunting. (27/11/2012)

Niamh's experience of teaching the live lesson during the second LS cycle was much less daunting and she felt "positive" about her experience:

> My experience of LS2 was very different to LS1. In the first place while I wasn't exactly looking forward to teaching opposite my colleagues I was definitely less nervous than I had been the first time. This was partly due to the fact that I had done it before, but due more so to the subject matter of the lesson itself. The actual format of the lesson was more typical of the kind of teaching I would normally do than in the first lesson study. (17/04/2013)

Following her observation of the first live lesson, in LS3 Niamh was able to write the following: "I felt that this being our third lesson study everyone was very relaxed and generally at ease over it. This is a remarkable departure when compared to the first day" (TLL, 05/06/2013). After she had taught the same lesson, the final live lesson of the intervention, she remarked:

> I felt a sense of exhilaration after this lesson. There were so many surprises in student thinking that I don't think we expected or anticipated, especially since this was the best group. At the end of the lesson I know that my teaching of maths will never be the same again. (11/06/2013)

Her final reflection on the matter was as follows:

> If you take the LS, I absolutely felt that the first lesson was about me and my teaching, now I really did. That was why I was in such a state over it. I think that teaching in front of, opposite your peers, is a huge hurdle to get over. Now maybe if we did more of it as teachers we might be less apprehensive. (27/06/2013)

This narrative mirrors findings by others (e.g. Bocala, 2015), discussed earlier, whereby teachers, as they become more experienced in LS, tend to focus less on themselves and the LS routine and more on student learning.

The literature informs us that it is the live lesson that sets LS apart from other models of CPD (e.g. Corcoran, 2007). It is the key to putting an end to the professional isolation and traditional norms prevalent in teaching, in a structured and non-threatening way. Through our engagement in observation of live lessons we teachers in this study were conscious of how our professional learning was enhanced:

> I learned loads. I mean some of which we had heard before, say the zone of proximal development (ZPD) I certainly had come across that in previous courses…a huge amount of it we hadn't come across before…it certainly widened our knowledge. (Niamh, 27/06/2013)

In particular, for me as teacher/researcher, the live lesson observation was a privilege that helped link theory and practice:

> It's such a privilege I think to be able to get into someone else's class because we never do. The three of us are so long teaching that we probably do a lot of this anyways, but it's actually great just to see it. I've been reading all this theory…Shulman came up with PCK in 1987 and then Deborah Lowenberg Ball came up with MCK to tweak what he was saying …today regarding PCK, your knowledge of the students, the content and the students, your knowledge of the content and the teaching, and then our knowledge of the curriculum when we were doing our plan…it all became visible. (LS Reflection Meeting, 10/04/2013)

We, the teachers, gained deeper insights regarding our students' learning as my comment at the intervention reflection meeting illustrates:

> I learned so much from LS, through observing ye, because you can't observe to the same degree what children are thinking and doing when you are teaching them yourself. (27/06/2013)

With observation of the live lesson being such an integral part of the whole LS process it undoubtedly encouraged and facilitated collaborative work.

Encouraging collaboration

As a result of our increased collaboration, we felt less isolated than heretofore e.g.:

> The collaboration was such a big thing. It's such an isolated profession in general that you know anything when you're working with people that you can have a laugh with as well apart from anything else, it really does make life an awful lot more interesting. (Niamh, Intervention Reflection Meeting [IRM], 27/06/2013)

Engaging in LS gave us the incentive and opportunity to plan collaboratively, a process we enjoyed, and which we believed also enhanced our learning. My comment typically illustrates this:

> I've thoroughly enjoyed it. It's been an amazing experience. I think it's rejuvenated my teaching as well in that the kids really enjoyed it so much and I think the collaboration has been amazing…I just really enjoyed meeting…I just think that we all learned so much and I found it a very positive experience. (IRM, 27/06/2013)

I elaborated:

> It was like good peer pressure because you kind of feel that you have to try this within the next fortnight because we're going to meet again in two weeks time and the others will have done it so I need to do it.

The following comment from Eimear probably best demonstrates why the three of us considered LS an effective tool for deprivatising teaching and encouraging collaboration:

> The lesson study discussion and collaboration we have done together has been very interesting and worthwhile, and getting the opportunity to observe each other teaching is so enlightening. Most of the time each teacher is teaching in isolation, so it's really wonderful to collaborate and support each other, to reflect and share. Thumbs up for lesson study! (14/12/2012)

The next section explores how engaging in the collaborative practice of LS acted as a catalyst to help change classroom practice in mathematics.

LS: A process that impacts mathematical thinking, learning and classroom practice

Post-intervention, the three teachers involved in this study believed that engaging in LS had impacted our classroom practice and our mathematical thinking and learning, and the mathematical thinking and learning of our students. The literature review highlighted that LS is particularly suited to mathematics teaching (Corcoran, 2008), and suggested that involvement in the LS process increases teachers' MKT, a particular type of PCK. Pre-intervention the three of us felt confident about teaching mathematics and believed that our PCK of mathematics was very good. However, following my teaching of the lesson in LS2, I commented:

> Concerning my own teaching of maths, I always felt that my PCK was very good. Nevertheless, I have also learned new things, especially when we get together to plan and discuss our maths lessons. The lesson study process has been particularly good in this regard, as I've had the privilege of observing my colleagues teaching maths and witnessed them going through the various dimensions of Rowland's Knowledge Quartet: foundation, transformation, connection and contingency. All of these things are definitely benefiting my teaching of maths. (19/04/13)

This hints at the potential benefits of teachers being open to the idea of lifelong learning and brings to mind Wiliam's (2011b) assertion that "the job of teaching is so difficult, so complex, that one lifetime is not enough to master it" (p.29).

As part of the intervention, we had planned mathematics together for the whole year and, while we saw great value in planning together, we acknowledged that we sometimes found it difficult to stick to the particular live lesson plan, especially when we felt students were struggling with the content of the lesson. This comment of mine reflects our common thinking: “I think that if this was not a live lesson I would have abandoned the lesson plan and gone back over the areas of difficulty in fractions and decimals” (17/04/2013). Nevertheless, it was clear from Niamh’s comment that collaborative planning was an enjoyable activity:

> I just think everything about the fact that we were learning, like we were working together as a team, just made the planning of maths so much more enjoyable than any other subject area, because it just takes the isolation out of it; and you know bouncing ideas off each other, like everybody knows something more about some things than the other person. We all have something to learn is really what I’m saying, and I think we all really did benefit from LS. (27/06/2013)

As discussed earlier, engaging in LS provided us with a unique opportunity to observe live lessons. While effective teachers regularly observe their students in order to help them fulfill their potential, teachers in this study were aware that being observers at the live lesson enabled us to notice student thinking and learning in a way that would not otherwise be possible. Because we were solely observing the lesson we were not under pressure to help weaker students, ensure students were on task, or clarify misconceptions. As Eimear commented following the first teaching of LS2:

> It’s very interesting just watching them…there was great learning and sharing…the amount of thinking and discussion and the buzz and the engagement, particularly the second part of the lesson. They [the children] really enjoyed it. (10/04/2013)

Instead, as observers we had time to interact more with students, probe their thinking and notice things that we might otherwise have missed e.g. when asked by an observer what she had learned one student replied: “I should have worked it out before I drew the diagram. It was harder to work it out without a proper plan. I thought it would have been easier but it was harder.” (Eimear, 10/04/2013). All three of us mentioned the fact our involvement in LS helped us focus more on student thinking. Niamh's comment typified this:

> LS does give you an idea of how clear their thinking is, on the whole thing. Certainly, I'd say there's a lot of work to be done on fractions in general, on equivalences and other things. (10/04/2013)

However, Eimear's advice following the intervention, agreed by all, is also worth noting: "Don't presume anything…what they're thinking or what they're working out is not necessarily the way we think they're going to" (27/06/2013). Notwithstanding, she had previously commented: "people interested in AfL would be delighted to see lessons like that…how engaged the children were, learning from their mistakes" (10/04/2013).

Involvement in LS also facilitated our engagement in reflective practice and inquiry stance. As early as the first LS cycle, Eimear could state: "how wrong answers allow us to identify and challenge learners' misconceptions" (14/12/2012), while I commented:

> I was happy with the maths lesson. The children were very engaged throughout the lesson and there was a great buzz in the room. They obviously loved the lesson and thought that it was great fun. It just goes to show how well active learning works. It definitely helped that we had meticulously prepared the lesson together. (26/11/2012)

Not only do the teachers' comments echo the literature regarding LS facilitating reflective practice they also demonstrated we were able to utilise knowledge and theory gained during our meetings when reflecting on the live lessons:

> For me today the exciting thing was the theory, you could actually see it in practice. You [Niamh] were interpreting what the children were saying. You were transforming foundational knowledge in order to make it accessible to the children. You had to explain it in language they would actually understand. When they were giving you misinformation today you had to probe and prompt them and ask them questions in order to help them. And then connection…you were connecting fractions and decimals and real life, so you were making connections the whole time. Then contingency, that's actually thinking on your feet, when something totally unexpected that we

> hadn't anticipated came up you dealt with it ….so for me today it was just the complexity of teaching. You also have the knowledge of the content and the students because you're teaching that group so you know them really well, so you'd immediately look down at student x because you'd know that she might need extra help or if you really wanted someone to explain something really clearly you know who to ask for that. (Ann Marie, 10/04/2013)

Additionally, and similar to research elsewhere (e.g., Lewis et al., 2012), Eimear could articulate how she believed LS helped build teacher motivation e.g.:

> The maths teaching this year has been different. We would have said before that you would have always thought that you did make an effort to teach maths differently. We put a lot more into this year. (27/06/2013)

Some other issues that came up during LS reflection meetings are also worth discussing briefly since they impact classroom practice and mathematical thinking and learning. Firstly, my teaching of the second lesson in LS2 seemed to precipitate concerns regarding the school's streaming policy for mathematics, which I articulated as follows:

> For me personally, I have very much pushed for streaming in maths in our school and once we started streaming approximately six years ago I would have said it was great, regardless of all the literature which does not recommend it. However, today was the first day where I really saw that you need some bright sparks in the class when doing problem solving to get the ball rolling or just to spark some ideas. This lesson has really caused me to reflect on and even question the appropriateness of streaming for maths in our school. (16/04/2013)

Secondly, I also highlighted students' apparent inability to utilise their mathematical knowledge in differing contexts and ultimately the impact this can have on their performance in standardardised tests:

> I just really feel that they can't transfer the maths knowledge that they're learning in maths class to knowledge outside of that, or even the Sigma-T ultimately at the end of the year. (16/04/2013)

Eimear, who taught the SEN group, raised similar concerns following her teaching of the third lesson in LS3: "I learned how difficult it is for my group to transfer or apply what they know, even with lots of scaffolding" (25/04/2013). Earlier, similar to views expressed by teachers elsewhere (INTO, 2013; 2015), she had made a very pertinent point regarding the appropriateness of standardised tests:

> In fairness, you know the way we're always talking about differentiating for them or scaffolding their learning or whatever but sure they all have to do the same standardised test in the end which seems in lots of ways unfair too, because they do know more than they show. It's just that it doesn't transfer. (16/04/2013, p.4)

Another interesting discussion took place following the teaching of the top stream during LS2. This lesson in particular had provided evidence regarding the dangers of proceduralising mathematics and, additionally, the apparent inadequateness of solely relying on standardised test scores to assess students' mathematical ability. It was obvious from how students approached the problem solving activities that many lacked any in-depth understanding of area and perimeter and had obviously proceduralised both. One student with a STen[8] of nine, when told that she couldn't use a ruler for the second part of the lesson, started measuring with her finger instead of utilising the information she had gained previously. Niamh, who taught the lesson, commented:

> I found it mesmerising that x with her STen of 9…the way they were approaching the problems was just mind-boggling. It was just incredible that they couldn't see they had all the information. They had done all the measuring already…I was just gobsmacked by so many things. Like multiplying, that's something we're going to have to be careful of…I'd put my hand up and say that I say 'area you multiply…perimeter you add' and I have drummed it into them, so I mean you can't blame them for multiplying. (11/06/2013)

Eimear and I agreed and admitted we had at times used similar strategies when teaching area and perimeter. Ultimately a valuable and worthwhile lesson was learned

[8] A STen score is a standard score with a distribution of 1-10 (STen means "standard-ten").

by all three of us, not just regarding area and perimeter but regarding how we should approach our teaching of mathematics in general. Appendix Z gives an example of one lesson plan mid-way through lesson study cycle two.

To conclude, both quantitative and qualitative findings pertaining to the third hypothesis suggest that LS is a feasible, worthwhile, efficient and effective model of providing CPD in AfL that can improve teachers' skills knowledge and use of AfL and their attitudes and beliefs towards AfL as a form of assessment; ergo the third hypothesis is accepted.

Conclusion

This chapter explored and analysed the findings from one school-based intervention that investigated the effectiveness of using LS to leverage change in teachers' AfL practices and attitudes towards FA as a form of assessment; and the concomitant impact AfL practices had on students' mathematical disposition and achievement. An in-depth account and analyses of the quantitative and qualitative findings pertaining to each of three research hypotheses was presented in turn. The next chapter synthesises these findings and analyses and discusses any implications they may have and uses them as the basis for making recommendations for future research and practice. It also looks at the various limitations of this study before reaching its final conclusions.

CHAPTER FIVE: CONCLUSIONS

Introduction

In this concluding chapter, I begin by revisiting the purpose of this study which was twofold: to investigate the effects of using assessment for learning (AfL) practices on students' learning of mathematics in fourth class in a girls-only primary school in Ireland; and to explore the potential of peer-to-peer learning, especially lesson study (LS), to impact teachers' knowledge of, and skills using, AfL principles, strategies and techniques. In particular, the study investigated whether the use of AfL practices improved the standardised mathematics scores of participating students in comparison to a similar cohort not involved in the intervention. It also examined if the intervention impacted students' dispositions towards mathematics, specifically by exploring the effects on students' mathematical confidence, their engagement with, and attitudes towards, mathematics. In short, the research could be viewed as an exploration of the interplay between a professional learning community (PLC), continuing professional development (CPD) in AfL and mathematics teaching and learning. A convergent parallel mixed methods design was used to achieve the aims of the research.

The next section begins with a brief synopsis of the various significant findings pertaining to each of three research hypotheses, in light of which conclusions are drawn. Following this, the different limitations of the study are discussed. Subsequently, any implications the findings have for research, policy and practice are considered and recommendations for future research presented. The chapter concludes with an epilogue, which provides a reflexive account of the intervention.

Summary of Key Findings and Conclusions

Since the principal findings from this research study have been discussed and analysed in depth in Chapter Four, this section presents a synthesis of the more salient findings and offers some conclusions. These are presented forthwith, taking each hypothesis in turn.

Hypothesis One

A nine-month, school-based intervention employing Assessment for Learning principles, strategies and techniques will improve the standardised mathematics results of participating students in comparison to a similar cohort not involved in the intervention.

The data set for hypothesis one was quantitative and comprised students' scores from two standardised mathematics tests specifically designed for the Irish context, Standardised Irish Graded Mathematics Attainment Tests (SIGMA-T) and the Drumcondra Primary Mathematics Test Revised (DPMT-R). A series of parametric and non-parametric statistical tests did not indicate any statistically significant difference between the mathematical achievement of the comparison and intervention groups in the Sigma-T, pre- or post-intervention, or in the DPMT-R when comparing both cohorts at the end of fourth class. Furthermore, analysis by stream indicated that, post-intervention, the scores achieved by students in the top, middle and SEN intervention groups were not statistically significantly different when compared to their peers in the comparison group. In sum, it is clear that findings were consistent across tests and across streams and so it was concluded that the first hypothesis should be rejected. Nonetheless, it should be noted here that the three teachers were surprised by these results. It was their professional opinion, from their classroom observations and daily interactions with both cohorts over the course of one academic year, that students in the intervention group had made greater progress in their mathematics learning in fourth class than had their peers in the comparison group. Perhaps, as previously highlighted, the intervention was too short to make a quantitative difference on standardised test scores or, alternatively, it might suggest that standardised tests in their current format cannot adequately measure the effects of AfL. Indeed, findings from recent INTO (2015) focus groups on curriculum suggest that primary teachers are concerned about an over-reliance on standardised tests and have expressed dissatisfaction with these tests arguing they do not reflect what teachers are teaching, take group and pair work into consideration or take account of students with English as an additional language (EAL). Additionally, it is worth remembering Eisner's (2005) remarks, in a different context, when he cautioned "what we measure we focus on" and stated:

> Assessment should try to provide a more complete picture of the developing child. Put most simply, we need to be concerned about more than the measurable. Not everything that matters is measurable, and not everything that is measurable matters…We need forms of assessment that help us better understand how to nourish the children we teach. (p.18)

This prompts the question as to how the learning that AfL practices are trying to inculcate can be adequately measured.

Hypothesis Two

The use of AfL strategies and techniques, and the adoption of AfL principles, will enhance children's mathematical confidence, and improve their engagement with, and attitudes to, mathematics.

Two instruments in particular, the Attitudes to Mathematics Questionnaire (ATMQ) and the Children's Assessment for Learning Audit Instrument (CAfLAi), were used to collect quantitative data to test this hypothesis. Beginning with the ATMQ, analysis of the data suggests that the children believed their attitudes to mathematics were substantially more positive following the intervention. They felt more confident doing mathematics, and they concluded that their engagement with, and motivation to do, mathematics had improved. Regarding the CAfLAi, and mindful of the aforementioned limitations of this instrument, statistical tests revealed a statistically significant increase in the children's use of AfL strategies and techniques by the end of the intervention. These findings from the quantitative data were corroborated by qualitative findings from the children's and teachers' learning logs and from the children's focus group interviews. It was concluded, therefore, from both quantitative and qualitative evidence, that the second hypothesis should be accepted.

In addition, it is worth reiterating that, by the end of this intervention the children readily used the language of AfL, engaged in self- and peer-assessment and showed early signs of self-regulation and metacognition. Furthermore, they clearly articulated their ideas and opinions regarding learning, teaching, AfL, LS and mathematics. Therefore, the data also suggest that children, even at primary level, are capable of engaging in AfL practices and playing an active role in their own learning.

Hypothesis Three

Peer-to-peer professional learning is a feasible, worthwhile, efficient and effective model of CPD in AfL and will improve teachers' skills, knowledge and use of AfL and their attitudes and beliefs towards AfL as a form of assessment.

Results from the AfLAi reflect those of the CAfLAi in that they indicated that the three teachers, similar to their students, believed their use of AfL principles, strategies and techniques had improved significantly following the intervention. The teachers reported their use of the first three strategies, sharing learning intentions and success criteria (LISC), questioning and classroom discussion (QCD), and feedback (FB), had moved from sporadic, pre-intervention, to established afterwards, while their use of peer- and self-assessment (PSA) had moved from sporadic to emerging, although they still had some reservations about using peer- and self-assessment (PSA) unless it was scaffolded. These results mirror findings from Lysaght and O'Leary's (2013) national survey of Irish primary teachers that also found PSA to be the least embedded AfL strategy.

Analysis of the qualitative data from the teachers' learning logs, various reflection meetings and the researcher's journal, corroborate quantitative findings from the AfLAi in that the teachers articulated how they considered their skills, knowledge and use of AfL, and their attitudes and beliefs towards AfL as a form of assessment, had improved as a result of their participation in the intervention. Furthermore, they believed they were on their way to embodying the *spirit* of AfL and viewed AfL as a positive leverage for change in their classrooms. The teachers concluded that it was their participation in the LS process that had facilitated this change in their AfL practices and in their attitudes to AfL as a form of assessment. Additionally, they maintained that this had a concomitant positive impact on their students' mathematical disposition and achievement. They considered LS, as a model of CPD, had been more effective that previous approaches they had engaged in ,such as workshops or one-week summer courses, particularly since it was on-site, ongoing, had increased collaboration and deprivatised classroom practice, while also improving mathematical thinking and learning. In sum, the teachers believed their involvement in the intervention had been very worthwhile and rewarding, albeit challenging at times. Considering the evidence, it was concluded that the third research hypothesis should be affirmed.

Limitations of the Study

While this research contributes new insights regarding AfL, LS and mathematics, it is important to consider its findings within the context of the study's various limitations: methodological, contextual and temporal. The fact that this research was small-scale, took place in one single-sex school and employed a case study strategy with convenience rather than non-probability sampling, can be considered the first limitation of this study. Fears regarding generalisability and replicability are usually the primary concerns under such circumstances. Nevertheless, using a case study strategy was considered the most appropriate approach for this research since it enabled in-depth exploration of each hypothesis, facilitated "thick description", and encouraged "relatability" so that readers could compare this study with their own situations and potentially benefit from its findings. Scholars such as Yin (2009) argue that generalisations can be established if a study is replicated several more times in different circumstances and so the extent to which findings from this study can be generalised requires further investigation.

A second limitation of the study is the fact it was conducted in my place of work. While scholars such as Gray (2014) and Robson (2011) acknowledge there are various advantages to undertaking "insider" research, for example, having intimate knowledge and experience of the context of the study, it also brings limitations such as maintaining objectivity or dealing with the dual role of researcher and colleague

(Robson, 2011). Thus, as advocated by Gray (2014), I endeavoured to maintain a sense of detachment, remained mindful of the danger of role confusion (Asselin, 2003) and, as detailed in Chapter Three, remained open to contrary findings and tried to guard against my own biases. In this way, being an "insider" did not unduly influence the research in a negative way but, instead, enabled unique insights that may not have been possible otherwise.

A further limitation of the study is that none of the teachers had previously completed specific CPD in AfL, although the I had access to online material from the *Assessment for Learning* Masters Programme run by St. Patrick's College, Dublin, and in addition had researched widely on the subject. Furthermore, the teachers had never engaged previously in LS and so this too may have impacted the study's findings in that the quality of the LS process, especially in the first LS cycle, may have been enhanced had the teachers been familiar with LS, or had access to a Knowledgeable Other or someone with LS experience.

Similar to many empirical studies, the research presented here was also limited by the measures used, which included questionnaires, interviews and observations. In this study, the limitations of the ATMQ, and in particular the CAfLAi, have been discussed previously, as have the steps taken by me to ensure the trustworthiness, validity and reliability of the research. To reiterate, in order to alleviate the limitations of the instruments utilised in this study, each was piloted at least once and, as recommended by research scholars (e.g., Miles et al., 2014), there was triangulation by data source (e.g., teachers and students), method (e.g., questionnaires and focus groups) and data type (quantitative and qualitative).

Finally, some temporal limitations need mentioning. It was highlighted in Chapter Two that it takes time and perseverance if teachers and students are to become proficient users of AfL (e.g., Bennett, 2011) and engage in AfL practices in the *spirit* in which it is meant (Marshall & Drummond, 2006). Similarly, scholars (e.g., Cajkler et al., 2015) have emphasised it requires time and sustained commitment for LS to take root and for teachers to reap the full benefits to their thinking and practice. Thus, the intervention may just have been too short to impact student achievement in a way that is measurable using the standardised tests currently available, or to significantly enhance teacher practice through engagement in LS. In sum, if the substantial rewards promised by both AfL and LS are to be fully affirmed, it will be necessary to carry out similar studies in multiple sites but over a longer period of time.

Implications and Recommendations

Findings from this study have particular implications for policy, practice and research, especially in areas such as assessment, teacher professional development and student voice. These implications are discussed forthwith, beginning with implications for policy.

Implications and Recommendations for Policy

Assessment

As argued (DES, 2012a), arrangements for evaluation and assessment at policy level should be structured and planned. Not only that, scholars (e.g., Cooper & Cowie, 2009) posit that countries need better alignment or interconnection between the macro or policy-related factors and the micro or school-related level to ensure more consistent use of assessment throughout their education systems. In the Irish context, the key policy document at primary level regarding assessment remains the

National Council for Curriculum and Assessment (NCCA) authored *Assessment in the Primary School Curriculum: Guidelines for Schools* (2007). However, few teachers received the promised Department of Education and Skills-funded CPD in assessment following its publication, possibly explaining why at policy level there is an apparent acknowledgement that teachers' assessment literacy and assessment practices, in both AfL and Assessment of Learning (AoL), need to be improved. A recent submission by the DES (2012a) to the OECD for the *Review on Evaluation and Assessment Frameworks for Improving School Outcomes* stated that "student assessment in terms of its function and implementation remains an issue in primary schools" (p.14) and highlighted that there is "limited use of assessment for formative purposes, despite various interventions and support in this area" (p.13).

While there is no clear overall evaluation and assessment policy in the Irish context, Hislop (2013), chief inspector with the DES, recently described the *Literacy and Numeracy Strategy* (DES, 2011a) as "groundbreaking" since it is "perhaps the only statement of public policy on how evaluation and assessment arrangements are intended to work together in the Irish school system" (p.8). The strategy makes clear all teachers are expected to be assessment literate. It highlights the need "to use a continuum of well-considered assessment approaches to determine the next steps in learning and in planning approaches to teaching" (p.32) and "to combine good *assessment for learning* practice with appropriate *assessment of learning* approaches" (p.74). However, while it states, "*AfL* should be used to inform all teaching", the strategy also emphasises "It is not used sufficiently widely in our schools and we need to enable teachers to improve this practice" (p.74). Nevertheless, while the strategy mentions "access" to approved CPD in literacy, numeracy and assessment should be provided to primary teachers, five years after the publication of the *Literacy and*

Numeracy Strategy no such CPD has been forthcoming. The challenging economic circumstances in Ireland since the NCCA guidelines were sent to individual teachers in early 2008 possibly contributed to this lack of CPD in assessment, with other areas being prioritised. However, it remains to be seen if investment in quality, ongoing professional development in assessment will be forthcoming.

Heritage and Wylie (2013) argue that the effective implementation of AfL, even its very existence, depends on appropriate policy support. Therefore, my first recommendation regarding policy is that an evaluation and assessment framework should be developed which would outline, for example, the knowledge, concepts, goals, standards, processes, outcomes and resources necessary to guide assessment in the Irish context. Indeed, Hislop (2013) has argued similarly in a recent speech:

> We in Ireland could find it beneficial to examine deliberately how we want the essential components of an evaluation and assessment framework to develop in a coherent way that will support the sort of student learning to which we aspire in the school system of the 21st century. (p.8)

As envisaged, such a framework should provide answers for questions such as the following:

- What are our goals for assessment? (nationally, at class, individual and school-level);
- How will we achieve our goals?
- How do we best align assessment research, policy and practice?
- What training, supports, resources and funding will we provide?
- What will assessment look like in our schools in 2018, 2020 and subsequently?

Furthermore, an evaluation and assessment framework would explicate the government's vision and plans for a balanced assessment system and should improve assessment practices as well as making them more widespread, systematic and consistent. As is, the status quo in Ireland with regard to assessment policy contrasts with other jurisdictions, for example Scotland, New Zealand and Australia, where policymakers have committed significant resources to support the implementation of formative assessment, mainly focused on the provision of teacher CPD to enable teachers to use AfL as an overall approach to teaching and learning (Heritage, 2013).

Continuing professional development

Findings from this study suggest how difficult it would be to implement AfL in the *spirit* in which it is meant without the appropriate assessment skills or disciplinary knowledge. The teachers believed it was through their engagement in ongoing CPD, LS in this case, that they acquired the necessary skills and assessment literacy to successfully implement AfL principles, strategies and techniques in their classrooms. They admitted it had been a challenging process that required time and sustained effort to effect change in their assessment practices. Scholars (e.g., Wiliam, 2011b), too, argue that teachers require CPD and ongoing support to enable them to implement and embed AfL effectively and consistently in their classrooms. If the Irish government is serious about improving teachers' assessment literacy, and making AfL a reality in all classrooms, then teachers require CPD in assessment as a matter of urgency. Findings from this research imply that LS is an effective and non-expensive model of CPD in AfL. Therefore, I recommend that the DES commission a number of Irish experts in AfL and LS so they can collaborate on the development of an assessment resource pack similar to the *Keeping Learning on Track* (KLT; Wiliam & Thompson, 2007) professional development programme in the US, *the Assessment is For Learning (*AifL; Hayward, Priestley & Young, 2004) programme in Scotland or

Embedded Formative Assessment (EFA, Wiliam & Leahy, 2015) in the UK and USA. This assessment resource pack would be specifically geared to support the development of formative assessment in the Irish context and would include sufficient materials for two years of monthly meetings after which teachers should have developed enough understanding to produce their own materials for subsequent years or, alternatively, return to year one. It would contain not only content regarding AfL theories, principles, strategies and techniques but also would detail how teachers could use the LS process as a model of CPD to implement and embed AfL practices in their schools, rather than the TLC/PLC process used in the other programmes just mentioned. As envisaged, this assessment resource pack would adopt the same "tight but loose" framework of the KLT (Wiliam & Thompson, 2007) programme so that it would be sufficiently flexible to allow it to be adapted to local circumstances (the "loose" part) but sufficiently rigid to ensure that any modifications made would maintain fidelity to the original design (the "tight" part). Once completed, the experts would trial the assessment kit with a small number of selected schools and, after one initial face-to-face presentation with the teachers involved, would then make use of webinars for follow-up meetings and discussions. By using webinars there would be less demand on the experts' time than if they had to visit individual schools, teachers' participation in the project would not be determined by geographical location and it should ultimately encourage teachers to embrace a leadership role in embedding AfL practices within their own schools through the use of LS while at the same time having recourse to experts or KOs through the webinars. If successful, this idea could then be extended to other schools and so it is, ultimately, scalable.

