Lollipops and Traffic Lights

An investigation into the academic merits of introducing Assessment for Learning questioning methods at a selective, independent, girls' school

Darian R. Evans

The Cardiff School of Education
University of Wales Institute Cardiff

The dissertation is being submitted to the University of Wales in partial fulfilment of the requirements of candidature for the degree of MA Education Leadership and Management

October 2011
DECLARATION

This work is being submitted in partial fulfilment of the requirements for the degree MA Education Leadership and Management and has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed: 

Date 1st October 2011

STATEMENT 1

This dissertation is the result of my own work and investigations, except where otherwise stated. Where correction services have been used, the extent and nature of the correction is clearly marked in a footnote(s). Other sources are acknowledged by footnotes giving explicit references. A bibliography is appended.

Signed: 

Date: 1st October 2011

STATEMENT 2

I hereby give consent for my dissertation, if accepted, for photocopying and for inter-library loan, for deposit the title and summary may be available to outside o

Signed: 

Date: 1st October 2011

Approved by UWIC

Signed: 

Date: 1st October 2011
Abstract

This piece of research attempts to establish whether there would be academic merit in a medium-sized, independent, girls' school introducing AfL principles. This was done primarily by testing examination performance through experimentation, whilst also assessing whether the school’s academic environment would be suitable for the implementation of AfL strategies. Unlike other studies, such as Rowe (1974) and Black and Wiliam (2002), this research focuses entirely on older students completing the first year of their A Level studies. The main focus is on AfL questioning methods, and through a range of quantitative and qualitative data collection, the research examined Year 12 over a period of a year along with scrutinizing a small scale pilot scheme.

There were a number of significant findings. There was a strong and consistent correlation between AfL and AS Level performance when students were exposed to AfL questioning in a pre-test/post-test- non-equivalent group design; candidates’ AS UMS increased by up to 15 marks (1 ½ grades) across a range of students, when GSCE results and/or TSA baseline data were used to normalise the sample.

Notwithstanding the attainment improvements, the study also found that the school, with its present cultural environment, may not be one entirely conducive to the changes proposed by AfL. Firstly, staff and students’ opinions on questioning were diametrically opposed and some Sixth Form students preferred not to participate in lessons; many have become accepting and adept at disengaging in lessons. Secondly, a large number of students had an overwhelming preference for basic classroom questioning and were mostly averse to rich questioning which involves development, evaluation and synthesis. In addition to this, the questions students preferred, and those they found most useful, were consistently at odds and often polarised. Thirdly, the negative feelings expressed by students about class participation and the richness of questioning were linked to student ability, indicated by GCSE results.

Therefore, based purely on the quantitative experimental data, there was clear evidence to suggest that the school should launch AfL and fully embrace its principles across all year groups. However, the qualitative data were less clear and outlined some caveats concerning students’ largely negative feelings towards participation and inclusion. If care is not taken this could counteract any potential attainment improvement by adversely affecting students’ confidence, self-esteem and enjoyment of lessons. In order to circumvent some of these issues, there were several recommendations offered to the School’s Senior Management Team.
Contents

Declaration and Statements ................................................. ii
Abstract .................................................................................. iii
Contents .................................................................................. iv-v
Table and Figures ...................................................................... vi
Acronyms .................................................................................. vi

Chapter 1: Introduction

1.1 Research Context .......................................................... 1
1.2 Research Statement .......................................................... 3
  1.2.1 The Role of Questioning ................................................. 5
  1.2.2 Student Participation ...................................................... 7
  1.2.3 The Use of ‘Wrong’ Answers ......................................... 8
  1.2.4 Rich Questioning ......................................................... 8
  1.2.5 Wait Time ................................................................. 9
1.3 Background and Rationale ................................................ 10
1.4 Research Questions ......................................................... 15

Chapter 2: Literature Review

2.1 An Introduction to Assessment for Learning (AfL) .............. 16
2.2 The Importance of Effective Questioning ......................... 20
2.3 Student Participation ....................................................... 25
2.4 The Use of ‘Wrong’ Answers .......................................... 27
2.5 Rich Questioning ........................................................... 29
2.6 Wait Time ................................................................. 30
2.7 Criticism of Assessment for Learning ............................... 33

Chapter 3: Methodology

3.1 Introduction ......................................................................... 35
3.2 Research Methods ........................................................... 39
  3.2.1 Sampling ...................................................................... 39
  3.2.2 Questionnaires ........................................................... 39
  3.2.3 Interviews .................................................................... 41
  3.2.4 Experimentation ......................................................... 42

3.3 Reliability and Validity ....................................................... 44
3.4 Triangulation .................................................................... 45
3.5 Ethical Considerations ....................................................... 46
3.6 Evaluation of Methodology .............................................. 48
3.7 Research Time Line .......................................................... 49
Acronyms

AfL  Assessment for Learning
ALIS  Advanced Level Information System
AS/A Level  Advanced Subsidiary/Advanced Level
BERA  British Educational Research Association
BBC  British Broadcasting Corporation
CEM  Curriculum, Education & Management Centre
DCELLS  Department for Children, Education, Lifelong Learning and Skills
DfES  Department for Education and Skills
ESRC  The Economic and Social Research Council’s
GCSE  General Certificate of Secondary Education
GSA  The Girls School Association
HMC  Headmasters' and Headmistresses' Conference
INSET  In Service Training
KMOFAP  King’s-Medway-Oxfordshire Formative Assessment Project
KS3/KS4  Key Stage 3/4
NQT  Newly Qualified Teacher
OFSTED  The Office for Standards in Education, Children's Services and Skills
SD  Standard Deviation
SMT  Senior Management Team
TDA  Test of Developed Ability
TES  Times Educational Supplement
TLRP  Teaching and Learning Research Programme
UMS  Uniform Mark Scale

Tables and Figures

Figure 1: CEM Added Value Residuals 2002-2010 .................................................. 12
Figure 2: Experimental Design for RQ3 ................................................................. 43
Figure 3: Current Use of AfL Questioning Techniques (n=80) ............................. 53
Figure 4: AfL Questioning – Student Preferences (n=30) .................................... 59
Figure 5: Preferred vs. Actual Questioning Experience (n=80) ............................ 64
Figure 6: Question Type: Differential – Preferred/ Worthwhile (n=80) .............. 69
Figure 7: Staff Perception of AfL Benefits (n=30) ............................................... 81
Figure 8: Student Perception of AfL Benefits (n=15) .......................................... 81

Table 1: Students 1 & 2 GCSE Normalised (n=30) ........................................... 76
Table 2: Students 3&4/5&6 GCSE Normalised (n=30) ........................................ 77
Table 3: Students 7&8 GCSE Reversal (n=30) ...................................................... 78
Table 4: Students 9 & 10, 11 & 12 TDA Normalised (n=30) ............................ 78
Table 5: Group Comparisons GCSE Normalised (n=30) ................................... 79

Figure A1: Impact of Changes (n=80/n=82) ....................................................... 97
Figure A2: Knowledge of AfL (n=80) ................................................................. 97
Figure A3: Student Experience of Questioning (n=80) ........................................ 98
Figure A4: Current Use of AfL Questioning Techniques (n=30) ........................ 98
Figure A5: Desirable Wait Time – Staff vs. Students (n=80/n=30) .................... 99
Chapter 1: Introduction

1.1 Research Context

Assessment for Learning (AfL) has been *de rigueur* in the state sector for over a decade and has been discussed in academic circles for at least the past two – Black and Wiliam (1998, 2002), Brown and Wragg (1993, 2001), Hargreaves (2004). Since the turn of the century there has been extensive take up of AfL at primary and secondary schools in the maintained sector in England and Wales, with its pedagogical benefits being championed by both national and local governments. Assessment *per se* has generally become more sophisticated in this time and there have been a large number of initiatives relating to formative assessment, particularly in the maintained sector.

However, despite its fairly recent popularity, many schools in the independent sector have not taken up AfL and have, instead, focused even more intensively on summative assessment in the wake of increased parental pressure, league table positions and the pursuit of ever more elusive university places. According to a spokesperson for the Headmasters’ Conference of Schools (HMC) which represents over 250 schools, ‘we are certainly aware of the work of the Assessment Reform Group’ Evans (2012); they also went on to say that there have been ‘discussions of AfL at HMC’s Academic Policy Sub-Committee’. So, although they are clearly not averse to AfL - apparently having had talks by Dylan Wiliam and Gordon Stobart at their Annual Conference (Bournemouth 2007) - if HMC is reflective of the independent sector, then they would appear to be lagging behind the maintained sector by some distance.
On the whole, the take up of AfL principles has been widespread outside the independent sector in Wales, particularly in south Wales Secondary Schools. The Welsh Assembly Government describes formative assessment on their website as ‘a continuous process of assessment for learning which examines the learner’s attainment which feeds into future planning strategies for improving their learning outside of level descriptors and curriculum targets’ (http://wales.gov.uk, no date). In contrast, they class summative assessment as ‘the statutory assessment of learning which assesses and records learners’ work at the end of a key stage, thereby monitoring their overall attainment’ (DCELLS Ensuring consistency in teacher assessment, January 2008: 4).

Black and Wiliam (1998:2) explain the distinction between summative and formative assessment in their seminal paper *Inside the Black Box* as follows:

The term assessment refers to all those activities undertaken by teachers, and by their students in assessing themselves, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged. Such assessment becomes formative assessment when the evidence is actually used to adapt the teaching work to meet the needs of the student.
1.2 Research Statement

Essentially, this piece of research attempted to establish whether there would be academic and wider pedagogical benefits in introducing AfL as a whole school policy approach. It attempts to do this by firstly examining the current state of questioning in the classroom, including both staff and students’ approaches and attitudes. Secondly, it will explore the learning preferences of students and will explore how AfL has been received so far. Thirdly, it will build on the students’ experiences hitherto and will determine what academic benefits can be achieved by using AfL questioning. It will do this by scrutinizing examination data and by also seeking staff and students’ opinions.

The small pilot scheme used in this study included a small number of A level departments between 2010-11. In order to cover subjects traditionally considered to be Arts or Sciences, the following A Level departments were approached about introducing AfL: Economics, Modern Foreign Languages, English, Biology and Psychology.

The research findings are intended to help the Director of Teaching and Learning, along with other members of the SMT, to examine the practical hurdles involved with introducing AfL before a ‘school-wide’ take up. Quantitative data will be generated through testing and examination; qualitative data will also be generated with the reactions of students, parents and staff being considered.
The latter will be of particular importance due to the pressure of implementing process changes in this largely product-driven environment. AfL would therefore not only involve changes in how students are taught and assessed in lessons, but could also signal a cultural shift away from the competitive environment of results and performance (something seemingly at odds with the extremely high expectations of fee paying parents).

Summarised below are the four main principles of AfL, as suggested by Black and Wiliam et al (2003:30). The research will consider the application of the first principle, the role of questioning, in the context of Sixth Form teaching.

1. Questioning
2. Feedback
3. Sharing Criteria
4. Self-assessment

The role of questioning is an extremely important principle that has the potential to not only increase academic performance, but also to fundamentally change the teaching and learning experience of students in the classroom.
1.2.1 The Role of Questioning

One of the most important aspects of AfL is that of questioning in the classroom (Black and Wiliam 1998, 2002). Through asking questions in the classroom the teacher will attempt to create the correct classroom atmosphere whilst also trying to assess the students’ knowledge and understanding; questions should also be concerned with developing the students’ thinking skills – something particularly important in this piece of research. However, how much of this is currently undertaken in the School’s AS/A Level lessons?

The interaction achieved between students and teachers during a questioning session is therefore crucial for learning and, whilst the reasons for the importance of the use of questioning can be deduced from the above functions of questions, a crucial aspect will be the technique used for asking the questions and for eliciting responses from the students (these are explained in more detail in the Methodology). Questioning is one of the main tasks of a teacher and most non-educationalists would assume that formulating and asking questions would be second nature. However, how many lesson plans contain a list of the precise questions to be asked in that lesson i.e. how many teachers prepare questions in advance of the lesson? How often do teachers really think about how to phrase a question to elicit the best response, or think about what the question is supposed to achieve?

It is all too easy for busy, yet experienced teachers to leave questions to instinct and some will, of course, be naturally situational; generating questions ‘organically’ as the lesson progresses, is integral to an effective lesson and many questions that are asked in this way simply cannot be formulated beforehand.
However, the theory of AfL advocates that the teacher should ‘bridge the gap’ between where the student is now and where they need to be. For this to happen it is important that questions are thoughtful and structured, and should be an integral part of the lesson planning.

Black et al. (2003) argue that effective questioning is crucial for advancing student understanding. For AfL to be successful, the teacher must share criteria with the students; they indicate that the students cannot be expected to advance to the next stage of their learning if they do not know what that next stage involves. If a student is to take ‘responsibility for their own learning and advancement’, they must be able to see how a certain question will enable them to make that leap.

Black, at King’s College, headed a team which has carried out research and subsequently published several works on AfL. In their research, questioning was one of the areas they cited as most important, to both the teacher and the student. They identified where students are now, where they need to be and how they are going to get there (an area identified in the School’s Inspection Report). In Assessment for Learning: Putting it into Practice (2003:32-42), Black et al. identify four aspects of questioning which need to be considered if questioning is to help with the formative assessment process and these will be the areas of primary focus in this investigation:

1. Student Participation
2. The Use of ‘Wrong’ Answers
3. Rich Questioning
4. Wait Time
1.2.2 Student Participation

Student participation refers to the situation where the whole class is engaged and active in the lesson at all times, rather than just the keen minority; the principle is that if students are targeted randomly they will have no choice but to be included. Black et al. (2003) provide several examples of lessons in which certain students did not participate at all. This could of course be for a number of reasons - laziness, lack of interest, inability to answer or lack of confidence. However, whatever the reason, it is the teacher’s responsibility to include all students in any question session.

