An examination of how teachers operationalise formative and summative assessment with students against a backdrop of high stakes testing.

By Jennifer Peadon

A Thesis Submitted for the Degree of Doctor of Education

School of Education
Durham University
2020

Abstract

Assessment has always had a fundamental role to play in education and is carried out in educational establishments to serve a range of purposes. This study examined how teachers in a Sixth Form College in the North East of England operationalise formative and summative assessment with their students in light of the emphasis placed upon high stakes summative testing in the English education system. High stakes testing is frequently associated with a range of unintended consequences, namely the narrowing of the curriculum, teacher-centred pedagogy, and teaching to the test, as students become passive recipients of knowledge in the learning process. In this enquiry, a case study approach is adopted which incorporates 30 lesson observations across a range of A Level and BTEC subjects and 2 focus group interviews with teaching staff. Three key themes have been identified: formative and summative approaches to assessment and their associated challenges, student actions and responses and changing examination structures. Findings from this study suggest teachers value the embedding of formative assessment into day-to-day teaching practice which should be supported by regular summative assessment. Although time has been identified as an area of difficulty in this regard. Questioning has been highlighted as a key formative assessment tool while other formative assessment strategies such as peer assessment are less popular amongst staff. Teachers attach much value to the promotion of independence amongst their students, particularly following the transition to linear examinations. Furthermore, this research supports previous findings through the identification of a value-practice gap in which teachers work.

Contents

List of Figures:

List of Tables:

List of Appendices

Glossary:

AQA:	Assessment and Qualifications Alliance. An awarding body in England, Wales, and Northern Ireland. It compiles specifications and holds examinations in various subjects at GCSE, AS and A Level.
BTEC:	Business and Technology Education Council. They are designed for young people who have an interest in a particular sector or industry. BTECs are broken down into three main levels of study: 1. BTEC Firsts offer an introduction to work in a vocational sector. 2. BTEC Nationals are a similar standard to A Levels and are available at level 3. 3. BTEC Apprenticeships are available at levels 2-5.
Command Words	The words and phrases used in exams and other assessment tasks that tell students how they should answer the question, e.g., analyse, argue.
CPD:	Continuing Professional Development. This refers to the activity's professionals engage in to develop and enhance their abilities.
Eduqas:	A brand of the British examination board WJEC which was created as a response to differing education systems in England and Wales. From September 2015 all reformed qualifications offered by WJEC in England are branded as Eduqas, whereas those in Wales remain WJEC. Qualifications which were being delivered in England

	continued to be branded WJEC until they were reformed. Eduqas produce specifications and hold examinations at GCSE, AS and A Level.
EFA	Embedding Formative Assessment is a two-year programme based around embedding formative assessment into teacher's day-to-day practice, drawing on resources authored by Dylan Wiliam.
Kahoot:	A freely available, online gaming platform which can be used to generate games to support learning or revision. Students can participate using their phones.
Kerboodle:	A range of resources in digital format to support learning in class, examples include auto-marked tests, videos, animations, and podcasts.
Learning Walk	A brief classroom visit (usually by a member of the Senior Leadership Team) which provides an opportunity to reflect on teaching and learning strategies being used, student interaction with the lesson content and levels of student engagement. Learning walks usually take place at regular intervals throughout the school year to provide a snapshot over time.
Levelled (banded) questions:	A specific number of marks is attached to an expected outcome. For example, to score 1-2 marks (band 1) out of a possible 6, limited or fragmented knowledge may be demonstrated compared to 5-6 marks (band 3) where a much wider range of knowledge should be applied. A

	student must demonstrate that their answer successfully meets these outcomes to be awarded the marks.
Linear examinations:	Students follow a linear path of progression which sees them sit their examinations at the end of a programme of study.
Modular examinations:	Course content is broken down into units (modules). Students can sit examinations based on each unit twice yearly, in January and May/June, thus providing the opportunity for resits.
Properly marked:	The marking of a piece of work by a teacher as opposed to a fellow student (as part of peer assessment).
Quality marked:	The provision of detailed feedback. Teachers comment on positive aspects of student work before identifying areas for improvement.
SLT:	Senior Leadership Team. Within a school or college this term incorporates the Head Teacher, Deputy Head Teacher(s), and Assistant Head Teacher(s).
SSAT	Schools, Students and Teachers Network is an educational membership organisation which provides support and training to schools and teachers in four main areas: teaching and learning, curriculum, networking, and leadership development.
Upskilling	A term utilised by teachers in this study to denote students developing their academic skills in readiness for high stakes summative assessment.

WJEC:	Formerly the Welsh Joint Education Committee. An examination board providing examinations and educational resources to schools and colleges in Wales, England, and Northern Ireland.

1 Introduction

1.1 Background

Approaches to student assessment have been well documented in relation to secondary school classrooms in the UK. Teachers operationalise both formative and summative assessment strategies to meet a variety of learning goals, with the former the domain of the individual teacher and the latter often externally guided and implemented. Thus, the co-existence of formative and summative assessment, coupled with the pressures of high stakes summative assessments, can create assessment backwash. The focus of my research is centred on an exploration of how teachers operationalise formative and summative assessment with students in post 16 education, in the face of pressures exerted through a reliance on high stakes summative assessments in the English education system.

High stakes assessment can shape teaching and learning practices both positively and negatively, the term assessment backwash or washback is often used to describe this (Cheng and Curtis, 2012; Green, 2013). However, backwash effects are largely seen as negative (Biggs, 1998), promoting a surface approach to learning which in turn leads to a student using a lower level of cognition. In the event that test items correspond with curriculum objectives, the backwash effect is likely to be positive. Nevertheless, it can be argued that backwash resulting from high stakes testing is often viewed negatively as it is the assessment itself which defines what is taught in the classroom and ultimately the curriculum (Biggs, 1998). This in turn can lead to unintended consequences, such as the narrowing of the curriculum and greater emphasis on teacher-centred pedagogy (Au, 2007; Harlen, 2007) which will be discussed in greater detail in Chapter 3. In contrast, as will be explored in Chapter 2, formative assessment effects are often considered to be positive (Black and Wiliam, 1998) with feedback from formative assessment promoting a deep approach with students using a level of cognition appropriate to the task. Indeed, Biggs (1998) argues that "significant gains are to be found as much in mitigating or reversing backwash as by enhancing feedback" (106). Given the prevalence of high stakes testing in the English education system and teachers having to guide their

students towards external examinations, the extent to which formative assessment and its purported benefits are marginalised can be questioned.

This chapter aims to set the scene for my research through a consideration of my biography as a researcher and the path which led me to undertake study at doctoral level. This is significant as decisions which I have taken as part of this process have affected the shape of my thesis, borne out of my experiences. I will then discuss my research questions and begin to introduce the methodological approach I have adopted alongside the ethical considerations I have encountered. Finally, I will set out the structure of the remaining thesis chapters.

1.2 My Biography as a Researcher

I embarked upon the Doctorate in Education programme at Durham University on a part-time basis in October 2014 while employed full-time as a Teacher at a secondary school in the North East of England. I selected to study at Durham University after having completed both my PGCE and masters degree here in 2008 and 2011, respectively. Furthermore, I had made the decision that I did not want to remain in the teaching profession for much longer and wanted to complete doctoral level study to enhance my career options in the future. While balancing the demands of full-time work and part-time study was difficult, I was determined to pursue my studies and develop my skillset as a researcher. The Doctorate in Education has enabled me to research a specific issue which ignited my interest as a teaching professional in depth, this will be addressed in the following section where I address my motivations for undertaking this piece of research.

Following an extremely negative professional experience I took the decision to leave teaching in October 2016. This was a particularly difficult time, and I invested much time in applying for jobs. In November 2016 I secured a part-time position at a Higher Education Institution. Initially, the part-time nature of this position gave me time to focus on my thesis. Despite this, there was still a need for me to acquire further work to increase my income, at this point I was still applying for full-time jobs but having little success. I made the difficult decision to take on some supply work, this was a decision borne out of necessity as opposed to want. Finally, in October 2018 my perseverance paid

off and I secured a full-time job supporting students with their Academic Skills in Higher Education. Since taking up this role I have obtained Associate Fellowship of the Higher Education Academy. I have since secured a new full-time position working for an Awarding Organisation. This role enables me to utilise and apply my knowledge of assessment to the production of valid, robust, and reliable assessments to support a range of qualifications.