Standardised tests

While the *Literacy and Numeracy Strategy* (DES, 2011a) acknowledged that "standardised tests cannot measure the progress students have made in achieving many important learning outcomes" and admitted that "the aggregated results will not tell the whole story" (DES, 2011a, p.75), it nevertheless increased the number of times primary schools must send aggregated data from standardised tests to the DES from two to three (end of 2^{nd}, 4^{th} and 6^{th} classes). While I fully acknowledge the value of using standardised tests, especially for diagnostic purposes, I believe that the introduction of representative sampling of students across grades at primary level to access what students know would have been a better alternative to the increase in mandated testing and reporting. Additionally, I suggest that the fact standardised test scores have to be reported to parents, the Board of Management and the DES, could have unintended consequences such as exerting pressure on teachers to teach to the test or result in children being labelled with comments such as "She's a six" (i.e. Sten score). Indeed, various children in this study demonstrated anxiety regarding standardised tests with some asking, towards the end of the intervention, "Is this on the test?", demonstrating how conscious they were of the impending standardised tests. Indeed, most could remember what Sten they got in previous years, thus enabling comparison with their peers if so desired. I am not recommending the abandonment of standardised tests. Rather, I would like to see more emphasis and value placed on classroom assessments. This is because teachers can utilise a range of assessment methods to assess their students' understanding of a particular topic and can analyse these assessments and react immediately, or in the next lesson, to meet individual student's needs. Additionally, students would be less anxious about classroom assessments and more likely to demonstrate what they know. Thus, I recommend there should be less emphasis from a DES perspective on the reporting of standardised test results and greater attention paid to formative assessment, especially at primary level, with teachers having primary responsibility for this. Nevertheless, in order to make this a reality most teachers urgently need CPD in assessment.

Student Voice

Scholars (e.g. Heritage, 2013) have highlighted the importance and value of involving students in their own learning and engagement. However, it is interesting to note that policy documents in the Irish context are largely silent on this issue with little expectation of student involvement, albeit the *Literacy and Numeracy Strategy* (DES, 2011a) does mention "conversations with the learner" (p.77). Indeed, AfL seems to be viewed as an approach that helps teachers optimise their instruction rather than one that helps pupils decide where they are in their learning. Thus, policy wise, more thought is needed about how to include children in the assessment process. As has been seen from this study, young students, even at primary level, are capable of expressing their opinions about learning and of engaging in AfL practices, even self- and peer-assessment. Listening to students' voice is not only underpinned by children's rights legislation (Elwood, 2013; Hopfenbeck, 2013) but also offers unique insights into teaching and learning that remain largely untapped. Therefore, I recommend that future policy decisions, at national and local level, incorporate students' perspectives from primary through to tertiary level, since it is students who are best placed to provide informed perspectives of how they experience learning and assessment at the various stages of their education journey. This could be achieved through the establishment of student councils at school and district level, but would succeed only if the process is not tokenistic and students feel that their input is valued and they are given real opportunities to discuss issues that matter, such as learning, teaching and assessment.

Linked to the idea of listening to the student voice, and taking account of the centrality of the learner, I also recommend the development of a reliable instrument that would measure children's AfL practices, perhaps by building on the CAfLAi. This instrument would not only reveal the student perspective on AfL but would also offer feedback to teachers regarding AfL practices in their classrooms. Additionally, when used in conjunction with the teachers' AfLAi, such an instrument would provide a more complete picture of AfL practices within a school, since the perspectives of both teachers and students would be considered.

Implications and Recommendations for Practice

As highlighted in Chapter Two and in a previous section discussing policy regarding CPD in assessment, various scholars (e.g., Leahy & Wiliam, 2012; Moss & Brookhart, 2009) agree that AfL, when used effectively, is a warranted strategy that can improve student learning and achievement. However, researchers (e.g. Wiliam, 2011b) also concur that implementing AfL effectively is complex and challenging for many teachers. Thus, researchers (e.g., Wiliam, 2016) argue that sustained, practice-focused, job-embedded professional development is needed to support teachers in their development of formative assessment and particularly highlight the critical role of school-based professional learning communities in advancing teachers' AfL practices (e.g., Birenbaum, Kimron & Shilton, 2011; Opfer & Pedder, 2011).

Since this is a small-scale research project with just three teachers and fifty-one students, it is not feasible to generalise from its findings, although it is possible to affirm findings from previous research regarding AfL and CPD. Similar to other studies (e.g., Harrison, 2013), and highlighted earlier, teachers in this study found it challenging to successfully embed AfL in their classrooms in the *spirit* in which it is meant. However, because of the collaborative process they engaged in through LS and with support from the principal, by the end of the intervention the teachers felt they had achieved a much better understanding of AfL principles and practices and so were better able to successfully implement them in their classrooms. They particularly highlighted the benefits of collaboration onsite in a sustained way, and articulated their beliefs regarding the positive impact this had on their AfL practices. Therefore, until CPD in assessment is forthcoming, I recommend that teachers should form PLC or LS groups when trying to develop AfL practices in their classrooms. Additionally, when and if teachers receive CPD in assessment I recommend that consideration be given by the DES to using LS as the model of CPD in AfL since it embodies many of the core features of effective CPD such as active teacher participation, sustained teacher learning linked to practice and collaboration (e.g Darling-Hammond, 2009; Desmione, 2009). While there are some publications available to support teachers in their use of AfL such as *Assessment in the Primary School Curriculum* (NCCA, 2007), the guidelines in *Aistear, Supporting Learning and Development through Assessment* (NCCA, 2009b) and the *Action* section of the NCCA website, the extent to which AfL practices are being enacted in primary classrooms in Ireland in accordance with the *spirit* rather than the *letter* of AfL through such practices as the promotion of metacognition or a change on the power balance between teachers and students, remains an open question.

As argued (Hattie, 2012; Hayward, 2012), engaging in effective AfL would also necessitate change in how teaching and learning is generally perceived and would ultimately impact classroom practices, roles and relationships. Teachers sometimes struggle with the changes that engaging in AfL practices bring. Indeed, engaging in AfL can be challenging as teachers in this study admitted, for example when they found it difficult to relinquish control when students were engaging in PSA. However, Wiliam (2010) cautions against the idea that the role of the teacher is eroded, arguing that it is just different:

> Teachers do not create learning; only learners can do this and so many have called for a shift in the role of the teacher from the "sage on the stage" to the "guide on the side". The danger with such a characterisation is that it is often interpreted as relieving the teacher of the responsibility of ensuring that learning takes place. What I propose here is that the teacher be regarded as responsible for "engineering" a learning environment, both in its design and its operation. (p.152)

In such classrooms children would become more active in their own learning and assessment and adopt more responsibility for it, whereas teachers would adopt a more facilitatory role. The importance of including children's perspectives at policy level has been highlighted and findings from this study have shown how articulate and capable children can be regarding AfL practices. Indeed, scholars (e.g., Heritage, 2013; Elwood & Lundy, 2010) are increasingly looking at assessment from a children's rights approach. As Heritage (2013) argues, surely children's rights include access to learning and deserve a voice in matters that affect their futures. Thus, in the context of this research, it is recommended that pupils be given a greater role with regard to their own learning and assessment, particularly through the use of regular self- and peer-assessment.

Implications and Recommendations for Research

Replicate the current study

Scholars (e.g., Makel & Plucker, 2014; Warne, 2014) argue there is merit in replicating educational research since it makes findings more reliable and trustworthy, helps shape educational policy and practice and, ultimately, improves the lives of students. While findings from this intervention appear positive, as mentioned earlier, they are not generalisable since it is only a single case. Therefore, I recommend this study be replicated in the Irish context in multiple sites, preferably serving similar populations of pupils. Then, findings from this study could be corroborated or disconfirmed, thus contributing further to the knowledge base regarding AfL. The resulting data, along with similar studies like Lysaght's (2009), might then expedite large scale research into AfL in the Irish context or, at the very least, contribute to the development of sound assessment policy and practice.

Pilot Lesson Study

Teachers in this study indicated they had enjoyed participating in LS and found it had cultivated professional dialogue and reflection. They particularly liked the time they spent collaborating and believed that their engagement in LS as a model of CPD had a greater impact on subsequent teaching and learning in their classrooms than had the other models of CPD in which they had participated, such as one-week summer courses or one-day workshops. They explained this was probably due to the fact it was possible to implement what they learned immediately in their classroom and over a sustained period. Thus, I recommend that a pilot study into the use of LS as a model of CPD in AfL in mathematics be carried out in a small number of schools in Ireland over a three-year period. It would then be possible to explore the effects of sustained LS as a model of CPD in AfL and these findings could be used to guide more broadscale research into LS in the Irish context, and also future policy and practice regarding CPD. This resonates with Corcoran's (2008) recommendation whereby, instead of participating in a summer course, teachers would dedicate twenty hours communally during the school year to run four to five cycles of LS and, in exchange, would receive three days of extra personal vacation, as they do on completion of a summer course.

One further point warrants mentioning here. When preparing the research lessons it was difficult and time-consuming to locate appropriate mathematics problems that demanded sufficiently high levels of critical thinking, and allowed for differentiation, collaboration and active learning. Delaney (2012) has also highlighted that many problems in Irish mathematics textbooks are of poor quality, while anecdotally, many teachers concur and are also critical of the lack of readily available problems, especially ones geared to the Irish context. Primary teachers must teach a broad range of subjects and so few have sufficient time to source suitable mathematical problems to meet their needs on a regular basis. Therefore, I recommend that the NCCA compile and make available online a bank of *quality* mathematics problems for each class level. With the introduction of a new mathematics curriculum at primary level imminent, the availability of such problems is important to ensure less reliance by teachers on textbooks.

Dualism's dividends

Popham (2013) recently used the title, *Dualism's Dividends,* when discussing the different and normally separate worlds of academics and practitioners, arguing that practitioners "make things work" while "academics increase knowledge about how things work" (p.vii). Acknowledging Popham's (2013) argument, Black (2015) posits that this is a problem underlying many innovations and calls on academics in education:

> To find ways to build fruitful interactions between their world and the world of practising teachers if they are ambitious to explore, and to learn how to implement, the potential benefits of their work. (p.174)

As mentioned previously, at one particular point in this study, when struggling with LISC, we were fortunate to have recourse to professional expertise and thereafter made greater progress. Consequently, while acknowledging some partnerships between schools and third level institutions already exist, my penultimate recommendation is that more academics and practitioners should work together to secure dualism's dividends in the Irish context. This could be mutually beneficial and has the potential to impact the learning of many more students.

Other research

As suggested at the beginning of this chapter when revisiting hypothesis one, and also by Lysaght (2009), perhaps standardised tests in their current form might not be the best way to measure the impact AfL has on learning. Therefore, my final recommendation is that researchers in the Irish context try to develop assessments that are congruent with contemporary theories of learning and that probe students' higher order thinking skills and understanding rather than memorisation of facts. As part of this research, consideration should also be given to the exploration of ICT-based assessments in our schools and, linked to this, investigation into the equivalence of test outcomes across digital and pen-and-paper modes of delivery. Thus, the establishment of the new Centre for Assessment, Research, Policy and Practice (CARPE) in Dublin City University (DCU) is to be welcomed with regard to assessment research in the Irish context, as is its recently announced joint research project with the INTO into standardised tests.

Synthesis of Recommendations

To conclude this section, I reiterate the need for research, programmes and initiatives that incorporate AfL and I now present a summary of the recommendations arising from this research and discussed above, once again organised under the headings of policy, practice and research.

Policy:

- Develop an assessment and evaluation framework to provide a vision for assessment, and link policy and practice in the Irish context;
- Invest in mandatory CPD in assessment to improve teachers' assessment literacy;
- Redesign standardised tests and use alternative measures;
- Include students' perspectives.

Practice:

- Encourage teachers to try using AfL strategies and techniques and to engage in school-based, sustained, teacher-led CPD;
- Use children's perspectives. Give them more responsibility and autonomy and get them more involved in their own learning.

Research:

- Replicate this study in multiple sites in the Irish context serving similar populations of students, with a view to generating data regarding AfL that can be used to inform assessment policy and practice;
- Conduct a pilot of lesson study nationally;
- Increase collaborative research between practitioners and academics (dualism's dividends);
- Develop third generation assessments congruent with contemporary views of learning and investigate the value and feasibility of digital modes of assessment.

In conclusion, findings from this study suggest that employing AfL practices had an affective impact on students' learning of mathematics but did not have a statistically significant effect on students' standardised mathematics scores. Findings also indicate that, following their participation in three cycles of LS, teachers considered it an effective model of CPD that helped improve their skills, knowledge and use of AfL, and their attitudes and beliefs towards AfL as a mode of assessment. The implications of these findings for policy, practice and research have been discussed and it has been highlighted that further research is required before any generalisations about the findings can be made.

Formative developments following the intervention

This research acted like a catalyst for various formative developments that have taken place in Scoil na nAingeal since its completion. As a result of our participation in the intervention, the other two participating teachers and I have embedded formative assessment practices in our classrooms, not just in mathematics. This use of AfL strategies and techniques, coupled with our participation in three cycles of lesson study, has improved our teaching and has resulted in greater refinement of learning outcomes in mathematics. Additionally, we have increased our collaborative planning of mathematics lessons and continue to use the skills we learned from the live lessons to focus more on student thinking and anticipated responses. Other changes in our teaching include improved formative use of standardised tests, giving students greater responsibility for their own learning and assessment, as well as greater use of collaborative learning and problem solving approaches in mathematics lessons. Linked to this, our observation of live lessons highlighted the need to have students of all ability levels in our mathematics classes and prompted a review of Scoil na nAingeal's policy of streaming for mathematics. As a result, streaming in mathematics is currently being phased out in favour of mixed ability groupings and in-class support. From our live lesson observations it was also evident to each of us that for lesson study to be at its most effective, it was imperative that the student groups being taught for each research lesson should be as similar as possible with regard to mix of ability levels, articulateness etc. Otherwise you were not just making adaptations to improve the lesson plan per se but were also have to take into account making changes for the differing ability levels of each stream, thus changing the dynamics of the lesson study process. One final significant impact of this research is also worth noting. Since the intervention lesson study has become the primary vehicle for CPD in our school. I have been involved in lessons study groups with various teachers, exploring literacy at infant level and mathematics at first class level. Indeed, this school year a group of six other teachers and I are in a lesson study group investigating coding for junior infants to first class and we have completed two live research lessons to date. To conclude, I believe that this research, and particularly our use of lesson study has contributed to the deprivatisation of classroom practice within Scoil na nAingeal, and this is to be welcomed.

Epilogue

At this point of the research journey, it is opportune to offer some final reflections regarding the intervention. Assessment for learning, CPD, especially LS, and mathematics are areas of national and international interest and so the focus of this study was timely, topical and of particular interest to the three teachers involved. On a professional level, it afforded the opportunity to continue our journey of lifelong learning in a job-embedded, sustained, collaborative, supportive and active way. In retrospect, prior to the intervention, while each teacher felt confident teaching mathematics, our knowledge of AfL was rudimentary and we had no previous experience of LS. Nonetheless, by the project's conclusion my colleagues and I were convinced our skills using, and knowledge of, AfL had greatly increased and we believed we were implementing AfL practices in the *spirit* of which they are meant. Furthermore, by the end of the intervention, AfL was integral to teaching and learning in our classrooms and had become a vital part of the overall assessment process, not just in mathematics. Regarding participating students, by the end of the project it was clear that they readily engaged in AfL strategies and techniques and easily articulated their ideas about learning. Once again, this was the case not just in mathematics but increasingly in other areas too, for example history, art, homework, etc. Indeed, the children's growing autonomy and their capacity to engage in discussions about learning was a particular highlight of the intervention since I have always valued their perspectives. The fact that the children seemed to experience an intrinsic enjoyment or fun in their learning was also noteworthy.

Notwithstanding, while there were many highlights over the course of the study, it was the live research lessons that particularly stand out. Teaching is such an isolated profession that it was a privilege to have this unique opportunity to observe colleagues, whom I admire and respect, teach a lesson we had planned carefully together. It was even better to have the time and freedom to closely observe students' thinking, interaction and collaboration, in real time, during mathematics, without having to deal with the normal classroom pressures when teaching oneself. In fact, we all agreed LS was the best model of CPD we had ever engaged in, particularly because it was collaborative, sustained and onsite, and had positively impacted our teaching and our students' learning. Of course, the project was not without its challenges, especially organisational, but none of these proved insurmountable. Personally, the greatest challenge was probably trying to balance my role as teacher, researcher, participant, and colleague, while also acting as deputy principal of a very busy school. Persevering with the intervention was hard work at times, but definitely worthwhile.

Undoubtedly, the intervention helped us realise the value of using AfL practices and engaging in LS; and the potential of both to impact mathematics learning, teaching, and achievement. Ultimately, however, this study was about learning. Through participation in the project, students enhanced their learning by engaging in AfL practices, while teachers learned together and improved their practice through engagement in LS – a veritable community of learners. It is important to note, however, that embedding AfL effectively and developing quality LS is a slow process that takes time before reaping any dividends. Additionally, one needs to be prepared for changes in classroom practices, roles and relationships. To conclude, there is still much to be learned regarding AfL, LS and mathematics, especially in the Irish context but, hopefully, this study will have made some contribution to the knowledge base.

APPENDICES

Appendix A: Popular AfL Techniques

(Many of these techniques come from *Embedded Formative Assessment* (Wiliam, 2011b)

1. **W.A.L.T. (We are learning to…).**

   This technique was used in most lessons and refers to the specific learning intention/s of that particular lesson. For younger children the acronym W.A.L.T. (we are learning to) can be used.

2. **W.I.L.F. (What I'm looking for…).**

   Similarly, this technique was used in every lesson and is often used with children to refer to the success criteria which the teacher, children, or both, decide are required to be successful in their learning. As above, the acronym W.I.L.F (What I'm looking for) is often used with younger students.

3. **Two stars and a wish**

   Students use this technique when engaging in self- and/or peer-assessment. They choose two things they like about the piece of work and one thing they think could be improved on.

4. **Think-pair-share**

   Children are encouraged to think about their learning/answer before sharing this information with another child. They could share two/three things that they learned, things they found easy/difficult or what they might like to study further. It works better if children are partnered with someone they feel happy to share with. Subsequently, children will then share their ideas with the whole class.

5. **Thumbs-up, thumbs-down**

   Children indicate their feelings about or their level of understanding of a particular concept by using their thumbs. If the students agree with an answer or idea, they hold up their thumb; if they disagree they hold their thumb downwards; and if they are unsure they hold their thumb sideways.

6. **Fist to Five**

   This technique is similar to 'thumbs' but instead of three levels of understanding or feelings, there are now five. Ranging from holding up all fingers which indicates full understanding of a concept, all the way down to holding up their fist for no understanding, children indicate where they are in their learning. This is a useful technique for using with the whole class as it quickly gives teachers an understanding of where their students are in their learning of a particular concept or topic.

7. **Brainstorming**

   Children come up with as many ideas as possible.

8. **Learning logs**

   Learning logs are a useful way of getting students to document their self-assessment of, and reflection on, their learning/written work at the end of a lesson. Asking students to respond to two or three of the following prompts can help ensure a more thoughtful reflection process:

   - Today I learned …
   - I was surprised by …
   - The most useful thing I will take from this lesson is …
   - I was interested in …

- What I like most about this lesson is …
- One thing I'm not sure about is …
- The main thing I want to find out more about is …
- After this session, I feel …
- I might have gotten more from this lesson if … (Wiliam, 2011b).

9. ABCD cards/corners

Each student has a set of cards: A, B, C and D. AM see my CPDs for example. There can be one or more correct answers and it is easy for the teacher to see at a glance how students are doing in their learning. ABCD cards can also be used to attain views about various topics, e.g. Which of the following do you believe? (a, b, c or d) or can be used as a bridge between lessons. One disadvantage of ABCD cards is that teachers are usually required to have planned the questions ahead of time and so they are not appropriate for spontaneous discussion.

10. Rubrics

"A rubric is an assessment tool which describes varying levels of quality in a specific piece of work" (NCCA, 2007, p.24). Rubrics can be designed by the teacher alone or in collaboration with the students. Rubrics have two essential features:

- A list of criteria which clarify the important elements of the work
- Levels of quality which explain the quality of work expected at each level

11. K.W.L.

A know, want to know, learned grid is completed by the children over time. At the start of the topic, the child focuses on what he/she already knows, then what he/she would like to learn and finally what new things he/she has actually learned.

12. Traffic lights/Coloured cups (NCCA)

Traffic lights are used by many teachers to help activate students as owners of their own learning. The teacher shares the learning intentions and success criteria with the children and then they assess their level of understanding. Traffic lights can be used at any stage in the lesson and are similar to thumbs. They are easily used with young children and can be used for group work and class discussions too. Using traffic lights, the children can indicate their level of understanding, or their feelings by holding up the appropriate coloured card or by putting it at the top of the pile if the cards are left in the desk:

- Green Card = I fully understand or I feel confident about my learning
- Orange Card = I'm still not sure yet
- Red Card = I don't understand or I have not learned what was intended
- As an adaptation of traffic lights, children can draw red, green or orange dots at the top of the page to indicate their level of understanding and this can be very useful for revision purposes when doing tests.

13. Sharing exemplars/best samples

Sharing exemplars provides an opportunity to talk about 'quality' work. |One can use samples of work that exemplify excellence or what you are looking for, that clarify the next steps students can take.

14. Find the errors and fix them

This technique is particularly suited to mathematics. Say, for example, if a student does ten mathematics questions and gets three wrong, instead of putting an x beside these answers one just says to the student, "Three of these are wrong. I want you to find them and fix them".

15. Preflight checklist

Preflight checklist is used when a task has to satisfy multiple requirements before submission. Before a student hands up an assignment, it is must be signed off by a buddy who checks that all requirements are satisfied. If, when a teacher corrects it, items on the preflight checklist are not up to standard, it is the buddy who is taken to task.

16. Phone a friend/50-50/ask the audience

This technique is similar to the game show 'Who wants to be a millionaire' in that the students can seek help from their peers.

17. I-You-We checklist

Following a group activity each student records one way in which s/he has contributed, one way in which another student has contributed, and an evaluation of the work of the group as a whole.

18. Hot-seat questioning

The teacher asks s student a question and this is followed by further questions which probe the student's knowledge or ideas in depth.

19. Learning portfolios

Learning portfolios don't just display the latest and best student work. Rather, they provide an incremental view of a student's ability in that students can look back at earlier work and see how their work has progressed. By seeing how they have improved, students understand how further improvement is possible. Additionally, by focusing on improvement, the learner is more likely to have a growth mindset. Students can start developing learning portfolios when they are in infants.

20. Mind/concept maps

These are excellent for getting students to focus on key ideas and to illustrate and visually organize these concepts. Tony Buzan has done great work here.

21. PMI

Children are encouraged to assess their own work by identifying elements that they think worked well (plus), did not work so well (minus) and elements they found interesting. PMI elements can be written or presented in graph or table format.

22. Comment Only Marking

Based on research which recommends that the children receive comments as feedback on their work rather than grades or a combination of both (e.g. Butler, 1988; Wiliam, 2011b).

Appendix B: Seven Principles of Good Feedback

Nicol and Macfarlane-Dick (2006) proposed the following seven principles of good FB practice in relation to the development of self-regulation, stating that good FB practice:

1. Helps clarify what good performance is (goals, criteria, expected standards);
2. Facilitates the development of self-assessment (reflection) in learning;
3. Delivers high quality information to students about their learning;
4. Encourages teacher and peer dialogue around learning;
5. Encourages positive motivational beliefs and self-esteem;
6. Provides opportunities to close the gap between current and desired performance;
7. Provides information to teachers that can be used to help shape the teaching. (p. 205).

Appendix C: The LS Cycle in further detail

At the initial stage of the LS process, teachers in the LS group come together with a shared goal; analyse the curriculum, examine research and explore resource materials; and over time, work collaboratively to design a detailed research lesson, while anticipating student thinking and learning (Lewis & Hurd, 2011; Murata, 2011). It is this pervasive concern with student thinking and learning that binds the various parts of the LS cycle together and distinguishes LS from other types of professional development activities such as action research (Cerbin & Kopp, 2011; Corcoran, 2007; Murata, 2011). The idea here is not to plan the 'perfect' or 'best lesson', since there can never be a 'perfect' or 'best lesson', but to test a teaching approach and build a community of practice, where members actively and imaginatively work together to seek deeper understanding of their work (Corcoran, 2007: Murata, 2011).

Once the planning of the research lesson is complete, the next step in the LS cycle is for one teacher to teach the lesson while other group members directly observe the research, live or study lesson in real time. This makes teacher thinking and practice visible (Murata, 2013) and offers a unique learning opportunity for teachers within a developing professional community (Lewis, Perry et al., 2012; Murata, 2011). The teaching itself is not being observed but, instead, students' learning is closely observed in the context of being taught (Dudley, 2013). This type of observation is in marked contrast to what teachers in Ireland have typically experienced e.g., inspections for evaluative purposes. Observing teachers make annotations on their copy of the research lesson plan and notice various aspects of teaching and learning that might not be apparent in professional development settings that are artificially replicated or outside the classroom setting (Lewis Perry, Friedkin & Roth, 2012; Murata, 2011). Most LS scholars agree that the research lesson is central to the whole LS process and is what sets LS apart from other types of CPD (e.g., Corcoran, 2007; Lewis, Perry et al., 2012; Murata, 2011). In Japan, and increasingly in the US now too, people from outside the LS group are invited to observe the live lesson (e.g., Xu & Pedder, 2015). Initially these visitors may be from within one's own school and subsequently further afield. In this way, the knowledge and skills learned by LS groups can be shared with many others.

Following the live lesson, group members meet to discuss and share their observations and reflections on the lesson, with the result that their professional knowledge is shared and exposed in a special way, from multiple perspectives (Dudley, 2012a; Murata, 2011). It is generally advised that these colloquia take place as soon as possible after the live lesson with comments based on observed evidence and focused on student responses to the 'showcase' lesson rather than the teaching (Corcoran, 2007; Dudley, 2012a; Takahashi, 2005). Frequently, the LS group revises and improves the lesson, which is then taught by a different member of the group with different children, while other group members again observe. While reteaching the research lesson is optional, it is recommended (Lewis & Hurd, 2011; Yoshida, 2011). Emphasis in LS is therefore on the process rather than the product (Corcoran, 2011; Perry & Lewis, 2009) and this process is repeated as required.

Appendix D: The Knowledge Quartet

The KQ is a four-dimensional practice-based framework for the observation, analysis and development of mathematics teaching, with particular emphasis on the development of teachers' MKT. Beliefs and knowledge evidenced in mathematics teaching can be seen in the four dimensions, termed Foundation, Transformation, Connection and Contingency, with each domain being made up of subcategories, which can overlap. Foundation underpins the other three dimensions and consists of teachers' knowledge, beliefs and understanding about mathematics and mathematics pedagogy. Transformation pertains to the way teachers transform their own knowledge to make it accessible to learners through the use of analogies, examples, explanations and demonstrations. Connection concerns the coherence of planning and teaching across lessons and includes the ordering of tasks and exercises and judgments about conceptual complexity. It also encompasses the knowledge which teachers display when they make connections within and between mathematical ideas. Finally, Contingency pertains to the way teachers respond to unanticipated classroom events and could be seen as the ability 'to think on one's feet' (Corcoran, 2011; www.knowledgequartet.org/introduction, accessed September 6th 2015). The KQ "provides a repertoire of ideal types that provide a heuristic to guide attention to, and analysis of, mathematical knowledge–in–use within teaching" (Ruthven, 2011, p.85).

Appendix E: Assessment for Learning Audit Instrument (Lysaght & O'Leary, 2014)

COLÁISTE PHÁDRAIG
ST PATRICK'S COLLEGE
DROIM CONRACH | DRUMCONDRA

Assessment for Learning Audit Instrument

Any assessment where the primary purpose is to improve teaching and learning is called Assessment for Learning (AfL) or Formative Assessment.

This audit instrument you are being asked to complete was constructed with the intention of evaluating the extent to which teachers use four key AfL strategies:

1. Sharing learning intentions/success criteria
2. Questioning/Classroom discussion
3. Feedback
4. Peer- and self- assessment.

For each strategy, you are asked to respond to a number of statements describing different way in which the strategy can be implemented in the classroom. For each statement, you should indicate the extent to which the statement reflects how you work currently, using the following six-point scale:

- The practice is *Embedded* = This happens approximately 90% of the time
- The practice is *Established* = This happens approximately 75% of the time
- The practice is *Emerging* = This happens approximately 50% of the time
- The practice is *Sporadic* = This happens approximately 25% of the time
- This *Never* happens
- I *Do Not Understand* what the statement means.

When completing this audit please be as honest as possible about your current AfL practices. This will ensure that the feedback to the school is accurate and useful. The intention is that the instrument will be used in the process of identifying the professional development needs of teachers in AfL.

You should be aware that data from the instrument may be used for research purposes. Please b assured that all responses will be treated confidentially and individual respondents or schools will not be identified.

Please note: Your completion of this instrument confirms that you understand the purpos of this study and that you freely consent to participate in it.