Some popular AfL techniques were used in the pilot. Firstly, use of random questioning and the ‘no hands up’ rule. This was recently featured in Dylan William’s 2010 BBC two-part documentary ‘The Classroom Experiment’, where the teacher used lollipop sticks; each student has their own ‘stick’ which is held by the teacher and pulled out at random to determine who answers each question. Secondly, use of a traffic light system, which requires students to continually monitor their understanding of topics being covered, or individual questions asked of them; the student will display green if they fully understand; amber if they are becoming unsure; and red if they no longer understand.
1.2.3 The Use of ‘Wrong’ Answers

When a student offers a ‘wrong’ answer, the teacher normally tells the student that it is incorrect and quickly moves on to the ‘right’ one, either by providing it themselves or asking someone else (usually a student that they know can offer the correct answer).

It is very rare for an answer to be so incorrect that some additional learning could not take place as a result of it. A discussion can develop with the teacher not only asking for reasons why the answer was insufficient, but also asking other students how they might improve it. Therefore, by using the term ‘wrong’, the teacher (and the class) can augment the understanding, not only of the student who initially answers, but also of other members of the group. However, care must be taken here as there is nothing more embarrassing and demoralising for a student than being told they are ‘wrong’.

1.2.4 Rich Questioning

A rich question is an open-ended, higher order question that requires learners to link or apply ideas, give reasons, summarise or evaluate. Sometimes the questions encourage learners to ask further questions to qualify what the question is actually asking them to explain. These questions generally require extended answers; they should prompt a range of possible answers from a number of learners, which in turn raise a list of smaller questions that need to be answered before the answer to the ‘big’ questions can be formulated i.e. those that require a more sophisticated and refined intellectual engagement with the subject matter.
All too often, a question is merely a test of a student’s memory and, due to the large body of knowledge required for A level, it is all too easy to allow students to recall facts by asking questions requiring knowledge acquisition, rather than those reflecting understanding. In contrast, rich questioning requires students to use information they already have to draw conclusions and inferences and to further enhance their knowledge using logic and sequential strategies.

Rich questioning can, therefore, inform the teacher and student, helping the latter to understand how to use them and guiding them towards the next stage in their understanding. A student is unlikely to benefit from or ‘advance his learning’ as a result of a question which simply requires recall of fact; the situation is even less productive if the student does not know that particular fact.

1.2.5 Wait Time

The research of Black et al. (2003:32) also shows that the average amount of time a teacher gives a student to answer a question is between 6 and 9 seconds. After this point, the teacher moves on to another student, rephrases the question or answers the question himself. If the aforementioned rich questioning is to be used i.e. questions that set in motion a train of thought, the students must be given sufficient time for that thought process to feed through. Failure to do so will leave students feeling inadequate, or not as ‘clever’ as the student who gave the answer. Naturally, this is not a matter of intelligence, but rather a matter of understanding the different thought processes which may occur in different students who have different learning styles.
As Dylan Wiliam proposes in his BBC documentary “smart is not what you are...smart is what you get” Dylan William (The Classroom Experiment, 2010). Proponents of AfL such as Paul Black and Dylan Wiliam suggest that by allowing all students a certain amount of thinking time, their confidence can be boosted as they realise that they are capable of answering the questions. This will, in turn, encourage students to participate in the future in the knowledge that they are able – spawning a ‘virtuous cycle’ of performance and confidence.

Again, there are various ways in which this wait time can be lengthened but, without adopting a strategy to increase wait time, the teacher cannot possibly know whether students did not know the answer, or whether they were simply not given sufficient thinking time.

1.3 Background and Rationale

This small piece of research focused on a secondary school in south Wales. The school is a medium-sized independent girls’ school with just over 600 students, aged from 7-18. It is part of the GSA group and is both fee-paying and selective.

For many years, the school has achieved excellent GCSE and A Level results and the result of intensive coaching for GCSE has been a proliferation of A* grades. The pressure to perform at GCSE and A Level seems to be amplified in the independent sector where there is pressure from all stakeholders to achieve excellent results.
This pressure not only affects the students, but also the teaching staff, middle managers, through to the senior management team; the pressure will also be felt by external stakeholders, such as the governing body and parents.

At this Girls’ School this issue is often debated, with staff members often referring to the school as an ‘exam factory’. As Satterly (1981: 34) points out, summative assessment ‘favours the middle class and the majority view’; he also states that it limited thinking skills in learners as they learn tricks to ‘pass exams in a very mechanical fashion’.

The need to move away from this teacher-centered, spoon-feeding environment is known all too well by staff and senior managers, but there are conflicts with the need to produce excellent final examination results. The move could be seen by many stakeholders as a high risk strategy due to the perceived link between the teacher-led, product-based environment and examination performance. In terms of both examination results and university entry, as with many independent schools, the environment is one where expectations are high. There is also an increasing need, faced with strong local competition and an uncertain economic climate, for the school to achieve extremely high GCSE and A Level results, thereby ensuring healthy numbers of students continue to enter the school.

Figure 1 overleaf shows the School’s performance in adding value (attainment) from GCSE to A level over the past nine years. The added value residual across all A Level subjects has been at best low, but has actually been largely negative. Although 2010 saw an improvement, the added value is still worrying at A Level – although the scope for improvement could be hampered somewhat by high performance at GCSE Level – the school colloquially being ‘hoist by its own petard’.
Another reason for an overhaul in assessment in all year groups was the verdict reached by inspectors (Estyn, 2009). One of the inspection responses was to Key Question 2; the inspection team found that there were ‘high levels of subject knowledge’ and ‘a keen awareness of how best to use this knowledge to support students’ learning’; they also found that teachers’ ability, through their own knowledge, enthusiasm or highly effective demonstrations, ‘engaged students well’. However, there were several shortcomings that prevented the school from achieving the top grade for this question. A summary of the points raised, pertinent to assessment and learning, are as follows:

- ‘throughout the school, the quality of marking varies and does not always include constructive comments that help students know what they must do to improve’

- ‘the pace of the lesson reduces the quality of students’ learning, for example when there is a slow start to the lesson or lack of clear time limits for particular activities’
• 'there is a lack of challenge so that students do not extend their learning enough’

• 'tasks are overly prescriptive and do not provide enough opportunities for students to learn independently’

• 'there is overuse of closed questions that neither extend students’ thinking nor reinforce learning’

School Inspection Report (Estyn), December 2009

In order to address the above issues, the school’s Director of Teaching and Learning is, after exploring a range of possibilities, considering the introduction of Assessment for Learning (AfL). It would be particularly advantageous to introduce this new initiative now as the school has recently had an influx of new staff, a great many of whom are NQTs; the new staff have already shown interest in AfL and also have experience of its principles, having worked with them in the state sector for many years; in fact, many new staff were surprised that the school had not already adopted them.

The school’s SMT knows that change is needed and that any policy should be global; however, the Headmistress has decided to firstly pilot AfL in a small number of A Level departments before considering the launch of AfL across the school. As the Director of Teaching and Learning has pointed out to staff: “AfL will only bring about improvement if it is underpinned by a whole-school belief that we can and must improve no matter how good we are now – you don’t have to be ill to get better”.

13
The school is therefore considering AfL, which would move it from delivering a largely *product-driven* curriculum, to one focused more on *process*. As Avis, Fisher and Thompson (2010:107) point out, citing Kelly (2009), 'curriculum as process must focus more on the process of learning rather than the outcome…the art of the teacher is to take the position not of an expert, but to be cast in the role of the learner'.

This large pedagogical change being considered by SMT, should be focused around four simple questions, first offered by Tyler (1949), cited by Avis *et al.* (2010: 107-108). This should allow SMT to establish the grounds for change (through consideration of the first three questions).

Once presented with the right evidence, the management team could then consider the last question relating to logistics and planning:

1. What educational purposes should the school seek to attain?
2. What educational experiences can be provided that are likely to attain these purposes?
3. How can we determine whether these purposes are being attained?
4. How can these educational experiences be effectively organised?
1.4 Research Questions

There were initially four research questions, developed after consideration of the rationale and research statement. Over time, these were reduced to just three, which were progressively honed over the research process, including a retrospective review after development of the methodology. The final research questions are as follows:

1. What is the current standard of classroom questioning used in the Sixth Form?

2. To what extent would AfL questioning methods suit the students’ teaching and learning preferences?

3. What appear to be the academic benefits generated by the AfL questioning pilot scheme?
Chapter 2: Literature Review

2.1 An Introduction to Assessment for Learning (AfL)

When approaching the literature there needed to be a clear focus on questioning and, more broadly, the way in which AfL has been embraced nationally by academic staff and implemented across secondary schools in England and Wales. There were extensive resources – by local authorities or independent, profit making organisations – on how teachers can best implement strategies related to formative assessment. However, the academic literature on AfL was surprisingly narrow. There are a few notable providers of research in this area, such as George Brown and Ted Wragg, along with one of the best-known educational writers, David Hargreaves. However, most textbooks and journal papers that have been developed on AfL are dominated by academics from Kings College, London.

By far the most prolific contributors to AfL are Professor Paul Black and Professor Dylan Wiliam. They have carried out an extensive research on AfL over the past decade and wrote the seminal paper, *Inside the Black Box* (1998). This revolutionised the way in which formative assessment was viewed and, moreover, their work pioneered many of the changes that have been made ever since in the maintained sector, in both England and Wales. *Inside the Black Box,* was effectively a review that drew upon 250 research journals and publications between 1988 and 1997.
The journey of AfL in its current form began in 1988, when a government task force formally underlined the significance of formative assessment in teaching and learning. The Assessment Task Group (now known as the Assessment Reform Group) asked Black and Wiliam to review the literature on formative assessment and gave the initial impetus for the King’s-Medway-Oxfordshire Formative Assessment Project (KMOFAP) project. This began in 1999 and involved schools in Oxfordshire and Medway LEAs, including forty eight teachers.

The findings of this project are captured in the aforementioned paper’s sequel, Working Inside the Black Box (2002). Black and Wiliam’s research spanned all age groups (from 5-year-olds to university graduates), subjects and nationalities. However, the main evidence generated via the KMOFAP project only illustrated learning gains in students up to Year 10 and there is no reference to empirical studies of Sixth Form students taking A Levels.

In Inside the Black Box, Black and Wiliam (1998: 5) criticised government policy as placing too much emphasis on ‘high-stakes external tests’ that do little to promote AfL, providing only ‘overall summaries of achievement rather than helpful diagnosis’. Black and Wiliam also agreed with the government’s own findings that formative assessment was ‘seriously in need of development’. Black and Wiliam likened the classroom to a ‘black box’, with government initiatives focusing on the box’s input and output, but not what went on inside it; they became concerned that formative assessment was increasingly sidelined in favour of summative assessment – the recording of data for certification and evaluation.
Black et al. (2003: 2) define AfL as 'any assessment for which the first priority is to serve the purpose of promoting students’ learning... it is usually informal, embedded in all aspects of teaching and learning, and conducted by different teachers as part of their own diverse and individual teaching'.

The Assessment Reform Group define AfL as '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' (Assessment Reform Group, 2002). With AfL, feedback is therefore used to modify learning activity - this might take the form of teachers giving comments on how a student can improve their work instead of awarding grades, or students assessing themselves and determining where their weaknesses are. It is based on the idea that all students can improve, and fosters a spirit of cooperation in the classroom, with students being actively involved in their own learning.

Black and Wiliam (1998) had demonstrated, via a wide-ranging review of the literature, that such an approach could improve both learning and exam results. However, they and their colleagues at King's realised that they needed to provide practical examples of its use if AfL was to be widely adopted and embraced by UK schools – enter the KMOFAP project. Their findings were that AfL not only improves student performance, but also advances the learning and understanding of students. Moreover, their research showed that implementation of AfL raises GCSE grades by an average of 1.5. Although this was based on younger students in their first year of GCSE, the Black and Wiliam study provided an interesting benchmark and helped to forge one of this study’s research questions, focusing on academic performance.
Since the publication of *Inside the Black Box*, a large of amount of intellectual debate in education has ensued, focusing on the benefits of AfL and whether it can indeed raise standards in the classroom. As Weeden, Winter and Broadfoot (2002: 75) suggest, the paradigm in which AfL operates is really where 'emphasis and responsibility [is] being placed on the learner for their own learning’. Another commentator, Clarke (2005: 3), suggests that it is about the ‘involvement of students in the learning process, beyond anything traditional teaching has previously allowed’. An alternative explanation of the principles of AfL is provided by James (1998:185). She suggested that formative assessment differs from summative assessment in three key ways:

- It is a central part of teaching which has positive aims and is continuous, whereas Summative is at particular intervals / at the end of a specific stage
- It is not simply criterion-referenced but incorporates the needs of individual learners
- Formative identifies inconsistencies in learning and uses them to inform future planning, whereas Summative may identify them as ‘faults/errors’
2.2 The Importance of Effective Questioning

The literature referenced below is focused specifically on classroom questioning. After a general examination of questioning was carried out, it was then crucial to examine more closely, each of the questioning elements outlined earlier: student participation, the use of ‘wrong’ answers, rich questioning and wait time.

Questioning is fundamental to helping students acquire knowledge at a basic level, or at a more advanced level, to encourage higher-order thinking skills and problem solving. All too often, questions posed by teachers merely require factual answers or are simply closed (requiring one word answers). An article in the TES (October 2003) also criticised classroom question sessions which can simply be ‘an invigorating rapid-fire exchange of recalled facts from previous lessons’.

Research has revealed that the total number of questions asked by an average teacher in their [working] lifetime is about two million (Brown and Wragg, 2001:32); by contrast, another study showed that each student in the classroom asks, on average, just one question per month (Dillon, 1988: 7); if extrapolated, and normalised for comparison, this would be approximately 360 questions over the same period. So, given the amount of time teachers spend asking questions, it is vitally important that those questions advance the students’ learning and provide teachers with important information about their students’ understanding to enable them to plan for future learning.
Good classroom dialogue should certainly allow the teacher to identify misconceptions and gaps in the students’ knowledge and when students learn that a simple one word answer is insufficient the ‘why’ will become embedded and will subsequently be transposed to their written work.

If what Black and Wiliam (1999) suggest is correct, then students should come to realise that learning depends less on their capacity to spot the right answer and more on their readiness to express and discuss their own understanding.

Often, the teacher is looking for a specific answer and, when this does not seem forthcoming, the teacher ‘lacks the flexibility or the confidence to deal with the unexpected’ (Black and Wiliam, 1998: 11). In regard to questioning and dialogue, teachers found that if they allowed more time for students to reply to their questions, more thoughtful answers were given. They also found that lessons became richer if they changed the focus of questions from testing students’ knowledge of facts to exploring their understanding. Once again, the assertion that questioning boosts understanding was crucial in determining whether this study should explore the same causation.