1.3 Motivation for my Research

As a teacher, the issue of assessment and students meeting their target grades had always been in my remit. However, as a Curriculum Leader it became an even more prominent feature of my day-to-day role. This was due to me no longer just being responsible for the performance of students within my classes, but also for the performance of students across the Department. One aspect of my role was to regularly update the Senior Leadership Team (SLT) on the progress of Year 11 students in relation to their target grades. At this time, we were all so focused on preparing students to sit their GCSE exams, to satisfy the demands placed on us by SLT. This subsequently led me to think about how I was preparing students for external high stakes summative assessments as part of my day-to-day practice. Furthermore, what seemed extraordinary to me at this time was the commitment made by the institution to focusing on formative assessment as part of staff Continuing Professional Development (CPD) through the Embedding Formative Assessment (EFA) programme provided through the SSAT (Schools, Students and Teachers Network). EFA is a two-year programme based around embedding formative assessment, drawing on resources authored by Dylan Wiliam. This enabled staff to share best practice with colleagues and reflect upon approaches to formative assessment. A number of formative assessment strategies were selected by the Senior Leadership Team within my institution, as shown in Figure 1.1 below, and subsequently shaped my experiences of the programme: “providing feedback that moves learner’s forward” and “activating students as owners of their own learning” (SSAT, 2019). This enabled me to provide pupils with more time to reflect on their learning (by allocating time for pupils to respond to teacher comments and make improvements to their work) and their progress and ultimately take

the necessary steps to close the gap in their learning. In addition, this approach encouraged a movement away from marking with grades, instead emphasising the value of teacher feedback.

Figure 1.1 - Five Key Strategies of the Embedding Formative Assessment Project (Wiliam, 2011)

<table>
<tr><th></th><th>Where the learner is going</th><th>Where the learner is</th><th>How to get there</th></tr>
<tr><td>Teacher</td><td rowspan="3">Clarifying, sharing and understanding learning intentions</td><td>Engineering effective discussions, tasks and activities that elicit evidence of learning</td><td>Providing feedback that moves learners forward</td></tr>
<tr><td>Peer</td><td colspan="2">Activating students as learning resources for one another</td></tr>
<tr><td>Learner</td><td colspan="2">Activating students as owners of their own learning</td></tr>
</table>

This led me to think about the extent to which teachers were utilising such formative assessment strategies as part of their day-to-day teaching when faced with the pressures of external summative assessments. I also considered whether formative assessment was simply a token effort. While the concept of formative assessment was deemed to be important, I questioned whether there was a disparity between what teachers themselves value and what they undertake in their classrooms to prepare their students for external summative assessments. This initial thinking has provided the basis for the development of my research. This ultimately led me to base my study around teachers' use of formative and summative assessment with students at two different stages of post-16 study, Years 12 and 13. I constructed my research around the following four research questions:

1. How do teachers use formative and summative assessment strategies with their students and why do they use them?
2. How does high stakes testing affect teachers' approaches to assessment?
3. How do teachers use formative assessment to prepare their students for the summative?

4. What are teachers' beliefs about assessment strategies? Is there a relationship between teacher values and day-to-day practice?

My experiences of the Embedding Formative Assessment Project (Wiliam, 2011) led me to consider whether teachers were applying and utilising formative assessment strategies as a means to move their students forward in their learning or having to side-line this approach due to the demands of high stakes testing. As will be explored further in Chapter 3, high stakes summative assessment in the English education system is closely linked to accountability (West, 2010), and examination results can determine students next steps, be it access to employment or further or higher education which it can be argued further increases pressures to succeed. Given this emphasis on accountability, the performance of schools and ultimately their teachers and students, I considered it pertinent to investigate whether high stakes testing affects teachers approaches to assessment. Indeed, it could be argued that as schools are subject to evaluation using a range of measures, such as Progress 8, teachers face pressures to meet such targets.

Alongside my professional teaching experience, my reading and reviewing of the assessment literature, which is discussed in greater detail in Chapters 2 and 3 respectively, has also informed and shaped the development of my research. Whether teachers integrate formative assessment into their day-to-day teaching, an approach which has been associated with positive outcomes for students (Black and Wiliam, 1998) is particularly significant given the unintended consequences which can result from simply teaching to the test (Au, 2007). Ingram, Elliott, Morin, Randhawa, and Brown (2018) argue that such approaches do not always best meet the needs of students. In the literature, some researchers argue that formative and summative assessment should be separate entities (Simpson, 1990) whereas others suggest that they are inextricably linked (Bennett, 2011; Winterbottom *et al,* 2008) with the formative summative distinction being somewhat blurred. However, this led me to consider whether formative assessment could have a role to play in supporting learning as teachers prepare their students for the summative.

Following my exploration of the assessment literature, particularly the work of James and Pedder (2006) and Winterbottom *et al,* (2008), who discuss the idea of a value-practice gap between teachers' ideals and values and what they are able to achieve on a day-to-day basis, I sought to determine the extent to which a value-practice gap in relation to teachers' use of assessment strategies was in existence.

1.4 The Research Context

My study has been undertaken in a mainstream, mixed Sixth Form College in the North East of England, which has been given the pseudonym of Mitford College for the purposes of anonymity. The student cohort of Mitford College is predominantly white British and the proportion of students with Special Educational Needs (SEN) and/or disabilities is lower than the national average. There is also a lower percentage of students from disadvantaged backgrounds than the national average. Students arrive at this institution having just completed high stakes summative assessments, their GCSEs, which culminate in the conclusion of their secondary school career. This establishment has a good reputation for study and students travel to attend from a host of other areas within the North East.

Students enter the college in Year 12 and embark upon a new programme of study, either A Level or BTEC courses, their entrance to which is contingent upon them meeting the entry requirements, both at the institution and subject-specific level. This includes achieving both Maths and English at grade 4 or above. Those who have not successfully passed these qualifications upon entry are required to take resits alongside their A Level or BTEC courses. Therefore, their performance in high stakes summative assessments at GCSE level has the potential to immediately affect their educational opportunities post-16. Moreover, high stakes testing forms the basis of much assessment at post-16 level, the outcomes of which can have a direct impact upon a students' future. It could be argued that this has been compounded by the transition to linear examinations, which is discussed in Chapter 5.

As will be explored in Chapters 2 and 3, the literature contains a number of theories and ideas about assessment and it is fair to say that there is much agreement and disagreement in this regard. Before discussing my research further, I think it is pertinent to clarify the meanings I intend to convey when I refer to particular concepts throughout my thesis. When I talk about formative assessment (or assessment for learning) I am referring to assessment carried out by teachers themselves in their classrooms as part of their day-to-day practice, as will be discussed later, this can take many forms. When discussing summative assessment (or assessment of learning) I will be referring to both internal and external forms of summative assessment with the former relating to teachers' use of summative assessment to check students' progress or for internal reporting purposes. With regard to external summative assessment, I am discussing those assessments which students sit at the end of a two-year programme of study, such assessments have high stakes attached as they help to determine students next steps in their educational career as well of the success of Mitford College against performance metrics and in league tables.

1.5 The Research Design

Securing access to a research site raised several important ethical considerations. Upon leaving the teaching profession I was potentially faced with the difficulties other outside researchers face in terms of gaining access to a site to study and negotiating with gatekeepers. While working in a school meant that I had access to a sample population, I have simply been required to engage in a different process from what I was initially expecting to negotiate access. Instead, I successfully used my contacts to gain access to a site to carry out my research, thus, removing any difficulties associated with gaining access to gatekeepers. Ethical issues are a particularly prominent part of the research process (Brooks, te Riele and Maguire, 2014). Consequently, an important part of my research design has encompassed a consideration of how to mitigate any potential ethical dilemmas to safeguard participants in my research. My approach to ethics has been guided by the regulatory framework of BERA (British Educational Research Association) and their associated principles of respect and responsibility. Initially, I obtained ethics approval from the Ethics Committee at Durham University. Upon gaining access to my research site, I provided participants

with a copy of both my ethics and research proposals, I therefore sought to use my positionality as a researcher to give participants as much awareness as possible of what my research entailed. I also informed participants of their right to withdraw at any point in the process and made available to participants any documents which I produced to enable them to express any concerns they may have had. Furthermore, the identify of both the research site and participants has been protected though the use of pseudonyms throughout this thesis.