Dr Zita Lysaght Dr Michael O'Leary

Teacher Biographical Data

**Please note that the information you provide here will not be used to identify you in any way.
It will be used for data interpretation only.**

****Please feel free to write comments about the instrument in the margins and beside the text. We would be grateful if you would not write on or near the bubbles as this will make the document unreadable during the electronic scanning process. Thank you.**

Blacken one circle in each row like this ●

Your Gender	Female	○
	Male	○
Years teaching experience	0-5	○
	6-10	○
	11-20	○
	20+	○
Current Teaching Role	Teaching Principal	○
	Mainstream Class Teacher	○
	Special Class Teacher	○
	Learning Support Teacher	○
	Resource Teacher	○
	Other (specify)	________

Embedded	**Happens 90% of the time**
Established	**Happens 75% of the time**
Emerging	**Happens 50% of the time**
Sporadic	**Happens 25% of the time**
Never	**Never happens**
Don't Understand	**I don't understand what the statement means**

Blacken one circle in each row like this ●

Sharing Learning Intentions and Success Criteria	**Embedded**	**Established**	**Emerging**	**Sporadic**	**Never**	**Don't Under.**
1. Learning intentions are shared with pupils at appropriate times during lessons (e.g., *Halfway through the lesson, the teacher might say*: "*Remember, we are learning to distinguish between 2D and 3D shapes*").	○	○	○	○	○	○
2. Learning intentions are stated using words that emphasise knowledge, skills, concepts and/or attitudes i.e., what the pupils are learning NOT what they are doing.	○	○	○	○	○	○
3. Pupils are reminded about the links between what they are learning and the *big learning picture* (e.g., "*We are learning to count money so that when we go shopping we can check our change*").	○	○	○	○	○	○
4. Pupils are provided with opportunities to internalise learning intentions by, for example, being invited to read them aloud and/or restate them in their own words.	○	○	○	○	○	○
5. Child-friendly language is used to share learning intentions with pupils (e.g., "*We are learning to make a good guess (prediction) about what is likely to happen next in the story*").	○	○	○	○	○	○
6. Success criteria related to learning intentions are differentiated and shared with pupils.	○	○	○	○	○	○
7. Pupils are involved in identifying success criteria.	○	○	○	○	○	○
8. Prompts are used to signal learning intentions and success criteria with pupils (e.g., *using WALTS and WILFs in junior classes*).	○	○	○	○	○	○
9. Success criteria are differentiated according to pupils' needs (e.g., *the teacher might say*, "*Everyone must complete parts 1 and 2…; some pupils may complete part 3*").	○	○	○	○	○	○
10. Samples of work are used to help pupils develop a *nose for quality*.	○	○	○	○	○	○
11. Assessment techniques are used to assess pupils' prior learning (e.g., concept mapping…).	○	○	○	○	○	○
12. Pupils are reminded of the learning intentions during lessons.	○	○	○	○	○	○
13. Learning intentions are available throughout lessons in a manner that is accessible and meaningful for all pupils (e.g., *written on the black/whiteboard and/or in pictorial form for junior classes*).	○	○	○	○	○	○
14. Pupils' progress against key learning intentions is noted and/or recorded as part of lessons	○	○	○	○	○	○
15. Pupils demonstrate that they are using learning intentions and/or success criteria while they are working (e.g., *checking their progress against the learning intentions and success criteria for the lesson displayed on the blackboard or flipchart, for example*).	○	○	○	○	○	○
16. Pupils are given responsibility for checking their own learning against the success criteria of lessons.	○	○	○	○	○	○

Embedded	**Happens 90% of the time**
Established	**Happens 75% of the time**
Emerging	**Happens 50% of the time**
Sporadic	**Happens 25% of the time**
Never	**Never happens**
Don't Understand	**I don't understand what the statement means**

Questioning and Classroom Discussion	**Embedded**	**Established**	**Emerging**	**Sporadic**	**Never**	**Don't Under.**
1. When planning lessons, key, open-ended questions are identified to ensure that pupils engage actively in lessons (e.g., "*If we put a coat on our snowman in the school yard, do you think the snowman last longer?").*	O	O	O	O	O	O
2. Assessment techniques are used to facilitate class discussion (e.g., *brainstorming*).	O	O	O	O	O	O
3. Questions are used to elicit pupils' prior knowledge on a topic.	O	O	O	O	O	O
4. During lessons, hinge questions are used to determine pupils' progress in lessons (e.g., "*We have been learning to sort 3D shapes that stack and roll. Now, if you were given a choice, would you build a tower with spheres or cubes?").*	O	O	O	O	O	O
5. Assessment techniques are used to activate pupils /get them thinking during discussions and/or questioning (e.g., *using think-pair-share or talk partners*).	O	O	O	O	O	O
6. Assessment techniques are used to encourage all pupils to engage with questions (e.g., *no hands up, names out a hat, etc.*).	O	O	O	O	O	O
7. Pupils are encouraged to share the questioning role with the teacher during lessons (e.g., *the teacher routinely invites pupils to question their peers' contributions to discussions*).	O	O	O	O	O	O
8. Assessment techniques are used to encourage questioning of the teacher by pupils (e.g., *using hot-seating or a Post-Its challenge*).	O	O	O	O	O	O
9. Questioning goes beyond the one right answer style (*where the focus is often on trying to guess the answer in the teacher's mind*) to the use of more open-ended questions that encourage critical thinking.	O	O	O	O	O	O
10. The pace of class discussions is deliberately slowed down to encourage pupils to think before responding (e.g., *using wait time*).	O	O	O	O	O	O
11. Pupils are asked to explore their own ideas with others, using *think-pair-share*, for example.	O	O	O	O	O	O
12. Individual answers to questions are supplemented by pupils taking an answer round the class so that a selection of responses from the pupils is used to build a better answer.	O	O	O	O	O	O
13. Pupils' incorrect responses are used to guide teaching and learning (e.g., *a pupil is asked to explain why he/she gave a particular answer*).	O	O	O	O	O	O
14. Pupils are asked to evaluate their peers' responses to questions (e.g., "*Fiona, do you agree with what Regina has said and why?*).	O	O	O	O	O	O
15. Pupils can explain to others what they are learning (e.g., *if a visitor came to the classroom, pupils could articulate what they are learning in terms that identify the knowledge, skills, concepts and/or attitudes being developed*).	O	O	O	O	O	O
16. Pupils are asked to explain why they are undertaking particular tasks (e.g., *the teacher might ask*, "*Why are we completing this worksheet/what are we learning by doing it"?*).	O	O	O	O	O	O

Embedded	**Happens 90% of the time**
Established	**Happens 75% of the time**
Emerging	**Happens 50% of the time**
Sporadic	**Happens 25% of the time**
Never	**Never happens**
Don't Understand	**I don't understand what the statement means**

Feedback	**Embedded**	**Established**	**Emerging**	**Sporadic**	**Never**	**Don't Under.**
1. Feedback to pupils is focused on the original learning intention(s) and success criteria (e.g., "*Today we are learning to use punctuation correctly in our writing and you used capital letters and full stop correctly in your story, well done John").*	○	○	○	○	○	○
2. Assessment techniques are used during lessons to help the teacher determine how well pupils understand what is being taught (e.g., *thumbs up-thumbs-down and/or two stars and a wish*).	○	○	○	○	○	○
3. Written feedback on pupils' work goes beyond the use of grades and comments such as "well done" to specify what pupils have achieved and what they need to do next.	○	○	○	○	○	○
4. Teachers' praise of pupils' work (e.g., *"that's excellent; well done*"), is deliberately and consistently supplemented with feedback that specifies the nature of the progress made (e.g., *"Well done Kate, this paragraph helps me to visualise the characters in the story because of the adjectives you use*").	○	○	○	○	○	○
5. Teacher-made tests are used diagnostically to identify strengths and needs in teaching and learning (e.g., *identifying common mistakes in the addition of fractions).*	○	○	○	○	○	○
6. Diagnostic information from standardised tests is used to identify strengths and needs in teaching and learning (e.g., *common errors in the comprehension section of the MICRA-T are identified and used in teaching*).	○	○	○	○	○	○
7. Pupils are involved formally in providing information about their learning to their parents/guardians (e.g., *portfolios or learning logs are taken home).*	○	○	○	○	○	○
8. Feedback focuses on one or two specified areas for improvement at any one time (e.g., *in correcting written work, punctuation errors may not be marked if the primary focus of the writing is on the use of adjectives*).	○	○	○	○	○	○
9. *Closing-the-gap*-feedback is used to focus pupils' attention on the next step in their learning.	○	○	○	○	○	○
10. When providing feedback, the teacher goes beyond giving pupils the correct answer and uses a variety of prompts to help them progress (e.g., *scaffolding the pupils by saying*: "*You might need to use some of the new adjectives we learned last week to describe the characters in your story*").	○	○	○	○	○	○
11. In preparing to provide pupils with feedback on their learning, the teacher consults their records of achievement against key learning intentions from previous lessons (e.g., *the teacher reviews a checklist, rating scale, or anecdotal record that s/he has compiled*).	○	○	○	○	○	○
12. Pupils are provided with information on their learning on a *minute-by-minute, day-by-day* basis rather than end of week/month/term.	○	○	○	○	○	○

Embedded	**Happens 90% of the time**
Established	**Happens 75% of the time**
Emerging	**Happens 50% of the time**
Sporadic	**Happens 25% of the time**
Never	**Never happens**
Don't Understand	**I don't understand what the statement means**

Peer- and Self-Assessment	**Embedded**	**Established**	**Emerging**	**Sporadic**	**Never**	**Don't Under.**
1. Pupils are given an opportunity to indicate how challenging they anticipate the learning will be at the beginning of a lesson or activity (e.g., *by using traffic lights).*	○	○	○	○	○	○
2. Pupils are encouraged to record their progress using, for example, learning logs.	○	○	○	○	○	○
3. Lessons on new topics begin with pupils being invited to reflect on their prior learning (e.g., *pupils complete a mind map or concept map or brainstorm a topic*).	○	○	○	○	○	○
4. Pupils are provided with opportunities to reflect on, and talk about, their learning, progress and goals.	○	○	○	○	○	○
5. Pupils assess and comment on each other's work (e.g., *they are taught how to use the success criteria of a lesson to judge another pupil's piece of work*).	○	○	○	○	○	○
6. Pupils are encouraged to use a range of assessment techniques to review their own work (e.g., *a rubric, traffic lights, thumbs up/down, two stars and a wish*).	○	○	○	○	○	○
7. A visual record of pupils' progress is maintained to celebrate pupils' learning and show areas of/for development (e.g., *a bulletin board displaying progression in story writing over a term*).	○	○	○	○	○	○
8. Time is set aside during lessons to allow for self- and peer-assessment.	○	○	○	○	○	○
9. Assessment techniques are used to create an environment in which pupils can be honest about areas where they are experiencing difficulty (e.g., *talk partners are used to facilitate conversations between pupils about the challenges they face in their learning*).	○	○	○	○	○	○
10. When pupils have difficulty in their learning, they are encouraged to draw on a range of self-assessment strategies and techniques to help them overcome the problem (e.g., *they consult with an exemplar on the bulletin board*).	○	○	○	○	○	○
11. Pupils use each other as resources for learning (e.g., *response/talk partners who comment on each others' work and discuss how it can be improved*).	○	○	○	○	○	○
12. Time is set aside during parent/guardian-teacher meetings for pupils to be involved in reporting on some aspects of their learning (e.g., *pupils select an example of their best work for discussion at the meeting*).	○	○	○	○	○	○
13. Pupils use differentiated success criteria to self- and/or peer-assess (e.g., *pupils can distinguish between what must be achieved to be successful on a task and what might be done to gain extra credit*).	○	○	○	○	○	○
14. Pupils have ready access to exemplar materials showing work at different levels of achievement across a range of subject areas (e.g., *pupils use examples of collage on the Art display board when advising peers on how to improve their work*).	○	○	○	○	○	○

THANK YOU FOR COMPLETING THIS ASSESSMENT FOR LEARNING AUDIT INSTRUMENT

Appendix F: Attitude to Mathematics Questionnaire (ATMQ)

Attitude Survey for 4th Class Students (ATMQ)

Name: ___________________________________ Date: _________________

Directions

In this survey, you will find questions about what happens during maths class. Read each question carefully and pick the answer you think is best. Be as honest as you can and remember that there is no right and wrong answer. You may ask for help if you do not understand something or are not sure how to answer. Each question is followed by number of answers. Colour in the circle next to the answer of your choice. If you decide to change an answer to a question, put an 'X' over your first choice and colour in the circle for your new choice.

Example 1:

*Colour in **one** circle for each line*

	Agree a lot	Agree a little	Disagree a little	Disagree a lot
	↓	↓	↓	↓
I enjoy painting a picture	①	②	③	●

Example 2:

	Agree a lot	Agree a little	Disagree a little	Disagree a lot
	↓	↓	↓	↓
a) I always get good results in school	①	②	③	④
b) I am not good at school	①	②	③	④
c) I am nervous about doing science	①	②	③	④
d) I feel confident about doing science	①	②	③	④

ATMQ-TIMSS

How much do you agree with these sentences about learning maths?

*Colour in **one** circle for each line*

	Agree a lot	Agree a little	Disagree a little	Disagree a lot
a) I usually do well in maths	①	②	③	④
b) I would like to do more maths in school	①	②	③	④
c) Maths is harder for me than for most other students in my class	①	②	③	④
d) I enjoy learning maths	①	②	③	④
e) I am not good at maths	①	②	③	④
f) I learn things quickly in maths	①	②	③	④
g) Maths is boring	①	②	③	④
h) I like maths	①	②	③	④

ATMQ-SCLM

How much do you agree with these sentences about maths?

*Colour in **one** circle for each line*

	Agree a lot	Agree a little	Disagree a little	Disagr a lot
	↓	↓	↓	↓
a) Maths is my least favourite subject	①	②	③	④
b) My mind goes blank when doing maths	①	②	③	④
c) Doing maths makes me feel nervous	①	②	③	④
d) When I hear the word maths, I start to daydream	①	②	③	④
e) It makes me nervous to even think about having to do a maths problem	①	②	③	④
f) Maths does not scare me	①	②	③	④
g) I am very confident when it comes to maths	①	②	③	④
h) Solving maths problems is easy for me	①	②	③	④
i) I expect to do fairly well in most of my maths classes	①	②	③	④
j) I get confused in my maths class	①	②	③	④
k) I learn maths easily	①	②	③	④
l) I think I am good at solving maths problems	①	②	③	④

ATMQ-MOT

How much do you agree with these sentences about maths?

*Colour in **one** circle for each line*

	Agree a lot	Agree a little	Disagree a little	Disagr a lot
a) I am confident that I could learn difficult maths	①	②	③	④
b) I would like to avoid using maths in secondary school	①	②	③	④
c) I would be willing to do extra maths	①	②	③	④
d) When I get into secondary school I would love to do lots of maths	①	②	③	④
e) I think maths is fun because you have to figure things out	①	②	③	④
f) I think I can do even the hardest maths if I keep trying different ways to find the answer	①	②	③	④
g) If I'm not one of the best in my maths class then I don't try at all	①	②	③	④
h) I try my best at maths because I want to learn new things	①	②	③	④
i) I try my best at maths when there's a reward	①	②	③	④
j) If I find maths difficult I give up straight away	①	②	③	④

Appendix G: Children's Assessment for Learning Audit Instrument (CAfLAi)

AfL Questionnaire for Fourth Class Students

Name:____________________ **Date:**________________

General Directions

In this questionnaire, you will find questions about what happens during maths class. Read each question carefully and pick the answer you think is best. Be as honest as you can and remember that there is no right and wrong answer. You may ask for help if you do not understand something or are not sure how to answer. Each question is followed by a number of answers. Colour in the circle next to the answer of your choice. If you decide to change an answer to a question, put an 'X' over your first choice and colour in the circle for your new choice.

Example 1

Do you go to school?

Fill in **one** circle only

Yes	●
No	○

Example 2

How often do you do these things?	**Never**	**Sometimes**	**Often**	**Always**	**I don't understand what this means**
1. I listen to music	○	○	○	○	○
2. I talk with my friends	○	○	○	○	○
3. I play sports	○	○	○	○	○

Colour in **one** circle only

Sharing Learning Intentions and Success Criteria

	During maths class:	**Never**	**Sometimes**	**Often**	**Always**	**I don't understand what this means**
1	It is clear to me what the teacher wants me to work especially hard at.	○	○	○	○	○
2	The teacher reminds me of what I am learning and what she is looking for in terms of my work.	○	○	○	○	○
3	I discuss with the teacher what the best way of doing the maths might be.	○	○	○	○	○
4	I look at examples of work that the teacher has shown us to find out what I did well and how I could do better.	○	○	○	○	○
5	I talk to my teacher and the other children in the class about what we are learning and how we could use this in everyday life.	○	○	○	○	○

Colour in **one** circle only

Questioning and Classroom Discussion

	During maths class:	**Never**	**Sometimes**	**Often**	**Always**	**I don't understand what this means**
1	I take part in class discussion.	O	O	O	O	O
2	I ask other students maths questions and explain how I got my answers.	O	O	O	O	O
3	I work in a pair or small group.	O	O	O	O	O
4	I share my ideas and I discuss possible answers.	O	O	O	O	O
5	Before starting something new I work with the other children in the class to find out what we already know about the topic e.g. brainstorming.	O	O	O	O	O

Colour in **one** circle only

Feedback

	During maths class:	**Never**	**Sometimes**	**Often**	**Always**	**I don't understand what this means**
1	I discuss with my teacher how to improve my work as I go along.	○	○	○	○	○
2	I think that the teacher really listens to what I have to say about my work because of the comments she makes and the questions she asks me about it.	○	○	○	○	○
3	When the teacher praises my work she explains what was good about it.	○	○	○	○	○
4	When my teacher looks at my work she uses different ways to show me what I have done well and where I might improve.	○	○	○	○	○
5	I make changes to my work after the teacher corrects it so that I learn from my mistakes and my work improves.	○	○	○	○	○

Colour in **one** circle only

Peer- and Self- Assessment

	During maths class:	**Never**	**Sometimes**	**Often**	**Always**	**I don't understand what this means**
1	I try to understand new ideas in maths by thinking about what I already know.	O	O	O	O	O
2	I think about what I did well before and use it to help me learn better this time.	O	O	O	O	O
3	I judge if my work is successful by looking at examples of really good work.	O	O	O	O	O
4	If I get stuck with something, I ask other students to help me.	O	O	O	O	O
5	I talk with a partner and I tell them how I think their work could be better.	O	O	O	O	O

Different ways of learning:

In your maths class how often do you use:

		Never	Sometimes	Often	Always	I don't understand what this means
1	W.A.L.T. (We are learning to…).	O	O	O	O	O
2	W.I.L.F. (What I'm looking for…).	O	O	O	O	O
3	Two stars and a wish	O	O	O	O	O
4	Traffic lights/Coloured cups	O	O	O	O	O
5	K.W.L.	O	O	O	O	O
6	I-You-We checklist	O	O	O	O	O
7	P.M.I.	O	O	O	O	O
8	Think-pair-share	O	O	O	O	O
9	Hot-seat questioning	O	O	O	O	O
10	Thumbs-up thumbs-down	O	O	O	O	O
11	Brainstorming	O	O	O	O	O
12	Mind/concept maps	O	O	O	O	O
13	Learning logs	O	O	O	O	O
14	Learning portfolios	O	O	O	O	O
15	ABCD cards/corners	O	O	O	O	O
16	Rubrics	O	O	O	O	O
17	Sharing exemplars	O	O	O	O	O
18	Find the mistakes and fix them	O	O	O	O	O
19	Preflight checklist	O	O	O	O	O
20	Phone a friend/50-50/ask the audience	O	O	O	O	O

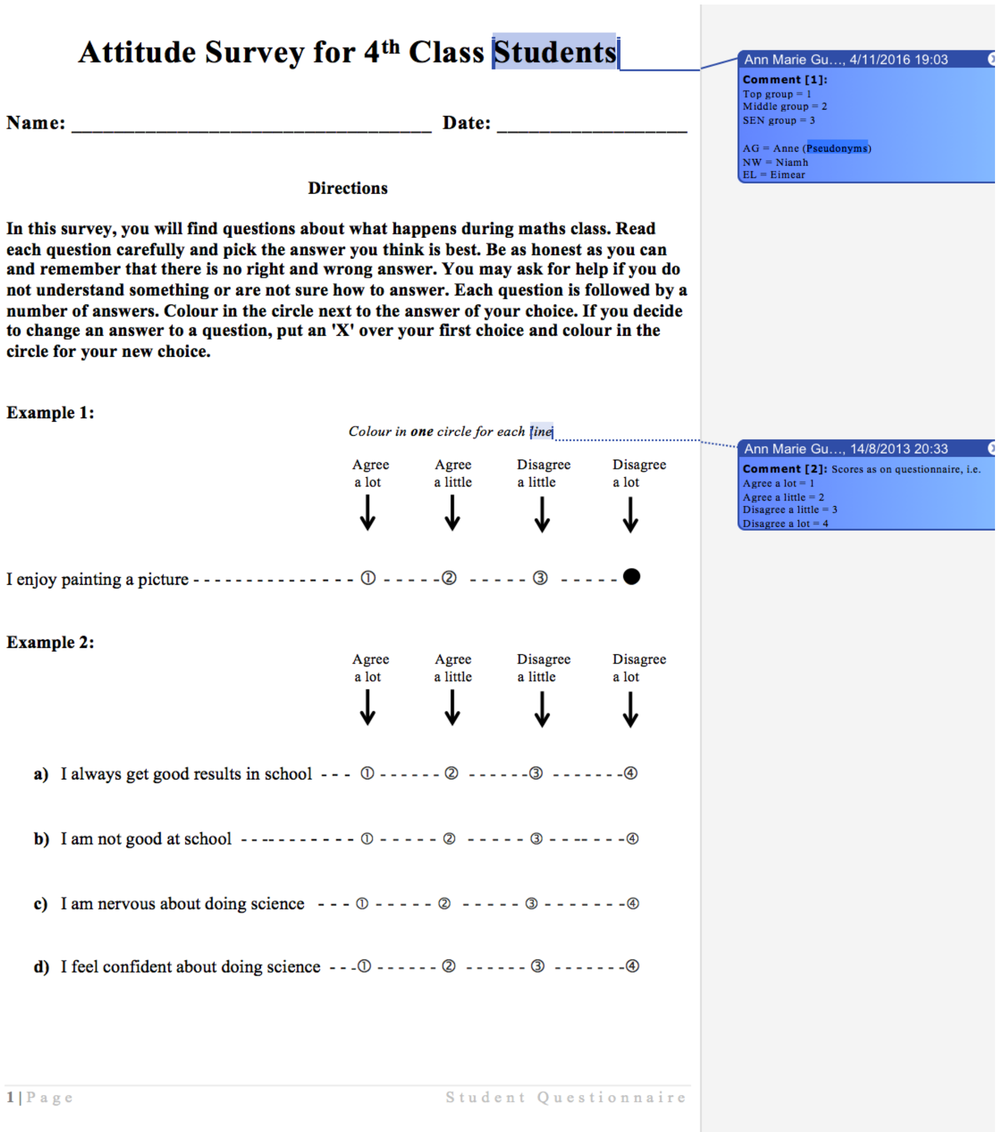
Thank you for your time, effort and thought in completing this questionnaire.

Appendix H: ATMQ coding (partial)

ID	NAME	STREAM	PRE TMS a	PRE TMS b	PRE TMS c	PRE TMS d	PRE TMS e	PRE TMS f	PRE TMS g	PRE TMS h	PRE SC a	PRE SC b	PRE SC c	PRE SC d	PRE SC e	PRE SC f	PRE SC g	PRE SC h	PRE SC i	PRE SC j	PRE SC k	PRE SC l	PRE MOT a	PRE MOT b	PRE MOT c	PRE MOT d	PRE MOT e	PRE MOT f	PRE MOT g	PRE MOT h	PRE MOT i	PRE MOT j	Post TMS a	Post TMS b	Post TMS c	Pos TMS
1	Names and Class groups redacted	1	2	1	3	2	3	2	4	2	3	3	4	3	2	2	2	2	2	2	2	2	3	3	2	3	2	3	4	1	1	4	1	2	4	1
2		1	2	1	3	1	3	2	4	1	4	3	4	2	3	2	2	3	2	2	2	2	3	3	1	4	1	2	4	2	2	3	1	2	4	1
3		1	2	1	3	1	3	2	4	1	3	3	3	3	3	2	1	3	2	2	1	2	3	1	2	2	1	1	4	1	1	4	1	3	4	1
4		1	1	1	3	1	4	1	4	1	3	3	4	3	4	1	1	1	1	4	1	2	1	4	2	2	1	1	4	1	1	4	1	1	4	1
5		2	2	3	4	3	3	4	2	3	1	4	4	3	2	2	3	2	1	1	3	2	4	1	4	4	4	2	3	1	3	3	2	2	3	2
6		1	1	1	4	1	4	2	4	1	4	4	4	4	4	1	2	2	1	4	2	1	3	4	1	3	2	2	4	1	1	4	1	2	4	1
7		1	3	4	2	3	2	4	3	3	2	3	1	3	3	2	4	3	3	2	4	3	4	4	4	4	2	3	4	2	1	2	1	2	3	2
8		2	3	1	3	1	2	2	4	1	3	2	2	2	4	1	3	3	1	2	3	1	3	4	2	3	2	1	4	1	1	4	1	1	4	[illegible]
9		2	2	4	2	2	3	3	2	3	2	3	4	2	2	3	3	3	3	2	3	3	3	4	2	3	2	2	4	2	1	3	2	1	3	[illegible]
10		3	2	3	1	2	1	3	2	2	2	3	1	1	2	4	2	4	2	4	3	2	4	1	4	4	3	4	2	3	1	3	2	3	4	3
11		2	2	2	4	2	4	3	4	1	4	4	3	4	2	1	4	3	2	2	2	1	4	4	2	3	3	3	4	1	4	4	1	3	4	2
12		3	3	4	4	2	1	3	1	4	2	1	2	1	1	3	4	4	3	1	4	3	4	1	4	4	4	4	4	1	1	3	1	1	4	1
13		3	2	2	4	1	4	2	4	1	2	4	4	3	4	1	2	1	2	2	2	3	2	4	2	3	1	2	4	1	1	4	2	3	2	1
14		2	2	3	3	3	4	2	3	2	3	3	2	2	3	2	2	2	2	3	2	2	4	2	3	3	3	2	4	1	1	4	2	2	4	2
15		1	1	2	4	1	4	2	4	1	4	4	3	4	2	1	2	2	1	3	3	3	3	4	2	3	1	2	4	1	1	4	1	2	4	1
16		2	3	1	4	2	4	2	4	1	2	2	2	4	2	1	2	3	4	2	4	2	4	2	3	2	1	1	4	1	4	4	2	3	2	2
17		2	2	3	1	2	1	3	4	2	2	2	2	3	1	3	3	4	2	2	3	4	3	2	3	4	3	3	3	1	1	4	1	2	4	1
18		1	1	2	4	1	4	2	4	1	4	4	4	4	3	1	2	2	1	3	3	2	3	3	2	4	1	2	4	1	1	4	2	1	4	1
19		2	1	1	3	1	4	2	4	1	4	3	3	4	4	4	1	2	2	3	2	3	3	1	2	2	1	3	4	1	3	4	1	1	4	1
20		2	2	3	3	1	3	4	4	1	2	3	4	2	4	2	1	2	2	2	3	2	2	4	2	3	2	1	4	1	3	2	2	3	4	2
21		2	2	2	4	1	3	2	4	1	4	4	3	4	4	1	2	2	1	3	2	1	2	4	1	2	1	1	4	1	1	4	1	1	4	1
22		2	1	4	4	2	3	3	3	2	3	3	3	4	3	2	2	3	2	3	4	2	4	2	4	4	3	3	4	1	1	4	1	2	3	1
23		1	1	1	4	1	4	2	4	1	4	4	3	4	1	2	2	2	3	3	3	3	3	4	1	1	1	3	4	1	1	4	1	2	4	1
24		1	2	1	4	1	3	2	4	1	4	4	4	3	4	1	2	2	1	3	2	3	3	4	1	1	1	2	4	1	1	4	1	2	4	1
25		1	1	2	3	2	3	2	3	2	3	2	3	2	3	3	3	3	2	3	2	2	1	2	3	2	1	1	4	1	2	4	1	3	3	2
26		2	2	3	1	3	2	3	2	2	1	2	1	1	2	1	3	3	2	2	4	4	3	3	2	4	2	3	2	1	1	3	2	2	2	2
27		1	2	3	4	1	4	2	3	2	3	3	2	3	2	1	2	2	1	2	2	3	3	3	2	3	2	2	4	3	1	2	1	3	4	2
28		1	2	1	3	2	3	2	3	2	2	3	2	3	3	3	2	3	2	2	3	2	3	3	3	4	3	2	4	2	1	3	1	3	4	2
29		2	2	4	2	3	1	4	1	3	1	2	2	3	1	3	4	4	2	2	4	4	4	3	4	4	4	2	4	2	1	3	2	2	4	[illegible]
30		3	3	1	1	1	2	4	4	1	4	1	4	4	3	1	3	4	1	3	4	3	4	4	1	4	1	2	4	1	1	4	1	1	4	1
31		1	1	1	3	1	3	2	4	1	4	3	4	3	4	1	1	1	1	3	2	1	2	4	1	1	1	1	4	1	4	4	1	2	4	1
32		1	2	1	4	1	3	2	4	1	3	4	4	4	4	1	2	2	1	2	2	1	2	3	1	2	1	1	4	1	2	4	1	2	3	1
33		3	1	2	3	1	4	3	4	1	4	4	4	4	3	1	1	1	1	4	2	2	2	4	2	3	2	3	4	1	2	3	1	1	4	1
34		1	3	3	3	3	3	3	3	2	2	3	2	3	4	3	2	1	3	2	3	2	3	2	3	4	2	2	4	2	1	3	2	3	2	1
35		3	2	1	2	4	2	2	2	4	2	2	2	1	2	1	2	2	2	2	2	2	1	2	3	3	1	3	99	2	2	1	2	1	1	1
36		2	3	4	3	4	2	3	2	4	1	1	2	2	3	4	4	3	3	1	4	4	3	1	4	4	4	3	2	2	1	4	3	3	2	2
37		2	1	3	4	2	2	2	3	2	1	4	2	2	4	1	2	1	1	2	2	1	3	3	4	4	2	1	4	1	1	3	1	2	3	1
38		1	3	4	3	4	3	3	1	4	2	3	4	1	3	3	4	2	3	2	3	3	4	2	4	4	4	3	3	4	1	3	1	1	3	1
39		2	2	3	1	1	3	3	2	2	2	1	2	3	2	2	3	3	2	1	3	2	3	4	4	4	2	3	3	1	1	4	2	3	2	1
40		3	4	4	1	4	1	4	2	4	1	2	1	3	1	3	4	4	3	4	4	4	4	4	4	4	4	3	4	1	4	3	2	3	3	2
41		2	2	1	3	1	4	2	4	1	2	4	4	3	3	3	1	4	2	2	1	4	3	3	2	2	1	1	4	1	1	3	2	3	3	2
42		2	1	1	2	1	4	2	3	1	3	2	3	2	3	4	2	2	1	2	3	2	3	3	2	2	3	1	4	2	1	3	1	1	4	1
43		1	3	1	2	2	3	3	2	3	3	2	3	3	2	2	3	4	2	2	3	4	3	3	2	3	2	3	4	1	1	3	1	2	4	1
44		1	1	1	4	1	4	1	4	1	4	4	4	4	4	1	1	1	1	4	1	1	1	4	2	1	1	1	4	1	4	4	1	2	4	1
45		2																																		
46		1	2	1	4	1	3	2	4	1	4	4	4	4	4	1	2	1	1	3	2	2	2	4	1	1	1	1	4	1	1	4	1	2	4	1
47		1	2	2	2	1	4	2	2	2	3	2	2	3	3	1	3	3	2	2	4	2	4	2	4	2	4	2	4	1	1	4	2	2	4	1
48		3	4	4	1	1	1	4	4	1	4	3	2	4	1	3	3	4	4	2	4	4	4	4	99	1	1	1	4	1	4	1	4	1	4	1
49		2	3	2	3	2	3	4	3	2	3	2	2	4	2	3	3	2	3	2	3	3	4	3	2	4	3	2	4	1	2	2	1	1	3	1
50		1	2	2	4	1	4	2	4	1	4	2	3	2	2	1	2	2	1	2	2	3	3	4	4	4	2	2	4	1	1	2	1	2	4	1
51		1	2	3	3	2	3	2	3	2	3	3	2	3	3	3	2	3	2	2	2	2	3	3	2	3	3	3	4	3	1	3	1	3	4	1

Appendix I: Coding, Collating and Cleaning of Raw Data ATMQ

Fifty participants from the intervention group completed the ATMQ, pre- and post-intervention. These data were coded using a codebook for example; the top group was assigned the number 1, the middle group 2, and SEN group 3 (See ATMQ codebook below). Responses were also coded, for example *Agree a lot* was assigned the number 1 up to 4 for *Disagree a lot*.