In 2008, an OFSTED report fed back on its AFL National Strategy support which, since 2006, moved from AFL across the school to focus instead on the core subjects of English and Mathematics. The findings of this report are based on inspections of assessment in English or Mathematics in 27 primary and 16 secondary schools between April and December 2007.
The effectiveness of AfL in the schools was mixed. A summary of the key finding follows:

- Discussion and well-focused questioning allow them to explore what they are learning and to try out ideas.
- Questions required only brief and usually factual answers without explanation or justification. Teachers often asked the question to a whole class and then selected a volunteer to answer.
- Teachers posed relatively closed questions or tasks which generated generally similar responses and, therefore, limited the potential for discussing the answers.


David Hargreaves (2004: 25) also criticises most question and answer sessions that occur across classrooms in the UK as being ‘oral checks on what students already know’. He goes on to imply that the teacher has a ‘predetermined answer’ in his head and the point of the exercise is to guess what that answer is. However, Hargreaves’ study did not include A Level students, focusing only on younger students where lessons should contain fewer high order questions that simply test recall - wouldn’t A Level students expect better?

Black and Wiliam (1998: 11) also express this opinion by suggesting that ‘over time the pupils get the message – they are not required to think out their own answers…the object of the exercise is to work out, or guess, what answers the teacher expects to see or hear, and then express it so that the teaching can proceed’.
Hargreaves' research flagged up some interesting ideas which were used in the design of the first research question – but just how prevalent would the use of closed questions, that only test memory, be in the Sixth Form?

Brown and Wragg (1993: 18) believe that teachers do not spend enough time preparing questions in advance and provide the following points as some of the most common errors made by teachers when questioning their students:

- Asking a question and answering it yourself
- Asking questions only of the brightest or most likeable
- Always asking the same type of questions
- Not using probing questions
- Not giving pupils the time to think
- Not correcting wrong answers
- Failing to build on answers

As already indicated, there are many reasons for asking questions, but one of the issues with traditional questioning is that it sometimes misses the target. As Wiliam suggested (1999:16) ‘by asking questions, teachers try to establish whether students have understood what they are meant to be learning. If students answer the question correctly, it is tempting to assume that the students’ conceptions match those of the teacher’.
This is not always so; they may give the right answer for the wrong reason. Lin and Mackay (2004: 5) also considered the use of effective questioning and found that in order to achieve the National Curriculum thinking skills requirements, students should 'know how', as well as 'know what'. Their conclusions were that for a question to have diagnostic value it should 'take the lid off how children think and learn, and free teachers to make informed choices about how and when they might best intervene to take learning forward.'

Kyriacou (1997: 39) concludes that questioning, by both teachers and students, should place emphasis on 'describing and explaining new information through direct teacher-pupil interaction'. His research links to Afl questioning as he goes on to suggest that the role of questioning in Afl should 'utilise listening and thinking...when students respond to what the teacher asks then discussion should be used to facilitate and explore pupil learning'.

Kyriacou develops the process of questioning by classifying questions as 'lower order', that simply require recall of facts and information, and 'higher order', that require some manipulation of the information provided. In agreement with Brown and Wragg (1993), Kyriacou (1997) also indicates that teachers use lower order questions too frequently, but advocates the importance of using both.

Kyriacou's findings were also essential in deciding how the methodology should generate the evidence for the second research question i.e. those relating to student questioning preferences which therefore formed an integral part of this study.
There are clear links to Bloom's Taxonomy of Educational Objectives here, Bloom *et al* (1956); higher order thinking should be used consistently at A Level, where students need to hone skills of evaluation and synthesis. However, how widespread would the use of such questions be in the Sixth Form?

2.3 Student Participation

One of the adaptations to teaching in the KMOFAP project and in Dylan Wiliam's *BBC documentary* 'The Classroom Experiment', was the use of random questioning – more specifically the 'no hands up' rule. In the latter, teachers used random questioning by using lollipop sticks with each 'stick' featuring a student's name. This was then pulled out in order to randomise who answers each question. The 'no hands up' rule attempts to make lessons more inclusive and to encourage whole class participation - bringing in students who would have normally been disengaged from the lesson, whilst limiting the domination of the class by the keenest and brightest students. This is certainly an issue in the highly competitive environment of a selective, independent girls' school.

The introduction of this method produced some interesting results. At first, both teachers and students felt incredibly uncomfortable; both parties had to deal with a completely new type of classroom organisation and also how to respond to replies that may not be right, whilst also having to deal with students who simply didn’t want to reply. This was echoed in the National Strategy document, *Pedagogy and Practice (Unit 7: Questioning)*, that indicates that when teachers introduce strategies such as no hands up they may feel 'self-conscious and students may find the approach unusual'.
However interestingly, over time, teachers developed their own way of managing the no hands up rule and students had to deal with either the discomfort of being in the spotlight, as opposed to being previously hidden, or no longer being centre of attention.

What the KMOFAP study found, along with the research undertaken at Hertswood School for the BBC documentary, was that the quality of the feedback becomes a critical part of the no hands up system of participation. As Black (2009: 3-5) points out ‘a pupil’s answer to a question can reveal how the pupil understands the issue, and the teacher can then respond to help develop that understanding’. The BBC documentary not only reiterated Black and Wiliam’s assertion that AfL boosts performance, but also provided some AfL techniques that would be used in the pilot - lollipops and traffic lights.
2.4 The Use of ‘Wrong’ Answers

This is another widely discussed strategy and one which naturally lends itself to the three other questioning methods. As Black (2009:2) puts it:

In composing a useful response, the teacher has to interpret the thinking and the motivation that led the pupil to express that answer. It helps if the teacher first asks the pupil to explain how he or she arrived at that answer, then accepts any explanation without comment and ask others what they think. This gives value to the first answer, and draws the class into a shared exploration of the issue.

The key is to have a discussion where it is explicitly stated that getting things wrong is a fundamental part of the learning process and where the teacher makes it clear that how a student answers a question demonstrates how much has been understood and what the teacher needs to do next. As Black et al. (2003: 39) point out:

the aim is not for discrete right answers to be celebrated, but for a discussion of the ideas to be explored...questioning is used to elicit student understanding and promote shared learning. An answer which demonstrates a misconception can often provide the opportunity to extend learning both for the student and for others in the class who may well share the same misconception.

However, using incorrect answers well is one of the most onerous elements to effective questioning. As Kyriacou (1997: 44), outlines, feedback is the ‘final skill’ to be honed and is critical, as ‘answering questions is often a high-risk strategy and an emotionally charged activity, in part because it is usually public and in part because it usually involves explicit teacher judgement’.

27
Brown and Wragg (1993:22) also agree that feedback is difficult, yet crucial to questioning and initiating dialogue: ‘responding moves are, in a sense, the lynch pins of a lesson. They are important, therefore, in sequencing and structuring a lesson’.

Marland (2002: 108) believes that a teacher’s response should always be ‘clear, colourful, and tactful’. His advice to teachers on this is as follows:

> despite your care, ‘wrong’ answers will frequently be given. Never mind. Use them for what they are – next steps to further thought – and take up the answer as a challenge for you to devise rapidly the next appropriate question…it will be a guide to the pupil’s misunderstanding. You need to deduce which route the pupil took to the wrong answer, and use that deduction to help him see for himself the better alternative.

Research undertaken by Brown and Wragg, and Black and Wiliam, was particularly useful and helped to phrase the questions in my questionnaire, as well as my interviews. The shared exploration of answers and the careful use of feedback by the teacher came across as vitally important in the literature, so it was thought critical that this was examined in the context of Sixth Form teaching – just how prevalent would the use of wrong answers be here?
2.5 Rich Questioning

Another key feature to effective questioning is the use of ‘rich questioning’. The Department for Education’s National Strategies defines a rich question as an ‘open-ended, higher order question that requires learners to link or apply ideas, give reasons, summarise or evaluate’. Sometimes these questions encourage learners to ask further questions to qualify what the question is actually asking them to explain and therefore often require extended answers’.

Black and Wiliam (1998: 7) uphold that ‘opportunities for pupils to express their understanding should be designed into any piece of teaching, for this will initiate the interaction through which formative assessment aids learning’. Closed questions that require factual recall are therefore unlikely to enable a pupil to articulate her understanding; this notion was also flagged up in research evidence provided by the ESRC’s Learning How to Learn project which indicated that ‘good questions are those that encourage learners to make explicit their thinking, not just to give ‘right’ answers’.

Kyriacou (1997: 44) believes that there are four key aspects inherent in effective questioning, with one of the most critical being the quality of the question itself ‘in terms of clarity and appropriateness for meeting its intended function, quality is clearly of importance…in part, this depends on the teacher’s ability to take account of the pupil’s perspective when asking the question.’
Marland (2002: 104) once again offers clear advice here, suggesting that 'the successful teacher's question is precisely one that can be answered, not one that cannot. The devising of the answer should be an intellectual gain, with the pupil using her past knowledge to develop new insights'. This is of course critical in the milieu of Sixth Form study – the depth of learning being developed and extrapolated to one of the highest levels before students embark on undergraduate study. The Assessment Reform Group's '10 principles of AfL' indicate very clearly that tasks and questions should prompt learners to 'demonstrate their knowledge, understanding and skills'.

As already outlined in the section on the Importance of Effective Questioning, the research on rich questioning was one of the most important areas guiding this research and was critical in the construction of the second research question; it also allowed the questions regarding students' preferences to be amended using more specific terms.

2.6 Wait Time

Giving students the opportunity to develop their answers also coincides with the other strategies; for instance, there is little use in asking a rich question unless students have time to think and reflect. There is a general consensus that increasing 'wait time' is beneficial and required in order to develop a risk-taking atmosphere.

The literature produced by Black and Wiliam (1998: 11) provides some statistics on the amount of time teachers usually wait before eliciting answers from their students. They
suggest that teachers ‘do not allow enough quiet time so that pupils can think out and offer an answer’.

A leading paper in regard to wait time was undertaken by Rowe (1974: 81-94). Her study of elementary school science classes in the US found that wait time was just 0.9 seconds. Rowe also purports that allowing a period of up to 8-9 seconds or more can encourage longer and more thoughtful responses which go beyond factual recall; in lessons where wait time was given she found that:

- failure to respond decreased
- responses were more confident;
- students challenged and/or improved the answers of other students
- more alternative explanations were offered.

Rowe’s findings, that are also referenced in Black and Wiliam (1998:5), were crucial as they determined how wait time was classified in this study; a wait time of 8 seconds was highlighted to staff and students as being a sufficient length of time for most students to answer most questions.

Hargreaves (2004: 33) estimates that teachers rarely wait more than two seconds, and suggests the reason being that ‘they fear a longer period of silence for thinking will create an opportunity for distraction or deviance’. However, Hargreaves’ study explored the wait time given to younger students (up to GCSE) where the level of questioning would necessarily be more straightforward than those put to students at A Level – but should the same amount of time be observed in the Sixth Form, where questions should be more sophisticated?
Brown and Wragg (1993: 20) indicate that student teachers often ask more questions than they receive answers to and suggest that the reason is due to insecurity on the part of the NQT, disliking pauses and feeling the need to fill them. They also propose that even experienced teachers may have a tendency to ‘favour, unconsciously, asking bright or knowledgeable pupils, if only because their answers are more rewarding’.

Hargreaves (2004:27 ) believes that extended wait time of just three seconds increases both the number of students who offer to respond and the thoughtfulness of responses ‘when students are not rushed into an immediate answer and are less likely to panic into either frozen silence or a wild guess’.
2.7 Assessment for Learning: A Critique

As indicated, the range of literature on Assessment for Learning was rather limited and there are few written critiques of AfL per se. This is quite possibly due to the fact that it is, after all, based on sound educational principles, developed over a number of decades. There were criticisms, however, of the way in which AfL has been recently packaged and implemented.

In an article published in the Times Educational Supplement (Ward, 2008), Bill Boyle - Professor of Educational Assessment at Manchester University - suggests that the AfL packs provided by the DfES ‘only pay lip service to the idea’ and suggested that schools were not changing their practice and were being sold ‘more of a glossy brochure than solid practice’. He found that around eighty per cent of the 480 heads surveyed said that AfL was a very high priority; however, follow-up visits to 24 schools found little evidence that teachers were using the technique’s principles in their day-to-day practice.

Professor Boyle goes on to suggest that AfL principles often get reduced to a ‘shopping list of things to do’. When Professor Boyle visited schools, he observed classrooms with learning objectives on the board and some peer assessment, but found that teachers still controlled the learning rather than giving pupils more active involvement.
Dylan William’s BBC documentary prompted Professor Boyle to write a letter to the TES, (Boyle, 2010), to criticise AfL’s seemingly over-simplified approach and hoped that it would not ‘deter those teachers who have moved on in their pedagogy’; he was also quoted as describing whiteboards, traffic lights and lollipop sticks that appeared in the documentary as simply ‘gimmicks and gadgets’. In a further article in the Times Educational Supplement (Bloom, 2010), the author refers to Dylan William as ‘an educational maverick…wielding a box of educational tricks’. 
Chapter 3: Methodology

3.1 Introduction

After establishing what needed to be investigated, as well as how the evidence should be organised, there was much deliberation over the form the research should take; this actually changed as the research questions and research methods were considered and constructed.

It was crucial to decide on the best kind of research methodology, before fully exploring the specific research instruments and methods of data collection. As indicated by Cohen and Manion (2007: 47) ‘the aim of methodology is to help us understand, in the broadest possible terms, not the products of the scientific inquiry but the process itself’. In selecting the appropriate methodology, the following questions were asked - as indicated by Bell (1999: 101):

- ‘What do I need to know and why?’
- ‘What is the best way to collect information?’
- ‘When I have this information, what shall I do with it?’

Due to the nature of this research, and the sample of students being used, it was not immediately obvious what format it should take. At first, action research seemed to be the most appropriate method, as part of the study involves making changes in the classroom, such as teaching practice (AfL questioning techniques).
However, the changes put in action, and those subsequently used in this study, were only a small part of this research project. The experimentation element of the research could have, however, easily fallen under the auspices of action research, as Blaxter. Hughes and Tight (2002:67) suggest:

action research is well suited to the needs of people conducting research in their own workplace and who have a focus on improving their own practices...the teacher who is concerned to improve performance in the classroom may find action research useful because it offers a systematic approach to the definition, solution and evaluation of problems and concerns.