As will be discussed in greater depth in Chapter 4, I have selected a case study approach for my research as this enabled me to place several boundaries on the case. Boundaries have taken the form of the context, one educational institution, the focus, encompassing teachers use of formative and summative assessment with students in Years 12 and 13 and data collection techniques, consisting of observation, focus groups and document analysis. Furthermore, a case study approach is the best fit for an inquiry that is situated within a single institution.

I have integrated multiple data sources into my research design as a form of triangulation. Furthermore, I aimed to generate a rich body of data to inform my understanding of the case and record data in the moment. Observation often constitutes a key element of case study research and I elected to use a structured non-participant approach alongside an observation schedule. This promoted a consistent approach across all observations which I carried out. This was supplemented using field notes which acted as an aide memoire. However, I took care to review my fieldnotes as soon after the observation as I could to ensure that detail was not forgotten (Hammersely and Atkinson, 2007). Field notes were subsequently transcribed and coded. Observation can be seen to compliment other research methods which aim to explore meanings behind observable behaviours.

Alongside observation I carried out focus group interviews to determine the rationale behind teachers' approaches to assessment. This approach also enabled me to collate data from multiple participants at the same time, thus mitigating against time pressures faced by participants. Moreover, focus groups provided participants with a forum to discuss their perspectives on assessment. During the focus group interviews I used an interview schedule to guide the topics of conversation but took opportunities to facilitate further discussion around issues of interest where

appropriate. In addition, the focus groups were audio recorded with the data later being coded and transcribed to identify key and recurrent themes. Further triangulation of data was also achieved through the analysis of documents, those published by examination boards such as marks schemes as well as those materials amassed during observations. This enabled me to gain further insight into teachers' use of assessment strategies with their students.

My experiences regarding formative assessment as a secondary school teacher have informed my positionality vis-à-vis the framework which underpins this thesis. For the purposes of this research, I subscribe to the position of providing feedback to students, as opposed to summative marks or grades. This can be attributed to my experience of engaging in a staff CPD programme, produced by Dylan Wiliam, entitled "Embedding Formative Assessment" (The Schools, Students and Teachers Network, 2019). This staff development programme actively encouraged the implementation of formative assessment strategies and provided staff with the opportunity to share best practice with colleagues and reflect on their approaches to formative assessment. My positionality in relation to the framework for this thesis therefore rests upon my personal experience of having first-hand witnessed the benefits of pupils improving their learning and making progress using teacher feedback.

1.6 Thesis Structure

I begin in Chapters 2 and 3 with a discussion and review of the literature on formative assessment, with a consideration of the work of a number of prolific researchers within the field of assessment, including Black and Wiliam (1998). Critiques of such works will also be explored, for example those of Dunn and Mulvenon (2009). I will then examine literature on summative assessment, particularly the role and purpose of high stakes testing and its effects on classroom assessment practice. I will also consider whether teachers work within a value-practice gap (James and Pedder, 2006). In Chapter 4 I present my research methodology, including a consideration of the research process itself, ethical issues will also be explored. Chapters 5 and 6 serve to present my findings through a discussion of key and recurrent themes identified through the process of analysis. In Chapter 7 I will

return to each of my research questions and provide an answer to each before finally in Chapter 8, presenting my final conclusions and reflections.

2 A Focus on Formative Assessment

Assessment has always had a fundamental role to play in education and is carried out in educational establishments to serve a range of purposes. Much of the literature is concerned with formative and/or summative forms of assessment. While in some cases they are viewed as discrete entities, there is evidence to suggest that they are inextricably linked and the lines between assessment for formative and summative purposes can become blurred.

This chapter aims to discuss and review literature sources on formative assessment. I begin with an examination of assessment purpose before moving on to consider a theoretical framework for formative assessment. Schools of thought relating to learning will be discussed with a particular focus on constructivist theory, upon which formative assessment is based. This is followed by discussion of how formative assessment is conceptualised within the literature and the issues associated with defining such an approach. The work of several prolific researchers and theorists in this field will be referenced, including Black and Wiliam (1998) whose seminal piece was regarded as a real turning point in the assessment literature. It is also pertinent to investigate the distinction made between formative and summative assessment in the literature and the links between them, particularly as the lines between formative and summative assessment can become blurred. Finally, this chapter will turn to explore examples of formative assessment, namely self/peer assessment, and teacher questioning.

It was Scriven (1967) who first put forward the idea of a distinction between summative and formative assessment. For Scriven (1967) and later also for Bloom (1969) summative assessment was concerned with determining the value of an educational programme whereas formative assessment was associated with programme improvement. However, some researchers suggest that summative forms of assessment can be used for formative purposes, for example Gipps (1996) states that “assessment for selection, monitoring and accountability can be assessment to support learning” (261) on the proviso that teachers are appropriately trained to use it for that purpose, and it involves the whole population as opposed to a sample. Despite this, others contest this and

instead suggest that one form of assessment cannot contribute to processes of accountability and the improvement of teaching and learning and favour the maintenance of a distinction between formative and summative assessment (Harlen, 2005; Simpson, 1990). Indeed, Taras (2005) contends that formative assessment is as a form of "good" assessment which teachers should be working towards, while summative assessment is a form of "bad" assessment which should be side lined in favour of formative assessment.

2.1 Why Assess? - The Purpose of Assessment

There is no single definition of assessment purpose, but there is a pretty broad consensus. Indeed, several researchers refer to three main purposes of assessment. According to Black (1998), these encompass supporting learning, reporting student achievements, and meeting the needs of public accountability. Newton (2007) coins the terms "judgement level", "decision level" and "impact level" to describe purposes of assessment. While Brown (2004, 2006) also cites three purposes which are concerned with moving teaching and learning forward, the accountability of students and the accountability of teachers and schools. Black (1998) argues that to enable effective teaching to take place, teachers require feedback relating to their students' performance so that difficulties can be identified, and remedial action implemented. Thus, using assessment formatively plays an important role in supporting learning insofar as teachers develop an understanding of their pupils' level of knowledge and understanding, can adapt their practice accordingly and implement appropriate interventions to aid pupil learning (Black, 1998; Dixson and Worrell, 2016). In addition, formative assessment cannot simply be bolted onto an existing scheme of work. Instead to provide appropriate support to students, it must be built into teaching plans. With so much emphasis placed upon summative tests, the ability of a teacher to regularly review and interpret pupil work to identify areas in need of attention may be in conflict with the demands placed upon them in the wake of summative assessment.

Newton's (2007) first purpose is concerned with the arrival at a "standards-referenced judgement" in the English education system, in other words, the allocation of grades in relation to performance. The "judgement level" is often referred to in official documentation such as those produced by the

Department for Education, purporting the purpose of a particular assessment in relation to the determination of a students' level of attainment in a subject.

The second purpose of assessment according to Black (1998) is that of reporting student achievements. As students move through their educational career and transfer to different classes and/or learning establishments, the responsibility for their learning moves with them to a new teacher. In such instances, information relating to student achievements should also be transferred to enable the new teacher to plan for the needs of each student. Newton's (2007) second purpose, the "decision level", relates to the way in which an assessment judgement is used, in particular what that enables, for example access to a programme of study. Newton (2007) deems this to be "the most significant usage of the term "assessment'" (150).

The third purpose is that of accountability (Black, 1998). Schools are required to provide evidence for public accountability purposes, which incorporates the "aggregation of the results of their pupils' performance in the various public examination systems which their pupils have to take for their own personal needs" (Black, 1998: 31). Nevertheless, such performance data alone is not enough to build a true picture of pupil performance. For example, schools whose pupils are from poor educational backgrounds in areas of social deprivation may publish results, which suggest that pupil performance is below national average. Therefore, it is important to consider other information such as achievements upon intake, catchment area and home backgrounds of pupils.