Attitude Survey for 4th Class Students

Ann Marie Gu…, 4/11/2016 19:03

Comment [1]:
Top group = 1
Middle group = 2
SEN group = 3

AG = Anne (Pseudonyms)
NW = Niamh
EL = Eimear

Name: ______________________ **Date:** ______________

Directions

In this survey, you will find questions about what happens during maths class. Read each question carefully and pick the answer you think is best. Be as honest as you can and remember that there is no right and wrong answer. You may ask for help if you do not understand something or are not sure how to answer. Each question is followed by a number of answers. Colour in the circle next to the answer of your choice. If you decide to change an answer to a question, put an 'X' over your first choice and colour in the circle for your new choice.

Example 1:

*Colour in **one** circle for each line*

Ann Marie Gu…, 14/8/2013 20:33

Comment [2]: Scores as on questionnaire, i.e.
Agree a lot = 1
Agree a little = 2
Disagree a little = 3
Disagree a lot = 4

	Agree a lot	Agree a little	Disagree a little	Disagree a lot
I enjoy painting a picture	①	②	③	●

Example 2:

	Agree a lot	Agree a little	Disagree a little	Disagree a lot
a) I always get good results in school	①	②	③	④
b) I am not good at school	①	②	③	④
c) I am nervous about doing science	①	②	③	④
d) I feel confident about doing science	①	②	③	④

1 | Page Student Questionnaire

These data were then collated and cleaned, similar to the previous Appendix, before being inputted into SPSS 21 for analysis. The initial stages of analysis involved an exploration of Frequency Tables for the original un-recoded data from each of the three scales (pre- and post-): ATMQ-TIMSS, Self-Confidence in Learning Mathematics (ATMQ-SCLM) and Motivation to do Mathematics (ATMQ-MOT). Subsequently, negative statements e.g., ‘Maths is boring’ were recoded and percentage scores for those who ‘Agreed a lot’ and ‘Agreed a little’ were amalgamated into combined ‘agree’ percentage scores. Similarly, scores for ‘Disagree a lot’ and Disagree a little’ were also combined. These data were then summarised as Tables 18-21, which are presented and discussed in Chapter Four.

Appendix J: Students' Learning Log

Students' Learning Log

Name: ______________________ **Date:** ______________________

Please respond to no more than three of the following prompts:

- Today I learned …
- I was surprised by…
- The most useful thing I will take from this lesson is…
- I was interested in…
- What I liked most about this lesson was…
- One thing I'm not sure about is…
- The main thing I want to find out more about is…
- After this session I feel…
- I might have got more from this lesson if…
- The skill I want to develop more is…

Appendix K: Focus Group Interviews

Choosing Focus Group Candidates

1. Paper-based coding of children's learning logs (mid-June 2013).
2. Determination of initial themes and identification of FG participants. Main themes identified were AfL and mathematics with subthemes of AfL strategies and techniques and enjoyment of mathematics and increased confidence doing mathematics and candidates for FGs were chosen according to these themes.

Prompts for researcher:

- Stress voluntary nature of participation and ability to withdraw at any stage
- Confidentiality and anonymity
- Permission to film/record
- Outline approximate duration and purpose of focus group

Questions:

1. Tell me about Maths this year.
2. Brief discussion regarding learning log entries:
    a. Strategies
    b. Techniques
    c. Liking maths/confidence
3. In your opinion, has using AfL made a difference to your learning? How?
4. How do you feel about maths now?
5. Was there anything you especially liked or disliked about doing maths this year? Is there any particular maths lesson that stands out as your favourite?
6. What, if anything, would you want next year's teacher to continue doing during maths class?
7. Is there anything else you'd like to add?
8. Are there any questions you'd like to ask me?

THANK YOU

Appendix L: Teachers' Learning Log

Teachers' Learning Log/Reflection Tool

Name: ____________________________ **Date:** ____________________________

Please respond to no more than three of the following prompts:

- ❖ Today I learned …
- ❖ I was surprised by…
- ❖ The most useful thing I will take from this meeting is…
- ❖ I was interested in…
- ❖ What I liked most about this meeting was…
- ❖ One thing I'm not sure about is…
- ❖ The main thing I want to find out more about is…
- ❖ After this session I feel…
- ❖ I might have got more from this meeting if…
- ❖ The skill I want to develop more is…

1. What worked well since the last meeting (Even better if…)?

2. How is the project affecting the students in your class?

3. What are your present concerns about the project?

4. What do you feel you are struggling with?

____

____

____

____

____

5. What are your greatest needs at this point?

____

____

____

____

____

6. Have you any ideas about something that would make the project work better?

____

____

____

____

____

7. Any other comments?

____

____

____

Appendix M: Lesson Study Observations

Lesson Study Experiences: Our Process-Research Lesson

REMEMBER: SINCE THE LESSON STUDY LESSON IS JOINTLY PLANNED BY THE GROUP IT IS JOINTLY OWNED BY THE GROUP. THIS MEANS THAT THE FOCUS FOR THE OBSERVERS IS LESS ON THE TEACHER AND MORE ON THE LEARNERS. DO NOT HELP THE STUDENTS OR OTHERWISE INTERFERE WITH THE FLOW OF THE LESSON.

Teacher observation

Here is a list of some of the data you may collect while observing the lesson:

- Questions that come to mind as you observe
- Types of questions the students asked
- Interesting questions students asked
- Types of questions the teacher asked
- Critical things happening in the classroom
- Evidence of higher level thinking
- Evidence of confusion
- Percent of students who raise their hands/level of participation
- Body language, "aha" moments, shining eyes
- Shifts in thinking that are evident
- Number of times students refer to and build on classmates' comments
- Evidence of engagement
- Following the lesson and deviations from the lesson
- Students' comments of note
- Barriers to learning
- Common patterns e.g. students misunderstand something in the same way
- Note extent to which what was planned was achieved
- Timing of lesson
- Initial thoughts about the research lesson process
- What did you learn?

NOTE THE TIME AGAINST EACH ANNOTATION IF YOU CAN

OBSERVATIONS:

<u>Critical Lenses for Lesson Study:</u>

"To benefit from lesson study teachers will need to apply *critical lenses* to their examination of the lessons" (Fernandez et al., 2003, p.171).

1. **The Researcher lens:**

In the article by Fernandez et al (2003, p.173) Japanese teachers encouraged their American counterparts "to see themselves as researchers conducting an empirical examination, organised around asking questions about practice and designing classroom experiments to explore these questions. In particular, the Japanese teachers emphasised four critical aspects of good research: the development of meaningful and testable hypotheses, the use of appropriate means for exploring these hypotheses, the reliance on evidence to judge the success of research endeavors, and the interest in generalizing research findings to other applicable contexts. (*Hypotheses are like predictions about how student weaknesses can be improved through the planned lesson -P.174).* Reflection is integral to this research process.

2. **The Curriculum developer lens:**

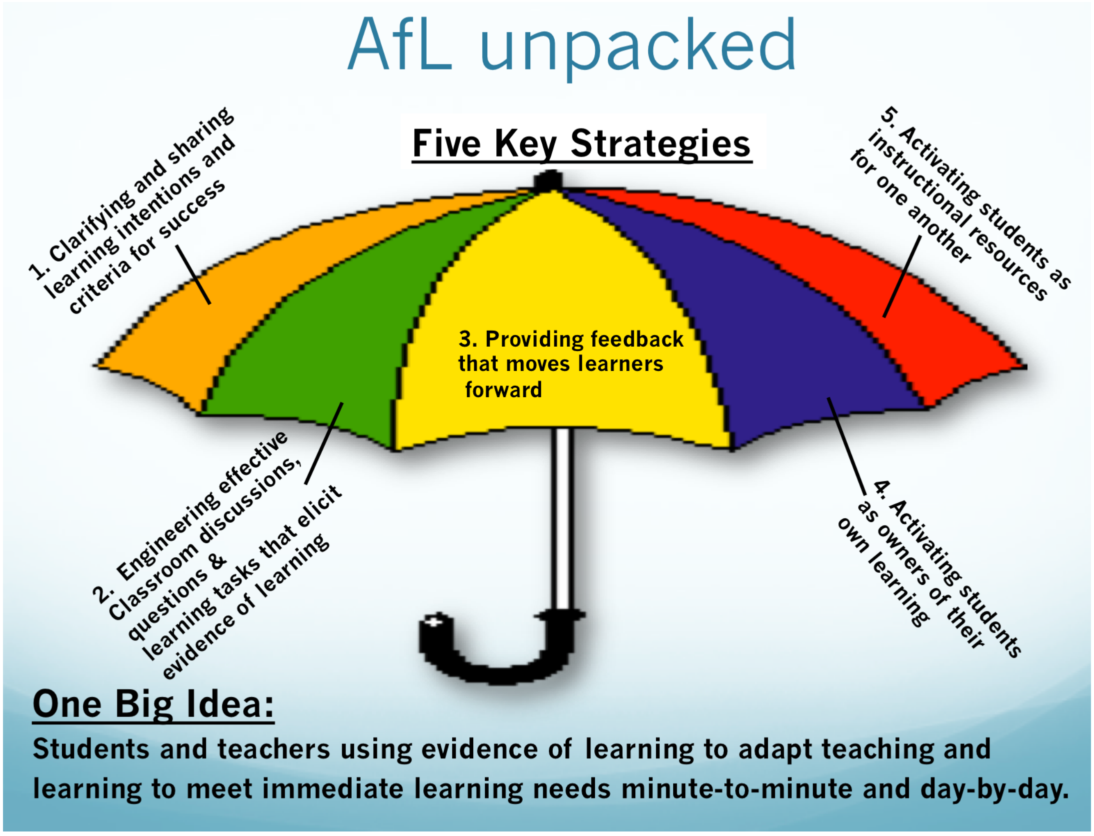
Importance of connecting study lessons to prior and future lessons (across and within lessons)-how to sequence and connect children's learning experiences.

"…it would be wise to select lessons on core topics because these topics always have antecedents and developments both within and across grades, thus facilitating curricular discussions" (Fernandez et al., 2003, p.177). (an understanding of how we want to structure the knowledge that we want the children to acquire during the lesson+how the lesson relates to content taught in other grades).

3. **Adopting the student lens:**

…to examine all aspects of a lesson through the eyes of our students. "…importance of teachers adopting the *students lens* by attempting to understand students' thinking, anticipate their behaviours, and determine how to use this knowledge to build students' understanding" (Fernandez et al., 2003, p.179)….using it to guide how we design lessons and evaluate them. "…to develop students' understanding, teachers need to consider not only what they want students to learn, but how they expect them to learn it and how they are going to help students reach that understanding. In particular, the Japanese teachers conveyed that the problems that teachers pose to students, as well as the precise way these problems are presented, affect how students will approach and think about the content of the lesson" (Fernandez et al., 2003, p.181).

Appendix N: One Big Idea, Five Key Strategies, Multiple Techniques



Appendix O: CPD Timeline

Term One

August 31st 2012:

- Meeting with participating teachers to discuss intervention.
- Collection of teacher consent forms
- Distribution and completion of AfLAi forms.

October 1st 2012:

- General introduction to the project and what's involved.
- TLC, AfL, LS in brief.
- AfL Strategy: Focus on Learning Intentions

October 15th 2012

- Brief Recap of previous session
- Learning Intentions continued
- AfL Strategy: Success criteria

October 22nd 2012:

- Introduction to Lesson Study.
- Looking at DVD re lesson study and how learning communities improve instruction (Lewis & Hurd, 2011).

November 5th and 19th 2012:

- Meeting to plan first research lesson and source resources

November 26th 2012:

- First Teaching of research lesson cycle one (Middle Group _ Ann Marie)
- Discussion of Research Lesson and review/adaptation

November 27th 2012:

- Second teaching of research lesson cycle one (Top Grouping _ Niamh)
- Discussion of Research Lesson and review/adaptation

December 14th 2012:

- Final Teaching of research lesson cycle one (Students with SEN _ Eimear)
- Discussion of Research Lesson and review

CPD Timeline Terms Two and Three

January 7th 2013:

- Review and recap of AfL to date
- Further work on Learning Intentions and Success Criteria
- Brief look at NCCA Assessment guidelines
- Some more AfL Techniques (No hands up, Wait Time, Fist to five, ABCD Cards, Hot Seat questioning)
- AfL Topic: Tuning into learners' minds

January 21st 2013:

- Why teachers matter
- AfL Strategy: Providing feedback that moves the learner forward
- Introduction to work of John Hattie (and Jo Boaler)
- AfL Techniques: Comment-only marking, plus minus equals.

January 30th 2013: Visit to St. Patrick's College, Drumcondra, Dublin:

- Workshop on learning intentions and success criteria with Dr. Zita Lysaght and Dr. Michael O'Leary.

February 12th 2013:

- Review of feedback.
- Topic: Marking Less to achieve more
- Further discussion re LISCs
- Initial look at possible topics/ lessons for Cycle 2 of lesson study

February 26th 2013:

- AfL Strategy: Peer and Self Assessment
- AfL Techniques: Learning logs; two stars and a wish; think, pair share revisited. Self-assessment pages leading to peer assessment. To look at and discuss what makes a good piece of work with the students.
- Further Planning for our next research lesson (Garden Plans)

March 5th and 12th 2013:

- Further discussion and selection of topic for next research lesson cycle
- Planning of Research lesson for lesson study Cycle 2

April 8th 2013:

Unscheduled meeting of Niamh and I to finalise plan for her doing first teaching on this Wednesday. (not taped)

April 9th 2013:

Niamh and I met again briefly (not taped)

April 10th 2013:

- First Teaching (Niamh) of research lesson beginning lesson study cycle two.
- Reflection meeting to discuss the first teaching and to adapt this lesson for second teaching next week.

April 16th 2013:

- Second Teaching of research lesson (Ann Marie)
- Reflection meeting to discuss the second teaching and to adapt this lesson for third teaching next week.

April 25th:

- Third Teaching of research lesson (Eimear)
- Reflecting Meeting

May 20th and 27th

- Beginning of cycle three of lesson study. Meeting to choose strand and topic and lesson planning.

June 5th

- First Teaching of research lesson beginning lesson study cycle three (Eimear)
- Reflection meeting to discuss the first teaching and to adapt this lesson for second teaching next week.

June 10th

- Second Teaching of research lesson cycle three (Ann Marie)
- Reflection meeting to discuss the second teaching and to adapt this lesson for third teaching next week.

June 11th

- Third Teaching of research lesson cycle three (Niamh)
- Reflecting Meeting

Summary of peer-to-peer professional learning community Scoil na nAingeal 2012-2013		
CONTENT *(What?)*	**PROCESS** *(How?)*	**CONTEXT**
AFL: *One Big Idea, Five Key Strategies and 100+ techniques* **One big idea:** Students and teachers use evidence of learning to adapt teaching and learning to meet immediate learning needs minute-by-minute and day-by-day. **Five Strategies:** 1. Share learning intentions and success criteria 2. Engineer effective classroom discussion 3. Provide feedback that moves learners forward 4. Activate students as the owners of their own learning 5. Activate students as instructional resources for one another **Techniques:** WALT; WILF; Think, Pair, Share; fist-to-five; two stars and a wish; rubrics; ABCD cards; learning logs; thumbs etc. **Japanese Lesson Study:** A collaboration-based teacher professional development originating in Japan, which helps develop teachers' pedagogical knowledge, is student-focused and improves students' learning. It is centers around teachers' interests and the "live" research lesson sets it apart. **Theory*:** • Blooms Taxonomy • Work of John Hattie: _'visible learning' but especially his work on feedback • Vygotsky's_zone of proximal development (ZPD) and scaffolding • Tuning into learners' minds • Good questioning techniques • Marking less to achieve more • Rowland's Knowledge Quartet (foundation, transformation, connection, contingency) • Pedagogical Content Knowledge especially Ball et al. (2008) Mathematical Knowledge for teaching (MKT) • Sociocultural learning theory * As focus of CPD was on AfL, theories were only covered in brief and on a need to know basis.	Combination of teacher learning community and Japanese lesson study vs. traditional/other forms of CPD. Various methods used: • Powerpoint presentations • Handouts • Video • Learning logs • Discussion • Reflections • Lesson planning • Observation • Lesson Study Cycle	• Peer-to-peer professional learning • School-based (S na nA.) • Job embedded • No 'expert' or 'knowledgeable other' • After school (own free time) • Year-long • WHY? • DES policy • Ourselves • Other

Appendix P: Synthesis of CPD

Appendix Q: A Worked Example of Thematic Analysis from a Transcript from one of the Children's Focus Groups

Phase 1: Familiarisation with the Data:

I began the qualitative analysis by immersing myself in the data and reading and re-reading the transcripts, making notes, jotting down early impressions and marking ideas for subsequent phases. Below is one example of a rough note I made on this particular transcript from Children's Focus Group 1:

> It's obvious from reading this transcript again that the children have been thinking a lot about feedback and how they'd use it and learn from it. I found it more interesting though when the discussion develops into talking about positive and negative feedback. (31/10/2014)

Phase 2: Generation of Initial Codes

Since, in this analysis, I was concerned with specifically addressing Research Hypothesis 2 this was a theoretical thematic analysis rather than an inductive one (Braun & Clarke, 2006; Maguire & Delahunt, 2009). I began by identifying interesting aspects and patterns and then coded each segment of data that captured something interesting relevant to this Research Hypothesis. I used open coding so I did not have preset codes although I had already developed preliminary ideas about codes from the first phase (familiarisation with the data). I did this coding manually, working through the transcript with highlighters, pens and post-its, collating data relevant to each code. As recommended by Braun and Clarke (2006), these codes continued to be developed and defined throughout the entire analysis.

Data extract	Coded for
I just find AfL is very very good and I'm really glad that we got to learn all these new things because it really helps you and I think the next class will really enjoy it as well. (Hollie).	Positivity towards AfL
I thought it made us much confidenter and much funner doing maths and much easier as well (Chloe). Ok (Researcher).	Affective impact on maths/positivity towards AfL.
I'd like if the teacher that we got next year did AfL with us because we'd be learning more things but it would be in 5th class so they'd be a bit harder (Mia). Ok (Researcher).	Positivity towards AfL. Impact on learning.
Well, I agree with Chloe. Like, it's just like, it makes maths so much more fun, 'cause like in third class I used to hate maths and them like now that there's all of these different strategies it just makes like maths so much easier (Sophie). Ok (Researcher).	Affective impact on maths. Positivity towards AfL.
I agree with Sophie and Chloe about what they said and I'd love to do AfL next year and with Mia because it's going to be harder next year and we're getting pushed so next year we're going to develop in our learning (Ruby).	Positivity towards AfL

Data Extract, with codes applied

Phase 3: Searching for Themes

Once the initial coding was complete I began to look for potential themes i.e., patterns in the data that were interesting or important in terms of Research Hypothesis 2 (Maguire & Delahunt, 2009) and that encapsulated what was being spoken about (Braun & Clarke, 2006). I tried to figure out how the various codes might combine to form overarching themes. Since parts of the transcript could be used for two or more codes I had multiple copies. I cut up the transcripts and organized the codes into theme-piles. I also used mind-maps. The following is a list of initial themes:

- Language of AfL
- Positivity towards AfL
- Using AfL strategies and techniques
- Understanding AfL
- Impact on Learning
- Feeling like a teacher
- Taking more responsibility
- Making maths easier and more fun
- Improving at maths
- Affective impact on maths (confidence)
- Better at learning (impact/next steps)
- Using feedback/next steps in learning
- Other/Interesting Insights

Phase 4: Reviewing Themes

Next, I had to review, modify and develop the primary themes that I identified in Phase 3 to ensure that they made sense. I gathered together all the data that was relevant to each theme. Once again, I had multiple copies of the transcripts and I cut these up and organized them into the theme-piles identified in Phase 3. I then read the data in each theme-pile to consider if it actually did support that particular theme. In this example, I'm just using one extract but in the final analysis there was more than this. Initially I was interested in each individual theme and then I wanted to see how the themes related to each other, and ultimately whether the themes worked in relation to the entire data set. For example, the Language of AfL did not really work as a theme and was eliminated and incorporated into 'Enjoying the AfL Journey' and 'PSA- A Highlight for Children'. Additionally, there were too many themes identified in Phase 3 and these were combined together to form the following five main themes:

1. Enjoying the AfL journey
2. Growing positivity and self-confidence in mathematics;
3. A changed classroom dynamic;
4. Peer- and self-assessment: - a highlight for children;
5. Unexpected insights.

Phase 5: Defining and Naming Themes

In Phase 5 I reviewed the five main themes I had identified in the previous phase and then as recommended by Braun and Clarke (2006), I tried to "…identify the 'essence' of what each theme is about" (p.92) and what's interesting and why.

1. Enjoying the AfL journey

   This theme included statements that showed positivity towards AfL, examples of where children had appropriated AfL in areas other than mathematics, instances where the children had shown and enjoyed using AfL strategies and techniques and the language of AfL. It also included their observations, reflections and recommendations regarding the use of AfL practices.
2. Growing positivity and self-confidence in mathematics;

   This theme included all references to children's beliefs about getting better at mathematics, feeling more confident, less nervous or more positive about mathematics.

3. A changed classroom dynamic;

   This theme reflected how children were less confused in their mathematics classes, how there was increased collaboration in their learning which they valued, their descriptions of their use of feedback to improve their learning and their recognition of the teacher as guide, as well as the need for them to take increased responsibility for their own learning.

4. Peer- and self-assessment: - a highlight for children;

   Time and again reference was made to PSA and the fact that they love using PSA.

5. Unexpected insights.

   This theme was about things that really surprised me as researcher and teacher in that I did not know until the end of the intervention how deeply the children had thought about AfL practices and enjoyed using them. It also contains items that didn't readily fit under the other themes such as data about Comment-Only marking, Rubrics and other noteworthy and interesting comments.

Phase 6: Producing the Report

Once I had fully worked out the themes from each item in the qualitative data set, I then finalised the overall themes that are discussed in detail the main text in Chapter Four. I selected compelling extract examples relating to each theme and embedded these within a scholarly narrative.

Appendix R: Five Validity Criteria (Anderson & Herr, 1999)

1. *Outcome validity*, i.e. "the extent to which actions occur which lead to a resolution of the problem that led to the study", was addressed by forming a PLC/LS group and using lesson study, endeavouring to improve participants' AfL literacy and mathematics;
2. *Process validity* was achieved by examining the adequacy and appropriateness of the various research methods and enquiry processes used throughout the study;
3. *Democratic validity*, was accomplished by working collaboratively with teachers and students in attempting to improve AfL literacy and mathematics achievement and by honouring participants' multiple perspectives and interests;
4. *Catalytic validity* "is the degree to which the research process reorients, focuses and energises participants towards reality in order to transform it". I was mindful of this at all times.
5. *Dialogic validity* was obtained by monitoring analyses through critical and reflective dialogue with peers, colleagues and with my supervisors.

Appendix S: Letter To Board Of Management

xxxxxxxxx,

xxxxxxx,

xxxxxx,

June 12th 2012.

Dear Xxxxxxx,

I am writing to you as chairperson of Scoil na nAingeal to ask you to bring before the Board my request for permission to undertake further research in the school. As you are aware, my previous research into school attendance was very beneficial for the school and contributed to greatly improved attendance rates. On this occasion, in partial fulfilment of my Doctoral studies, I wish to study the effects of Assessment for Learning (AfL) strategies on students' achievements in numeracy as mediated through a teacher learning community (TLC). AfL is acknowledged worldwide as a powerful tool for improving student learning and achievement while increasingly TLCs are being advocated as efficacious approaches to providing teacher professional development thus improving classroom practice. As you are aware the recent publication *Literacy and Numeracy for Learning and Life* (2011) by the Department of Education and Skills highlights the importance of schools focusing on improving numeracy as well as literacy skills. Therefore, I consider this project to be both timely and of value.

The study will commence by asking all the teachers to complete an AfL audit to ascertain current assessment practice and AfL knowledge within the school. Over the period of the next academic year (2012-2013), I will work in close collaboration with my colleagues at fourth class level to introduce AfL strategies and techniques when teaching maths. Much of the research will include activities that can be considered a normal part of teachers' work in schools. We will meet regularly to critically review the progress we are making with the aim of fine tuning our assessment skills so the children benefit to the optimum and will engage in professional dialogue, evaluation and reflection. Standardised maths tests, questionnaires, focus groups and samples of children's work will be used to assess the benefit to the children. Other sources of data to be used in the research will include teacher notes and reflections.

Parents/guardians and the students will be consulted prior to the commencement of the project and will be required to give their consent if the students are to participate. It is important to note that involvement in this research study is voluntary. Consequently, participants may choose to withdraw at any time. This research project will be undertaken in strict accordance with the ethical guidelines of St. Patrick's College Of Education, Drumcondra, with whom I am registered. I am in the initial stages of the dissertation process, the exact details of the study have yet to be worked out and finalised with my supervisors. However, I will of course keep Bernie informed, as principal, at all stages of my work.

Should the project prove beneficial I intend to share the new practices with my teaching colleagues and to encourage their use in other classes throughout the school, thus benefiting all children. I would therefore greatly appreciate your approval and support for the proposed study.

Yours sincerely,

Ann Marie Gurhy.

Appendix T: Plain Language Statement

Individual Teacher Consent Form

Ann Marie has explained the research project to me and I understand what is involved. She has given me the opportunity to ask questions and to answer any queries I might have. I understand that I am free to terminate my participation in the project or to withdraw any information traceable to me at any time up to the project's completion in June 2013 without giving a reason.

On this basis, I agree to participate in this research project.

- I understand that taking part in the project will involve the following:
    - Participation in professional development on AfL strategies and techniques and the use of lesson study.
    - Regular use of various AfL strategies and techniques when teaching maths.
    - Participation in the lesson study process where we plan, observe, reflect and revise "research lessons".
    - Attendance at review meetings with colleagues to reflect on and share my experiences of using AfL.
    - Meeting to plan "research lessons", to observe or teach them, to observe and collect data on student learning, to review the lesson, modify if necessary, re-teach if required, all as part of the lesson study process.
- I agree to complete an AfL audit of my understanding and use of AfL before the intervention starts and after its completion.