A more succinct description was offered by Griffiths (1998) cited by Blaxter. Hughes and Tight (2002:67), 'the purpose of action research is, always and explicitly, to improve practice'. So, due to the rationale of this research, and the reported benefits of AfL (Black and Wilam, 1998), the research could have certainly fallen under this model. However, as action research it wouldn’t have been entirely appropriate, as other parts of the research simply reported and reflected on experiences and what should happen, rather than making changes to ‘flawed practice’.

As a result, a case study was chosen eventually, as it best reflected the overall research intentions. However, the research did lay the foundations and therefore ‘paved the way’ for future action research - namely setting in motion the changes through AfL; after these have been fully implemented, there should be scope for analysis and evaluation of the benefits (See Chapter 5: Scope for Further Research).
As Cohen and Manion (2007: 85) suggest, a case study is appropriate when the research is required to:

- interpret the uniqueness of individuals and situations (Current practice – RQ1);
- contribute towards action and intervention (AfL experiment class – RQ3);
- use participant and non-participant observation (AfL questionnaires – RQ2);
- allow for the holistic treatment of phenomena (Inclusion of staff RQ1/2).

There also needed to be a balance of both quantitative data and qualitative data: quantitative data were generated from closed questions (from the questionnaires) and also from more scientific enquiry, using empirical evidence generated through AS Level examination data; qualitative data were produced through open questions (from the questionnaires) and also through one-to-one interviews. These were vital in order to investigate students' attitudes and opinions and created pertinent comments on the issues raised, which went on to enrich the study.

As indicated by one of the research questions, the case study did also present the opportunity to use some experimentation. This was introduced in order to establish a more scientific mechanism to test for any possible causation between AfL questioning techniques and academic performance, as found by Black and Wiliam (1998) in regard to GCSE results.
As Bowling (1997), cited in Blaxter, Hughes and Tight (2002: 74), suggest:

The experiment is a situation where the independent variable (the AfL intervention) is carefully manipulated by the investigator under known, tightly defined and controlled conditions (in the classroom)...the experiment consists of an experimental group, which is exposed to the intervention, and a control group which is not.

Given the nature of social science research, the testing method required a quasi-experimental design i.e. experimentation in a natural setting, but one where variables could be controlled, isolated and manipulated. The design used a pre-test/post-test- non-equivalent group design, which is one of the most commonly used in education research.

Cohen and Manion (2007: 84) also explain that experimentation should be used in situations where it:

- allows individuals and groups to be compared (Control vs. experiment group)
- measures achievement and potential (Correlation between AfL and UMS)
- diagnoses strengths and weaknesses (Strength of correlation)
- assesses performance and abilities (Consistent correlation – range of ability)
3.2 Research Methods

3.2.1 Sampling

In total, there were eighty students involved in this study, along with thirty members of staff. The former included all of Year 12 and meant that a range of ability and attitudes would be included across all subjects, along with a large mix of subject combinations. All thirty members of staff used in the study were involved in AS/A level teaching and made up exactly three quarters of all staff involved in Sixth Form teaching. The sample used was a non-probability sample as a particular group needed to be targeted: rather than trying to represent the wider population, it simply represented itself. More specifically, a purposive sample was used in order to select ‘knowledge people’ for a specific purpose.

3.2.2 Questionnaires (RQ 1-3)

Student Questionnaires were given to all Year 12 students (n=80) and included questions that would generate both quantitative and qualitative data. They were chosen as they could capture large amounts of data and would involve all Year 12 students, not just those involved with the pilot and experimentation. Although the questionnaires were anonymous, they were all numbered and logged for reference.
Questionnaires were responsible for the majority of the research findings for all research questions. These were a vital method of data collection, concerned with gathering evidence of current questioning practices in the Sixth Form, as well as student preferences for questioning and classroom management. As already indicated, some departments were experimenting with Afl principles, but at the time of research most had not fully engaged with them. The main instrument used in this was the student questionnaire, given to all members of the lower sixth; questionnaires were also given to all staff involved with AS/A Level teaching.

The student questionnaires (documented in Appendix III) included quantitative (closed) questions that allowed easier data capture and the production of tables and graphs for analysis. In addition to this, qualitative (open) questions were included to allow students to voice their opinions and feelings regarding the use of questioning in the Sixth Form. Staff questionnaires (also documented in Appendix III) only included quantitative questions that were analysed to ‘gauge staff opinion’, along with their use and knowledge of Afl.

The questionnaire was kept short and questions were closed in order to encourage staff to fill them in, therefore creating a greater return rate (out of forty questionnaires sent out, thirty were finally collected in). Quantitative questions used a Likert Scale to establish the learning preferences of students. These included either ‘1-4’ or ‘1-5’, measuring ‘like’, ‘importance’, and ‘usefulness’. Staff Questionnaires were given to approximately half the teaching staff (n=30) and only included questions requiring quantitative analysis; this was used to compare and contrast staff attitudes and procedures to that of the students (RQ 1-2).
A rating scale was once again used to provide more detail than simply ‘Yes’ and ‘No’. Staff were presented with a similar format; however, rather than just using ‘1-4’, they were given the option ‘Always/Often/Sometimes/Never’. Staff in the pilot also had to complete an additional section that included their thoughts on their AfL experiences to date. Student/staff responses to each question were subsequently logged, calculated and analysed using a spreadsheet.

3.2.3 Interviews (RQ 2-3)

Interviews were undertaken in order to glean specific information from those involved with the Sixth Form Pilot (n=30). They were therefore used to establish how the questioning methods had been received, juxtapose other lessons where AfL had not used. The interviews were critical as they involved students that were directly exposed to the pilot, as well as those connected to the experiment and allowed students to express their personal views and feelings on the changes made.

The interviews included a random sample of students from each subject in the pilot – equal numbers from Economics, Modern Foreign Languages, English, Biology and Psychology. The interviews were designed to glean more information – both quantitative and qualitative – from those students and staff involved. Students were asked to expand on some of the answers in the questionnaires and were asked additional questions about their experiences in the pilot, in order to provide more candid information about their feelings, anxieties and expectations (catalogued in Appendix IV). They were also able to provide their views on the potential benefits that have been, or could be, achieved with use of AfL at the School (the students are labelled 1-27 and are referenced in the Results Section).
3.2.4 Experimentation (RQ 3)

Analysis of AS Level Economics Examination Data (n=30) was used to establish whether there were any academic performance improvements through the implementation of AfL in the classroom. AS performance, along with GCSE and ALIS scores, was used to standardise the data in order to draw the most accurate conclusions.

This method was restricted to just one of the subjects in the pilot – AS Level Economics. This sample was limited in order to control the conditions of the experiment more easily i.e. to ensure that the test was not affected by other variables and to ensure correct procedures were implemented at all times, thereby safeguarding consistency and accuracy.

In order to quantitatively test the success of AfL, the research needed to establish a way of measuring ‘academic benefits’. The main measure used here was final examination performance (at AS Level) – not just the grade achieved, but also the UMS (Uniform Mark Scale), using average percentage scores. The following grades are allocated to UMS scores: A = 80+, B = 70+, C = 60+, D = 50+, E = 40+.

An experimental group (Year 12A Economics) was set up where AfL questioning was used for the whole academic year. Alongside this, to compare and contrast the effects of AfL, a matched control group (Year 12B Economics) was used where no changes were made to existing questioning techniques.
The experimental design is illustrated in Figure 2 below. The candidates would be contextualised, and data normalised, through the pairing of candidates (e.g. candidates 1 and 2) with identical GCSE scores and/or ALIS TDA test scores.

![Figure 2: Experimental Design for RQ3](image)

The TDA (Test of Developed Ability) is a baseline test taken by students in Year 12, which acts to benchmark student ability, independently from any other summative assessment.

The experiment involved thirty students – 14 in the experiment group and 16 in the control group. In order to accurately determine causation, and to ensure reliability of any link between exam performance and exposure to AfL, other variables needed to be held constant: both groups contained only female students, had very similar class sizes, the same teacher, as well as students of similar academic ability (using average GCSE scores). Therefore, the only observed difference between the experiment and control class should be the use of AfL questioning methods.
3.3 Reliability and Validity

Bell (1999: 103) defines reliability as 'the extent to which a test or procedure produces similar results under constant conditions on all occasions'. She goes on to describe that validity 'tells us whether an item measures or describes what it is supposed to measure or describe'. Bell is essentially suggesting that if another researcher were to use your research instrumentation, they should get the same outcomes. The research questions were chosen carefully, as were the methods used to collect the data in each. The data were also collected over a period of one academic year, which should therefore allow for a reliable reflection of teaching and learning benefits.

The use of several methods of investigation, along with a fairly extensive and wide sample of both students and staff, should present the evidence as accurately as possible. All quantitative data were systemically collected and double-checked using several spreadsheet models, using both verification and validation at each stage. Where possible, quantitative data used scientific experimentation to eliminate any possible conflicts. The experimentation element used final AS Level examination data from WJEC, along with official data produced by CEM at Durham University. This supported the reliability and validity of this research, as the data used gathered evidence from long-standing and respected official bodies.
The students responded to the questionnaires accurately, without discrimination and were given sufficient time to respond to the question set. All questionnaires (both staff and students) were filled in anonymously, which allowed for more candid feedback; the questions were also set at the end of the year, so students could reflect on their whole time in Year 12. The interviews were carried out over the same period and students were promised anonymity and were also given as long as they needed to respond to the questions.

In order to aid validity, a pilot was used with both sets of questionnaire and interviews. Three members of staff and students were chosen at random and asked to complete the questions in order to test comprehension and ease of data capture. There were some issues with ambiguity and phrasing and, as a result, questions were either amended or omitted. Interview questions were also tested on a small random sample of students which helped to hone question design and determine interview length.

3.4 Triangulation

'Triangulation techniques in the social sciences attempt to map out, or explain more fully, the richness and complexity of human behaviour and studying it from more than one viewpoint, and in doing so, making use of both quantitative and qualitative data' (Cohen and Manion, 2007: 141)
Three methods of data capture were used in order to triangulate the data and provide a more accurate portrayal of the state of questioning. By involving students and staff, along with the use of summative exam data, the research increased in terms of validity and reliability.

All techniques reinforced the evidence collection by further scrutinising the sample. For instance, students filled in a questionnaire and then almost a third was interviewed; staff filled in a questionnaire, which also included those in the pilot. Some students were then examined further, through analysis of AS examination grades and UMS achievement in the experimental class.

3.5 Ethical Considerations

In accordance to BERA guidelines (BERA, no date) the research took into account the ethical issues involved at all times. Voluntary informed consent was sought from all participants and students and staff involved in the study had 'the right to freedom and self-determination', Cohen and Manion (2007). The content of the research, along with the potential benefits, was explained to all parties.
All parties were not only offered a chance to ‘opt out’ of the research, but were also told exactly what the study would involve and also what it was attempting to achieve. As Blaxter, Hughes and Tight (2002:146) remind the researcher, ‘ethics is about being clear about the nature of the agreement you have entered into with your research subjects’.

There were additional ethical considerations with this study as the students were directly involved. The use of experimentation in Research Question 3 was particularly sensitive as it necessitated the use of an experimental class and corresponding control class in the same year group. Once again, these participants had the opportunity to opt out of the study from the initial stages. However, in this fairly small sample, the experiment group were interested in learning more about AfL questioning; likewise, the control group were happy with the *status quo*.

Also, the year group that was to be tested was also chosen carefully; Year 12 were used rather than Year 13 due to reasons of practicality and fairness; if the changes made by the study produced unfavourable results, then students and staff would have had greater scope to put in place remedial action in order to rectify any issues before the final A Level examination; in addition to this, assessment outcomes would not directly affect university offers.
3.6 Evaluation of Methodology

Overall, the results generated as a result of the methodology appeared to be valid and completely pertinent to the rationale. However, the evidence would have been stronger if the sample had been extended. Although the sample of eighty students was fairly large, it only represented one year group, over a short period of time. Ideally the experimentation should now be expanded to examine performance in each subject and should involve a longitudinal study.

Another limitation of this study was the lack of classroom observations. This was certainly part of the initial methodology but, due to time constraints and other logistical problems, it was simply not possible. This was disappointing, as it would have acted as an interesting addition to the data collection methods, helping to triangulate what the staff and students said happened, with what actually happened.

The experimental element of the study was also narrow, and only focused on AS Level Economics. Although there were significant findings here, it would have been much more effective to have used a larger scale experiment over the range of subjects in the pilot. Part of the questionnaire focused on questions that required students to discuss their individual subjects. Although this did actually provide a good overview of student questioning experiences, on reflection these questions became largely redundant. This did take a large amount of time and effort to capture and analyse – time that could have been used to implement other methods. Some of the other questions also proved to be useless in light of an ex post review of the methodology and results, due to the research questions being reviewed and honed, and therefore making the data no longer relevant.
The interviews were generally effective but were undertaken at the end of a busy year, after AS Level examinations. Although a period of time was required for students to fully reflect on the pilot, the students were often tired and not fully engaged in the process, which meant that their responses were sometimes truncated. In addition to this, some students were also unable to fully judge some of the criteria - e.g. whether their understanding had increased - was this because they weren’t exactly sure how to measure it, or due to the relatively short period of time in which they had to appraise it?

3.7 Research Time Line

For a list of research tasks, timings and milestones – please see Appendix I (a&b)
Chapter 4: Results and Analysis

4.1 Introduction

This section draws on the data collected from the various research instruments detailed in the Methodology. In order to address the 'how' and 'why' questions, thereby achieving a meaningful evaluation of the data, there needed to be a balance of both quantitative data (generated from the closed questions), as well as pertinent comments on the issues raised by the students (via the questionnaires and one-to-one interviews). The closed questions allowed for a useful generation of graphs which are, for the most part, displayed in the Appendices.

However, for ease of reference some of the graphs that illustrate this study’s significant findings are contained within the main body of this section. The open questions also allowed there to be an informed narrative and an opportunity for the students to convey their feelings regarding questioning. Equally there were significant findings here which will be vital in determining how AfL is introduced, if at all.