Newton's (2007) final purpose, the "impact level", is "the intended impacts of running an assessment system" (Newton, 2007: 150), for example one intended impact is that students are taught a common core in each subject. However, such impacts relate more specifically to the design of the assessment rather than the educational system within which it sits.

While the key purposes of assessment are interpreted in distinct ways, there are several common themes in the literature, namely those centred on assessment for accountability purposes and assessment to facilitate teaching and learning. For example, Barnes, Fives and Dacey, (2015) discuss assessment purposes to meet goals of teaching and learning and accountability. Research

conducted by Remesal (2007) led to the identification of four dimensions of assessment which belong to a continuum of assessment purpose. This continuum ranges from the extremes of the pedagogical end to the accountability end. From the perspective of the pedagogical end, formative assessment is deemed to be a tool to improve teaching and learning. In contrast, the accountability end, a direct result of the outputs of summative assessment such as examination certificates and league tables, "is merely seen as an instrument of social control, as a means to certify the students' final results and thus, it is seen as a way of exposing to the public the teachers' professional activity" (Remesal, 2007: 31).

2.2 Formative Assessment

2.2.1 A Theoretical Framework for Formative Assessment

Prior to a consideration of how formative assessment plays out in a teaching and learning context, it is pertinent to provide a theoretical context upon which to base the impending discussion in this chapter. Learning theory has traditionally been rooted in psychology, however, in recent years the discussion has occurred in other areas, namely biology and social science, but also neuroscience, anthropology and sociology. Within these areas exist differing schools of thought relating to learning. Up until the 1980s, behaviourist psychology dominated research on learning, with other, conflicting schools of thought coming to the fore. Indeed, the movement to a constructivist paradigm was championed by several researchers (Brown, 1994; Cooper, 1993; Scheurman, 1995). Formative assessment rests upon a constructivist theory of learning, as opposed to the more traditional model predicated upon behaviourism, where the teacher is responsible for passing on knowledge. In constructivism, emphasis is placed upon the students' role in the learning process and this area of research builds upon the work of Piaget and Vygotsky (Müller, Carpendale and Smith, 2009; Vygotsky, 1978). Arguably, this shift to a constructivist framework underpins the National Curriculum, introduced by the Conservative Government in 1988 as part of the Education Reform Act. Teachers are encouraged to become facilitators and emphasis is placed upon "depth of understanding rather than superficial treatment of subject matter" (Hackmann, 2004: 698), which is in direct contrast to the approach often adopted by teachers to prepare pupils for high

stakes summative tests. Indeed, the work of Piaget emphasised the existence of cognitive constructivism with cognitive structures dominating children's thinking at differing developmental milestones. Development of Piaget's work has come to emphasise social constructivism based on "the premise that individuals must be socially engaged in learning, actively creating knowledge from their existing knowledge base, beliefs and personal experiences" (Hackmann, 2004: 697). Through their construction of knowledge, students take an active role in their learning as they seek "to find meaning in their experiences" (Boghossian, 2006). From a constructivist perspective, each individual experience is deemed to be as significant as the next, therefore "no one has an epistemically privileged viewpoint" (Boghossian, 2006: 714). As such, is not possible to determine what constitutes knowledge as knowledge can be interpreted differently as "no two people necessarily have the same constructions" (Boghossian, 2006: 714). Having the same constructions would be associated with inferences about the world which the constructivist perspective would not recognise.

The constructivist model has been subject to critique insofar as it is not easy to move away from the traditional teacher-centred approach predicated on the view that knowledge should only be transmitted by teachers to their students (Elkind, 2004) and develop one which is learner-centred. In addition, there is not a consensus regarding the constructivist view, although, there are some shared ideas. Firstly, prior knowledge is deemed to affect the learning process, "students must make connections between old knowledge and new information" (Bhattacharjee, 2015: 66) to enable meaningful learning to take place. Secondly, constructivism is concerned with active participation of learners in problem-solving and there is a focus on in-depth as opposed to superficial learning. In a constructivist classroom, the teacher takes on a facilitator role to "assist students in developing new insights, and to connect them with their previous learning" (Bhattacharjee, 2015: 70).

Social constructivism points to a growing emphasis on social learning contexts (Illeris, 2007). While the original focus has centred on the individual side of learning, conflicting theories pertain that learning does not only happen in the single individual. According to Illeris (2007), both

constructivism and social constructivism are required for learning to occur. Learning has an individual and a social side, however, neither constructivism nor social constructivism on an individual basis provide a full and "correct" understanding (Illeris, 2007). Illeris (2007) suggests that learning involves two processes, the interaction between individuals and their environment and acquisition. The former relates to social constructivist theory, which purports the societal context of learning which "provides impulses and sets the frames for what can be learned and how" (Illeris, 2007: 19), there are differences in the learning which occurs in school, working life and everyday life. The latter is concerned with constructivist theory and "typically has the character of a linkage between the new impulses and influences and the results of relevant earlier learning" (Illeris, 2007: 22).

2.2.2 Conceptualising Formative Assessment

Assessment has always had an integral role to play in classroom activities and is inextricably linked to teaching and learning. However, "our understanding of the three-way interaction between teaching, learning and assessment has been driven by contrasting theories, without universally accepted meanings of some of the concepts involved" (Gipps, 1996: 261), "formative assessment" and "feedback" are such examples. As previously discussed in Section 2.2.1, approaches to teaching, learning and assessment have been predicated on differing schools of thought which can be exemplified through the movement away from a behaviourist to a constructivist framework of learning, upon which formative assessment is based.

Black and Wiliam (2010) use the term assessment to encompass all activities which are carried out by teachers to elicit information to be used as feedback. The process only becomes formative when teacher instruction is modified as a result to meet the learning needs of pupils. Nonetheless, there is no consensus regarding an agreed upon term or definition of formative assessment. Bennett (2011) suggests that this is problematic as without being able to define an innovation, it becomes increasingly difficult to determine how effective it is. In addition, while Black and Wiliam (2010) advocate the use of formative assessment, they suggest that there is no one-size-fits all approach to its success. While all classrooms and pupils are different, it is possible to argue that this lack of

uniformity in approach could serve to compound understanding and expectations around what successful formative assessment involves. According to Dunn and Mulvenon (2009), this "directly contributes to the weaknesses found in the related research and dearth of empirical evidence" (2) as this impedes ones understanding of what is being studied and what the term formative assessment actually means.

Perie, Marion and Gong (2007) contend that assessment should be defined by its purpose. From this standpoint formative assessment relates to assessments used by teachers with the aim of adapting teaching and learning in contrast to interim assessment which serves to inform policymakers and educators. However, defining assessment in this way could be problematic as one assessment utilised by teachers and their pupils as part of the learning process could also be drawn upon by administrators to forge policy changes (Dunn and Mulvenon, 2009). It is also possible to argue whether the definitions necessarily support the intentions of formative assessment. For example, a test designed to provide formative feedback to students can only be described as formative in the event that the teacher uses it to give feedback to students. Definitions of assessment by use can also be problematic, while an assessment could be designed to be formative or summative, the extent to which an assessment is formative or summative is contingent upon "the actual methodology, data analysis, and use of the results" (Dunn and Mulvenon, 2009: 2). This can be exemplified through the work of Wininger (2005) who utilised a combination of formative and summative assessment, which he termed "formative summative assessment", to review exam performance with pupils. This involved the collection of both qualitative and quantitative feedback on their level of understanding of the test items. It can therefore be questioned where the 'formative' resides, be it in the tool or the use of the tool (Dixson and Worrell, 2016).

Dunn and Mulvenon (2009) suggest that the issue of assessment should be separated from the issue of evaluation, although they do acknowledge that it is a related concern. In English curriculum and policy discourse the term evaluation is often equated with quality assurance, while US curriculum discourse frequently refers to "evaluation", in England this is termed "assessment". They argue that while an assessment could be designed for one purpose, the resultant data could

be used either formatively or summatively; therefore, they employ the terms formative and summative evaluation. According to Dunn and Mulvenon (2009) summative evaluation relates to evaluating assessment-based data to establish the academic progress of pupils over a pre-determined period of time in relation to a series of identified criteria. In contrast formative evaluation is concerned with evaluating assessment-based evidence to enable the provision of feedback on teaching and learning (Dunn and Mulvenon, 2009).