Printed Name of Teacher: ______________________________

Signature of Teacher: ________________________ Date: __________

Appendix U: Plain Language Statement

Parents' Information Sheet (Fourth Class 2012-2013)

Dear Parent/Guardian,

My name is Ann Marie Gurhy and, as you probably know, I am a teacher in your daughter's school. I am currently studying for a doctorate in education in St. Patrick's College, Drumcondra, Dublin. As part of my studies I am doing a research project to see if children's mathematics skills and results can be improved by using assessment for learning (AfL) strategies. Most of the activities can be considered part of normal teaching and learning in primary school.

The Department of Education and Skills want to improve English and Mathematics in Irish schools and they view using AfL as an important way of doing this. A recent document by the Department explains that when teachers use AfL they "use assessment information on the progress that their students have achieved to date, they share this information with their students, and they use this information to plan the next steps in their teaching and their students' learning". Indeed, AfL is acknowledged worldwide as a powerful tool for improving student learning and achievement.

I have obtained the permission and full support of the Board of Management and the principal to carry out this research as they feel that it will be of great benefit to the school and the children involved. The teachers participating in the study are Mrs xxxxx, Mrs xxxxx and I and as part of the research process we will meet twice a month after school. At our first meeting we will learn about one of the AfL strategies and plan how we will use this strategy during mathematics class. At the next meeting we will talk about how things are going and discuss any changes we think we should make in order to improve things further. This process will continue over the full school year 2012-2013.

All of the children who are presently in fourth class in our school will be invited to take part in the project. In order to decide if the new strategies are improving mathematics learning I need to collect information before the study begins, for the duration of the study and at the end of the study. Information will be collected from the teachers about their teaching and from students to see how well they are learning. This will help us to find out if the project is working. The other teachers and I will be collecting information on a regular basis throughout the year to assess your daughter's progress in mathematics. I hope to take photographs of the children's work from time to time and may tape or video record some of them (on a voluntary basis) to capture their ideas/feelings about using AfL.

Thank you most sincerely for taking the time to read this information. This is a very worthwhile project which I hope will make a big difference to children's enjoyment of mathematics and to their mathematics skills and achievement. I hope that you will support it by giving permission for your daughter to participate. Should you have any queries or desire further information about this project please don't hesitate to contact me by email at xxxxx@gmail.com or phone (xxxx-xxxxxxxx).

Yours sincerely,

Ann Marie Gurhy.

PARENT/GUARDIAN CONSENT FORM: FOURTH CLASS 2012-2013

I, ……………………………………………………………………………….. am the legal parent/guardian of

……………………………………………………………………………. and I give my consent for her to take part in the research study on mathematics taking place in Scoil Mhuire gan Smál.

- I acknowledge that I have read the Parents' Information Sheet explaining the purpose of the study and what is involved.
- I understand that this project can be considered a part of normal teaching and learning in the primary school.
- I understand that it is entirely voluntary for my daughter to take part in the project and that I can withdraw my consent at any stage.
- I understand that my daughter can choose to participate or not in the project and is free to withdraw at any time.
- I agree to tests being given to assess my daughter's progress in mathematics and to information the school already has on my daughter's mathematics ability being used for the project.
- I agree to photographs of my daughter's work being taken during the project (without her face being shown).
- I understand that during the project tape or video recordings may be made (on a voluntary basis) to capture the children's ideas/feelings about using AfL.
- I understand that information obtained during the study about my daughter will be treated in the strictest confidence and in keeping with the school's assessment policy.

____________________________ ____________________________

Printed Name of Child Date

____________________________ ____________________________

Printed name of Parent/Guardian Signature of Parent/ Guardian

PLEASE SIGN THIS FORM AND RETURN IT TO ME BY THURSDAY SEPTEMBER 6TH, 2012.

Appendix V: Plain Language Statement

Parents' Information Sheet (Fourth Class 2011-2012)

Dear Parent/Guardian,

My name is Ann Marie Gurhy and, as you know, I am a teacher in your daughter's school. I am currently studying for a doctorate in education in St. Patrick's College, Drumcondra, Dublin. As part of my studies I am doing a research project to see if children's mathematics skills and results can be improved by using assessment for learning (AfL) strategies. Most of the activities can be considered part of normal teaching and learning in primary school.

The Department of Education and Skills want to improve English and Mathematics in Irish schools and they view using AfL as an important way of doing this. A recent document by the Department explains that when teachers use AfL they "use assessment information on the progress that their students have achieved to date, they share this information with their students, and they use this information to plan the next steps in their teaching and their students' learning". Indeed, AfL is acknowledged worldwide as a powerful tool for improving student learning and achievement.

I have obtained the permission and full support of the Board of Management and the principal to carry out this research as they feel that it will be of great benefit to the school and the children involved. The teachers participating in the study are Mrs xxxxx, Mrs xxxxx and I and as part of the research process we will meet regularly after school to learn about the AfL strategies before implementing them in our teaching of mathematics. This process will continue over the full school year 2012-2013.

All of the children who are presently in fourth class in our school are invited to take part in the project, so your daughter will not be directly involved this year. I realise that at the moment your daughter will not benefit directly from the project but I would expect that if the project is successful it may lead to AfL strategies being used more comprehensively throughout the school and therefore she may benefit next year. Indeed, it is hoped that in future years many of the students and teachers in our school will benefit from the findings.

In the meantime, in order to see if the project is successful in improving mathematics results, I am asking for your consent to use the information the school already has on your child's performance as indicated by her scores on the

Non-Reading Intelligence Tests, Sigma T and Drumcondra Primary Mathematics Tests from fourth class so that I can compare them with this year's results at the end of the study. Your daughter's name will not be used and her individual results will not be identifiable.

Thank you most sincerely for taking the time to read this information. I hope that you will support the project by giving permission for your daughter to participate. Should you have any queries or desire further information about this project please don't hesitate to contact me by email at smgscarlowdp@gmail.com or phone (059-9142705).

Yours sincerely,

Ann Marie Gurhy.

PARENT/GUARDIAN CONSENT FORM (FOURTH CLASS 2011-2012)

Please tick the box below to indicate whether you will allow your daughter's results to be used in the mathematics study.

☐ I grant permission for my child's results to be used in Ms. Gurhy's study on mathematics.

☐ I do not grant permission for my child's results to be used in Ms. Gurhy's study on mathematics.

Printed Name of Child ______________________ Date ______________________

Signature of Parent/ Guardian ______________________ Printed name of Parent/Guardian ______________________

PLEASE SIGN THIS FORM AND RETURN IT TO ME AS SOON AS POSSIBLE. THANKS.

Appendix W: Plain Language Statement

Children's assent form to participate in the research project

Please note that this form will first be read with the children and explained to them to enhance their understanding of its contents.

Name of Project: A project using new ways of teaching to help children learn better.

About the project:

At the moment, I am studying in a college in Dublin and as part of the course I have been asked to do a research study which is a bit like a science experiment where you try out different things to see if they work and to see which one works best. In my research study I am trying to learn more about new ways of teaching and learning and I need your help with this. I am going to look at how these new and interesting ways of teaching and learning might help you to become better at maths.

Who will be involved?

All the girls in this year's fourth class will be given the chance to take part in the study and there will be three teachers involved: Mrs ……, Mrs ……. and I.

What part will I be asked to play?

If you agree to take part in the study I will ask you to do some things over the next year.

- As we go through the year and learn new things in mathematics, I will ask you to either write down something about what you have learned or talk to me about it. Then I will try to improve the lessons and ask your opinion again.
- I may take photographs of your work but will not show your face so that when other people are looking at them they will not know whose work it is.
- Some of you may be asked to give your opinion individually or in groups about the new ways of teaching and learning. This may include tape or video recordings (on a voluntary basis).
- When you are working out a problem or sum during mathematics class, for example to see how you could divide two pizzas between five people, the other teachers and I might watch you and take notes.

All of this information will be treated confidentially and will be destroyed once the project is completed.

Who will the study help?

This study probably will help you, because if you participate, you will be working with the other children and the teachers to try out new ways of teaching and learning which should help you with your maths. It may also help other students to do well in their learning in the future.

Who should I talk to about the project?

Please talk to your parents/guardians about the research study before you decide whether or not you will take part. Your parents/guardians have already given you permission to take part in the study but even though your mam/dad/guardian said "yes" you are still free to decide not to take part. It's your decision.

What if I don't want to participate?

If you don't want to participate in this research study then you do not have to. If you agree to participate now and then change your mind at a later stage then that's fine too. Remember taking part in this study is up to you and no one will be upset if you choose not to. You will still be participating in normal class work and activities.

What if I have questions?

I love answering questions so if you can think of any questions that I have not answered at this stage then make sure that you ask me. If any other questions crop up as we go along, make sure that you ask me to answer them and remember that you can always withdraw later.

What now?

If you would like to take part in the project read this form with your parent or guardian and discuss it with them. Then, if you are happy to take part in the project sign your name below and return the form to me before Friday Sept. 14th, 2012.

STUDENT ASSENT FORM: FOURTH CLASS 2012 -2013

I ____ agree to take part in the project and I agree to:

a. Allow teachers to watch me and take notes.
b. Have photographs of my work taken, without my face being seen.
c. Allow my work to be used as exemplars.
d. Be tape or video recorded (on a voluntary basis).
e. Discuss and share my ideas about the project.
f. Answer questionnaires about what I think about maths and about how I learn.

Don't forget that I will keep all the information that you share with me in confidence; in other words I won't be telling someone your name and how you scored in the tests.

Printed Name of Student: ____

Signature of Student: ____

Date: ____

Please sign the form and return it to your teacher before Friday September 14th, 2012.

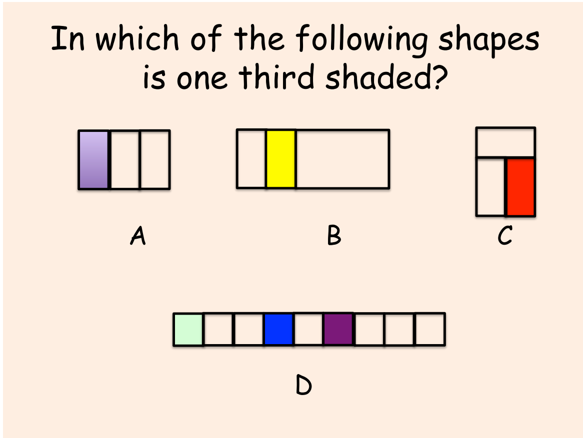
Thank you very much for your time.

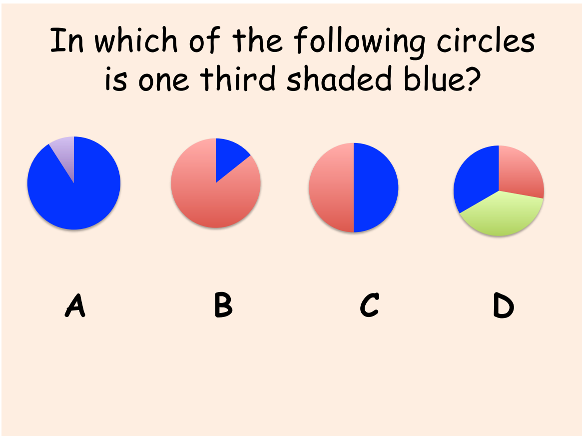
Appendix X: Samples of AfL Techniques from the Intervention

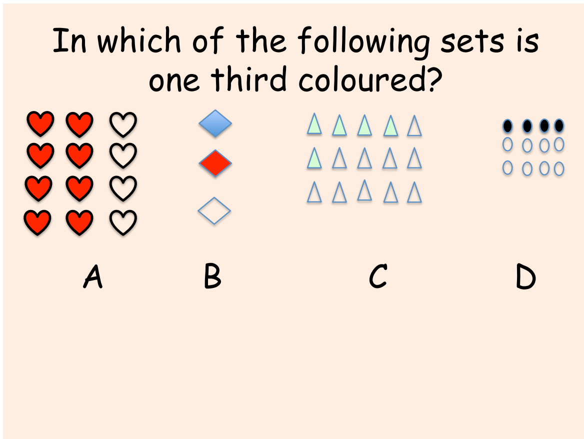
ABCD CARDS

ABCD Cards

Fractions







Which of the following answers are correct?

A. $\frac{1}{3}$ of 90 is 30

B. $\frac{1}{3}$ of 300 is 200

C. $\frac{2}{3}$ of 600 is 400

D. $\frac{2}{3}$ of 9000 is 7000

Which statements are true?

A. $\frac{2}{3}$ of 9 is 7

B. $\frac{1}{3}$ of 21 is 8

C. $\frac{1}{3}$ of 81 is 27

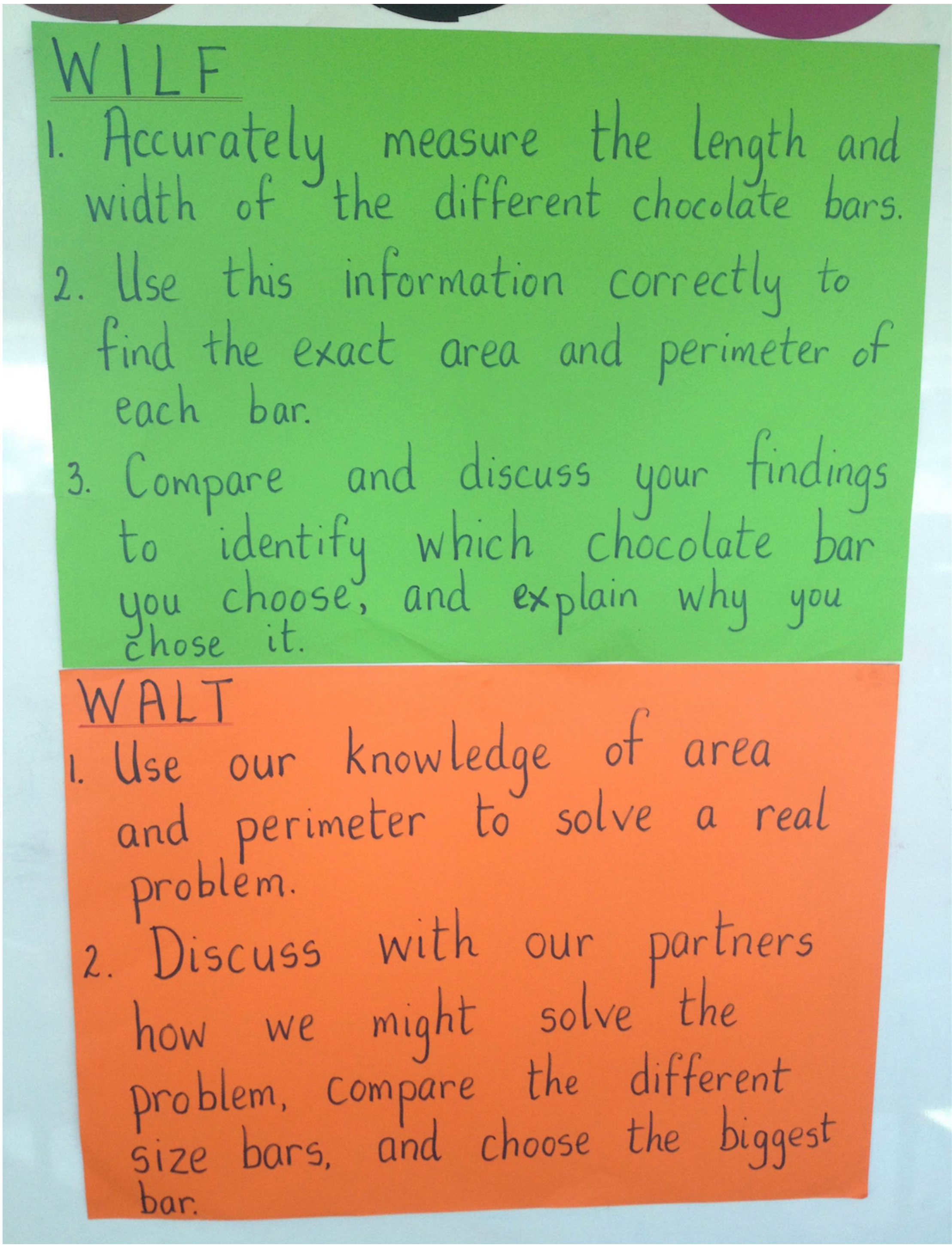
D. $\frac{2}{3}$ of 30 is 20

RUBRICS

Maths Rubric: Data	Student: ______________________			
	Excellent (4)	Very Good (3)	Good (2)	Fair (1)
Parts of Graph	Graph is correctly constructed and has all parts (title, bars, labels, numbers/scale).	Graph is correctly constructed but is missing one part (title, bars, labels, numbers/scale).	Graph is correctly constructed but is missing more then one part (title, bars, labels, numbers/scale).	Graph is not correctly constructed.
Neatness of Student's Work	Graph is neat, organised, and coloured. A ruler has been used in all instances.	Graph is neat and sized properly, a ruler was used consistently.	A ruler was not used consistently, graph is not sized properly.	Work is messy. A ruler was not used.
Questions/Answers	Answered all questions correctly.	One answer is incorrect.	Two answers are incorrect.	Three or more answers are incorrect.
Task Completion	The full task is completed.	3/4 of the task is completed.	At least 1/2 of the task is completed.	Less than 1/2 of the task is completed.
Total Score: ______________				

Maths Rubric: Chocolate Bars	Names: ______________________			
	Excellent (4)	Very Good (3)	Good (2)	Fair (1)
Neatness of Student's Work	Work is very neat and well organised.	Most of the work is neat and organised.	Some of the work is neat and well organised.	Work is messy.
Measurement	All lengths and widths are measured accurately.	Most lengths and widths are measured accurately.	Some lengths and widths are measured accurately.	No lengths and widths are measured accurately.
Calculation	All areas and perimeters are correct.	Most areas and perimeters are correct.	Some areas and perimeters are correct.	No areas or perimeters are correct.
Interesting Ideas discovered about area and perimeter	3 or more interesting ideas were discovered	2 interesting ideas were discovered	1 interesting idea was discovered	No interesting ideas were discovered
Task Completion	The full task is completed.	3/4 of the task is completed.	At least 1/2 of the task is completed.	Less than 1/2 of the task is completed.

WALT and WILF



SELF-ASSESSMENT

Two Stars and a Wish: Self Assessment

Name:

Topic: Time adding & subtracting

Use the two stars and a wish to tell me two reasons why you are pleased with this piece of work and one way in which it could be improved.

★ I only got one sum wrong and my task is completed.

★ You can see my writing clearly and it's very neat.

Next time I will check my rename. I would have gotten full marks.

Two Stars and a Wish: Peer Assessment

Name: [redacted]

Topic: Symmetry

Use the two stars and a wish to tell your partner two reasons why you think that their work is good and one way in which they could improve it.

Your flags are straight a ruler was always used. In the lines of symmetry most of them were right.

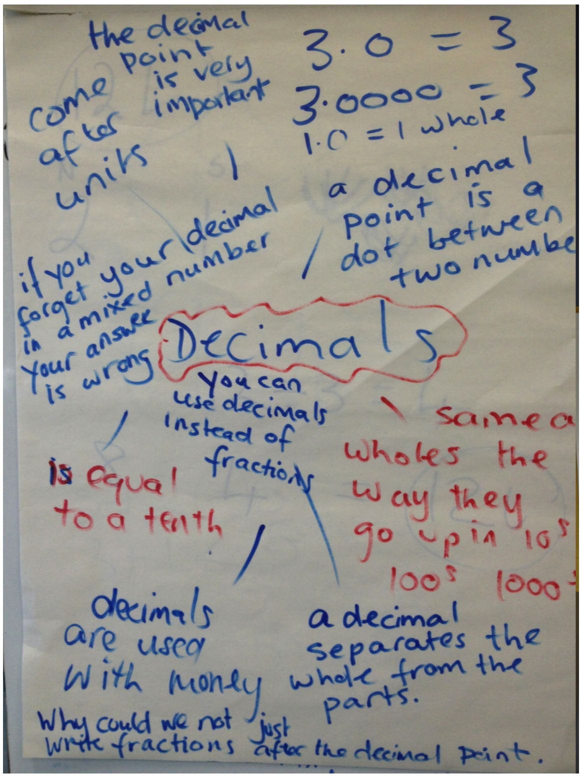
In number 3 the dots are perfect. Your work is excellent. Your colouring of the dots is also very neat. Well Done!

In number 3, some of the reflections are bigger than the other side. So just rember that for next time.

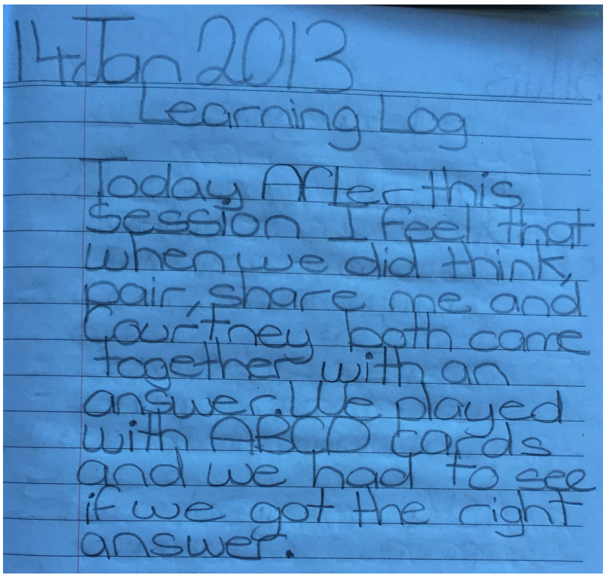
BRAINSTORMING and KWL

Topic:		Date:
K (What children know already)	W (What they want to know)	L (What they have learned)

Teacher's Notes

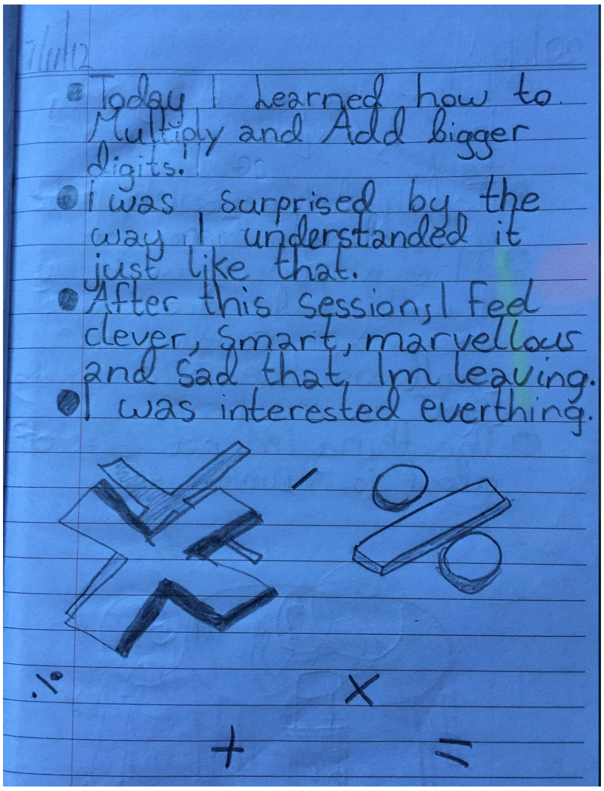


STUDENTS' LEARNING LOGS

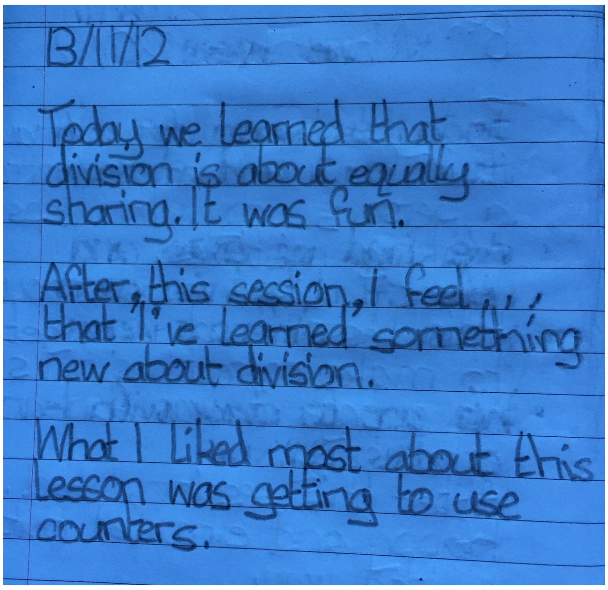

14 Jan 2013

Learning Log

Today After this Session I feel that when we did think, pair, share me and Courtney both came together with an answer. We played with ABCD cards and we had to see if we got the right answer.


7/11/12

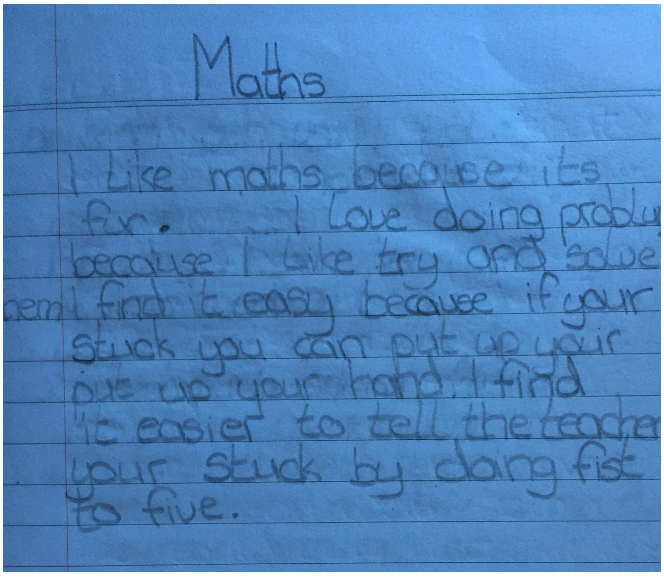
- Today I Learned how to Multiply and Add bigger digits.
- I was surprised by the way I understanded it just like that.
- After this session, I feel clever, smart, marvellous and sad that, Im leaving.
- I was interested everthing.


13/11/12

Today we learned that division is about equally sharing. It was fun.

After, this session, I feel... that I've learned something new about division.

What I liked most about this lesson was getting to use counters.


Maths

I Like maths because its fun. I Love doing problems because I Like try and solve them I find it easy because if your stuck you can put up your put up your hand. I find it easier to tell the teacher your stuck by doing fist to five.

Appendix Y: Summer Courses

Many Irish primary teachers voluntarily participate in the summer course model of professional development during their summer holidays. These courses can be face-to face, online, or blended and are of twenty hours duration. If teachers complete a summer course, they can avail of 3 Extra Personal Vacation (EPV) days the following academic year, subject to Board of Management approval. However, these days do not qualify for substitution cover and are usually at teachers' own expense. Since 2011, in keeping with the Literacy and Numeracy Strategy, the DES insists that all summer courses must place some emphasis on literacy and/or numeracy, and school self-evaluation and school improvement.

These courses cover a broad range of topics and subjects from core subjects such as English and Mathematics to areas such as mindfulness and

Summer courses have to meet strict criteria as set out by the DES and are provided by a range of approved providers such as education centres, the INTO, colleges of education etc….to ensure quality.

Appendix Z: Sample Reseach Lesson and Photographs

Mathematics Research Lesson Plan Two (2nd Teaching)

Title of Lesson: Garden planning! Date: April 16th

Strand: Number

Strand Unit: Fractions and Decimals

Goals:

1. To provide students with opportunities of working with peers to deepen their understanding of mathematics, in particular fractions and decimals.
2. To encourage students to use their existing knowledge of fractions and decimals so that they will design their own garden.

Learning Intentions/Success Criteria:

WALT:

1. Solve 'real' problems using our knowledge of fractions and decimals.
2. Use our knowledge of fractions and decimals to design our own gardens.

WILF:

Produce accurate diagrams recording your different garden designs.

AfL Strategies:

1. Clarifying and sharing learning intentions and success criteria.
2. Engineering effective classroom discussions, questions and learning tasks that elicit evidence of learning.
3. Providing feedback that moves the learner forward.

AFL Techniques: Thumbs; Fist to five; Think, Pair, Share; ABCD Cards; Brainstorming; learning logs.

Introduction (10 mins):

- **Explain to the children that the lesson is to do with number, especially fractions and decimals.**
- **Ask the children to use 'ABCD' cards to briefly revise what they have already learned about fractions and decimals and to prepare them for solving today's problems.**
- **Tell the children the learning intentions of the lesson and tell/elicit the success criteria.**

Comparing and discussing ideas arising:

Main Lesson: Today's problems (30 mins):

1. **We're going to look at Mrs Green's garden plan and try to figure out what decimals each section of the garden is equal to.**

Discuss Mrs Green's garden from the diagram on the IWB

Give out plans of Mrs. Green's garden and plenty of blank pages.

Questions:

e.g. 'What do you think will help you with this?'

Using 'think, pair, share come up with three ideas that will help you solve this question' etc.

Class discussion to try to solve the problem

2. **Now using what you have learned from solving the problem about Mrs. Green's garden, we're going to follow specific instructions and draw Mr. Brown's garden.**

Questions:

e.g. *Give out the blank 100 squares*

Questions such as:

What did you learn from the last problem that might help you with this one?

Where do you think would be a good place to start?

Any other ideas?

If you are finished doing this question early you may have a look at the question again to see if you could have designed Mr. Brown's garden in a different way.

Conclusion/Summing Up (10 mins):

1. What have students learned during the lesson?
2. **Ask students what they thought of the lesson-**

   Questions e.g.

   - **What have you learned today?**
   - **What helped you when something got tricky was…**
   - **What really made me think was…**
   - **What did you find easy/difficult/interesting?**
   - **What I would change about this activity to help another class learn is …**
   - **What would you do different the next time/advise someone else?**
3. Ask students to write in their learning logs.