The research evidence was based around the three research questions and specifically addresses them one by one. The qualitative data will be loosely reflected on in ‘themes’, in order to develop a meaningful reflection of the key issues. The comments are intended to be reflective of the sample and, due to the strength of their answers, some students’ comments are referenced more than once. In addition to this, there is synthesis via referencing of the appropriate secondary research, initially detailed in the Literature Review.
4.2 RQ1: What is the current standard of classroom questioning used in the Sixth Form?

In order to determine the scope for introducing AfL, it was crucial to ascertain the way in which questioning was organised across various subjects in the Sixth Form and how the small pilot had been received to date. One of the first questions put to students was whether they had noticed changes in the way questions had been asked this year. The figure generated was much higher than expected, with over half of Year 12 reporting that they had noticed a difference. However, when observing the accompanying qualitative data it was clear that the question was too ambiguous and, as a result, many students were simply reflecting on changes related to entering the Sixth Form – being treated more as adults etc. - rather than on AfL per se. Therefore, in order to obtain a figure that would relate specifically to AfL, Year 13 were set the same question and on this occasion 23% of students noticed changes. This figure was more in line with the number of students exposed to the pilot scheme (6 out of 23 subjects i.e. 26%). See Appendix II (Figure A1).

In addition to interpreting exposure to changes, it was also important to observe students’ awareness and understanding of the term ‘Assessment for Learning’. The term Assessment for Learning has become common parlance in the maintained sector, being used extensively over the past decade. However, as explained in the introduction, the term is not particularly well known in the independent sector. Due to in-house INSET, all staff at the school are now well aware of the term and there had been a ‘buzz’ reported in the staffroom about AfL and the work undertaken by some colleagues this year; however, how far had the term ‘AfL’ cascaded down to students?
It often takes time for learning pedagogy to be picked up by students (if at all); even when students are more capable of self-reflection, understanding of the term at this age can be limited. However, surprisingly, the term had clearly been used by staff and in some cases explained to students. Almost a third of students either knew what AfL meant, or had at least heard of the term, and 10% had a fairly clear knowledge of its objectives and methods (see Appendix II: Figure A2). This was reassuring, but clearly all students need to be aware of what AfL is and what it intends to achieve if the pilot is going to be expanded.

In the questionnaire, students were also asked what they felt was the main purpose of questioning. The results are shown in Figure 3 overleaf. Students were asked to rank their responses from 1 - 4 (1 being most important/4 being least important). As indicated, students are still far too concerned with the educational product, rather than the process. The focus on knowledge collection shown here, clearly demonstrates the students' desire for proving the extent of their intellectual prowess to their peers.

Almost two thirds of students agreed (by choosing ‘1’ or ‘2’) that questions should ‘test what you know’; the highest single frequency that was recorded in the ranking of preferences was ‘4’, with 38% of students feeling that ‘actively participating in lessons’ was the least important objective of classroom questioning. This was closely followed by ‘advancing learning’, with just over a third of students attaching to this the least importance.
These are two keys elements involved in AfL i.e. richness of questioning and student involvement. If AfL is to be successfully implemented then the perception and role of questioning clearly needs to be changed. As already suggested, students are not always able to fully appreciate the pedagogical benefits behind teaching strategies and this is, instead, left in the hands of teachers and educationalists. However, aversion to rich questioning and student participation are recurring themes that are developed in the next section and help to make up the key findings, developed later.

One of the first significant findings was the fact that the response of the staff largely disagreed with the students’ views on the role of questioning; with reference to ‘participation’, an overwhelming 92% of staff felt that this was of high importance; in contrast only 22% of students felt that ‘participation’ was highly important.
93% of staff suggested that making students ‘think more deeply’ was the most important aspect; in contrast only 25% of students felt that this was of high importance. The only area where students and teachers were in agreement was on the recalling of facts i.e. testing what they know, which was of high importance according to both.

If AfL is to be successful it therefore seems that staff also need to go through a cultural change in their pedagogical practice, moving away from their addiction to simply ‘collecting responses’ from students. As Black and Wiliam (1998) suggest ‘teachers often look for specific answers…with up to 80% being merely factual responses’. The authors cite that teachers often lack the flexibility or the confidence to deal with the unexpected; however, staff at the school are certainly flexible and highly confident - they have simply developed their own styles over time, sometimes decades. Pleasing annual exam results only act to reinforce such inward looking and entrenched practice, further strengthening the ‘if it ain’t broke, don’t fix it’ idiom.

In order to establish how various subjects in the Sixth Form manage and deliver questions, students were asked to provide evidence of their questioning experiences over the past academic year. The quantitative and qualitative evidence here was generated by collating data from the questionnaire, with each of the eighty students reflecting on their experiences in each of their four AS Level subjects (23 subjects in all).
As shown in Appendix II (Figure A3), the most prevalent method of handling questions across the Sixth Form, according to students, is when pupils are allowed to ‘shout out’ the answers. As suggested by the BBC Documentary (Wiliam, 2010), with this questioning strategy, students are able to hide and disengage from the learning in the knowledge that other students will always respond. Moreover, it does not allow students to think about their answer; some students simply require more time to grapple with the question in order to fully understand what has been asked of them.

If, as the data suggest, the brightest and or/confident students are shouting out over 60% of the time, then this will clearly hamper the learning process for the majority. This was similar to the findings of the report (OFSTED, 2008) on the AfL National Strategy which found that teachers, to the detriment of students, often ask only for brief, factual answers without explanation and ask questions to the whole class.

The problems associated with such a strategy are best illustrated through the thoughts of the students themselves. There were some strong opinions expressed when teachers allow students to ‘shout out’, with some clearly feeling annoyed when they were not allowed to get involved. Some students certainly need more time to think about the question and an opportunity to participate in a more structured way, particularly at AS Level where questions are necessarily more complex. Some common reactions from interviews included:

“Pupils may have tried hard to figure out the answer but, if someone shouts out, the chance is missed and your efforts aren’t recognised”. (Student 16)
“If someone shouts out it’s frustrating...people who need more time to think about the question don’t get a chance” (Student 19)

“I feel relieved when someone else is asked and not me”. (Student 11)

The second most prevalent experience reported was when teachers name pupils before they ask the question. Although, this is a more organised approach than simply shouting out, it encourages pupils to immediately ignore the question, as they already know that someone else is going to answer it. This strategy also allows the teacher to deliberately, or indeed subconsciously, choose students that he or she knows can answer the question. This does not provide an opportunity to judge class reaction to the question, but some students did like this – particularly those with weaker GCSE results. The evidence suggested that these students often lacked confidence and preferred to let stronger students - or those they perceived as being stronger - answer. This was reflected on in the last comment above; some students do not like pressure and instead prefer to have the burden (as they see it) taken away from them.

This, rather worrying, finding was also significant. Students who need to fully understand the rigours of the AS Level subjects are simply getting away with hiding in the shadows; moreover, their anxieties about their low ability (perceived or otherwise), along with not being able to answer questions, is being continually reinforced. They are simply not being given the chance to practise and improve - nor do they get a chance to receive positive feedback from their teachers and/or peers.
In order to appreciate the use of questioning in the Sixth Form, and gauge the extent of good practice, it was also vital to ascertain the views of the staff. The responses given by staff, reflecting on their own questioning approaches over the last year, are shown in Appendix II (Figure A4), as percentage responses. The columns indicate the percentage of staff who gave either 1 ‘Always’ or 2 ‘Often’ as their response. The most prevalent by far was their use of wait time. For the purposes of this study, it was suggested to students and staff that a wait time of 8 seconds was considered to be sufficient – the same figure provided by Rowe (1974). A large number of staff (84%) stated that they used wait time always or often (which was corroborated by 85% of students). Only 62% of staff indicated that they used an alternative to ‘hands up’, with only 23% of staff using it always. This meant that most of the time students are being allowed to disengage in lessons.

An even more concerning observation involved the richness of questioning. Only 62% of staff use what they consider to be rich questions in their Sixth Form lessons and only 16% of staff stated that they used them all of the time. Therefore, the aforementioned students, who are already reluctant to engage with rich questions, are not being given the opportunity to practise and hone higher order responses in class; teachers are therefore too focused on obtaining pithy answers, that only reflect knowledge (and memory) of previously taught material, rather than those that demonstrate true understanding. This seemed to support the research undertaken by Hargreaves (2004) who criticised most question and answer sessions as being ‘oral checks on what students already know’ and implied that teachers often have a ‘predetermined answer’ in their head.
The least prevalent questioning technique practised by staff was the use of wrong answers in order to improve student understanding. Less than 40% of staff use this method on a regular basis and only 12% admitted that they used it continuously. A third of the staff questioned admitted that they never used this technique. This could be due to the traditional methods of teaching and attitudes to learning at the school; the idea that students cannot be seen to make mistakes. It could also be that teachers do not want to make the girls feel stupid in front of their peers.

Summary of Significant Findings

Staff and student views on questioning were often diametrically opposed:

- 92% of staff felt that ‘participation’ was of high importance
- 22% of students felt that ‘participation’ was highly important
- 93% of staff felt that ‘thinking more deeply’ was the most important aspect.
- 25% of students felt that ‘thinking more deeply’ was most important

Some Sixth Form students are getting away with “hiding” and are serial non-participators:

- Issues about low ability (perceived or otherwise), along with not being able to answer questions, is continually being reinforced and is affecting confidence
- In a crucial year of study, students are not being given the chance to practise and improve

58
4.3 RQ2: To what extent would AfL questioning methods suit the students’ teaching and learning preferences?

As well as determining the standard of questioning over one academic year, it was also critical to determine - through both statistical aggregates and the feelings of students - the questioning preferences of the Sixth Form. In this section, the evidence is displayed through the four main questioning methods. In addition to this, student preferences are once again, on occasion, affected by academic ability according to GCSE results (using Summer 2010 exam data).

Asking for students’ GCSE results provided an opportunity to split students according to ability.

The scoring was constructed using the same system as Durham University’s Curriculum, Education & Management Centre (CEM): A* = 8; A = 6; B = 4; C = 2.
During the pilot, AS/A Level students were exposed to a variety of questioning methods. They were interviewed at the end of the year and were asked to reflect on their experiences of the pilot by simply stating ‘least’ and ‘most’ favourite. The results are illustrated in Figure 4 on the previous page. Wait time was by far the most popular technique; no students stated that they most preferred rich questioning.

The richness of questions set – particularly at AS/A Level - is an extremely important area, but the students did not seem to rate this very highly. The most popular change to questioning in the pilot was the use of wait time; students clearly liked the opportunity to think and reflect and this was expressed across the board. Although not universally liked, this was the most significant change made to teacher practice and was also the most effective. As indicated above, the use of wrong answers also proved to be popular, but at the moment Sixth Form students are unfortunately unable to enjoy the full benefits of this practice.

One of the least preferred methods was the use of random questioning. It was clear that students did not like being challenged and have strong feelings about looking ‘inferior’ in front of their peers. On closer examination of the questionnaire, and through one-to-one interviews, the preferences detailed generated some interesting data. The findings from each of the changes made, in each of the four AfL questioning methods, are now developed further.
4.3.1 Wait Time

As already indicated, 84% of staff suggested they did use wait time as part of their Sixth Form teaching which, given the size of the pilot, was extremely encouraging. The literature produced by Black and Wiliam (1998) indicated that there was a real deficiency in regard to wait time, which seems to be repudiated by the evidence generated here. The authors also cited US literature that suggested that, on average, wait time was just 0.9 secs. Both staff and students felt that wait time was of high importance, particularly in the Sixth Form. The graph in Appendix II (Figure A5) shows that, on average, students felt that at least 20 seconds was necessary in order to, as Black (2003) et al point out, ‘marshal their thoughts and to begin to assemble an answer... extend, elaborate’.

The minimum amount of time that students indicated they required was 5 seconds – which is still greater than that suggested by Hargreaves (2004). The spread of wait time indicated by students was fairly narrow; as the standard deviation demonstrates, there was considerable clustering around the mean – indicating a fairly strong consensus of opinion (the most popular wait time – expressed through the mode - was 10 seconds). Staff opinion on wait time agreed with that of the students – although they advocated even greater use of wait time with a mean time of 33 seconds, a figure more than four times greater than that advocated by Rowe (1974).
Staff and students therefore both realise the benefits of wait time which agreed with the body of research undertaken. Students indicated that, with more time, they were able to achieve some of the benefits suggested by Rowe (1974), namely increased confidence and an increase in alternative explanations offered.

Although staff indicated that a great deal of wait time was already used, it is not always apparent in Sixth Form lessons. Wait time is a vital element of questioning and AS students need this more than ever. As Black and William (1998:8) indicate ‘the dialogue between pupils and a teacher should be thoughtful, reflective, focused to evoke and explore understanding’. When they are not given such opportunity, strong emotions were generated. Firstly, a sizeable number of students felt real anger when they were not allowed to show peers and teachers their knowledge and understanding. The most frequent terms used in reaction to insufficient wait time were ‘annoyance’ and ‘disappointment’; the interview comments below typify the sentiment of the student body:

“I feel disappointed that I haven’t been able to show my full potential” (Student 1)

“I get annoyed…I never learn, I feel I’m wasting my time” (Student 6)

Secondly, and perhaps more crucially, was the effect on the students’ self-esteem and confidence. Although A Level classes should thrive on open dialogue and should access all students, a large number of students gave the response of feeling ‘inadequate’ or ‘stupid’.
Typical responses included:

"It doesn’t help me in any way...I feel useless and ignored" (Student 2)

"Degraded, depressed, stupid...I feel I want to cry" (Student 10)

"I often feel angry and humiliated" (Student 12)

Thirdly, some students added that they felt ‘dejected’ and had a feeling of ‘worthlessness’. By far the most common feedback from students was that in such situations they felt ‘relieved’ due to the fact that the ‘pressure was off’. Some students, through a mixture to personality, class politics, but most crucially a lack of confidence about their perceived ability, prefer to opt out of lessons and take an extremely passive role in the class. If AfL is to be launched across the board, then this is an area that will need careful consideration.