2.2.3 The Formative Summative Distinction

Some researchers (Stiggins, 2006) utilise the term "assessment for learning" to denote formative assessment and "assessment of learning" to signify summative assessment. However, this terminology can also be argued to be problematic insofar as it does not refer to the potential for summative assessment to support learning and fails to acknowledge the complex nature of the relationship (Bennett, 2011: 7). As previously stated at the beginning of Chapter 2, some researchers advocate the use of summative assessment for formative purposes, "summative assessment should fulfil its primary purpose of documenting what students know and can do but, if carefully crafted, should also successfully meet a secondary purpose of support for learning" (Bennett, 2011: 7). In practice, the lines between formative and summative assessment can easily become blurred, it is necessary for teachers to get to grips with both formative and summative forms of assessment. Although, the distinction between them may not be easily identifiable in the classrooms of experienced teachers (Winterbottom *et al,* 2008).

This split between types of assessment goes one step further to incorporate those who consider formative assessment to constitute an instrument and those who deem it to be a process. Regarding the former, formative assessment, for example as a diagnostic test, will "typically produce one or more scores, often claimed to have "diagnostic value"" (Bennett, 2011: 6). In contrast to this, those who view formative assessment as a process are concerned with students' understanding and one key element relates to "when the [results are] actually used to adapt the teaching to meet student needs" (Black and Wiliam, 1998: 140). However, Bennett (2011) argues

that simply defining formative assessment as either an instrument or a process, is an oversimplification as be it instrument or process, it needs to be fit for purpose.

Alongside the problem of definition, the research literature cites problems relating to the provision of feedback, the following section will focus on the nature of feedback and the impact of awarding marks and grades. Black *et al* (2003) argue that feedback should be comment only as opposed to providing marks and grades alongside comments. "Grades, marks and levels do not provide information about how to move forward; any information is too deeply encoded" (Stobart, 2006: 142). When pupils are provided with feedback along with summative ratings, they are less likely to use the feedback formatively and as such do not benefit from the feedback process. In addition, comments can lack specificity and often occur multiple times in the same students' exercise book, therefore suggesting that they have not been acted upon (Black *et al*, 2003), and ultimately implying that the feedback cycle is not working as it is intended. When awarded marks and grades, pupils often compare themselves to their peers, but it is possible to argue that the current educational system where high stakes tests dominate, promotes a culture of comparison. Work carried out by Butler (1988) found that the greatest learning gains were attributable to comment only marking as opposed to marks or a combination of marks and comments. However, Carless *et al* (2011) argue that there is need for assessment tasks and the associated feedback and comments to be both well-aligned and timely to enable students to benefit from such feedback. Black and Wiliam (2010) contend that "feedback to any pupil should be about the particular qualities of his or her work, with advice on what he or she can do to improve and should avoid comparisons with other pupils" (84). This is all well and good, however teachers are under ever increasing pressure, from both internal and external sources to prepare pupils to succeed in summative (high stakes) examinations and tests, this can influence how teachers approach assessment in their classroom. It is therefore possible to question the extent that formative assessment practice is distorted by policy. Tiknaz and Sutton (2006) argue that "teachers tend to be heavily influenced by the need to produce summative performance data to track progress for management, not necessarily learning purposes" (328).

Given the current education system in England and the overwhelming focus on summative, high-stakes assessment, it is pertinent to note that such assessments constitute only one form of summative assessment. Rather, summative assessment can take different forms and they have differing consequences and impacts. Considering the emphasis placed on summatively assessed examinations it is important to question whether components of formative assessment could affect pupil performance in external examinations. McDonald and Boud (2003) examined the impact of training final year high school students in self-assessment practices on their performance in external examinations. Practical interventions took the form of teachers incorporating self-assessment training into their day-to-day practice. McDonald and Boud (2003) found that the implementation of self-assessment practices was seen as a positive undertaking by both staff and students and was helpful in preparing for exams. Regarding examination performance, self-assessment training had a considerable impact on the performance of those students who had featured in the experimental as opposed to the control group and had thus utilised self-assessment practices, on average they performed better than those who had not. Therefore, the results are suggestive of a more promising view regarding the uptake of self-assessment in the classroom (McDonald and Boud, 2003). Despite this, it is noteworthy that such self-assessment interventions took place alongside a programme of training and support mechanisms for teachers themselves. Without a programme, it is possible to question whether such a positive impact would be seen in less favourable circumstances (McDonald and Boud, 2003).

There is some agreement amongst formative assessment writers that feedback should not just be the role of the teacher, students need to be involved and take some responsibility for the assessment of their work. According to Black and Wiliam (1998) “self-assessment by the student is not an interesting option or luxury; it has to be seen as essential” (54-55). In addition, while Black and Wiliam (1998) advocate the use of self-assessment through their promotion of formative assessment, they did not examine much of the literature on self-assessment in their 1998 review. McDonald and Boud (2003) assert that prior to their work much of the literature on self-assessment was centred on the ways in which it can be utilised in certain classes by individual teachers, self-

assessment practices and studies comparing the marks of students and teachers (211). Black and Wiliam (2010) have, in their work, referred to the importance of teachers' professional development and training to successfully implement formative assessment practices. In the work of McDonald and Boud (2003) teachers were "galvanised to participate in the programme because it offered them the prospect of training their students for life after school, as well as assisting their students indirectly in preparing for external examinations" (214). It is therefore possible to argue that this is an example of the dimension of assessment which James and Pedder (2006) termed "promoting learning autonomy", one which was valued by teachers but could not always be pursued in practice.

2.2.4 The Boundary between Formative and Summative Assessment

Formative and summative assessments can be viewed as being at opposite ends of a spectrum, at the formative end "the teacher conducts the assessment, makes the inferences, plans any action and will see and have to deal with the consequences" (Black, 1998: 117). Indeed, Cowie and Bell (1999) argue that the teacher has a fundamental role to play through the provision of feedback and their collation of knowledge regarding pupils' levels of progress. Gioka (2009) also contends that there is a clear role for the learner, namely their ability to examine their progress in relation to learning goals. While at the summative end "the conduct of the assessment, drawing of inferences and the planning of action go beyond the teacher and perhaps the school" (Black, 1998: 117). However, there are numerous examples of assessment, which fall between these two. While it can be argued that many assessments carried out by teachers are summative in nature, since the results are not used to inform and modify teaching and learning, such assessments could also be viewed as formative assessments which have simply been carried out badly or incorrectly. Indeed Black (1998) contends that there is no clear means to determine whether an assessment is formative or summative in nature, with this being contingent upon how they link to student work and the subsequent use of the results. There may be instances where the assessment in question can be utilised for either formative or summative purposes, but in some cases for both. Furthermore, an assessment and the resultant data may be used for a purpose for which they were not initially

designed, which in turn can lead to impacts that were not expected (Harlen, 2005). It can therefore be argued that this calls into question the polarisation of formative and summative assessment (Black, 1998).