Finding the solution to the problem:

Slow down the lesson and probe children's thinking.

Scaffold the children's learning where necessary.

Give students opportunities to explain their ideas – to explain their thinking and to make connections between different solutions.

Anticipated student thinking/learning:

1. **Mrs. Green's Garden:**
   - Difficulty with the 'strawberries' since they might not see that they add up to one tenth.
   - Similar difficulty with the 'path'
   - Some students may be ok finding the fraction for each section but may not be able to change this to a decimal in order to answer the question asked.

2. **Mr. Brown's Garden**
   - Some students may have difficulty finding $\frac{1}{4}$ or $\frac{1}{5}$ of the grid
   - Some students may have difficulty drawing a 'realistic' path and may just use what's left over to form the path
   - Some students may struggle when trying to use equivalent fractions and decimals.
   -

Resources:

Mrs Green's garden plan

Worksheet

Blank 100 squares (rectangular and square)

Pencils, rulers, colours, scissors, glue, blank paper

IWB, whiteboard,

Seating and pairing:

The children will work in pairs for most of the lesson and will be seated according to mixed ability groupings within the streamed class with the best child sitting with the weakest, the second best with the second weakest and so on.

Blackboard layout:

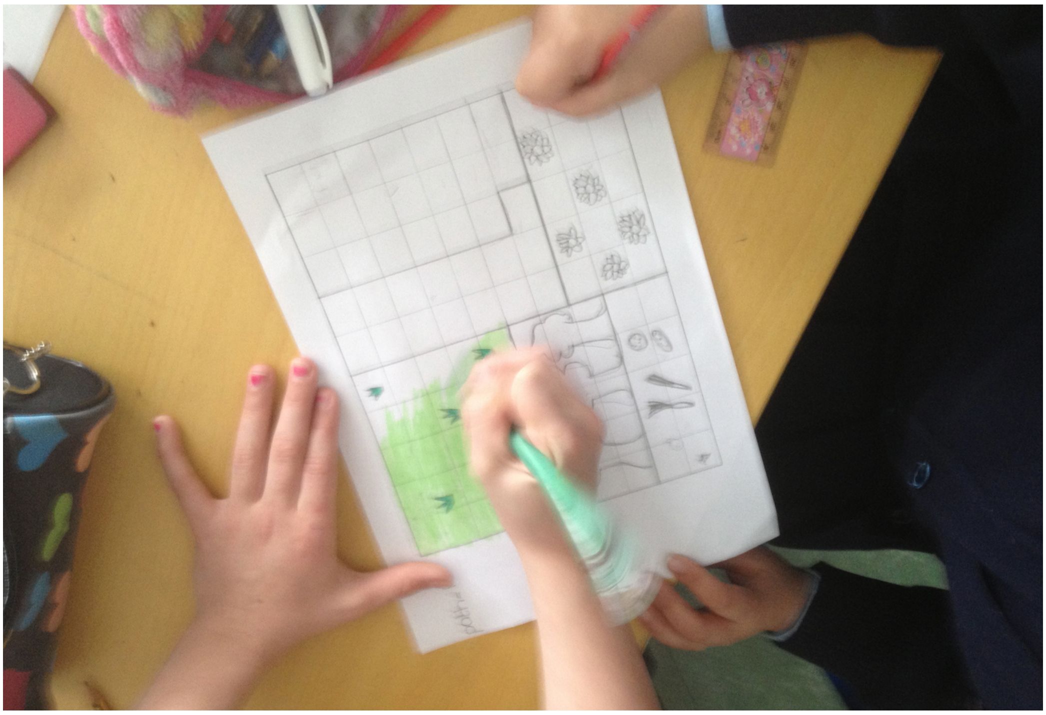
Evaluation/Assessment:

1. How did the students understand each problem?
2. What did the students understanding convey about how we can facilitate their understanding for the next time?
3. Were students able to answer the questions about Mrs Green's garden plan?
4. Were students able to use both fractions and decimals when developing their own garden plans and were they able to convert fractions to decimals and vice versa?
5. Were students able to recall and summarise what they learned and to write about it in their learning log?

General observations from the third teaching on Tues April 16th:

MAIN CHANGES WE DECIDED TO MAKE FROM LESSON STUDY SECOND TEACHING:

- Lesson was too long so leave out third problem
- ABCD Cards: too long and perhaps needs to be more differentiated for weaker groups. (Observers need to stand where they can see the answers being held up).
- Display the various correct options of Mr. Brown's garden at the end to show the children that there are various possible designs.
- Revise fractions and decimals with the ABCD Cards before discussing the WALT and WILF with the children for today's lesson.
- Put the problems on separate sheets and give out one at a time.



BIBLIOGRAPHY

Akiba, M. and B. Wilkinson (2016). Adopting an international innovation for teacher professional development: State and district approaches to lesson study in Florida. *Journal of Teacher Education*, *67*(1): 74-93.

Anderson, G. & Herr, K., 1999. The New Paradigm Wars: Is There Room for Rigorous Practitioner Knowledge in Schools and Universities*? Educational Researcher,* 28(5), 12-21.

Andersson, C. (2014). The Impact of a Comprehensive Teacher Professional Development Program in Formative Assessment on Teachers' Practice and Students' Achievement. *In:* Proceedings in Umeå, Sweden, November 5, 2014 Conference paper, Presentation (Refereed).

Andrade, H. (2010). Students as the definitive source of formative assessment: Academic self-assessment and the self-regulation of learning. In H. Andrade, & G. Cizek (Eds.), *Handbook of formative assessment.* (pp. 90–105). New York: Routledge.

Andrade, H. (2013). Classroom Assessment In The Context Of Learning Theory And Research. In J. H. McMillan (Ed.), *SAGE handbook of research on classroom assessment* (pp. 17-34). New York. SAGE.

Andrade, H., Du, Y., & Wang, X. (2008). Putting rubrics to the test: The effect of a model, criteria generation, and rubric-referenced self-assessment on elementary school students' writing. *Educational Measurement: Issues and Practices, 27*(2), 3-13.

Appleton, J. J., Christenson, S. L. and Furlong, M. J. (2008), Student engagement with school: Critical conceptual and methodological issues of the construct. Psychol. Schs., 45: 369–386.

Asselin, M. E. (2003). Insider research: Issues to consider when doing qualitative research in your own setting. Journal for Nurses in Staff Development, 19(2), 99-103.

Assessment Reform Group. (2002). *Research-based principles of assessment to guide classroom practice*. Retrieved: http://assessmentreformgroup.files.wordpress.com/2012/01/10principles_english.pdf

Baas, D., Castelijns, J., Vermeulen, M., Martens, R., & Segers, M. (2015). The relation between Assessment for Learning and elementary students' cognitive and metacognitive use. *British Journal of Educational Psychology*, 85, 33–46.

Baba, T. (2007). Japanese lesson study and education: An overview' (p.2-7) (NJ, USA: World Scientific Publishing) in book by M. Isoda, M. Stephens, Y. Ohara and T. Miyakawa (Eds) *Japanese Lesson Study in Mathematics: Its Impact, Diversity and Potential for Educational Improvement,* 2-7.

Baird, J. A., Hopfenbeck, T. N., Newton, P., Stobart, G. & Steen-Utheim, A. T. (2014). *Assessment And Learning: State Of The Field Review*. Oxford: Oxford University Centre for Educational Assessment.

Balan, A. (2012). *Assessment for Learning: A case study in mathematics education.* (Unpublished doctoral dissertation). Malmo University. Sweden.

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education,* 59(5), 389-407.

Banks, J., & Smyth, E. (2011). *Continuous professional development among primary teachers in Ireland.* Dublin: The Teaching Council and ESRI.

Barber, M. and Mourshed, M. (2007) *How the World's Best Performing Schools Come Out on Top.* London: McKinsey and Company.

Barron, B. J. S., Pea, R., & Engle, R. A. (2013). Advancing understanding of collaborative learning with data derived from video records. In C. E. Hmelo-Silver, C. A. Chinn, C. K. K. Chan, & A. M. O'Donnell (Eds.), *The international handbook of collaborative learning* (pp. 203–219). New York, NY: Routledge.

Bassey, M. (1981). Pedagogic research: On the relative merits of the search for generalisation and study of single events. *Oxford Review of Education*, 7(1), 73-93.

Bell, J. (2010). *Doing your research project: A guide for first-time researchers in education, health and social science.* Berkshire, England: Open University Press.

Bell, M., Cordingley, P., Isham, C., & Davis, R. (2010). Report of professional practitioner use of research review: Practitioner engagement in and/or with research. Coventry: CUREE, GTCE, LSIS and NTRP. Retrieved from www.curee-paccts.com/node/2303

Bell, L. & Nutt, L. (2002). Divided loyalties, divided expectations: research ethics, professional and occupational responsibilities. In Mauthner, M., Birch, M., Jessop, J. and Miller, T. (eds.) *Ethics in qualitative research* (pp. 70-90). London, Sage.

Bennett, R. E. (2011). Formative assessment: a critical review. *Assessment in Education: Principles, Policy & Practice*, 18(1), 5-25.

Berman, J. (2013). Utility of a conceptual framework within doctoral study: A researcher's reflections. *Issues in Educational Research*, 23(1), 1-18.

Bird, T., & Little, J. W. (1986). How schools organize the teaching occupation. *The Elementary School Journal*, *86*(4), 493–511.

Birenbaum, M., Kimron, H. & Shilton, H. (2011). Nested contexts that shape assessment for learning: School-based professional learning community and classroom culture. *Studies in Educational Evaluation, 37*, 35-48.

Black, H. (1986). *Assessment for Learning*. In D. L, Nuttall (Ed.), Assessing educational achievement, pp7-18. London: Falmer Press.

Black, P. (2015). Formative Assessment - An Optimistic But Incomplete Vision. *Assessment in Education: Principles, Policy and Practice*, 22(1), 161-177.

Black, P. (2016). The Role Of Assessment In Pedagogy – And Why Validity Matters. In D. Wyse, L. Hayward, & J. Pandya (Eds.), *The SAGE Handbook of Curriculum, Pedagogy and Assessment*. pp.725-739. New York: Sage Publications.

Black, P. & McCormick, R. (2010). Reflections and New Directions. Assessment and Evaluation in Higher Education, 35, 493-499.

Black, P. J., & Wiliam, D. (1998a). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy and Practice, 5*(1), 7-74.

Black, P. J., & 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. (2006) Developing a theory of formative assessment. In J. Gardner (Ed.), *Assessment and learning* (pp. 81-100). London: Sage.

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability, 21*(1), 5-31.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning. Putting it into practice.* Maidenhead: Open University Press.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan*, 86(1), 9–21.

Black, P., McCormick, R., James, M., & Pedder, D. (2006). Learning how to learn and Assessment for Learning: A theoretical inquiry. *Research Papers in Education, 21*(2), 119–132.

Blaxter, L., Hughes, C., & Tight, M. (2006). *How to research* (3rd ed.). Berkshire, England: Open University Press.

Bloom, B. S. (1969). Some theoretical issues relating to educational evaluation. In H. G. Richey & R. W. Tyler (Ed.), *Educational evaluation: new roles, new means: the 68th yearbook of the National Society for the Study of Education (part II)* (Vol. 68(2), pp. 26-50). Chicago, IL: University of Chicago Press.

Bocala, C. (2015). From experience to expertise: The development of teachers' learning in lesson study. *Journal of Teacher Education*, 66(4), 349–362.

Bogdan R. & Biklen, S.K. (2003). *Qualitative Research for Education: An Introduction to Theory and Methods.* Boston, Allyn & Bacon.

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher*, 33(8), 3-15.

Brannen, J. (2005). *Mixed methods research: A discussion paper.* London: Economic and Social Research Council, National Centre for Research Methods.

Brannick, T., & Coghlan, D. (2007). In defense of being "native": The case of insider academic research. *Organizational Research Methods*, 10(1), 59-74.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101.

Briggs, D. C., Ruiz-Primo, M. A., Furtak, E., & Shepard, L. (2012). Meta-analytic methodology and inferences about the efficacy of formative assessment. *Educational Measurement: Issues and Practice,* 31(4), 13-17.

Brookhart, S. M. (2009). *Exploring Formative Assessment* (The professional learning community series). USA: Association for Supervision and Curriculum Development.

Brookhart, S. M. (2011a). Educational assessment knowledge and skills for teachers. *Educational Measurement: Issues and Practice,* 30, 3–12.

Brookhart, S. M. (2011b). *The Use of Assessment to Support Learning in Schools – Formative Assessment. Executive summary of address presented to the Central Board of Secondary Education (CBSE), India and the Australian Council for Educational Research (ACER) International Educational Conference, New Delhi, India*. Retrieved http://www.acer.edu.au/documents/Susan.Brookhart.pdfFormative assessment and the design of instructional systems.

Brookhart, S. M. (2013). Classroom Assessment in the context of Motivation Theory and Research. In J. H. McMillan (Ed.), *SAGE handbook of research on classroom assessment* (pp. 35-54). New York. SAGE.

Brosnan, A. (2014). Introducing lesson study in promoting a new mathematics curriculum in Irish post-primary schools. *International Journal for Lesson and Learning Studies*, 3(3), 236 – 251.

Brown, J. D. (2014). Mixed Methods Research for TESOL. Edinburgh University Press Scotland.

Brown, J.D. (2016). Questions and answers about language testing statistics: Characteristics of sound mixed methods research in *Shiken 20*(1). June 2016.

Bubb, S., & Earley, P. (2007). *Leading and Managing Continuing Professional Development* (2nd ed.). London: Sage.

Buckley, C. A., & Waring, M. J. (2013). Using diagrams to support the research process: examples from grounded theory. *Qualitative Research, 13*(2), 148-172.

Burke, D. (2014). Audit of Mathematics Curriculum Policy across 12 Jurisdictions. National Council for Curriculum and Assessment, Dublin.

Burns, E., Fenwick, J., Schmied, V., & Sheehan, A. (2012). Reflexivity in midwifery research: The insider/outsider debate. *Midwifery, 28*(1), 52-60.

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): 1-14.

Cai, J. & Lester, F. (2010). *Why is teaching with problem solving important to student learning?* Reston, VA: National Council of Teachers of Mathematics.

Cajkler, W. & Wood, P. (2016): Lesson Study and Pedagogic Literacy in Initial Teacher Education: Challenging Reductive Models, *British Journal of Educational Studies, 64(4), 503-521.*

Cajkler, W., Wood, P., Norton, J., Pedder, D. & Xu, H. (2015) Teacher perspectives about lesson study in secondary school departments: a collaborative vehicle for professional learning and practice development*, Research Papers in Education,* 30:2, 192-213.

Cajkler, W., & Wood, P. (2015). Adapting 'lesson study' to investigate classroom pedagogy in initial teacher education: What student-teachers think. *Cambridge Journal of Education*, 46(1), 1-18.

Cajkler, W., Wood, P., Norton, J. & Pedder, D. (2014) Lesson Study As A Vehicle For Collaborative Teacher Learning In A Secondary School. *Professional Development in Education,* 40:4, 511-529.

Cameron, R. (2011). Mixed methods research: The five Ps framework. *The Electronic Journal of Business Research Methods*, 9 (2), 96-108.

Carroll, C., Masters 2013. Exploring the Impact of Lesson Study on the Theory-Practice Gap in Pre-service Teacher Education. Unpublished Masters in Education. Mary Immaculate College: Limerick.

Cauley, M. K. & James H. McMillan, J. H. (2010) Formative Assessment Techniques to Support Student Motivation and Achievement. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, 83:1, 1-6.

Cerbin, B. (2011). *Lesson study: Using classroom inquiry to improve teaching and learning in higher education.* Stylus publishing: Virginia.

Cerbin, B. (2013). Emphasizing learning in the scholarship of teaching and learning. *International Journal for the Scholarship of Teaching and Learning*, 7(1), Article 5.

Cerbin, W. & Kopp, B. (2011). *Lesson study guide*. Retrieved November 5, 2016 from http://www.uwlax.edu/sotl/lsp/guide.

Chappuis, S. (2005). Is formative assessment losing its meaning? *Education Week, 24*(44), 38.

Chappuis, S., Stiggins, R., Arter, J., & Chappuis, J. (2004). *Assessment for learning: An action guide for school leaders*. Portland, OR: Assessment Training Institute.

Chappuis, J. (2014). Clear purpose: Assessment for and of learning. In Chappius, J., Stiggins, R., Chappius, S., Arter, J. (Eds.), *Classroom assessment for student learning doing it right – using it well*, (pp. 21-44). Essex: Pearson Education Limited.

Chenoweth,K. (2000, December21). Inadequate teacher training puts U.S.students behind in math,sience.TheWashington Post, p.T09. Accessed August 29, 2015 http://www.washingtonpost.com/archive/local/2000/12/21/inadequate-teacher-training-puts-us-students-behind-in-math-science/e1aecd48-ec4e-4781-b06e-c44712eb91e3/

Cheung, W. M., & Wong, W. Y. (2014). Does Lesson Study work? A systematic review on the effects of lesson study and learning study on teachers and students. *International Journal for Lesson and Learning Studies*, 3(2), 137 – 149.

Chichiby, T. & Kihara, T. (2013). How Japanese schools build a professional learning community by lesson study. *International Journal for Lesson and Learning Studies,* 2(1), 12 – 25.

Chong, W. and Kong, C. (2012). Teacher collaborative learning and teacher self-efficacy: The case of lesson study. *Journal of Experimental Education,* 80(3), 263-283.

Clark, I. (2011). Formative Assessment: Policy, Perspectives and Practice inFlorida *Journal of Educational Administration and Policy*, 4(2), 158-180.

Clark, I. (2012). Formative assessment: Assessment is for self-regulated learning. *Educational Psychology Review,* 24(2), 205-249.

Clarke, S. (2008). *Active Learning Through Formative Assessment*. London: Hodder Education.

Clarke, S. (2014). *Outstanding Formative Assessment: Culture and Practice.* London: Hodder Education.

Clerkin, A., & Creaven, A-M. (2013). Pupil engagement. *In Eivers & A. Clerkin, (Eds), National schools, international contexts: Beyond the PIRLS and TIMSS test results*, 33-35. Dublin: Educational Research Centre. In The following appear as chapters in E. Eivers & A. Clerkin, (Eds), *National schools, international contexts: Beyond the PIRLS and TIMSS test results.* (2013). Dublin: Educational Research Centre.

Clerkin, A. (2013). Teachers and teaching practices. In E. Eivers & A. Clerkin (Eds.) *National Schools, international contexts: Beyond the PIRLS and TIMSS test results* (pp.77-105). Dublin: Educational Research Centre.

Cochran-Smith, C. & Lytle, S. (1993). *Inside/Outside: Teacher research and knowledge.* New York: Teachers College Press.

Cochran-Smith, C. & Lytle, S. (2009). *Inquiry as stance: Practitioner research for the next generation*. New York: Teachers College Press.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Cohen, L., Manion, L., & Morrison, K. (2010). *Research methods in education (6th ed.)*. Oxon, Great Britain: Routledge.

Collins, B. (2010). *The Impact Of Formative Assesment In A Socio-Cultural Learning Environment In The Visual Arts*. (Unpublished master's dissertation). St. Patrick's College, Drumcondra: Dublin.

Collins, B., & O'Leary, M. (2010). Integrating assessment with teaching and learning in the visual arts: a study in one classroom. *Oideas* 55, 54-62. Dublin: Department of Education and Science.

Connolly, L. (2013) Demographic Data in Research Studies MEDSURG Nursing;Jul/Aug2013, Vol. 22 Issue 4, p269

Conway, P. (2013). Overview of the Learning to Teach Study and RIITILS. Conference presentation: Professional Learning and Lesson Study Conference, UCC, 13 June 2013.

Conway, P.F., & Murphy, R. (2013). A Rising Tide Meets A Perfect Storm: New Accountabilities In Teaching And Teacher Education In Ireland. *Irish Educational Studies,* 32(1), 11-36. Oxon, UK: Routledge.

Conway, P., & Sloane, F. (2005). *International trends in post-primary mathematics education*. Dublin: National Council for Curriculum and Assessment.

Cooper, B. & Cowie, B. (2009). Collaborative research for assessment for learning. Teaching and Teacher Education, 26(4), 979-986.

Corcoran, D. (2007). Put Out into Deep Water and Pay out Your Nets for a Catch: Lessons learned from a pilot study in Mathematics Lesson Study. In S. Close, D. Corcoran & T. Dooley (Eds.), *Proceedings of Second Annual Conference on Research in Mathematics Education* (pp. 275-289). MEI 2: St. Patrick's College, Ireland.

Corcoran, D. (2008). *Developing Mathematical Knowledge For Teaching: A Three-Tiered Study Of Irish Pre-Service Primary Teachers.* (Unpublished doctoral dissertation). University of Cambridge, Homerton College, UK.

Corcoran, D. (2009). Mathematics Teaching as a Shared Enterprise: Learning by 'Doing'. In D. Corcoran, T. Dooley, S. Close & R. Ward (Eds.), *Third National Conference on Research in Mathematics Education (MEI 3)* (pp. 34 - 52). St. Patrick's College, Dublin.

Corcoran, D. (2010). Lesson Study: Researching Learning About Teaching From Research Mathematics Lessons pp.1-6. Http://Www.Nairtl.Ie/Documents/Research%20teaching%20linkages.Pdf

Corcoran, D. (2011a). Learning from Lesson Study: Power Distribution in a Community of Practice. In L. Hart, A. S. Alston & A. Murata (Eds.), *Lesson Study Research and Practice in Mathematics Education* (pp. 251-268). New York: Springer.

Corcoran, D. (2011b). Lesson Study in practice: Staff ownership of mathematics teacher professional development. In T. Dooley, D. Corcoran & M. Ryan (Eds.), *Fourth Conference on Research in Mathematics Education MEI 4* (pp. 130-142). St. Patrick's College, Drumcondra, Dublin 9.

Corcoran, D. (2012). Taking the teaching of mathematics seriously: making mathematics teaching public. St Patrick's College SMEC 2012.

Corcoran, D., & Pepperell, S. (2011). Learning to Teach Mathematics Using Lesson Study, in Mathematical Knowledge in Teaching. In T. Rowland & K. Ruthven (Eds.) *Mathematical Knowledge in Teaching.* London: Springer.

Corcoran, D., Reilly, M., Breen, S., Dooley, T., & Ryan, M. (2011). 'The application of lesson study across mathematics and mathematics education departments in an Irish third-level institution' From Informal Proceedings of the British Society for Research into Learning Mathematics (BSRLM), 31-1.

Cordingley, P. (2015b). The contribution of research to teachers' professional learning and development. *Oxford Review of Education*, 41(2), 234-252.

Cordingley, P., Bell, M., Thomason, S. and Firth, A. (2005) The impact of collaborative continuing professional development (CPD) on classroom teaching and learning. Review: How do collaborative and sustained CPD and sustained but not collaborative CPD affect teaching and learning? In Research Evidence in Education Library. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Cowie B., & Moreland, J. (2015). Leveraging disciplinary practices to support students' active participation in formative assessment. *Assessment in Education: Principles, Policy & Practice,* 22(2), 247-264.

Coyne, I. (2010), Research with Children and Young People: The Issue of Parental (Proxy) Consent. Children & Society, 24: 227–237.

Creswell, J. W. (2009). *Research Design: Qualitative, quantitative, and mixed methods approaches (3rd ed.).* California, United States: Sage.

Creswell, J. W. (2013). "Steps in Conducting a Scholarly Mixed Methods Study" (2013). DBER Speaker Series. Paper 48. http://digitalcommons.unl.edu/dberspeakers/48

Creswell, J., & Plano Clarke, V. (2011). *Designing and conducting mixed methods research.* California, United States of America: Sage.

Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research, 58*(4), 438–481.

Cumming, J.J., Maxwell, G.S., & Wyatt-Smith, C.M. (2016). School leadership inassessment in an environment of external accountability: Developing an assessment for learning culture. In G. Johnson, & N. Dempster (Eds.), *Leadership in diverse learning contexts, studies in educational leadership* (pp. 221-237). Dordrecht, The Netherlands: Springer.

Darcy, B. (2011). *The impact of 'Assessment for learning' on pupil writing.* (Unpublished master's dissertation). St. Patrick'c College, Drumcondra, Dublin.

Darling-Hammond, L. & Richardson, N. (2009). Teacher Learning: What matters? *Educational Leadership, 66(5), 46-55.*

Day, C. and Sachs, J. (2004) Professionalism, performativity and empowerment: discourse in the politics, policies and purposes of continuing professional development, in C. Day and J. Sachs, J. (Eds) International Handbook on the Continuing Professional Development of Teachers, Maidenhead, Birks, Open University Press.

Day, C., Sammons, P., Stobart, S., Kington, A., & Gu, Q. (2007). *Teachers matter: Connecting work, lives and effectiveness*. Milton Keynes: Open University Press.

De Corte, E. (2004) Mainstreams and Perspectives in Research on Learning (Mathematics) from Instruction, *Applied Psychology: An International Review,* 53 (2) 279-310.

Deevers, M. (2006, April). *Linking classroom assessment practices with student motivation in mathematics.* Paper presented at the annual meeting of the American Research Association, San Francisco, CA.

Delaney, S. (2005). Mathematics professional development for primary teachers: Looking back and looking forward. In S. Close, T. Dooley, & D. Corcoran (Eds.) *Proceedings of First National Conference on Research in Mathematics Education* (pp. 235-249). Dublin, Ireland: St. Patrick's College.

Delaney, S. (2008). *Adapting and using U.S. measures to study Irish teachers' mathematical knowledge for teaching*. (Unpublished doctoral dissertation). University of Michigan, United States.

Delaney, S. (2012b). Problems in Teaching Primary School Mathematics. Presentation Laois Education Centre, 18 October 2012.

DeLuca, C., Klinger, D., Pyper, J., & Woods, J. (2015). Instructional Rounds as a professional learning model for systemic implementation of Assessment for Learning. *Assessment in Education: Principles, Policy & Practice,* 22:1, 122-139.

Denscombe, M. (2008). Communities of practice: A research paradigm for mixed methods approach. *Journal of Mixed Methods Research*, 2(3), 270-283.

Denscombe, M. (2010). *The good research guide for small-scale social research projects (4th ed.)*. Berkshire, England: Open University Press.

Department of Children and Youth Affairs (2011). *Children First: National Guidelines for the Protection and Welfare of Children* (DCYA, 2011), Published by Government Publications, Dublin.

Department of Children and Youth Affairs (2012). *Guidelines for Developing Ethical Research Projects involving Children*. Published by Government Publications, Dublin.

Department of Education and Science (1999). *Primary School Mathematics Curriculum.* Dublin: The Stationery Office.

Department of Education and Science, (2006). Circular Letter 0138/2006

Department of Education and Skills (2010a). *Incidental inspection findings 2010. A report on the teaching and learning of english and mathematics in primary schools*. Dublin: Department of Education and Skills.

Department of Education and Skills (2010b). *The 2009 National Assessment of Mathematics and English Reading*. Dublin: Educational Research Centre. For the Department of Education and Skills.

Department of Education and Skills. (2011a). *Literacy and numeracy for learning and life: The national strategy to improve literacy and numeracy among children and young people 2011-2020.* Dublin: Department of Education & Science.

Department of Education and Skills (2011b). *Circular 0056.* Dublin: Department of Education and Skills.

Department of Education and Skills (2012a). Organisation for Economic Co-operation and Development (OECD) (2012). OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes: Country Background Report for Ireland. Paris: France: OECD.

Department of Education and Skills (2012b). WSE Report. Scoil na nAingeal.

Desmione, L. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher,* 38(3), 181-199.

Desmione, L. (2011). A primer of effective professional development. *Phi Delta Kappan,* 92(6), 68-71.

Desforges, C. (2015). Lesson Study as a Strategic Choice for CPD (Foreword). In Dudley, P. (ed.). *Lesson Study: Professional Learning for Our Time*. UK: Routledge.

Doig, B. & Groves, S. (2011). Japanese Lesson Study: Teacher Professional Development through Communities of Inquiry. LS *Mathematics Teacher Education and Development,* 13(1), 77–93.

Doig, B., Groves, S., & Fujii, T. (2011). The critical role of task development in Lesson Study. In L. Hart, A. Alston & A. Murata (Eds.), *Lesson study research and practice in mathematics education* (pp. 181–199). Dordrecht, The Netherlands: Springer.

Doig, B., & Groves, S. (2012). Japanese lesson study: Teacher professional development through communities of inquiry. Mathematics Teacher Education and Development, 13(1), 77–93.

Dolan, R. (2016). Initiation and Implementation: Changes to Teacher Education in Ireland. In Spector, J.M., Ifenthaler, D., & Sampson, D.G. (Eds.) *Competencies in Teaching, Learning and Educational Leadership in the Digital Age*: Papers from CELDA 2014 pp.33-46. Switzerland: Springer.

Dotger, S. (2011). Exploring and developing graduate teaching assistants' pedagogies via lesson study. *Teaching in Higher Education,* 16(2), 157–169.

Doyle, C. (2012). *Assessment for learning: an investigation of the knowledge, skills and attitudes of a teaching staff in an Irish primary school.* (Unpublished master's dissertation). St. Patrick's College, Drumcondra, Dublin.

Dudley, P. 2011. Lesson Study: A Handbook. Accessed 1 August, 2012. http://lessonstudy. co.uk/wp-content/uploads/2012/03/Lesson_Study_Handbook_-_011011-1.pdf.