The strongest responses at interview expressing such a ‘reprieve’ were also combined with feelings of failure:

“I feel relieved when the teacher moves on to another pupil or answers themselves, but also that I have failed” (Student 9)

“I often feel inadequate, but also thankful” (Student 15)
4.3.2 Student Participation

The way in which students are asked questions is key to the dynamic of any lesson, the classroom atmosphere, and also the way in which teachers relate to and connect with their pupils. *Research Question 1* introduced a graph that illustrated students' experience of questioning (Figure A4 in Appendix II). Figure 5 below presents the same data, but this time also incorporates the students' preferred questioning methods alongside.

![Student Questioning Preferences vs. Reported Experiences](image)

**Figure 5: Preferred vs. Actual Questioning Experience (n=80)**

Comparisons between what students get, and what they actually want, flagged up some interesting trends. Shown by the blue column, the data represent the percentage of students who supplied an answer of 1 (most preferred). One of the most striking deviations occurred with ‘hands up’, the popular method of questioning. This seems to be a practice entrenched in the teaching pedagogy and one which the students have become accustomed to since Year 7.
When questioned about this, different students seemed to like this system for different reasons: more academically able students like it as they are allowed to dominate and are given carte blanche to showcase their intellectual prowess; less academic students and/or those lacking in confidence are happy coasting, preferring classes where the teacher chooses students more capable of answering. Some of the most popular reasons provided for preferring a ‘hands up’ approach are shown below. Students are clearly reassured when someone else is able to answer and find this philosophy completely rational:

“Because people with their hands up are prepared to answer” (Student 2)

“It’s good to have some control over who answers” (Student 26)

“It’s better to ask someone who knows the answer” (Student 14)

It was encouraging that students were not given as much ‘hands up’ questioning as they would like. However, as the graph shows, it is a practice that is still overused and therefore one that will require specific attention. The second most popular method of questioning was where pupils ‘shout out’. This was the most commonly reported experience by students and is one which is also used too much.
When examining the mode (most popular response) given by each student their GCSE score seemed to affect the results: Students with relatively low GCSE average scores (less than 6) generated a mode of 5 (a strong disliking) for pupils shouting out; in stark contrast to this, students with relatively high GCSE results (greater than 6) generated a mode of 1 (a strong liking).

The least preferred techniques for choosing respondents were ‘teachers choosing pupils after asking questions’ and ‘random questioning’. The latter is strongly promoted by AfL, via the seminal literature, and has been publicised in a recent BBC documentary. Students had an extremely low preference for random questioning with only 5% indicating this as their favourite. If AfL is to be fully adopted then students’ aversion to this must be addressed in consultation with their concerns. Their reasons for disliking random questioning methods included:

“It takes time to understand some concepts... when I’m not ready I don’t want to embarrass myself by guessing... I usually frown to put off teachers from asking me” (Student 24)

“It is not fair to the pupils who don’t know the answer” (Student 20)

As indicated by Student 24, over time students have become accustomed to avoiding questions and have even developed body language to avoid what they see as ‘confrontation’. From both the questionnaire and the additional interviews with those exposed to the pilot, it is clear that random questioning is something that students are not yet wholly comfortable with.
Displayed below are some of the feelings that were expressed when students are ‘put on the spot’, via some method of random choosing of the respondent. Students with lower relative ability become particularly entrenched here and are frightened to contribute in case ‘they get in wrong’:

“I get a big sinking feeling in my stomach ...I hope the teacher will leave me alone...don’t want to face next lesson” (Student 1)

“It’s OK, but I often get the questions I can’t answer and someone who does isn’t asked” (Student 13)

“Hate it – it takes you by surprise” (Student 7)

“I don’t like it – they should ask people with their hands up as they know the answer” (Student 10)

“I feel under pressure – if you knew the answer you’d put your hand up” (Student 14)

This produced very similar results to the KMOFAP project - although it included students up to Year 10 (GCSE). In this project, documented in Black and Wiliam (2003), both teachers and students felt incredibly uncomfortable as both parties had to deal with a completely new type of classroom organisation and also how to respond to replies that may not be right.
However, the student view was by no means one-sided. Again, in the interview, students were more positive about the potential benefits and the changes witnessed so far. Some of the more positive comments included:

“It’s good as it makes everyone listen” (Student 8)

“Good as it gives everyone a chance” (Student 6)

“It makes me pay attention more, but I don’t like being put on the spot” (Student 15)

“I concentrate more as I really want to get the answer right” (Student 18)

“It encourages you to constantly think of answers and also shows teachers who doesn’t understand” (Student 11)

These feelings are certainly more commensurate with the evidence provided in the literature and shows that at least some students see the immediate benefits of being more focused in the classroom. However, interestingly, all of the positive responses were provided by individuals possessing higher than average GCSE scores (greater than 6).
4.3.3 Rich Questioning

Some of the most interesting data were witnessed when students were asked about the richness of questions put to them in class.

The Richness of Questioning
Student Preferences: Liked vs. Useful

One of the most striking observations made through the student questionnaires was the (often polarised) disparity between the questions they liked and those they recognised as having educational benefits.

Detailed in Figure 6 above are the three types of questions where the responses recorded were the most disparate i.e. where the difference between like and usefulness was the greatest (in each case, the results show the percentage of students choosing 1 or 2). As indicated, almost three quarters of students liked questions that require only one word answers, but only 10% thought that they were in any way useful.
The students’ preference for basic questions was further confirmed by the popularity of questions simply requiring “yes” or “no”. Here, half of the eighty students questioned liked this method, but only 2% found it useful in educational terms.

Conversely, students rated more complex, rich questions much lower and appeared to have aversion to any question requiring engagement and reflection. “Why” questions were rated the lowest with only 1 in 5 students liking this type of question. Interestingly, a much greater number - almost two thirds – thought that they were beneficial to the learning process. Exactly half of the students questioned had a differential between ‘like’ and ‘usefulness’ of 3 (e.g. Like 5, Usefulness 2), which further acted to confirm the disparity.

This was once again a significant finding that shows a polarisation between what students liked and what they know is educationally beneficial. Although some students preferred to take the easiest route, they did acknowledge the positive effect of rich questioning as suggested by Black and Wiliam (1998). They felt closed questions simply require factual recall and were therefore unlikely to enable them to ‘articulate their understanding’.

As with their participation preferences, students’ attitude towards question richness was also affected by the academic background of the student. Students with relatively low GCSE scores, had the most popular response of 1 (a strong liking) for ‘one word answers’. Once again, in stark contrast to this, students with relatively high GCSE scores had a mode of 5 (a strong disliking) for brief answers.
Students with higher GCSE results did not like one word answers as they do not let them explore topics in depth and minimise the time in which they can ‘hold forth’; conversely, those with lower GCSE results liked these questions as they ‘let them off the hook’ and do not require full understanding of the topics being learned; similarly, wrong answers can be dismissed quickly, as the teacher moves onto the next question.

As indicated previously, student confidence is clearly a factor determining question preferences; fear and embarrassment seemed to affect attitudes towards being placed ‘on the spot’ and this was also affected by academic background and achievement.

When asked why students liked one word answers, similar themes are witnessed such as wishing to get it’ out of the way’ quickly, without being made to feel awkward in front of their peers:

“Easier – they make me feel more confident when I get it right” (Student 22)

“It’s quick…less pressure” (Student 7)

“[With Y/N] you have a 50/50 chance – less pressure to get it right” (Student 8)

“[Y/N] doesn’t require much thinking and you have 50/50 chance” (Student 12)
4.3.4 Use of Wrong Answers

As indicated in Research Question 1, the use of wrong answers was currently the least prevalent questioning technique used in the school. The Girls’ School is an incredibly busy environment, for both students and staff, and the Sixth Form is no exception. A seemingly over-crowded A level curriculum in many subjects, coupled with the ensuing race to complete the specification in just two short terms before the ever earlier AS and A Level examinations, means that wrong answers are seldom used to their full potential. However, through interviewing those involved in the pilot, it was one of the most popular with students (shown in Figure 4). This was another significant finding that, if fostered appropriately, could vastly benefit the school.

Some of the students also felt that, if handled in the right way, the use of other students can create a culture of ‘everyone’s in it together’ and that understanding of difficult topics could be increased. Some of the popular comments at interview included:

“It’s great it feels like a group effort” (Student 13)

“If it’s a very good answer it helps me and shows me how I can improve” (Student 27)

“Yes, if my answer is built on by others then it helps me understand” (Student 11)
In addition to this, students found that something they would usually do outside the class - i.e. revising by talking between themselves, could help not only generate a more positive culture in the classroom, but should also help their own cognition and appreciation of the subjects; students of the same age are able to articulate certain words and phrases, often inaccessible to the teacher. Such reconciliation of the theory is critical at Sixth Form level and should be encouraged as much as possible. Some of the thoughts of students relating to this included:

“I like it because the knowledge of my peers adds to my own…they might explain something differently than the teacher” (Student 23)

“It’s good and often they can explain something you didn’t and get a different angle” (Student 4)

“It’s good hearing what other pupils think rather than just the teacher” (Student 10)

This is certainly an area that needs to be developed and should be a fundamental part of the learning process. However, as Kyriacou (1997) points out, feedback is the ‘final skill’ to be honed and needs to be handled carefully as it is often a ‘high-risk’ and ‘emotionally charged’, involving explicit teacher judgement in a public arena. This is therefore an area requiring extensive training – if the use of wrong answers are handled badly the consequences could be more damaging than their absence.
This was typified by some of the comments below, which portray how some students feel when they get answers wrong and/or if wrong answers are handled badly:

“It’s not fair when you’re used as a ‘bad example’ – it’s harsh” (Student 9)

“I feel stupid and confused….exasperated” (Student 3)

“Annoyed at myself for getting wrong and embarrassed if there’s lots of other people in the class” (Student 5)

Summary of Significant Findings

A large number of students have an overwhelming preference for basic questions whilst disliking rich questions and class participation:

- 73% of students liked questions that require ‘one word’ answers, but only 10% thought they were useful.
- 50% of students preferred ‘yes or no’, but only 2% found it useful.
- Only 20% liked “why” questions, but almost 65% thought that they were beneficial to the learning process.
- Students do not like being challenged and have very strong feelings about looking ‘inferior’ in front of their peers.

Use of Wrong Answers

- Use of wrong answers is currently the least prevalent questioning technique used in the school.
- It was clearly one of the most popular with students
4.4 RQ3: What appear to be the academic benefits generated by the AfL questioning pilot scheme?

The term ‘academic benefit’ was used deliberately so that the analysis would be broad enough to incorporate a variety of factors. In terms of whether AfL should be used across the school, this question was of paramount importance, not only the quantitative academic data gleaned through experimentation, but also through gathering the thoughts and feelings of staff and students, through interviews and questionnaires i.e. the perceived academic benefits.

4.4.1 AfL and Academic Performance

When considering academic performance, the most effective measure available was the students’ examination results. The students in this smaller sample group were the 30 students who studied AS Level Economics from 2010-2011. In order to establish whether there is a correlation between use of AfL questioning and exam performance, the AS Level grades and UMS marks of the experimental group were compared to that of the control group.

In order to limit the amount of other independent variables affecting the experiment, only students with identical GCSE results were used for comparison; in order to reinforce causation, students with near identical ALIS TDA scores were also used. On examination of the two groups, there was a total of six students with the same GCSE score, therefore allowing for three comparative pairings (Students 1&2, 3&4, 5&6).
The results of the experiment are detailed below; in each case, the candidate achieving the highest AS Level UMS mark is displayed first (the experiment group is shown in pink each time and the control group in blue).

Table 1 below shows the first grouping - Students 1 and 2 - with both students having a very high GSCE score of 7.8. Although both candidates achieved A grades, the UMS mark achieved by Student 1 in the experiment group (exposed to AfL) was almost 6 marks higher than Student 2 (just over 1/2 grade equivalent). This student pairing allowed for a particularly strong comparison as each student also achieved almost identical ALIS scores of 118.

Table 1: Students 1 & 2 GCSE Normalised (n=30)

<table>
<thead>
<tr>
<th>Candidate</th>
<th>Class</th>
<th>UMS (%)</th>
<th>Grade</th>
<th>GSCE Score</th>
<th>ALIS TDA Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student 1</td>
<td>Economics A</td>
<td>100</td>
<td>A</td>
<td>7.8</td>
<td>118.1963</td>
</tr>
<tr>
<td>Student 2</td>
<td>Economics B</td>
<td>94.5</td>
<td>A</td>
<td>7.8</td>
<td>118.0817</td>
</tr>
</tbody>
</table>
The same comparison was undertaken for the remaining pairs of students with identical GCSE scores (Table 2); in each case the same performance increase was witnessed with the candidate in the experiment group; the most striking effect was observed with Students 3 & 4, with a differential of 12 UMS marks - more than one whole grade (equivalent).

Table 2: Students 3&4/5&6 GCSE Normalised (n=30)

<table>
<thead>
<tr>
<th>Candidate</th>
<th>Class</th>
<th>UMS (%)</th>
<th>Grade</th>
<th>GCSE Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student 3</td>
<td>Economics A</td>
<td>92</td>
<td>A</td>
<td>6.4</td>
</tr>
<tr>
<td>Student 4</td>
<td>Economics B</td>
<td>80</td>
<td>A</td>
<td>6.4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Candidate</th>
<th>Class</th>
<th>UMS (%)</th>
<th>Grade</th>
<th>GCSE Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student 5</td>
<td>Economics A</td>
<td>92.5</td>
<td>A</td>
<td>7</td>
</tr>
<tr>
<td>Student 6</td>
<td>Economics B</td>
<td>87</td>
<td>A</td>
<td>7</td>
</tr>
</tbody>
</table>

The causation between AfL and exam performance was also witnessed when a student in the control group achieved a lower GCSE grade. Table 3 overleaf shows that Student 7 actually achieved higher UMS marks than Student 8, even though she scored 0.6 points lower at GCSE. In this case, the UMS differential was actually the greatest recorded - 15 UMS marks (the equivalent of 1½ grades).
Although the GCSE differential between Student 7 and 8 was only observed once, it was not witnessed in the opposite direction i.e. there was no evidence to suggest that, when a control group candidate achieved a lower GCSE score than someone in the experiment group, they achieved higher UMS.