While teachers are involved in both formative and summative assessments, it is possible to question the extent to which tensions exist between the two. "The teacher has to hold the boundary between the different requirements of the two roles" (Black, 1998: 120), one which is centred on the needs of their pupils, the other which is based upon a framework of high stakes testing. Indeed, Black (1998) suggests that the traditional predominance of the summative leads to the formative struggling for positionality. Summative assessment can hamper the growth of formative assessment as teacher assessments are often based upon external tests (Black, 1998). It can also be argued that emphasis on summative assessment and high stakes testing can lead to teachers frequently conducting practice tests with their students, training them to answer particular question types and adopting teacher-centred pedagogies (Kellaghan *et al,* 1996; Linn, 2000; Stiggins, 1999). While there are a range of arguments and perspectives on formative assessment in the research literature, it may be that regardless of perspective, teachers have a fundamental role to play in the process. It is certainly questionable, however, whether there is such a role for teachers in assessments conducted for certification or accountability purposes. Formative assessment makes use of a diverse set of data for the purpose of the "modification of the learning work to adapt to the needs that are revealed by the evidence" (Black, 1998: 105). The use of formative assessment aims to develop thoughtful and active learners, predicated on a constructivist model of learning, as opposed to passive recipients of knowledge, and should draw on a wide range of assessment activities. As such formative assessment can elicit a greater level of evidence relating to pupil attainment than external assessments and generate, it could be argued, a broader picture of pupil attainment. In addition, Black and Atkin (1996) suggest that teachers also consider not only the result, but also the way a piece of work is produced. Therefore, it can be contended that "any inferences drawn from the assessment result about the students' skills and understanding may be

more trustworthy" (Black, 1998: 106). Not only does the teacher have prior knowledge of the pupil, he/she is also privy to the context of the performance.

2.2.5 How Far can the Impact of Formative Assessment be Ascertained?

The publication of Black and Wiliam's (1998) review of assessment literature marked a significant turning point in this area of research. This has largely been attributed to the ardent claims they made in their 1998 paper regarding the potential for formative assessment to significantly increase learning gains. It is therefore pertinent to discuss key concepts raised by Black and Wiliam (1998) through their work and examine the associated critiques. Black and Wiliam (2010) argue that educational policies liken the classroom to a black box. Inputs, such as pupils, teachers, and resources go into the box and certain outputs are expected to ensue, improved test results, more knowledgeable pupils for example. Yet, no attention is paid to what happens inside the black box, teachers are simply expected to meet, if not surpass expectations.

Black and Wiliam (1998) argue that formative assessment practices can produce significant improvements in student outcomes and achievement. They state that in all the studies they reviewed, innovations associated with "strengthening the practice of formative assessment produce significant and often substantial learning gains" (Black and Wiliam, 1998: 83). Learning gains are measured through the comparison of test scores of control and treatment group participants. Black and Wiliam (1998) cite effect sizes of formative assessment experiments between 0.4 and 0.7 and note that they are bigger than those usually found following educational interventions. Thus, a recorded effect size of 0.4 would mean that "the average pupil involved in an innovation would record the same achievement as a pupil in the top 35% of those not so involved" (Black and Wiliam, 1998: 83). Effect sizes provide a useful means of "quantifying the difference between two groups" (Coe, 2002, para.2), and can accentuate the size of the difference. From an educational perspective, effect sizes can be useful to quantify how effective an intervention has been and can be interpreted in several ways. Coe (2002) depicts the conversion of effect sizes to percentiles. With a reported effect size of 0.4, an average person in the experimental group would be 0.4 standard deviations above an average person in the control group. In addition, this would

involve the surpassing of the scores of 66% of the control group. An alternative interpretation is that of Cohen (1969), who describes an effect size of 0.2 as “small”, 0.5 as “medium” and 0.8 as “large”. Some researchers have issue with this approach, for example Glass *et al* (1981) contend that the extent to which an intervention can be deemed effective is linked to the interpretation of other interventions which aim to have the same impact. It is also important to note that in the context of education the associated costs and benefits of an effect should be considered. Even a small change, which could raise student achievement by an effect size of 0.1, could indicate the potential to lead to a very significant improvement.

Several researchers have presented arguments which critique Black and Wiliam’s (1998) construction of formative assessment. Dunn and Mulvenon, (2009) argue that much of the research literature which Black and Wiliam (1998) reviewed to inform their paper consisted of problematic methodologies, and issues relating to both sample sizes and the quality of the effect sizes. Despite this, there are several studies which cite effect sizes similar to those reported by Black and Wiliam (1998). Work by Wang and Baker (1986), which was concerned with the provision of mainstream versus special education for primary age disabled students, noted an effect size of 0.44 and Hembree (1988) whose work examined therapy for test anxiety stated an effect size of 0.42. Notwithstanding, many of the meta-analyses from which these results were derived have been subject to critique. While the use of a meta-analysis can produce an overall “average” effect size, it is important to refer to the original studies in order to identify factors, which may have contributed to differences between those with large and small effects (Coe, 2002).

As reported by Bennett (2011), meta-analysis is often deemed to be synonymous with methodological rigour, owing to the combining of results from a set of corresponding studies. Despite this, he proposes that meta-analyses can generate meaningless results, an argument which he puts forward in relation to the work of Black and Wiliam (1998). Furthermore, he argues that research which Black and Wiliam (2010) drew upon is not comparable due to its disparate nature, referring to studies concerned with feedback, peer and self-assessment, teacher questioning behaviour, teacher use of tests and teacher choice of assessment task and therefore cannot be

summed up in a meaningful way through meta-analysis (11). Indeed, Bennett (2011) suggests that such a range of factors cannot be effectively amalgamated using one mean effect size statistic. Furthermore, Coe (2002) states that "if they are effect sizes from experiments that differ significantly in terms of the outcomes measures used, then the results may be totally meaningless" (para.29). In addition, the data cited by Black and Wiliam (2010) are not the result of their own quantitative work, they did not conduct a meta-analysis of their own. While they assign effect sizes to formative assessment, a source for those values is not provided. Although Bennett (2011) acknowledges that the review process itself produced an important qualitative synthesis.

Other examples drawn on by Black and Wiliam (1998) include the work of Fuchs and Fuchs (1986) whose study focused on classroom assessment work for children largely in special education. This study is particularly pertinent to mention as according to Dunn and Mulvenon (2009), Black and Wiliam (1998) utilised the work of Fuchs and Fuchs (1986) to make the greatest contributions to their conclusions. As 83% of participants were mildly handicapped it is possible to question the extent to which this study is representative and ultimately generalisable to the wider student body. Dunn and Mulvenon (2009) also cite issues relating to the quality of the effect sizes quoted by Black and Wiliam (1998). While Black and Wiliam (1998) cited an effect size of circa 0.63 for non-handicapped students, this study raised several methodological issues. A total of 96 effect sizes were incorporated into their analyses, 19 of which were labelled as good quality, 69 as fair quality and 8 as poor quality. Therefore, as almost 72% of the effect sizes incorporated in the analyses were deemed to be of fair quality, it is important to consider the issues associated with the studies. Those effect sizes of fair quality were categorised as such due to having no more than two less serious methodological problems. This consisted of "the use of technically inadequate dependent measures, uncontrolled examiner expectancy, unchecked fidelity of treatment, the employment of inappropriate statistical unit of analysis and inadequate teacher training" (Fuchs and Fuchs, 1986: 202). Therefore, the combination of poor and fair quality effect sizes accounted for 80% of the effect sizes in this study. Accordingly, Dunn and Mulvenon (2009) argue that "80% of the effect sizes

that contributed to the mean effect size of 0.7 from the 23 studies examined came from research that was methodologically unsound" (5).

Additionally, those studies, which were deemed to be of good quality, had a limit of one "less serious" problem. A problem of a statistical or measurement nature calls into question the robustness of the 0.70 average effect size found by Fuchs and Fuchs (1986). While an effect size of 0.70 is astonishing, the question of generalisability and the research quality casts doubts upon the ability of this work to show that the use of formative assessment has a positive impact upon achievement. Dunn and Mulvenon (2009) refer to other pieces of research, which Black and Wiliam incorporated into their 1998 review. Black and Wiliam (1998) cite the work of Fontana and Fernandes (1994) who studied 25 Portuguese teachers and 246 students aged 8-9 and 10-14 respectively. As part of the study, the teachers received training to enable them to support students' undertaking of daily self-assessment to raise performance in Maths. The twenty teachers assigned to the control group took part in an alternative professional development course. The results of the study showed that it was only the maths scores of the younger students who displayed significant improvement in relation to the control group. In response to this, the authors suggested that the pre-test was too basic to elicit the true difference in gains between the two groups. Although, according to Dunn and Mulvenon (2009), the authors failed to recognise that the absence of a statistical difference between control and treatment groups in older pupils could be attributed to the effect of professional development or maths performance within the control group. Additionally, it is difficult to draw conclusions relating to the effectiveness of formative assessment based on a study involving 25 Portuguese teachers who focused only on self-assessment, in the one subject with students aged 8-14 (Dunn and Mulvenon, 2009:6). This is particularly pertinent as Black and Wiliam (1998) cite other elements, which they deemed to contribute to effective formative assessment, self-assessment is only one facet. Although the effect sizes cited by Black and Wiliam (1998) have been called into question by several researchers, other sources referring to the size of the effects which result from the use of formative assessment need to be carefully considered. This can be exemplified through Rodriguez's (2004) observational study which

examined the link between teachers' assessment practice and student achievement in mathematics. Rodriguez (2004) cites several variables which are attributed to variations in maths achievement, namely the self-efficacy and effort levels of students in relation to mathematics. Considering this, it can be argued that caution should be applied when making claims about the success of formative assessment.