Dudley, P. (2012a) in Lesson Study: Making a Difference to Teaching Pupils with Learning Difficulties edited by Jeff Jones, Brahm Norwich, Abigail Paterson, Chapter 2 p.15-33: The general Rationale and Underlying Principles of LS. Bloomsbury Publishing: London.

Dudley, P. (2012b). Lesson Study development in England: from school networks to national policy. *International Journal for Lesson and Learning Studies, 1*(1), 85-100.

Dudley, P. (2013). Teacher Learning In Lesson Study: What Interaction-Level Discourse Analysis Revealed About How Teachers Utilised Imagination, Tacit Knowledge Of Teaching And Fresh Evidence Of Pupils Learning, To Develop Practice Knowledge And So Enhance Their Pupils' Learning. *Teaching and Teacher Education, 34*(1), 107-121.

Dudley, P. (2015). *Lesson Study: Professional Learning for Our Time*. UK: Routledge.

Duerr, S.R. & Harlow, L.L (2013). Mulitsample latent variable model of math achievement across international country clusters. *Multivariate Behavioural Research*, Vol 48 (1), 96-116.

Dunn, K. E., & Mulvenon, S. W. (2009). A Critical Review of Research On Formative Assessment: The Limited Scientific Evidence of The Impact of Formative *Assessment In Education. Practical Assessment, Research & Evaluation*, 14(7), 1-11.

Dunphy, E. Dooley, T & Shiel, G. (2014). *Mathematics in Early Childhood and Primary Education. Research Report 17,* National Council for Curriculum and Assessment, Dublin.

Dweck, C. S. (2000). *Self-theories: Their role in motivation, personality and development*. Philadelphia, PA: Psychology Press.

Dwyer, S. & Buckle, J. (2009). The Space Between: On Being an Insider-Outsider in Qualitative Research. *International Journal of Qualitative Methods*, 8 (1), pp. 54-63.

Earl, L. M. (2013). *Assessment as learning: Using classroom assessment to maximize student learning.* Corwin: Thousand Oaks, CA.

Earl, L. & Timperley, H. (2014). Challenging conceptions of Assessment. in C. Wyatt-Smith, V. Klenowski & Colbert, P. (Eds). Pp325-336. Designing Assessment for Quality Learning. Dordrecht: Springer.

Educational Research Centre (2007). The Drumcondra Primary Mathematics Test Revised (DPMT-R). ERC: Dublin.

Edwards, R. (2012). Critical Friend Commentary on Chapter 15 (Issues and strategies for student engagement through assessment in transnational higher education by Glenda Crosling). In Clouder, L. Broughan, C., Jewell, S. & Steventon, G. (eds.) *Improving Students engagement and development through assessment: Theory and Practice in Higher Education* (p.196-208). Oxon: Routledge.

Eisner, E. (2005). The Whole Child: Back to Whole. Educational Leadership. Vol 63 No.1. pp.14-18.

Eivers, E., & Clerkin, A. (2013). *National Schools, international contexts: Beyond the PIRLS and TIMSS test results.* Dublin: Educational Research Centre.

Elliot, J. & Lo Mun Ling. (2011). Editorial. *International Journal for Lesson and Learning Studies*, Vol 1, Issue 1.

Elwood, J (2006) Formative Assessment: possibilities, boundaries and limitations, *Assessment in Education*, 13(2), pp. 215-232.

Elwood, J. 2013. The Role(s) of Student Voice in 14–19 Education Policy Reform: Reflections from Students on What They Are, and What They Are Not, Consulted about. London Review of Education.

Elwood, J & Lundy, L 2010, 'Revisioning assessment through a children's rights approach: implications for policy, process and practice' *Research Papers in Education*, vol 25, no. 3, pp. 335-353.

Elwood, J., & Murphy, P. (2015). Editorial. Assessment systems as cultural scripts: A socio-cultural theoretical lens on assessment practice and products.

Assessment in Education: Principles, Policy & Practice, 22(2), 182–192.

Feilzer, M. Y. (2010). Doing mixed methods research pragmatically: Implications for the rediscovery of pragmatism as a research paradigm. *Journal of Mixed Methods Research*, 4 (1), 6-16.

Ferguson, L. M., Yonge, O., & Myrick, F. (2004). Students' involvement in faculty research: Ethical and methodological issues. International Journal of Qualitative Methods, 3(4), 1−14.

Fernandez, C. (2002). Learning from Japanese approaches to professional development: The case of lesson study. *Journal of Teacher Education, 53*(5), 393-405.

Fernandez, C., Cannon, J., & Chokshi, S. (2003). A US-Japan lesson study collaboration reveals critical lenses for examining practice. *Teaching and Teacher Education, 19*(2), 171-185.

Fernandez, C. (2005). Lesson study: A means for elementary teachers to develop the knowledge of mathematics needed for reform-minded teaching? *Mathematical Thinking and Learning, 7*(4), 265-289.

Fernandez, C. & Yoshida, M. (2004). *Lesson study: A Japanese approach to improving mathematics teaching and learning.* Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Filsecker, M. & Kerres, M. (2012). Repositioning formative assessment from an educational assessment perspective: A response to Dunn & Mulvenon (2009). *Practical Assessment, Research & Evaluation*, 17 (16), 1-9.

Flórez, M. T. & Sammons, P. (2013). *Assessment for learning: effects and impact.* England: Oxford University, Department of Education.

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, 407-417.

Fraenkel, J.R., & Wallen, N.E. (2006). *How to design and evaluate research in education*. New York: McGraw-Hill.

Fuchs, L. S., & Fuchs, D. (1986). Effects of systematic formative evaluation—a meta-analysis. *Exceptional Children, 53*(3), 199-208.

Fullan, M. G. (1991). Professional development of educators. In M. G. Fullan, *The new meaning of educational change* (pp. 315-344). New York: Teachers College Press.

Furlong, M. J. and Christenson, S. L. (2008), Engaging students at school and with learning: A relevant construct for *all* students. Psychology in the Schools., 45: 365–368.

Galanouli, D. (2010). School-Based Professional Development: The General Teaching Council of Northern Ireland.: Northern Ireland.

Gall, M. D., Gall, J. P., & Borg, W. R. (2003). *Educational research: An introduction (7th ed.).* Boston, United States of America: Allyn and Bacon.

Gardner, J. (2012a). Quality assessment practice. In J. Gardner (Ed.), *Assessment and learning* (2nd ed.), (pp. 103-121). London: Sage Publications.

Gardner, J. (2012b). *Assessment and Learning* (2nd ed.). London: Sage Publications.

Gardner, J., Harlen, W., Hayward, L. & Stobart, G. (2008). *Changing assessment practice process: Principles and standards.* Assessment Reform Group.

Gardner, J., Harlen, W., Hayward, L. & Stobart, G. (2010). *Developing teacher assessment.* England: Open University Press.

Gardner, J., Galanouli, D., Devlin, G., Magee, M., McSweeney, K., McHenry, M., McVeigh, I. & Mitchell, S. (2012). Exploring Japanese Lesson Study as a Model of Peer-to-Peer Professional Learning in *Three reports for the Standing Conference on Teacher Education North and South (SCoTENS)*. Northern Irelands: SCOTENS.

Geertz, C. (1973). Thick Description: Toward an Interpretive Theory of Culture. In *The Interpretation of Cultures: Selected Essays*. New York: Basic Books, pp.3-30.

Gerring, J. (2007). *Case study research: Principles and practices.* Cambridge, United States of America: Cambridge University Press.

Gersten, R., Taylor, M.J., Keys, T.D., Rolfhus, E., & Newman-Gonchar, R. (2014). *Summary of research on the effectiveness of math professional development approaches* (REL 2014-010). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southeast. Available at: http://ies.ed.gov/ncee/edlabs/projects/project.asp?projectID=391.

Gómez-Blancarte, A., & Miranda, I., (2014). Communities of Practice: A theoretical Framework to design for teachers' statistical learning: Conference ICOTS9 Invited Paper.

Government of Ireland, (1998). Education Act. Dublin, Stationery Office.

Government of Ireland (1999a) *Primary School Mathematics Curriculum* (Dublin, Stationery Office).

Government of Ireland (1999b) *Primary School Mathematics Teacher Guidelines* (Dublin, Stationery Office).

Grant, C. M. (1996). Professional development in a technological age: New definitions, old challenges, new resources. In A. Feldman (Ed.). Technology infusion and school change: Perspectives and practices. Cambridge, MA: TERC Research Monograph.

Gray, D. E. (2014). Doing Research in the real world. Sage: London, UK.

Greene, J.C. & Caraceili, V.J. (1997). Defining and describing the paradigm issue in mixed-method evaluation. *New Directions for Evaluation*, 5-17.

Greenstein, L. (2010). *What teachers really need to know about Formative Assessment.* United States: Association for Supervision and Curriculum Development.

Gurhy, A. M. (2008). As láthair arís: 'Every day counts in a child's education' – or does it? *A case-study investigation of attendance in an Irish primary school. Unpublished Masters dissertation. NUI Maynooth.*

Guskey, T. R. (2000). *Evaluating professional development*. Thousand Oaks, CA: Corwin.

Gutierez, S. B. (2015). Collaborative Professional Learning: Implementing Inquiry Based Teaching Through Lesson Study. *Issues in Educational Research*, *25*(2), 118-134.

Hammerness, K., Darling-Hammond, & Bransford, J. (2005). How teachers learn and develop. In L. Darling-Hammond & J. Bransford (Eds.), *Preparing teachers for a changing world: What teachers should learn and be able to do* (pp. 358-389). San Francisco, CA: Jossey- Bass.

Hargreaves, D. (2004). *Personalising learning – 2: Student voice and assessment for learning*. London: Specialist Schools Trust.

Hargreaves, A., & Fullan, M. (2012). *Professional capital: Transforming teaching in every school*. Teachers College Press.

Harlen, W. (2012) On the relationship between assessment for formative and summative purposes. In J. Gardner (ed.) *Assessment and Learning* (pp.87-102). London: Sage.

Harlen, W. (2014). *Assessment, standards and quality of learning in primary education*. York: Cambridge Primary Review Trust.

Harlen, W., & Johnson, S. (2014). *A review of current thinking and practices in assessment in relation to the Primary Years Programme.* The Hague, NL. International Baccalaureate Organization.

Harlan, W., & Winter, J. (2004). The development of assessment for learning: learning from the case of science and mathematics, *Language Testing, 21*(3), 390-408.

Harris, L. R., Brown, G., & Harnett, J.A. (2015). Analysis of New Zealand primary and seconday student peer- and self-assessment comments: applying Hattie and Timperley's feedback model. *Assessment in Education: Principles, Policy and Practice*, 22 (2), 265-281.

Harris, L.R., Gavin, T. L., & Harnett, J. (2012). Student pictures of feedback 1. Paper presented at AERA 2012, Vancouver Canada. Student pictures of feedback: Feedback is for learning and from teachers. The University of Auckland. Paper presented at the April 2012 AERA Annual Meeting of the American Educational Research Association conference, Vancouver, BC to the Classroom Assessment SIG.

Harrison, C. (2013). Testing Times: Reforming Classroom Teaching Through Assessment. In Jonathan Clifton (ed.) *Excellence And Equity Tackling Educational Disadvantage In England's Secondary Schools* (pp71-77). UK: Institute for Public Policy Research.

Hart, L. C., & Carriere, J. (2011). Developing the Habits of Mind for a Successful Lesson Study Community. In H. C. Lynn, A. Alston, & A. Murata (Eds.), Lesson Study Research and Practices in Mathematics Education. Dordrecht: Springer.pp.27-38.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement.* New York: Routledge.

Hattie, J. (2012). *Visible Learning for teachers: Maximising impact on learning*. New York: Routledge.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*, 81-112.

Hayward, L. (2012). Assessment and learning: The learner's perspective. In J. Gardner (Ed.), *Assessment and learning* (2nd ed.), (pp. 125-139). London: Sage Publications.

Hayward, L. (2013). From Why to Why Not? The Conundrum of Including Learners' Perspectives: A Response to this Special Issue. *London Review of Education*, 11(2), 184-189.

Hayward, L., Higgins, S., Livingston, K., Wyse, D. & Spencer, E. (2014). Special issue on assessment for learning, *The Curriculum Journal*, 25:4, 465-469,

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), 397-416.

Heritage, M. (2013). *Formative Assessment in Practice: A Process of Inquiry and Action* (Assessment, Accountability, & Achievement Series). Cambridge, Massachusetts: Harvard Education Press.

Heritage, M. & Wylie, C. (2013). Policy Support for Formative Assessment. Heritage, Margaret. In Formative Assessment in Practice: A Process of Inquiry and Action. Cambridge, pp.117-133. MA: Harvard Education Press.

Hesse-Biber, S.N. & Leavy, P.L. (2011). *The Practice of Qualitative Research*. United States, Sage.

Hickey, D. T. (2015). A situative response to the conundrum of formative assessment. *Assessment in Education: Principles, Policy & Practice*, 22(2), 202-223.

Hill, M., Laybourn, A. and Borland, M. (1996) 'Engaging with Primary-aged Children about their Emotions and Well-being: Methodological Considerations', *Children and Society*, 10, pp 129–144.

Hinton, P.R. (2008). Statistics Explained. 2nd ed. Routledge: UK.

Hirsh, S. (2009). Creating Effective Professional Learning Systems to Bolster Teaching Quality and Student Achievement. In Professional Learning in the Learning Profession: A Status Report on Teacher Development in the United States and Abroad (Preface). In Linda Darling-Hammond, Ruth Chung Wei, Alethea Andree, Nikole Richardson, and Stelios Orphanos (eds.) *Professional learning in the learning profession: A status report on teacher development in the United States and abroad.* Dallas: National Staff Development Council.

Hislop, H (2013). Applying an evaluation and assessment framework: An Irish perspective. *Presidency Conference: Better Assessment and Evaluation to Improve Teaching and Learning.* (pp.1-22). Dublin: Department of Education and Skills.

Hodgson, C. & Pyle, K. (2010). *A Literature Review of Assessment for Learning in Science.* Research Report. Slough: National Foundation for Educational Research.

Hodgen, J., & Wiliam, D. (2006). *Mathematics inside the black box: Assessment for learning in the mathematics classroom*. London, UK: NFER-Nelson.

Hogan, P. (2015). Lesson study East and West: Identifying some key issues, *International Journal for Lesson and Learning Studies*, 4(2), 178–181.

Hopfenbeck, T. (2013). Students' voice, aspirations, and perspectives: international reflections and comparisons. *London Review of Education*, 11(2), pp.179-183.

Hopfenbeck, Petour, & Tolo, (2015). Balancing tensions in educational policy reforms: Large-scale implementation of Assessment for Learning in Norway. *Assessment in Education: Principles, Policy and Practice,* 22(1), 44-60.

Hord, S. M. (2008). Evolution Of The Professional Learning Community. *Journal of Staff Development*, 29(3), 10-13.

Irish National Teachers' Organisation, (2010). *Assessment in the Primary School Discussion Document and Proceedings of the Consultative Conference on Education 20*08. Dublin: Irish National Teachers' Organisation.

Irish National Teachers' Organisation, (2011). *Literacy in a Changing World.* Discussion Document Education Conference Athlone. INTO.

Irish National Teachers' Organisation, (2013). *Numeracy in the Primary School: A discussion paper*. Dublin: INTO.

Irish National Teachers' Organisation, (2015). *Curriculum: A discussion paper*. Dublin: INTO.

James, M. (2006). Assessment, Teaching and Theories of Learning. In J. Gardner (Ed.), *Assessment and learning* (pp. 47-60). London: Sage.

James, M., (2008). Assessment and Learning. In S. Swaffield (ed) *Unlocking Assessment (*pp. 20-35). London: Routledge.

James, M. & Lewis, J. (2012). Assessment in Harmony with our understanding of Learning: Problems and Possibilities. Chapter 12 pp.187-205. in Gardner, J. (Ed.) *Assessment and Learning* (2nd Edition). London: Sage.

James, M., Black, P., McCormick, R., Pedder, D., & Wiliam, D. (2006). Introduction. Learning how to learn, in classrooms, schools and networks: Aims design and analysis. *Research Papers in Education*, 21(2), 101-118.

Jewitt, C. (2012). *An Introduction to Using Video for Research*. Institute of Education, London.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, 33 (7), 14-26.

Jonsson, A., Lundahl, C., & Holmgren, A. (2015). Evaluating a large-scale implementation of Assessment for Learning in Sweden. *Assessment in Education: Principles, Policy & Practice,* 22(1), 104-121.

Kane, E. & de Brun, M. (2005). *Doing your own research*. London, Great Britain: Marion Boyars Publishers.

Kanuha, V. K. (2000). “Being” native versus “going native”: Conducting social work research as an insider. *Social Work*, 45 (5), 439-447.

Keeley, P. (2016). *Science Formative Assessment*: *75 Practical Techniques For Linking Assessment, Instruction And Learning* (2nd Ed). USA: Corwin Press and NSTA Press joint publication.

Kelly, A. E., & Sloane, F. C. (2003). Educational research and the problems of practice. *Irish Educational Studies, 22(1).* Oxon, UK: Routledge.

Kieran, C., Krainer K., & Shaughnessy, J. M. (2013). Linking research to practice: Teachers as key stakeholders in mathematics education research. In K. Clements, A. Bishop, C. Keitel, J. Kilpatrick, and F. Leung, (Eds.), *Third International Handbook of Research in Mathematics Education* (pp. 361 – 392). New York: Springer.

Koh, K.H. (2011). Improving teachers’ assessment literacy through professional development, *Teaching Education*, 22(3), 255-276.

King, F. (2012). *Developing and Sustaining Teachers' Professional Learning: A Case Study of Collaborative Professional Development.* (Unpublished doctoral dissertation). University of Lincoln, UK.

Kingston, N.M. & Nash, B. (2011). Formative Assessment: A Meta-Analysis and a call for Research. *Educational Measurement: Issues and Practice*, 30(4), 28–37.

Kingston, N.M., & Nash, B. (2015). Erratum. *Educational Measurement: Issues and Practice,* 34: 55.

Klenowski, V. (2009). Assessment For Learning Revisited: An Asia-Pacific Perspective. *Assessment in Education: Principles, Policy and Practice, 16*(3), 263-268.

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*, 254–284.

Krainer, K. (2011). Teachers as stakeholders in mathematics education research. In Ubuz, B. (Ed.) Proceedings of the 35th Conference of the International Group for the Psychology of Mathematics Education. Vol. 1, pp. 47-62. Ankara, Turkey: PME.

Lawrence, C.A. & Chong, W.H. Asia Pacific Educ. Rev. (2010) 11: 565. doi:10.1007/s12564-010-9103-3.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. NY: 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.

Leahy, S., & Wiliam, D. (2012). From Teachers to Schools: Scaling up Professional Development for Formative Assessment. In Gardner, J. (2012). *Assessment and Learning* (2nd ed.) pp. 49-71. London: Sage Publications.

Leavy, A.M. (2010). Preparing pre-service teachers to teach informal inferential reasoning. *Statistics Education Research Journal*, 9(1), 46-67.

Leavy, A., & Hourigan, M. (2014). Motivating inquiry in statistics and probability in the primary classroom. *Teaching Statistics*, 37(2), 41-47.

Leavy, A., Hourigan, M., & Carroll, C., (2014). Lesson study in mathematics: Authentic Assessment of Inquiry Learning. SAILS Conference: *Thinking Assessment in Science and Mathematics.* Dublin City University, Dublin Ireland.

Lenski, S. J., Caskey, M. M., & Anfara, V.A. (2009). Using the Lesson Study Approach to Plan for Student Learning. *Middle School Journal*, 40(3), 50-57.

Lesham, S., & Trafford, V. (2007). Overlooking the conceptual framework. *Innovations in Education and Teaching*, 44 (1), 93-105.

Lesson Study Group at Mills College. (2000). *Can You Lift 100kg?* [DVD]. Available from http://www.lessonresearch.net/canyoulift1.html

Lesson Study Group at Mills College. (2003). *To Open a Cube.* [DVD]. Available from http://www.lessonresearch.net/opencube1.html

Lesson Study Group at Mills College. (2005). *How Many Seats?* [DVD]. Available from http://www.lessonresearch.net/howmanyseats.html

Lesson Study Research Group (2008) Available online at: http://www.tc.edu./lessonstudy/ (Accessed 28 September 2008).

Levine, T. R. and Hullett, C. R. (2002), Eta Squared, Partial Eta Squared, and Misreporting of Effect Size in Communication Research. Human Communication Research, 28: 612–625.

Lewis, C., (1995). Educating Hearts and Minds. UK: Cambridge University Press.

Lewis, C. (2000) Lesson Study: The Core of Japanese Professional Development. Invited Address to the Special Interest Group on Research in Mathematics Education American Educational Research Association Meetings, New Orleans, April 28, 2000, Session 47.09.

Lewis, C. (2002a). *Lesson Study: A Handbook of TeacherLed Instructional Change.* Philadelphia: Research for Better Schools.

Lewis, C. (2002b). What are the essential elements of Lesson Study? *California Science Project, 2*.

Lewis, C. (2005). Instructional Improvement through Lesson Study: Progress and Challenges. In Learning across Boundaries: In *U.S. – Japan Collaboration in Mathematics, Science and Technology Education.* National Science Foundation, USA.

Lewis, C (2006) Professional Development Through Lesson Study: Progress and Challenges in the U.S., in D. Corcoran and S. Breen (Eds) Proceedings of Second International Science and Mathematics Education Conference, Dublin, St Patrick's College, 90-102.

Lewis, C. (2008). How can Lesson Study Build System-Wide Improvement? Presentation in Sacramento, US on April 16th 2008. Retrieved from: http://www.csus.edu/mase/calessonstudy/2008/docs/proceedings/catherine_lewis.pdf

Lewis, C. (2011). Impact of lesson study on lesson materials, teachers, and teacher community over time. (From the article Response of Part iv: Seeing the Whole Iceberg- The Critical Role of Tasks, Inquiry Stance, and Teacher Learning in Lesson Study, 2011, p.238).

Lewis, C. (2009). What is the nature of knowledge development in lesson study? *Educational Action Research*, 17(1), 95-110.

(Lewis, 2011). (from Lewis personal correspondence). 'Lesson Study Theoretical Model: How Lesson Study Produces Instructional Improvement'.

Lewis, C. What have we learned about lesson study outside Japan? In I. Maitree & P. Wang-Iverson (Eds.), *Lesson study: Challenges in Mathematics Education*, World Scientific Publishing, 141-152.

Lewis, C., & Hurd, J. (2011). *Lesson Study Step by Step: How Teacher Learning Communities Improve Instruction*. USA: Library of Congress Cataloging-in-Publication Data.

Lewis, C., & Perry, R. (2014). Lesson study with mathematical resources: A sustainable model for locally-led teacher professional learning. *Mathematics Teacher Education and Development, 16*(1), 22-42.

Lewis, C., Perry, R., & Hurd, J. (2009). Improving mathematics instruction through lesson study: A theoretical model and North American case. *Journal of Mathematics Teacher Education, 12*(4), 285-304.

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.

Lewis, C.C., Perry, R.P., Friedkin, S., & Roth, J.R. (2012). Improving teaching does improve teachers: Evidence from lesson study. *Journal of Teacher Education*, 63(5), 368–375.

Lim & Chapman (2013) Development of a short form of the attitudes toward mathematics inventory Lim, S.Y. & Chapman, in *Educational Studies in Mathematics* (2013) 82: pp.145-164.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.

Little, J. W. (2007). Teachers' accounts of classroom experience as a resource for professional learning and instructional decision making. *Yearbook (National Society for the Study of Education)*, (1), 217–240.

Looney, A. (2016). Ireland: The Intersection of International Achievement Testing and Educational Policy Development. In Volante, L. (Ed.) *The Intersection Of International Achievement Testing And Education Policy: Global Perspectives On Large-Scale Reform* (pp.75-90). New York: Routledge.

Looney, J. (Ed.). (2005). *Formative assessment: improving learning in secondary classrooms*. Paris, France: Organisation for Economic Cooperation and Development.

Lui, A., & Andrade, H. (2014). Student peer assessment. In Gunstone, R.(ed) *Encyclopedia of Science Education* (pp.1-3). Dordreacht: Springer.

Lysaght, Z. (2009). *From Balkanisation to Boundary Crossing: Using a Teacher Learning Community to Explore the Impact of Assessment on Teaching and Learning in a disadvantaged School*. (Unpublished doctoral dissertation). St. Patrick's College, Dublin City University, Dublin.

Lysaght, Z. 2010. Assessment for learning and teacher learning communities: Warranted strategies worthy of consideration. *InTouch* 112: 49-51.

Lysaght, Z. (2012). *Feedback: The Sine Qua Non of assessment for learning.* Invited paper presented at the Special Conference of the Irish Learning Support Association (ILSA), St. Patrick's College, Dublin.

Lysaght, Z. (2015). Assessment *for* Learning and *for* Self-Regulation. *The International Journal of Emotional Education*, 7 (1), 20-34.

Lysaght, Z. & O'Leary, M. (2013). An Instrument To Audit Teachers' Use Of Assessment For Learning. *Irish Educational Studies*, 32(2), 1-16. Oxon, UK: Routledge.

Ma, X. (1999). A meta-analysis of the relationship between anxiety toward mathematics and achievement in mathematics. *Journal for Research in Mathematics Education, 30*(5), 520-540.

Makel, M. C., & Plucker, J. (2014). Facts are more important than novelty: Replication in the education sciences. *Educational Researcher*, 43, 304-316. DOI: 10.3102/0013189X14545513.

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.*

Maxwell, J. A. (2005). Qualitative Research Design: An Interactive Approach (2nd Ed.). Thousand Oaks, CA: Sage.

Maxwell, J. (2013). *Qualitative Research Design*. United States: Sage.

McCormick, R. and Murphy, P. (2008). Curriculum: The case for a focus on learning, in P. Murphy and K. Hall (eds), *Learning and Practice Agency and Identities*. London: Sage. pp. 3–18.

McCoy, S., Smyth, E. & Banks, J. (2012). Learning in Focus: The Primary Classroom: I nsights from the Growing Up in Ireland Study. Dublin: ESRI and NCCA.

McDonnell, M.J. (2013). Assessment for learning as mathematical problem solving: an investigation in the primary classroom. Unpublished M.Ed theisis. St Patrick's College, Drumcondra.

McLoone, M., (2011). Lesson study and its impact on participating teachers' attitudes in relation to teaching mathematics. In T. Dooley, D. Corcoran & M. Ryan (Eds.), *Proceedings of Fourth Annual Conference on Research in Mathematics Education: Mathematics teaching matters* (pp. 272-284). Dublin: St. Patrick's College.

McMillan, J. H., & Hearn, J. (2008). Student Self-Assessment: The Key to Stronger Student Motivation and Higher Achievement. *Educational Horizons*, 87(1), 40-49.

McMillan, J.H., Venable, J.C., & Varier, D. (2013). Studies of the Effect of Formative Assessment on Student Achievement: So Much More is Needed. *Practical Assessment, Research & Evaluation*, 18(2). Available online: http://pareonline.net/getvn.asp?v=18&n=2

Meng, C. & Sam, L. (2011). Encouraging the Innovative Use of Geometer's Sketchpad through Lesson Study. *Creative Education, 2,* 236-243.

Mertens, D. (1998). Research methods in education and psychology: Integrating diversity with quantitative and qualitative approaches. Sage Publications: USA.

Miles, M. B., Huberman, A. & Saldana, J. (2014). *Qualitative data analysis: An expanded sourcebook (3rd ed.).* California, United States: Sage.

Miller, T. & Bell, L. (2002) Consenting to what? Issues of access, gate-keeping and 'informed' consent. In Mauthner, M., Birch, M., Jessop, J. & Miller, T. (2002) *Ethics in Qualitative Research* London: Sage.

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.

Mohan, R. (2011). *Teacher education.* PHI Learning: New Delhi.

Morris, A., & Hiebert, J. (2011). Creating Shared Instructional Products: an alternative approach to improving teaching. *Educational Researcher*, 40(1), 5-14.

Morrow, V. & Richards, M.P.M. (1996). The ethics of social research with children: an overview. *Children & Society*, 10, 90-105.

Moss, C., & Brookhart, S. (2009). *Advancing Formative Assessment in Every Classroom: A Guide For Instructional Leaders.* New York: ASCD.

Mullis, I. V. S., Martin, M. A., & Foy, P. (2008). *TIMSS 2007 International Mathematics Report: Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades.*Chestnut Hill, MA: Boston College: TIMSS and PIRLS International Study Centre.

Mullis, I. V. S., Martin, M. A., Foy, P., & Arora, A. (2012). *TIMSS 2011 international results in mathematics.* Chestnut Hill, MA: Boston College: T IMSS and PIRLS International Study Centre.

Murata, A. (2011). Introduction: Conceptual Overview of Lesson Study. In L. C. Hart, A. S. Alston & A. Murata (Eds.), *Lesson Study Research and Practice in Mathematics Education* (pp. 1-12). USA: Springer.

Murata, A. (2013a). Keynote address: Lesson Study as a guide for ITE re-design and deepening engagement with pedagogy. Reform and Re-design of Initial Teacher Education: Deepening Engagement with Pedagogy Western Gateway Building, University College Cork Thursday 13th June 2013.

Murata, A. (2013b). Lesson Study: A Model of Professional Learning. Signature Pedagogy and Professional Learning Seminar . June 13, 2013 Reform and Re-design of Initial Teacher Education: Deepening Engagement with Pedagogy Western Gateway Building, University College Cork Thursday 13th June 2013

Murchan, D., Loxley, A., & Johnston, K. (2009) Teacher Learning and Policy I ntention: Selected Findings from an Evaluation of a Large-Scale Programme of Professional Development in the Republic of Ireland, *European Journal of Teacher Education,*32(4), 455 – 471.