Table 3: Students 7&8 GCSE Reversal (n=30)

<table>
<thead>
<tr>
<th>Candidate</th>
<th>Class</th>
<th>UMS (%)</th>
<th>Grade</th>
<th>GSCE Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student 7</td>
<td>Economics A</td>
<td>85</td>
<td>A</td>
<td>5</td>
</tr>
<tr>
<td>Student 8</td>
<td>Economics B</td>
<td>70</td>
<td>B</td>
<td>6</td>
</tr>
</tbody>
</table>

As it was only possible to compare a total of eight students using GCSE grades, pairs of candidates with the same ALIS score were also used. The figures in Table 4 also support the aforementioned casual relationship between performance and AfL. In each case, the TDA scores were extremely close and once again the experiment students achieved between 11 and 15 UMS marks higher than their non-AfL peers.

Table 4: Students 9 & 10, 11 & 12 TDA Normalised (n=30)

<table>
<thead>
<tr>
<th>Candidate</th>
<th>Class</th>
<th>UMS</th>
<th>Grade</th>
<th>ALIS TDA Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student 9</td>
<td>Economics A</td>
<td>95</td>
<td>A</td>
<td>110.6362</td>
</tr>
<tr>
<td>Student 10</td>
<td>Economics B</td>
<td>81</td>
<td>A</td>
<td>110.6362</td>
</tr>
<tr>
<td>Student 11</td>
<td>Economics A</td>
<td>81</td>
<td>A</td>
<td>105.5389</td>
</tr>
<tr>
<td>Student 12</td>
<td>Economics B</td>
<td>70</td>
<td>B</td>
<td>104.6798</td>
</tr>
</tbody>
</table>
In order to triangulate the data further, the average UMS achieved by the entire control and experiment groups was calculated. This was a valid comparison as the data were contextualised; on average, the students in each group possessed very similar GCSE scores (6.4 and 6.5).

Table 5: Group Comparisons GCSE Normalised (n=30)

<table>
<thead>
<tr>
<th>Group</th>
<th>Class</th>
<th>Avg. UMS</th>
<th>Avg. Grade</th>
<th>Avg. GCSE Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment</td>
<td>Economics A</td>
<td>88.4</td>
<td>A</td>
<td>6.5</td>
</tr>
<tr>
<td>Control</td>
<td>Economics B</td>
<td>69</td>
<td>B</td>
<td>6.4</td>
</tr>
</tbody>
</table>

At first glance, Table 5 above shows the difference in UMS achievement, and it appears to be extremely high, with the experiment group achieving close to two grades (equivalent) more than the control group. However, although there was still a marked difference in performance, some students achieved low UMS in the control group and therefore artificially affected the average. This problem is indicated in the groups’ standard deviation (SD): the SD in the experiment group was 8.5 UMS, indicating a close clustering of AS performance around the mean, whereas the control group’s SD was much more disparate at 24 UMS.

The evidence here is therefore convincing and did conclusively find a correlation between AfL and AS level performance, albeit in a fairly small sample. Overall, AfL use seems to cause candidates’ AS UMS to increase by 10-15 marks (1 ½ grades), across a range of students and GCSE results. This is very similar to that reported by Black and Wiliam (1998) who found that AfL use also raised performance by an average of 1 ½ grades – although their study focused on first year GCSE students.
4.4.2 AfL and its Perceived Benefits

The academic benefits were also studied more broadly, thorough an investigation into the feelings of the staff and student body. This not only allowed the perceived performance benefits to be measured, but also permitted reflection of the following factors, which are also important to the learning experience:

- Understanding
- Independence
- Enjoyment of Lessons

Figure 7 overleaf shows the thoughts of staff involved in the pilot; regarding the effects of AfL used to date, almost 85% felt that one of the main benefits was to student independence, an approach to learning much less reliant on staff. In addition to this, over 60% felt that students' understanding of their subjects had improved (or could improve). However, only half the staff felt that student performance would improve, which largely repudiated the findings in the previous section.
In contrast to this, Figure 8 below shows the students’ views on the pilot. Overall, they were more pessimistic about the benefits/potential benefits of AfL. They did not feel that it would produce the same degree of academic benefits as the staff and also indicated that their enjoyment of lessons had not improved. The response most aligned with the staff’s view regarding independence, with over 70% of students believing that it had increased (or could increase) their attitude to work and motivation.
Summary of Significant Findings

Strong and consistent correlation between AfL and AS level performance:

- AfL candidates’ AS UMS increased by 10-15 marks (1 ½ grades), across a range of students when GSCE results were normalised (relationship even stood up when GCSE results were lower).
- AfL candidates’ AS UMS also increased when the TSA baseline data were used to contextualise the data.
Chapter 5: Conclusions and Recommendations

5.1 Summary of Results

As indicated, this research was designed to inform SMT at a selective, independent girls’ school whether AfL should be implemented across the whole school. In order to inform senior managers on the possible pedagogical improvements, the methodology embraced both quantitative and qualitative data and was focused on girls in the Sixth Form, across academic disciplines. Although AfL comprises four elements, the research only targeted student questioning.

There were several key findings in this research in the sense that some of the evidence generated was stronger than anticipated and part of the research explored new ground - namely the effect of AfL on Sixth Form examination results. Most of the empirical evidence gathered by Black and Wiliam (2002) and the Assessment Reform Group involved KS3 and KS4 classes in Mathematics and Science and did not cover A Level teaching. This research also provided an opportunity to explore the impact of AfL in a Humanities subject.
5.2 Appraisal of Significant Findings

When examining the results of the questionnaires, and subsequent interviews, the case for introducing AfL across the school was strong, but by no means conclusive. Notwithstanding some of the positive effects of the pilot, it was clear from the evidence gathered from the students that, at present, the students’ classroom experiences were not commensurate with the way in which students wished to be taught. The most frustrating experience of the students was the use of ‘shouting out’, that prevented them from getting involved in the lessons and increased the chances of them becoming excluded and disengaged. Although students of lower relative ability enjoyed the sanctuary of this practice, it is these students who need to be more engaged, having continual interaction with the teacher; such avoidance strategies will clearly affect the effectiveness of A Level lessons and may potentially damage long term attainment.

When examining the way in which students view certain types of question, and the way in which they are asked, there were clearly reservations concerning the changes already made through the pilot. There was a clear conflict between what students liked and what they knew was educationally beneficial for them; students did not like rich questioning and had an even greater aversion to increased student participation i.e. the ‘no hands up’ rule.

At times, students had displayed passionate, yet negative, feelings towards any practice that put them on the spot, with recurring themes ranging from fear of failure and low confidence to embarrassment and low self-worth. The antipathy towards both of these strategies was also affected by ability, identified by the students’ GCSE scores.
Although GCSE score was a fairly blunt instrument, it allowed the year groups to be split according to academic attainment. Again, the evidence here was much stronger than anticipated and suggested that students with lower GCSE scores would find the transition to AfL questioning much more difficult. However, despite this, students did understand the benefits of such strategies, which did augur well for further AfL use; it is these issues that will be key to implementing a successful AfL strategy.

The two remaining AfL strategies – wait time and use of wrong answers - were better received. Wait time was the most successfully witnessed method, which was both used effectively by teachers and fully acknowledged by students. This was extremely encouraging as it created a culture suitable for fostering both rich questioning and teacher-student dialogue. The use of wrong answers was also popular with students in the pilot; it generated some very positive comments, but unfortunately it is hardly used by staff.

Although only analysing one subject with a relativity small sample of students (n=30), the use of AfL does seem to overwhelmingly indicate performance improvements as suggested by Black and Wiliam’s (2002) KMOFAP study. They suggested that AfL use increases GCSE attainment up to 1 ½ grades, but in the KMOFAP experiment all AfL methods were implemented. Therefore, as this study only focused on questioning, it does reinforce its importance. In all comparisons, whether contextualising pairs through identical GCSE results or TDA score, use of AfL questioning consistently delivered higher UMS than those witnessed in the control group.
This not only reflected greater performance *per se*, but also showed increased understanding, particularly those with 90%+ UMS (equivalent to an A* grade at A2). The causation witnessed was much stronger than anticipated, not only by the degree of increase in UMS, but also the frequency; UMS scores increased in all comparative tests, dramatically in some cases, and even increased where a student exposed to AfL possessed a *lower* GCSE score.

However, in social science research it is impossible to completely control the conditions and, as a result, there are always other factors that could corrupt any experimentation, weakening causation. However, due to the scale of the improvements observed there does seem to be prodigious academic merit in using AfL and there is no question that it should be adopted throughout the school.

5.3 Recommendations to SMT: How to Manage AfL with Staff and Students

Based on the results of the experiment, there is clear academic merit in embracing a whole school AfL policy. However, based on the reactions of the Sixth Form, rather than implementing *all* aspects of AfL, it would be sensible to administer change in phases. The advice would be to adopt questioning in the first year and then, after careful evaluation and reflection, gradually introduce the other elements of AfL - *Feedback, Sharing Criteria and Self-assessment.*
The exam improvements are only part of the picture at this fee paying, product focused school. Examination results and university entrance will remain paramount and it is therefore even more important to pay attention to the reactions of students to the changes made, whilst also placating other stakeholders - such as parents and governors.

One of the main challenges of implementing change, based not only on the evidence witnessed here, but also through an extensive review of the literature, will be the required culture shift; students and staff will need time to adapt to a more process-driven environment where students take ownership of their learning. Convincing students of their own abilities and that of their peers, along with changing the philosophy away from the ‘teacher knows best’ to empowering students, will take patience and perseverance.

Changes to teaching and learning pedagogy will be made even more difficult due to expectations of spoon-feeding and an over-reliance on staff to ‘do all the work’. However, over time the values and beliefs of both students and teachers should change and those coming up the school should find that the various questioning methods become second nature. Over time, they will hopefully begin to fully appreciate the educational benefits and witness an increase in confidence and self-esteem i.e. worrying less about being wrong and spending more time in the limelight.

In order to circumvent some of these problems – at least in the short term - the school should implement AfL, but do so vigilantly. The eight-point plan overleaf is to be proposed to the School’s SMT:
**CHANGE INVOLVING STUDENTS**

- Provide students with a detailed explanation of how and why AfL works, providing evidence from previous studies, including the performance benefits witnessed here.

- Pay particular attention to students who are struggling – by using established AfL techniques, such as traffic lights, and support weaker students by pairing them with stronger students.

**CHANGE INVOLVING STAFF**

- Ensure that teachers understand why AfL is being implemented and how effective questioning improves pupil achievement, whilst not creating work for them.

- Develop teacher skills in targeting questions to challenge pupils’ understanding and ensure they prompt students to explain and justify their answers individually or in small groups to boost confidence e.g. use of hot seating and matched pairs etc.

- Set up an AfL committee to discuss how AfL is progressing and allow teachers to learn from each other and keep abreast of new initiatives.

- Establish a collection of AfL leaders (from the pilot or otherwise) that share good practice providing INSET and continual support and advice to staff.

- Increase the use of observations by SMT to ensure practice is consistent and also through peer assessment to allow colleagues to share good practice.

- Actively promote the use of wrong answers in all lessons and provide more training and allow time for teachers to use this method more often.
5.4 Scope for Further Research

The AfL questioning experiment catalogued here will now be expanded to include the control group (Economics Group 12B) who, up until the time of writing, had not had any exposure to any AfL questioning methods. It will be interesting to record any correlation here, comparing the overall A Level results (AS + A2) of the original control group with that of the original experiment group.

The experimental element of the study only focused on AS Level Economics. Although there were significant findings, it would have been much more effective (and potentially more reliable) to have used a larger scale experiment and control group, over a range of subjects in the pilot. This should now be the focus of further investigation and could determine if AfL success is more likely in some A Level subjects than in others.

The study was necessarily narrow as it could only include the effects of AfL questioning on female students. Although not possible here, it would be fascinating to examine the effects of AfL questioning on male students. The study could not only examine possible differences in performance, but also the way in which boys take to, and feel about, increased student participation, use of rich questioning etc., compared to the girls.
In due course, there may also be scope for research into the benefits of the other three elements of AfL. Given the culture of examination results and focus on grades, the most interesting experiment would be how feedback - or more specifically ‘comment only marking’ - is received, as well as measuring any performance benefits. As indicated, this research has already shaped future opportunities for action research - studying in more depth, the effects of the changes made introducing all AfL principles.
References