As Sebatane (1998) notes, Black and Wiliam (2010) omitted a consideration of high stakes assessment (more specifically those dealing with certification and selection) from their 1998 paper, although Black and Wiliam (2010) do acknowledge that the threat of accountability has the potential to influence teachers' assessment practices. Black and Wiliam (1998) refer to "ecological validity" of studies included in their review. This is deemed to be problematic as teacher assessment utilised for one function can affect how it is used for other purposes. Given this statement, Sebatane (1998) is not therefore in agreement with the omission of a consideration of high stakes testing.

A number of researchers have advocated the positive impact of formative assessment on educational achievement. One such study is that conducted by Wininger (2005). A treatment group of 34 students were given feedback from the teacher and fellow students, they also received guidance on self-evaluating their own performance. In contrast, a control group of 37 students did not receive feedback nor guidance on the process of self-evaluation. After the second administration of the original test the treatment group performed considerably better than the control group. However, it is possible to call into question the small sample size along with the potential for researcher bias as the researcher used his own students to partake in the study. A further example relates to the work of Sly (1999) whose hypothesis was built around the idea that those students who undertook practice tests would do better than those who did not in first- and second-unit exams on a college Economics course. While Sly's (1999) hypothesis was proved to be correct, there were a few methodological issues. Namely, students were able to self-select whether they were part of the treatment or control groups, in other words whether they took the practice tests or not. This was not accounted for in the methodological design. While students who took

practice tests did do better than those who did not, their scores were recorded as being only 4 and 5 points higher. This represented a mean mark of 72.72 for those who had sat the practice tests compared to 67.56 for those who had not (Sly, 1999: 341). Arguably, the impact of this approach can be called into question. While it can be argued that the work of both Wininger (2005) and Sly (1999) is suggestive of formative assessment practices having a positive impact upon educational outcomes, methodological issues abound (Dunn and Mulvenon, 2009), which calls into question the impact of such approaches.

The discussion will now turn to a consideration of examples of specific pedagogies, namely peer and self-assessment and teacher questioning which can be viewed as modes of formative assessment. Given the extensive theoretical debate, it is important to consider the practical application of formative assessment as a key part of teachers' day-to-day practice.

2.2.6 Teachers' use of Formative Assessment

Following an examination of research into formative assessment, several researchers (Crooks, 1988; Black, 1993a) suggest that teachers' use of formative assessment is often associated with weak practice. Key features identified by the aforementioned researchers relate to classroom evaluation practices being synonymous with rote learning based around the "recall of isolated details" (Black and Wiliam, 1998: 17), an overemphasis on grades and a lack of reflection on assessment questions and materials. This could be attributed to the emphasis on external testing, which often takes precedence owing to the value attached to their results (Black, 1998). Despite this, some researchers argue that there is a need for higher quality classroom assessment undertaken by teachers (Baird, Hopfenbeck, Newton, Stobart and Steen-Utheim, 2014; Shepard, 2000). As such, much research has centred on the capabilities of teachers to present good quality assessment tasks to their students and interpret evidence accordingly to inform student learning (Looney, Cumming, van Der Kleij and Harris, 2018). The aforementioned capabilities are often referred to within the context of assessment literacy. Stiggins (1991) contends that to understand assessment literacy it is first important to consider what it is not. Those who Stiggins (1991) describes as assessment illiterate do not understand "what it takes to produce high-quality

achievement data and so do not evaluate critically the data they use" (535). While those who are deemed to be assessment literate recognise the significance of high-quality assessment, this is only of value however, when problems with unsound assessments are identified and dealt with.

Fullan and Watson (2000) take this definition further to encompass "the capacity of teachers- alone and together- (a) to examine and accurately understand student work and performance data and correspondingly, (b) to develop classroom, and school plans to alter conditions necessary to achieve better results" (457). Looney *et al* (2018) argue that owing to the significance of teachers' beliefs relating to assessment, assessment literacy alone cannot illustrate the extent and complex nature of such factors. There are other factors which should not be overlooked, for example the issue of amassing and documenting evidence as part of teachers' day-to-day practice, alongside difficulties associated with the need to differentiate teaching in light of assessment data (Black, 1998). Research by Cizek *et al* (1995) and Hall *et al* (1997) suggest that teachers do not make use of assessment results. According to Sun, Przybylski and Johnson (2016) "teachers lack the basic skills to understand, interpret and analyse data […] and address the weaknesses reflected from data analysis results" (5). They contend that this can be attributed to a lack of time available to undertake CPD on using and analysing data as well as collaborate and share best practice with other teachers. In addition, Stiggins *et al* (1989) argue that assessments are based around low-level aims such as recall. Thus, the assessments themselves do not inform teachers about their students' learning. Despite this, Sun, Przybylski and Johnson (2016) argue that teachers expressed a preference for short albeit informal instances of formative assessment as it gave them a "pulse check of where everybody was each day so as to adjust and plan teaching on a day-to-day basis" (15). Wylie and Lyon (2015) argue that while much work has focused on defining the concept of formative assessment and examining the success of formative assessment programmes, significantly less work has been concerned with teachers' implementation of formative assessment.

2.2.7 Students as Active Participants- Peer and Self-Assessment

Some researchers argue that one key element of successful formative assessment is self-assessment as this can serve to enhance student learning (Black *et al*, 2003). In recent years there

has been a movement away from students simply being involved in a process of acquisition to them playing an active part in knowledge and skill construction, a shift which is associated with the constructivist turn. Without this shift and formative assessment remaining in the hands of teachers, students are not being given the opportunity to become empowered and develop skills of self-recognition (Nichol and Macfarlane-Dick, 2006). Self-assessment encourages students to become self-directed learners (Heritage and Wylie, 2019). In addition, it is often assumed that students can easily decipher feedback from teachers, although this is not always the case. Black and Wiliam (1998) contend that a pupil who simply follows the teachers' direction without an understanding of the reasoning behind it will not learn. According to Chanock (2000) and Hyland (2000), this inability to make sense of and internalise feedback inhibits the student from constructing meaning to make productive improvements. Furthermore, Carless (2006) contends that the level of specificity of teacher feedback to one assignment can make it difficult for students to develop their learning in future assignments. As such, some researchers, for example Boud (2000) advocate the importance of improving self-assessment skills. The concept of developing students' skills was further discussed by Boud and Soler (2016) who coined the term "sustainable assessment" to question "the contribution of assessment to learning beyond the timescale of a given course" (400). In other words, such assessments can satisfy both formative and summative requirements but have the capacity to develop skills applicable to other future learning contexts. Boud (2000) originally coined this term in an earlier paper, describing it as assessment that "meets the needs of the present and [also] prepares students to meet their own future learning needs" (151). Despite this, in some educational settings students are taught the means to carry out self-assessment practices, thus it can be argued empowering them to "guide their own learning" and by "frequently applying standards and criteria to their own work they internalise them" (Nitko, 1995: 327). However, Nitko (1995) questions which students this process aims to benefit and in which classroom scenarios this is effective. Han and Fan (2020) further suggest that in order to make the process of self-assessment meaningful, students themselves should have a role to play in the creation of assessment criteria for the purposes of increasing transparency and emphasising their responsibility for their own

learning. Although, it is possible to question the feasibility of such an approach given the demands ensuing from high stakes summative assessment.