National Council for Curriculum and Assessment. (1999). *Revised Primary Curriculum*. Dublin: Stationary Office.

National Council for Curriculum and Assessment (2005) *Primary Curriculum Review: Phase One, Final Report.* Dublin: NCCA.

National Council for Curriculum and Assessment (NCCA) (2007). *Assessment in the primary school curriculum: Guidelines for schools*. Dublin, Ireland: National Council for Curriculum and Assessment.

National Council for Curriculum and Assessment (2008) *Primary Curriculum Review: Phase Two, Final Report with recommendations.* Dublin: NCCA.

National Council for Curriculum and Assessment (NCCA) (2008). (http://www.ncca.ieAccessed 3 August 2008).

National Council for Curriculum and Assessment (2009). *Aistear: the Early Childhood Curriculum Framework.* Dublin: Author.

National Council for Curriculum and Assessment (2014a). Research Report 17. Research Report No. 17 Mathematics in Early Childhood and Primary Education (3–8 years) Definitions, Theories, Development and Progression Elizabeth Dunphy, Thérèse Dooley and Gerry Shiel With Deirdre Butler, Dolores Corcoran, Miriam Ryan and Joe Travers International Advisor: Professor Bob Perry

National Council for Curriculum and Assessment (2014b). Research Report No. 18 Mathematics in Early Childhood and Primary Education (3–8 years) Teaching and Learning. Thérèse Dooley, Elizabeth Dunphy and Gerry Shiel With Deirdre Butler, Dolores Corcoran, Thérèse Farrell, Siún NicMhuirí, Maura O'Connor, and Joe Travers International Advisor: Professor Bob Perry

National Council for Curriculum and Assessment (2016). Background Paper and brief for the development of a new Primary Matheatics Curriculum.

National Council of Teachers of Mathematics, (2013). Formative Assessment A position of the National Council of Teachers of Mathematics. https://www.nctm.org/uploadedFiles/Standards_and_Positions/Position_Statements/Formative%20Assessment1.pdf [accessed June 28th 2016]

Natriello, G. (1987). The impact of evaluation processes on students. *Educational Psychologist, 22*(2), 155–175.

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.

Noyce, P. & Hickey, D.T. (2011). Conclusions: Lessons learned, controversies, and new frontiers. In P. Noyce & D. T. Hickey (eds.). *New frontiers in formative assessment* (pp. 223-238) Cambridge, MA: Harvard Education Press.

Ní Shúilleabháin, A. (2015). *Developing Mathematics Teachers' Pedagogical Content Knowledge through Lesson Study: A Multiple Case Study at a Time of Curriculum Change.* (Unpublished doctoral dissertation). Trinity College, Dublin.

Nyquist, J. B. (2003). *The Benefits Of Reconstruing Feedback As A Larger System Of Formative Assessment: A Meta-Analysis*. (Unpublished master's dissertation). Vanderbilt University, Nashville, TN.

Ó Breacháin A. and O'Toole, L. (2013). Pedagogy or Politics?: Cyclical Trends in Literacy and Numeracy in Ireland and Beyond. *Irish Educational Studies,* 32, 4, 401-419.

Ó Donnchadha, G. & Keating, Y (2013). Developments in External Evaluation and School Self-Evaluation: Perspectives on the future of external and internal school evaluation. IPPN Annual Conference, Dublin.

OECD (2005) Teachers Matter Education and Training Policy Attracting, Retaining, Developing Effective Teachers, Paris: OECD Publishing.

OECD (2013) Assessment for Learning – The case for formative assessment.

OECD/CERI International Conference "Learning In The 21st Century: Research, Innovation and Policy" Assessment for Learning Formative Assessment

OECD (2010), *PISA 2009 Results: What Students Know and Can Do – Student Performance in Reading, Mathematics and Science (Volume I)*

O'Leary, M. (2006). Towards a balanced assessment system for Irish primary and secondary schools. *Oideas*, 52, 7-24.

Olsen, J.F, Martin, M, & Mullis, I. (2008) TIMSS 2007 Technical Report Chestnut Hill, MA: Boston College: TIMSS and PIRLS International Study Centre.

Onwuegbuzie, A. J. & Johnson, R.B. (2006). The validity issue in mixed methods. In Research in the Schools, Vol 13 (1), pp.48-63).

Opfer, V. D., & Pedder, D. (2011). Conceptualizing Teacher Professional Learning. *Review of Educational Research, 81(3), 376–407.*

O' Shea, J., Teague, S., Jordan, G., Lang, J., & Dudley, P. (2015). Leading Lesson Study in Schools and Across School Systems. In Dudley, P. (ed.). *Lesson Study: Professional Learning for Our Time* (pp.59-85). UK: Routledge.

O'Sullivan, H. (2011). Leadership and Managing Professional Learning in Schools. In O'Sullivan, H. & West-Burnham, J. (eds.) *Leading and Managing Schools (*pp.111-125). SAGE: London.

Pachler, N., & Field, K. (2004) 'Continuing professional development' In Capel, S.; Heilbronn, R.; Leask, M. and Turner, T. (eds) *Starting to teach in the Secondary School (*pp. 233-246). Second edition. London: Routledge Falmer.

Pallant, J. (2013). *SPSS Survival manual: A step by step guide to data analysis using IBM SPSS (5th ed.).* Maidenhead, UK: Mc Graw Hill.

Pedder, D. (2015). Prospects for further development of Lesson Study. in). In Dudley, P. (ed.). *Lesson Study: Professional Learning for Our Time*. UK: Routledge.

Perdue NH, Manzeske DP, Estell DB. (2009). Early predictors of school engagement: Exploring the role of peer relationships. *Psychology in the Schools*. 46(10), 1084–1097.

Perrenoud, P. (1998). From formative evaluation to a controlled regulation of learning. Towards a wider conceptual field. *Assessment in Education: Principles Policy and Practice,* 5(1), 85-102.

Perry, R. R., & Lewis, C. C. (2009). What is successful adaptation of lesson study in the US? *Journal of Educational Change, 10*(4), 365-391.

Perry, R. R., & Lewis, C. C. (2011). *Improving the mathematical content base of lesson study. Summary of results*. Retrieved August 12, 2015, from http://www.lessonresearch.net/IESAbstract10.pdf.

Pintrinch, P.R. (1995). Understanding self-regulated learning. *New Directions for Teaching and Learning,* 63, 3-12.

Pintrich, P. R. (2000). Multiple goals, multiple pathways: The role of goal orientation in learning and achievement. *Journal of Educational Psychology, 92*, 544–555.

Popham, W. J. (2008). *Transformative Assessment*. Virginia, USA: Association for Supervision and Curriculum Development.

Popham, W. J. (2011). Formative Assessment-a process, not a test. *Education Week* 23/02/2011. Vol 30 (21).

Popham, W. J. (2013). Dualism's Dividends. (Foreword) in Formative Assessment in Practice: a process of inquiry and action by Margaret Heritage. Harvard Education Press: USA.

Princeton Review. (2006). Formative assessment. Framingham, MA: Princeton Review.

Pring, R. (2004). Philosophy of Educational Resarch (2nd edition). London: Continuum.

Punch, K.F. (2014). Introduction to Social research: Quantitative and Qualitative Approaches. 3rd Ed. UK: Sage.

Ramaprasad, A. (1983). On the definition of feedback. *Behavioral Science*, 28, 4-13.

Rawlins, P. (2010). Student Participation in Formative Assessment for NCEA *New Zealand Journal of Educational Studies*, Vol. 45, No. 1, 2010: 3-16.

Ravitch, S. M., & Riggan, M. (2012). *Reason and rigor: How conceptual frameworks guide research.* California, United States: Sage.

Ricks, T. E. (2011). Process reflection during Japanese lesson study Experiences by prospective secondary mathematics teachers. *Journal of Mathematics Education,* 14(3), 251–267.

Roberts, S., & Pruitt, E. (2009). *Schools as professional learning communities,* Thousand Oaks, CA: Corwin Press. .

Robinson, V., Hohepa, M., & Lloyd, C. (2009). *School leadership and student outcomes: Identifying what works and why.Best evidence synthesis iteration* (BES). Wellington: Ministry of Education.

Robson, C. (2002). *Real World Research: a resource for users of social research methods in applied settings* (2nd ed.). Oxford: Blackwell Publishing.

Robson, C. (2011). *Real World Research: a resource for users of social research methods in applied settings* (3rd ed.). UK: John Wiley & Sons Ltd.

Rosales, A. (2015). Mathematizing: *An emergent Math Curriculum Approach for Young Children*. Redleaf Press: United States.

Rowland, T., Huckstep, P., & Thwaites, A. (2005). Elementary teachers' mathematics subject knowledge: The knowledge quartet and the case of Naomi. *Journal of Mathematics Teacher Education, 8*(3), 255-281.

Rowland, T. Turner, F., Thwaites, A., & Huckstep, P. (2009) *Developing Primary Mathematics Teaching, Reflecting on Practice with the Knowledge Quartet* (London, Sage).

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*, 119-144.

Sadler, D. R. (1998). *Letting students into the secret: Further steps in making criteria and standards work to improve learning.* Paper presented at the Annual Conference for State Review Panels and District Review Panel Chairs, July 1998.

Sahlberg, P. (2011*). Finnish Lessons: What Can The World Learn From Educational Change In Finland?* New York: Teachers College Press.

Salkind, N.J. (2011). Statistics for People who think they hate statistics. USA: Sage.

Saito, E. (2012). Strategies to promote lesson study in developing countries, *International Journal of Educational Management*, 26(6), 565–576.

Sarantakos, S. (1997) *Social Research*, 2nd edn, London: Palgrave Publishers

Schunk, D. (2012). Learning Theories: An educational perspective (6th edition). Pearson Education Inc.: USA.

Schunk, D. & Zimmerman, B. (2012). *Motivation and Self-Regulated Learning: Theory, Research and Applications.* Lawrence Erlbaum Associates: New York.

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.

Searle, J. R. (1995). *The construction of social reality*. London, UK: Allen Lane, The Penguin Press.

Sebba J, Crick RD, Yu, G, Lawson H, Harlen W, Durant K (2008) Systematic review of research evidence of the impact on students in secondary schools of self and peer assessment. Technical report. In: *Research Evidence in Education Library*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Shepard, L. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4-14.

Shepard, L. (2005). Formative assessment: Caveat emptor. ETS Invitational Conference 2005. *The Future of Assessment: Shaping Teaching and Learning* New York October 10-11, 2005.

Shepard, L.A. (2006). Classroom assessment. In R. L. Brennan's Education Measurement, Fourth Edition (Pp. 623-646). Westport, CT: Greenwood Publishing Group.

Shepard, L. A. (2008). Formative assessment: caveat emptor. In C. A. Dwyer (Ed.), The future of assessment: shaping teaching and learning (pp. 279-303). Mahwah, NJ: Lawrence Erlbaum Associates.

Shepard, L. A. (2009). Commentary. Evaluating the validity of formative and interim assessment. *Educational Measurement: Issues and Practice*, *28*(3), 32–37.

Shepard, L. A. (2011). Foreword. In Noyce, P. & Hickey, D.T. *New Frontiers in Formative Assessment* (pp. vii-xi). Cambridge, MA: Harvard Education Press.

Shepard, L. A. (2013). Foreword. In J. H. McMillan (Ed.), *SAGE Handbook of Research on Classroom Assessment* (pp.xix-xxii.). New York. SAGE.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4-14.

Shulman, L. (1987). Knowledge and teaching: Foundations of the New Reform. *Harvard Educational Review, 57* (1), 1-23.

Shulman, L. (2005). *The signature pedagogies of the professions of law, medicine, engineering, and the clergy: Potential lessons for the education of teachers.* Paper presented at the Math Science Partnership Workshop: Teacher education for effective teaching and learning, Irvine, CA.

Shute, V. (2008). Focus on Formative Feedback. *Review of Educational Research;* 78 (1), 153-189.

Shute, V. J., & Kim, Y. J. (2013). Formative and stealth assessment. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of Research on Educational Communications and Technology (4th Edition).* (pp. 311-323). New York, NY: Lawrence Erlbaum Associates, Taylor & Francis Group.

Sibbald, T. (2009). The Relationship Between Lesson Study and Self-Efficacy. *School Science and Mathematics*, 109(8), 450–460.

Sloane, F. (2005). Japanese Teachers' Lesson Study in Mathematics Education as an exemplar of Design Research. In S. Close, T. Dooley & D. Corcoran (Eds.), *First National Conference on Research on Mathematics Education (MEI 1)* (pp. 29-41). St. Patrick's College of Education, Drumcondra, Dublin 9.

Smith, C., & Gillespie, M. (2007). Research on professional development and teacher change: Implications for adult basic education. *Review of Adult Learning and Literacy, 7.* Retrieved June 29, 2016, from http://www.ncsall.net/fileadmin/resources /ann_rev/smith-gillespie-07.pdf

Stake, R. E. (2005). Qualitative case studies. In Y. N. K. Denzin (Ed.), *Handbook of qualitative research*. California: Sage.

Stankov, L., Lee, J., Luo, W. & Hogan, D. J. (2012). Confidence: A better predictor of academic achievement than self-efficacy, self-concept and anxiety? *Learning and Individual Differences*, 22(6), 747-758.

Stiggins, R. J. (2006). Assessment for learning: A key to motivation and achievement. *Edge*, 2(2), 3-19.

Stiggins, R. J. (2007). Assessment through the student's eyes. *Educational Leadership, 64(8), 22-26.*

Stiggins, R. (2009, February). Assessment for learning in upper elementary grades. Phi Delta Kappan, 90, 419-421.

Stiggins, R. & Chappuis, J. 2008, "Enhancing Student Learning: Create profound achievement gains through formative assessments", *The District Administrator,* January: <http://www.districtadministration.com/viewarticle.aspx?articleid=1362&p=1#0>.

Stigler, J. W., & Hiebert, J. (1999). *The teaching gap: best ideas from the world's teachers for improving education in the classroom*. New York: Free Press.

Stigler, J. W., & Hiebert, J. (2009). *The teaching gap: best ideas from the world's teachers for improving education in the classroom*. New York: Free Press.

Stobart, G. & Hopfenbeck, T. N. (2014). Assessment for learning and formative assessment. In Baird, J. A., Hopfenbeck, T. N., Newton, P., Stobart, G. & Steen-Utheim, A. T., (eds.) *Assessment and learning: State of the field review* (pp.30-50). Oxford: Oxford University Centre for Educational Assessment.

Sugrue, C. (Ed.). (2004). *Curriculm & Ideology: Irish Experiences, International Perspectives.* Dublin: The Liffey Press.

Sugrue, C. (2006). Irish teachers' experiences of professional learning: Implications for policy and practice. *Journal of In-Service Education, 28*(2), 311 - 338.

Sugrue, C. (2011). Irish teachers' experience of professional development: performative or transformative learning? *Professional Development in Education, 37 (5):793-815.*

Suratno, T. (2013). Promoting teaching as the lifelong learning profession through lesson study in Indonesia. *Journal of Southeast Asian Education*, 1, 25-48.

Sutton, R. (1995). *Assessment for learning*. Salford, UK: RS Publications.

Swaffield, S. (2011). Getting to the heart of authentic Assessment for Learning. *Assessment in Education: Principles, Policy and Practice* 18 (4) 433-449.

Tabachnick, B.G. & Fidell, L.S. (2013). Using Multivariate Statistics, 6th Edition. Pearson Education Limited: UK.

Tapia, M. M. (2004). An instrument to measure mathematics attitudes. *Academic Exchange Quarterly*, 8(2), 16-21.

Taras, M. (2007). Assessment for Learning: Understanding theory to improve practice. *Journal of Further and Higher Education*, 31(4), 363-371.

Teddlie, C., & Tashakkori, A. (2009*). Mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioural sciences.* California, United States: Sage.

Teddlie, C., & Tashakkori, A. (2010). Overview of contemporary issues in mixed methods research. In A. T. Tashakkori (Ed.), *Handbook of mixed methods in social and behavioural research (pp. 1-41).* California, United States of America: Sage.

The Teaching Council, (2011). The Teaching Council. (2011). *Policy on the Continuum of Teacher Education*. Kildare: The Teaching Council.

The Teaching Council (2016). Cosán: Framework for teachers' learning. Dublin: The Teaching Council.

Thompson, M., & Goe, L. (2009). *Models for effective and scalable teacher professional development* (Research report RR-09-07). Princeton, NJ: Educational Testing Service.

Thompson, M., & Wiliam, D. (2008). Tight but Loose: A Conceptual Framework for Scaling up School Reforms. In E. C. Wylie (Ed.), *Tight but loose: scaling up teacher professional development in diverse contexts* (Vol. RR-08-29, pp. 1-44). Princeton, NJ: Educational Testing Service.

Timperley, H S, 2011, Realising the power of professional learning, Maidenhead, Open University Press

Trends in International Mathematics and Science Study (TIMSS) (2007). Student Questionnaire Grade Four. Retrieved from: http://timss.bc.edu/timss2007/PDF/T07_Questionnaires_G4.pdf Published by International Association for the Evaluation of Educational Achievement © Copyright IEA, 2007

Topping, K.J. (2009) Peer Assessment. *Theory Into Practice*, 48(1), 20-27.

Topping, K.J. (2010). Peers as a source of Formative Assessment. In Andrade, H., & Cizek, G.J. (Eds.), *Handbook of formative Assessment* (pp. 61-74). New York: Routledge.

Travers, J., Balfe, T., Butler, C., Day, T., Dupont, M., McDaid, R., O'Donnell, M, & Prunty, A. (2010). *Addressing the Challenges and Barriers to Inclusion in Irish Schools.* Dublin: Special Education Department, St Patrick's College.

Tsui, A., & Law, D. (2007). Learning as boundary-crossing in school-university partnership. *Teaching and Teacher Education*, 23(8), 1289-1301.

Tunstall, P., & Gipps, C. (1996). Teacher feedback to young children in formative assessment: A typology. *British Educational Research Journal,* 22(4), 389-404.

Turner, R. (2012). Mathematical literacy – are we there yet? "Mathematical literacy: are we there yet?" ICME-12, Topic Study Group 6: Mathematics Literacy. Seoul, Korea. Jul. 2012. Available at: http://works.bepress.com/ross_turner/22

Vandevelde, S., Van Keer, H. & Rosseel, Y. (2013). Measuring the complexity of upper primary school children's self-regulated learning: A Multi-component approach. *Contemporary Educational Psychology,* 38, 407-425.

Van Driel, J. (2014). Professional development through lesson study: teaching the derivative using GeoGebra (Professional Learning of Science Teachers). In *Topics and Trends in Current Science education.* Springer: New York.

Verhoef, N., Tall, D., Coenders, F., & van Smaalen, D. (2014).The Complexities of a Lesson Study in a Dutch Situation: Mathematics Teacher Learning. International Journal of Science and Mathematics Education, Vol 12(4), pp.859-881.

Voerman, L., Meijer, P. C., Korthagen, F. A. J., & Simons, R. J. (2012). Types and frequencies of feedback interventions in classroom interaction in secondary education. Teaching and Teacher Education, 28(8), 1107–1115.

Vrugt, A., & Oort, F. J. (2008). Metacognition, achievement goals, study strategies and academic achievement. *Metacognition and Learning*, 30, 123–146.

Wainer, H. (2011). *Uneducated guesses: Using evidence to uncover misguided education policies*. Princeton, NJ: Princeton University Press.

Wake, G., Swan, M. & Foster, C. (2015) Professional learning through the collaborative design of problem-solving lessons. Journal of Mathematics Teacher Education, 1-18. ISSN 1386- 4416.

Waldron, Fionnuala, (2006) 'Children as Co-Researchers: Developing a Democratic Research Practice with Children', in *Research Methods in Social Studies Education: Contemporary Issues and Perspectives*, Keith C. Barton (ed.), Charlotte: Information Age Publishing

Wall, E. & Burke, D. (2007). The Standardised Irish Graded Mathematics Attainment Tests (SIGMA-T). Mary Immaculate College: Limerick.

Walsh, T. (2016) 100 Years of Primary Curriculum Development and Implementation in Ireland: A Tale Of A Swinging Pendulum, *Irish Educational Studies,* 35:1, 1-16.

Wang-Iverson, P., & Yoshida, M. (2005*). Building our Understanding of Lesson Study*. Research for Better Schools: USA.

Warne, R. T. (2014).Two Additional Suggested Reforms to Encourage Replication Studies in Educational Research *Educational Researcher*, 43(9), 465.

Warwick, P., Shaw, S. & Johnson, M. (2015). Assessment For Learning In International Contexts: Exploring Shared And Divergent Dimensions In Teacher Values And Practices. *The Curriculum Journal*, 26(1), 39-69.

Watanabe, T. (2002). Learning from Japanese lesson study. *Educational Leadership, 59* (6), 36-39.

Watanabe, T. (2011). A closer look: the role of outside specialists in Japanese lesson study in Lesson Study Step by Step, p.34—38 in Lewis & Hurd 2011. Heinemann: USA.

Watkins, C. (2003). *A sense-maker's guide*. London: Association of Teachers and Lecturers.

Wei, R. C., Darling-Hammond, L., Andree, A., Richardson, N., Orphanos, S. (2009). *Professional learning in the learning profession: A status report on teacher development in the United States and abroad.* Dallas, TX. National Staff Development Council.

Weißenrieder, J., Roesken-Winter, B., Schueler, S., Binner, E., & Blömeke, S. (2015). Scaling CPD through professional learning communities: development of teachers' self-efficacy in relation to collaboration. *ZDM Mathematics Education*, 47, 27–38.

Wenger, E. (1998). *Communities of Practice: Learning, Meaning and Identity*. NY: Cambridge University Press.

Wenger, E. (2015). Introduction to Communities of Practice. Retrieved from: http://wenger-trayner.com/introduction-to-communities-of-practice/

White, B. Y., & Frederiksen, J. R. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition and Instruction, 16*, 3-118.

White, B. (2009). Student perceptions of what makes good teaching. Refereed paper presented at *Teacher education crossing borders: Cultures, contexts, communities and curriculum* the annual conference of the Australian Teacher Education Association (ATEA), Albury, 28 June – 1 July.

Wiliam, D. (2005). Keeping learning on track: Formative assessment and the regulation of learning. In M. Coupland, J. Anderson, & T. Spencer (Eds.), *Making mathematics vital: Proceedings of the twentieth biennial conference of the Australian Association of Mathematics Teachers* (pp. 26–40). Adelaide, Australia: Australian Association of Mathematics Teachers.

Wiliam, D. (2006). *Assessment for Learning: why, what and how*. Presentation to Cambridge Assessment Network, Cambridge UK: Institute of Education, University of London.

Wiliam, D. (2007). Keeping learning on track: classroom assessment and the regulation of learning. In F. K. Lester Jr (Ed.), *Second handbook of mathematics teaching and learning* (pp. 1053-1098). Greenwich, CT: Information Age Publishing.

Wiliam, D. (2010). An integrative summary of the research literature and implications for a new theory of formative assessment. In H. L. Andrade & G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 18-40). New York, NY: Taylor & Francis.

Wiliam, D. (2010a). An Integrative Summary of the Research Literature and Implications for a New Theory of Formative Assessment. In H. L. Andrade & G. J. Cizek (Eds.), *Handbook of Formative Assessment* (pp. 18-40). New York, NY: Taylor & Francis.

Wiliam, D. (2010b). The role of formative assessment in effective learning environments pp. 135-160. In the Nature of Learning: using research to inspire practice. OECD: France.

Wiliam, D. (2010b). What counts as evidence of educational achievement? The role of constructs in the pursuit of equity in assessment. In A. Luke, J. Green & G. Kelly (Eds.), *What counts as evidence in educational settings? Rethinking equity, diversity and reform in the 21st century* (pp. 254-284). Washington, DC: American Educational Research Association.

Wiliam, D. (2011a). What is assessment for learning? *Studies in Educational Evaluation, 37*(1), 2-14.

Wiliam, D. (2011b). *Embedded Formative Assessment*. Bloomington, IN: Solution Tree.

Wiliam, D. (2013). Assessment: The Bridge between Teaching and Learning. *Voices from the Middle,* 21(2), 15-20.

Wiliam, D. (2014). Formative Assessment And Contingency In The Regulation Of Learning Processes. Paper presented in a Symposium entitled *Toward a Theory of Classroom Assessment as the Regulation of Learning* at the annual meeting of the American Educational Research Association, Philadelphia, PA.

Wiliam, D. (2016). *Leadership for Teacher Learning: Creating a Culture Where All Teachers Improve So That All Students Succeed.* Learning Sciences International, United States.

Wiliam, D. & Leahy, S. (2015). *Embedding Formative Assessment: Practical Techniques for K-12 Classrooms*. Learning Sciences International: USA.

Wiliam, D., & Thompson, M. (2007). Integrating assessment with instruction: What will it take to make it work? In C. A. Dwyer (Ed.), The future of assessment: Shaping teaching and learning (pp. 53–82). Mahwah, NJ: Erlbaum.

Willey, K., & Gardner, A. (2010). Investigating the capacity of self and peer assessment activities to engage students and promote learning. *European Journal of Engineering Education,* 35(4), 429–443.

Williams, J. A. (2010). 'You know what you've done right and what you've done wrong and *Assessment in Education: Principles, Policy & Practice*, 17(3), 301-314.

Williams, J., Ryan, J., & Morgan, S. (2014). Lesson Study in a Performative Culture. In McNamara, O., Murray, J., & Jones, M. (eds.). *Workplace Learning in Teacher Education: international practice and policy (*pp.151-167). UK: Springer.

Willis, J. (2007). Assessment for Learning-Why the theory needs the practice. *International Journal of Pedagogies and Learning*, 3(2), 52-59.

Willis, J. (2011). *Towards Learner Autonomy: An Assessment for Learning Approach.* Unpublished Doctoral thesis. Available from Queensland University of Technology: Australia.

Willis, J., & Cowie, B. (2014). Assessment as a Generative Dance: Connecting Teaching, Learning and Curriculum. In C. Wyatt-Smith, V. Klenowski, & P. Colbert (Eds.) *Designing Assessment for Quality Learning* (pp. 23-37). Netherlands: Springer.

Wilson, V. (1997). Focus groups: A useful qualitative method for educational research? *British Educational Research Journal*, 23(2), 209-224.

Wood, K. (2015). Deepening Learning through lesson and learning study. In Wood, K. & Saratha Sithamparam (Editors). *Realising Learning: Teachers' Professional Development Through Lesson and Learning study (pp.1-24)*. UK: Routledge.

Wood, P. & Cajkler, W. (2013). Beyond Communities of Practice: investigating and developing the professional learning of trainee teachers through Lesson Study.

Wylie, & Lyon, (2015). The fidelity of formative assessment implementation: issues of breadth and quality. *Assessment in Education Principles Policy and Practice,* 22(1) 140-160.

Wylie, E. C., Lyon, C. J., & Mavronikolas, E. (2008). *Effective and scalable teacher professional development: A report of the formative research and development*. (ETS Research Report No. RR-08-65). Princeton, NJ: ETS.

Xu, H., and Pedder, D. (2015) Lesson Study: an international review of the research, in Dudley, P (Ed.) Lesson Study: Professional Learning for our time, London, Routledge, pp. 29-58.

Yin, R. (2003) Case Study Research: Design and Methods. 3rd edn, California.: Sage Publications

Yin, R. K. (2009). *Case Study Research: Design and Methods* (4th ed.). USA: Sage.

Yin, R. K. (2014). *Case Study Research: Design and Methods* (5 ed.). USA: Sage.

Yin, Y., Shavelson, R. J., Ayala, C. C., Ruiz-Primo, M. A., Brandon, P., Furtak, E. M., Tomita, M.K., & Young, D.B. (2008). On the impact of formative assessment on student motivation, achievement, and conceptual change. *Applied Measurement in Education, 21*(4), 335-359.

Ylonen, A. & Norwich, B. (2012), "Using Lesson Study to develop teaching approaches for secondary school pupils with Moderate Learning Difficulties: teachers' concepts, attitudes and pedagogic strategies", *European Journal of Special Needs Education*, Vol. 27 No. 3, pp. 301-317.

Ylonen, A. & Norwich, B. (2015) Lesson Study as a means to help teachers of pupils with specific needs or difficulties to identify and overcome barriers to learning, in Dudley, P. (Ed.) *Lesson Study: Professional Learning for our time* London: Routledge.

Yoshida, M. (1999). *Lesson Study: An Ethnographic Investigation Of School-Based Teacher Development In Japan*. Unpublished doctoral dissertation, University of Chicago.

Yoshida, M. (2012). Mathematics lesson study in the United States: Current status and ideas for conducting high quality and effective lesson study. *International Journal for Lesson and Learning Studies,* 1(2), 140-152.

Yoshida, M. & Jackson, W.C. (2011). Ideas for Developing Mathematical Pedagogical Content Knowledge Through Lesson Study in In L. Hart, A. S. Alston & A. Murata (Eds.), *Lesson Study Research and Practice in Mathematics Education* (pp. 279-288). New York: Springer.

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13-39). San Diego: Academic Press.

Zimmerman, B., (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal,* 45(1), 166-183.

Zimmerman, B. J., & Schunk, D. H. (Eds.). (2001). *Self-regulated learning and academic achievement: theoretical perspectives* (2nd ed.). Mahwah, NJ: Erlbaum.