Estyn (2009), School Inspection Report

Evans, D. (2012) Telephone conversation with the Secretary of HMC, 14th March


Appendices
### Appendix I (a) – Research Timeline

<table>
<thead>
<tr>
<th>Research Activity</th>
<th>Date</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brainstorm ideas</td>
<td>5&lt;sup&gt;th&lt;/sup&gt; October</td>
<td>Took far too long to hone</td>
</tr>
<tr>
<td>Finalise ideas (shortlist)</td>
<td>11&lt;sup&gt;th&lt;/sup&gt; October</td>
<td>Very onerous process</td>
</tr>
<tr>
<td>Select Title and generate initial Research Questions</td>
<td>10&lt;sup&gt;th&lt;/sup&gt; November</td>
<td>Spent far too long on Research Questions</td>
</tr>
<tr>
<td>Start Literature Review</td>
<td>1&lt;sup&gt;st&lt;/sup&gt; December</td>
<td>Couldn't find criticism of AFL</td>
</tr>
<tr>
<td>Devise Methodology</td>
<td>18&lt;sup&gt;th&lt;/sup&gt; February</td>
<td>Difficulty in deciding on model</td>
</tr>
<tr>
<td><strong>Workload Delayed Dissertation - But Milestone Achieved</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Devise questionnaire</td>
<td>5&lt;sup&gt;th&lt;/sup&gt; April</td>
<td>Difficult – took longer than anticipated</td>
</tr>
<tr>
<td>Give out questionnaire to pilot audience (students)</td>
<td>30&lt;sup&gt;th&lt;/sup&gt; April</td>
<td>Random sample</td>
</tr>
<tr>
<td>Give out questionnaire to pilot audience (staffs)</td>
<td>4&lt;sup&gt;th&lt;/sup&gt; May</td>
<td>Only one amendment</td>
</tr>
<tr>
<td>Modify questionnaire (in light of pilot)</td>
<td>6&lt;sup&gt;th&lt;/sup&gt; May</td>
<td>Only a few amendments</td>
</tr>
<tr>
<td>Complete balanced Literature Review</td>
<td>10&lt;sup&gt;th&lt;/sup&gt; May</td>
<td>Still no real criticism found - only delivery</td>
</tr>
<tr>
<td>Start to collect and analyse data</td>
<td>18&lt;sup&gt;th&lt;/sup&gt; May</td>
<td>Some questions ended up being N/A</td>
</tr>
<tr>
<td>Transpose questionnaire to main Excel spreadsheet</td>
<td>30&lt;sup&gt;th&lt;/sup&gt; May</td>
<td>Took longer than expected</td>
</tr>
<tr>
<td>Convert questionnaire data to bar charts</td>
<td>5&lt;sup&gt;th&lt;/sup&gt; June</td>
<td>Smoother process than expected</td>
</tr>
<tr>
<td>Fine tune/add to Methodology</td>
<td>18&lt;sup&gt;th&lt;/sup&gt; June</td>
<td>Methodology changed ex-post results</td>
</tr>
<tr>
<td>Compile final Methodology</td>
<td>26&lt;sup&gt;th&lt;/sup&gt; June</td>
<td>Focus on methods and model</td>
</tr>
<tr>
<td><strong>Workload Caused Severe Delay - Milestone Not Achieved</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Give out questionnaire to staff</td>
<td>2&lt;sup&gt;nd&lt;/sup&gt; July</td>
<td>40 given out - 30 received back</td>
</tr>
<tr>
<td>Revise and amend Research Questions</td>
<td>4&lt;sup&gt;th&lt;/sup&gt; July</td>
<td>Reduced to three RQs – more focus</td>
</tr>
<tr>
<td>Start to write up findings in Results Section</td>
<td>14&lt;sup&gt;th&lt;/sup&gt; July</td>
<td>Difficulty expressing qualitative data</td>
</tr>
<tr>
<td>Analyse AS Level exam data</td>
<td>20&lt;sup&gt;th&lt;/sup&gt; August</td>
<td>Late, but required 2011 AS Level results</td>
</tr>
<tr>
<td>Finalise Results and Analysis section</td>
<td>29&lt;sup&gt;th&lt;/sup&gt; August</td>
<td>Some significant findings discovered</td>
</tr>
<tr>
<td>Complete Conclusions and Recommendations</td>
<td>4&lt;sup&gt;th&lt;/sup&gt; September</td>
<td>Included scope for further research</td>
</tr>
<tr>
<td>Final improvements</td>
<td>20&lt;sup&gt;th&lt;/sup&gt; September</td>
<td>Ended up being rushed</td>
</tr>
<tr>
<td>Check for continuity</td>
<td>28&lt;sup&gt;th&lt;/sup&gt; September</td>
<td>Conflict with tenses</td>
</tr>
<tr>
<td>Final grammar and spell check</td>
<td>2&lt;sup&gt;nd&lt;/sup&gt; October</td>
<td>A few errors found and corrected</td>
</tr>
<tr>
<td>Print and bind</td>
<td>7&lt;sup&gt;th&lt;/sup&gt; October</td>
<td>Print shop – soft binding</td>
</tr>
<tr>
<td><strong>Submit</strong></td>
<td>7&lt;sup&gt;th&lt;/sup&gt; October</td>
<td></td>
</tr>
</tbody>
</table>
Appendix I (b) – Dissertation Milestones

1. **Title**
   - Start: Oct 05, 10

2. **Rationale**
   - Start: Oct 11, 10

3. **Research Statement**
   - Start: Oct 20, 10

4. **Research Questions**
   - Start: Oct 30, 10

5. **Milestone 1**
   - 11th November

6. **Updated RS post Tutor comments and added in more detail including literature references**

7. **Updated RQs after reflection of research goal and discussions with DoIT&L and Tutor**

8. **Milestone 2**
   - 5th January

9. **Updated Lit Review to include Critique and Reflection after Tutor consultation**

10. **Milestone 3**
    - 4th March

11. **Updated Methodology to include Reliability and Triangulation methods. Also, explain methods for each RQ in more detail**

12. **Milestone 4**
    - 24th June

13. **Results and Conclusions**
    - Delayed due to workload - results collected but not analysed. Results completed and emailed to Tutor

14. **Milestone 5**
    - 8th September

15. **Extra and Checking**

16. **Final checking post review by Tutor and critical friend completed 30th Sept. Final version bound and submitted 8th Oct**
Appendix II – Graphical Data

Have students noticed any change in questioning

Year 12
- Yes: 35%
- No: 65%

Year 13
- Yes: 21%
- No: 79%

Figure A1: Impact of Changes (n=80/n=82)

Student awareness of AfL: familiarity with the term

- Explained by Teacher: 10%
- Used by Teacher: 20%
- Neither: 68%

Figure A2: Knowledge of AfL (n=80)
Student Questioning - Reported Experiences

Figure A3: Student Experience of Questioning (n=80)
NB: Students could choose more than one response
Percentage Responses (%)

Teachers' Use of AfL Questioning

Figure A4: Current Use of AfL Questioning Techniques (n=30)
Percentage Responses (%)
Figure A5: Desirable Wait Time – Staff vs. Students (n=80/n=30)
Appendix III – Questionnaires

Sixth Form Student Questionnaire May 2011
The Use of Questioning in Class

1. Please list the AS Level subjects you are studying:

   1 ............................................................
   2 ............................................................
   3 ............................................................
   4 ............................................................

2. Please summarise your GCSE results e.g. 2A*s 6As 2Bs ........................................

Questions 3-6 relate to your experience of questioning since being in the Sixth Form

3. Have you noticed any difference in the way some of your teachers have been asking questions in class? Please circle Y/N If you’ve answered yes, what has changed?

   ........................................................................................................................................
   ........................................................................................................................................
   ........................................................................................................................................
   ........................................................................................................................................
   ........................................................................................................................................
   ........................................................................................................................................
   ........................................................................................................................................

4. Have your teachers explained or used the term Assessment for Learning (AfL)?

   Explained ☐
   Used ☐
   Neither ☐

5. Listed below are some of the reasons why teachers ask questions in class. Could you please rank the reasons in order you feel is most important to your learning 1-4 (1 being the most important and 4 being least important):

<table>
<thead>
<tr>
<th>Reason</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>To test what you know</td>
<td></td>
</tr>
<tr>
<td>To show you what you need to do next to advance your learning</td>
<td></td>
</tr>
<tr>
<td>To make you think more deeply</td>
<td></td>
</tr>
<tr>
<td>To encourage you to participate more actively in the lesson</td>
<td></td>
</tr>
</tbody>
</table>
6a. Below are some types of questions. In the first column, could you please indicate your preference by ranking each 1 - 5 (1 being the type of question you like best and 5 being the type of question you like least). In the second column, could you please also rank 1 - 5 this time on how useful the type of question is to your learning (1 being the most useful and 5 being the least useful):

<table>
<thead>
<tr>
<th>Type of Question</th>
<th>Like</th>
<th>Usefulness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Requiring one word answers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Can you explain? Give me a definition of.....etc</td>
<td></td>
<td></td>
</tr>
<tr>
<td>What would happen if.....?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Requiring 'Yes' or 'No'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Why.....?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6b. For types you've ranked 1 and 5 for 'Like', could you please give your reasons for doing so?

1. ................................................................................................................

5. ................................................................................................................

Questions 7-9 mostly relate specifically to each of your AS Level subjects. At the top of each column, please specify the A Level subject and then answer the respective questions in each case.

7a. This question is about wait time i.e. how much time the teacher waits before expecting an answer, before moving on. Please answer the following with Y/N in each case.

<table>
<thead>
<tr>
<th>AS Level Subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Are you ever given thinking time?</td>
</tr>
<tr>
<td>Are you ever allowed to discuss an answer with a friend before giving it?</td>
</tr>
</tbody>
</table>

7b. How do you feel if a teacher moves on to someone else or gives the answer themselves before you have had time to think?

..................................................................................................................................................................................................................................................................................................................................................................................
7c. In your opinion, what is a reasonable amount of time for a teacher to leave before expecting an answer to a challenging question?

---

8a. The question below is concerned with how questions are handled in class. Could you indicate (by placing a tick) how teachers usually choose students to answer questions in class:

<table>
<thead>
<tr>
<th>Question Management</th>
<th>AS Level Subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chooses someone with their hand-up</td>
<td></td>
</tr>
<tr>
<td>Names a pupil before asking the question</td>
<td></td>
</tr>
<tr>
<td>Names a pupil after asking the question</td>
<td></td>
</tr>
<tr>
<td>Pupils shout out</td>
<td></td>
</tr>
<tr>
<td>Uses a method of random questioning</td>
<td></td>
</tr>
</tbody>
</table>

8b. Listed below are the same questioning methods as above. This time, please indicate those you prefer by placing numbers 1-5 (1 being your most favourite and 5 being your least favourite)

<table>
<thead>
<tr>
<th>Question Management</th>
<th>Preference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chooses someone with their hand-up</td>
<td></td>
</tr>
<tr>
<td>Names a pupil before asking the question</td>
<td></td>
</tr>
<tr>
<td>Names a pupil after asking the question</td>
<td></td>
</tr>
<tr>
<td>Pupils shout out</td>
<td></td>
</tr>
<tr>
<td>Uses a method of random questioning</td>
<td></td>
</tr>
</tbody>
</table>

8c. For methods you've ranked 1 and 5, could you please give your reasons for doing so?

1. ........................................................................................................................................

........................................................................................................................................

........................................................................................................................................

5. ........................................................................................................................................

........................................................................................................................................

........................................................................................................................................
The Use of Questioning in Class Assessment for Learning (AfL)

Staff Questionnaire July 2011

Subject Taught:

1. Have you, personally received any training on Assessment for Learning (AfL)?  Y/N

2. To what extent have you used the AfL Questioning principles outlined below, this year? Please circle the most applicable, or delete as appropriate:

<table>
<thead>
<tr>
<th>Use of Wait Time</th>
<th>Always/Often/Sometimes/Never</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Hands Up (Random Questioning)</td>
<td>Always/Often/Sometimes/Never</td>
</tr>
<tr>
<td>Rich Questions</td>
<td>Always/Often/Sometimes/Never</td>
</tr>
<tr>
<td>Use of Wrong Answers</td>
<td>Always/Often/Sometimes/Never</td>
</tr>
</tbody>
</table>

3. Listed below are some possible reasons teachers ask questions in the classroom. In the first column could you please indicate (with a tick ✓) whether each of the following applies to your teaching? If you have ticked a particular purpose, could you also number the extent to which you feel it is important (1 being of high importance and 4 being of low importance).

<table>
<thead>
<tr>
<th>Purpose of the Question</th>
<th>Do you use?</th>
<th>Importance?</th>
</tr>
</thead>
<tbody>
<tr>
<td>To establish an orderly start to the lesson</td>
<td></td>
<td></td>
</tr>
<tr>
<td>To check recall of facts from a previous lesson</td>
<td></td>
<td></td>
</tr>
<tr>
<td>To stimulate pupils to ask questions of themselves and of others</td>
<td></td>
<td></td>
</tr>
<tr>
<td>To check understanding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>To develop reflection and comment by pupils on the group’s responses</td>
<td></td>
<td></td>
</tr>
<tr>
<td>To establish existing knowledge in order to proceed to the next stage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>To diagnose specific difficulties</td>
<td></td>
<td></td>
</tr>
<tr>
<td>To arouse interest and curiosity concerning a topic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>To inform yourself of the next stage in a pupil’s learning</td>
<td></td>
<td></td>
</tr>
<tr>
<td>To provide an opportunity for pupils to assimilate and reflect upon information</td>
<td></td>
<td></td>
</tr>
<tr>
<td>To make them think more deeply</td>
<td></td>
<td></td>
</tr>
<tr>
<td>To encourage them to participate more actively in the lesson</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4. Do you prepare specific questions for your lessons in advance? Always/Often/Sometimes/Never

5. When asking questions, do you allow students thinking time? Always/Often/Sometimes/Never
6. What would you consider to be a **reasonable amount of time** to give a student to consider a fairly challenging verbal question?  

7. Please tick which of the following strategies you use when choosing **who will answer** questions in class?

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Always</th>
<th>Often</th>
<th>Sometimes</th>
<th>Never</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asking an individual pupil by name (before asking q)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asking an individual pupil by name (after asking q)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Choosing a pupil who has his/her hand up</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Choosing a pupil who does not have his/her hand up</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A no-hands up policy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Talking partners</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hot-seating</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Traffic lights</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Giving pupils time to write answer before reading out</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other (please specify):</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*THE FOLLOWING QUESTIONS ARE FOR STAFF INVOLVED IN THE AFL PILOT*

9. To what extent do you feel AFL questioning techniques will be successful in increasing the **understanding** of A Level students, as well as improving their knowledge? Please rank 1-4 (1 being a great deal and 4 being no effect whatsoever)

   1  2  3  4

10. To what degree do you think AFL questioning techniques will improve students’ **enjoyment of lessons**? Please rank 1-4 (1 being a great deal and 4 being no effect whatsoever)

   1  2  3  4

11. How do you think the introduction of AFL questioning will increase student **independence and criticality**? Please rank 1-4 (1 being a great deal and 4 being no effect whatsoever)

   1  2  3  4

13. To what extent do you think use of AFL questioning in school will increase the **overall academic performance** of A Level students? Please rank 1-4 (1 being a great deal and 4 being no effect whatsoever)

   1  2  3  4
The Use of Questioning in Class

Student Interview Questions

Have your teachers ever used other students to expand or improve your answer? Y/N
If yes, how do you feel about this? Does it help your understanding of difficult topics?

How do you feel when you get an answer wrong? What, if anything, do you learn from a wrong answer?

Considering the main changes in Questioning this Year in (insert relevant subject) which did you find the most beneficial?

- Student Participation
- The Use of 'Wrong' Answers
- Rich Questioning
- Wait Time

Given the changes made this year, could you please indicate whether you strongly agree (1), agree (2), disagree (3), or strongly disagree (4) with the following statements?

Use of ATL will:

- Improve your Understanding
- Increase your Enjoyment of Lessons
- Enhance your Independence
- Boost your Academic Performance