While there are a number of varying definitions in the research literature of self-assessment, for example some relate to student self-concept, whereas others are aligned with teacher assessment practices, there is a commonality amongst researchers regarding the purpose of self-assessment. According to Blatchford (1997), self-assessment is the "ability of a student to judge his/her performance" (2), while Panadero *et al* (2016a) suggest it is a "wide variety of mechanisms and techniques through which students describe and possibly assign merit or worth to the qualities of their own learning processes and products" (804). In addition, there are a number of researchers who align the use of self-assessment with improved student performance, for example McDonald and Boud (2003) and Sharma *et al,* (2016) argue that engagement in self-assessment can lead to improvements in academic performance, which supports the work of Black and Wiliam (1998).

According to Sadler (1989), there are three pre-requisites to enable a student to positively gain from feedback. Firstly, it is important for the student to be aware of what constitutes good performance, secondly the way in which their present performance compares to good performance and lastly, what is needed to close the gap. Sadler (1989) suggests that for this to be possible, students require knowledge of evaluative skills possessed by their teachers. As such, peer assessment can be a useful precursor to the development of self-assessment, insofar as students hone the skills they require to assess their own work. Indeed, DeGrez, Valcke and Roozen (2012) discuss the association of significant learning gains with engagement in peer assessment.

However, this approach is not welcomed by all teachers as it requires letting go of the traditional idea that students are the recipients of knowledge, and themselves, the custodians, therefore implementation in the classroom can be problematic. This represents the shift to a constructivist model at work in the classroom. In addition, there are often concerns amongst teachers relating to the value and accuracy of self-assessment. Wylie and Lyon (2015) argue that this process is only considered to be effective when pupils clearly understand the targets of their learning. In practice, this can be difficult to achieve, as emphasis on high stakes testing and pupil performance in

examinations has in many cases led to pupils becoming passengers in their learning, passive recipients of knowledge as teachers teach to the test, an area which will be discussed later in Section 3.1. Moreover, external examinations have long been used as a mechanism to test students' knowledge and understanding; "an activity to which students were subject rather than one in which they have roles as active agents" (McDonald and Boud, 2003: 211). This is reinforced by Kearney (2013) who contends that when it comes to traditional assessments it is the teacher who exerts control over the process. To compound the issue, pupils do not always understand why they are learning something or what a good piece of work looks like. In addition, self-assessment may not be a successful strategy for some students as they are often not aware of their own difficulties as "they lack the metacognitive skills needed to evaluate the quality of their own progress" (Nitko, 1995: 327). Also, high stakes assessment situations may not elicit honest responses from students when assessing their own work, students may have "inflated perceptions of their accomplishments" (Ross, 2006: 1). This comes in complete contrast to the claim by Black and Wiliam (2010) that formative assessment can greatly improve the performance of low achievers in particular. Furthermore, Kearney (2013) argues that control should be "relinquished by the 'guardians of knowledge' to allow authentic learning to occur" (878). According to Black and Wiliam (2010), "the fact that such gains have been achieved by a variety of methods that have as a common feature, enhanced formative assessment, suggests that this feature accounts, at least in part, for the successes" (83). Despite this, they do acknowledge that achieving such gains on a larger scale may not be straightforward. According to Ross (2006), agreement between student and teacher assessments has a tendency to be higher in the following scenarios, firstly, when pupils have been taught the means by which to assess their work and secondly when students are aware that their marking will be compared to that of their peers or their teacher. Agreement is often more closely aligned with peers than teachers which could be attributed to the way students and teachers interpret assessment criteria (Ross, 2006).

2.2.8 The Importance of Teacher Questioning

As I will discuss later in Section 7.3, teacher questioning constitutes a key formative assessment strategy which has been exemplified through the findings of my study. Teacher questioning can be utilised to establish students' current level of learning which can in turn inform teachers' pedagogical decisions. Mehan (1979) discusses three stages in the questioning process: initiation, response, and teacher evaluation/feedback. It is also worthy to note that according to Black and Wiliam (2009) formative assessment also involves a three staged approach, eliciting, interpreting, and using evidence. Therefore, it can be argued that such replication of the three staged approach reinforces questioning as a fundamental element of formative assessment. While teacher questioning can be described as an assessment tool, it can be argued that it can also be seen as a teaching tool, for example when it is used to ignite student interest. For questioning to serve as an effective formative assessment tool, the questions asked should seek to develop students' understanding, responses should inform teacher decision-making and the actions which ensue should move students towards their learning goals (Jiang, 2014). In other words, if questioning is utilised to diagnose learning, but further action to facilitate learning is omitted, it cannot be categorised as formative assessment (Jiang, 2014). Questioning can make positive contributions to towards the effectiveness of teaching and learning (Black *et al,* 2003). This has been reinforced by the work of others (Ruiz-Primo and Furtak, 2006; 2007), however these studies were limited insofar as they were carried out exclusively in science education. In addition, Chen, Hand, and Norton-Meier (2017) found that teacher questioning has the potential to increase the level of cognitive responses from students.

Research has also shown that low cognitive questions are more widely used in teachers' day-to-day practice (Jiang, 2014). According to Bloom's (1956) taxonomy, lower cognitive questions relate to those which require the recollection of factual knowledge. Furthermore, Jiang (2014) argues that low cognitive questions encourage rote learning and inhibit critical thinking and encourage students to be passive learners. Other issues regarding teacher questioning relate to student responses as the way students respond to questions is not always a demonstration of their thoughts. Jiang (2014)

examined teacher questioning as an assessment tool and focused on how teachers used questions to "stimulate student thinking, uncover students' current level of learning and allow responses to inform pedagogic decisions" (287). It was found that 81% of questions asked by teachers were lower cognitive, only 7% were higher cognitive questions (Jiang, 2014). This supports the work of Tan (2007) who explored classroom questioning behaviour and the extent to which this affected student development. Tan (2007) reported that 87% of questions were classified as lower cognitive. Thus, it can be argued that teachers' use of questioning does not constitute an assessment tool in all contexts, despite being cited as a key component of effective formative assessment (Black and Wiliam, 1998).

2.2.9 Formative Assessment- Concluding Thoughts

Black and Wiliam (2010) suggest that for formative assessment to be a success, slow but radical change to classroom practice is required along with continuous professional development and support. Indeed, Perrenoud (1991) contends that to implement formative assessment teachers are required to adapt their pedagogy to offset the learned habits of pupils. However, Sebatane (1998) argues that such a radical reform in pedagogy is not needed to impact upon student achievement, citing the work of Kellaghan *et al* (1982). This work, centred on the provision of ''diagnostic" information based on the performance on standardised tests of pupils in primary schools, compared to the provision of norm-referenced information only" (Sebatane, 1998: 124), led to increases in student achievement, without the need for radical interventions.

According to Bennett (2011), "the term 'formative assessment' does not yet represent a well-defined set of artefacts or practices" (5). While there is some evidence to suggest that formative assessment can positively impact learning, the diversity in terms of definitions and resultant range of implementation ultimately leads to variation in the effects of formative assessment. Shute (2008) argues that "within this large body of feedback research, there are many conflicting findings and no consistent pattern of results" (153). Therefore, it is possible to suggest that not all the literature is fully supportive of formative assessment. Nevertheless, it can be argued that the bigger problem relates to the lack of consensus regarding an agreed definition.

This chapter has set out to explore and review literature on formative assessment. The variety of literature which has been drawn upon is representative of the complex interplay of issues at work vis-à-vis assessment, this can certainly be exemplified through the differing interpretations of assessment purpose. Notwithstanding, there is some commonality with respect to accountability and the facilitation of teaching and learning (Black, 1998; Brown, 2004, 2006). However, regarding the concepts of formative and summative assessment it is somewhat difficult to set them apart from each other, although there are researchers who attempt to distinguish between them (Simpson, 1990; Harlen, 2005). There is also a lack of consensus relating to the definition of formative assessment with some researchers, for example Bennett (2011) and Dunn and Mulvenon (2009) deeming this to be problematic. According to Black and Wiliam (2010), assessment is only regarded as truly formative when it is used to inform learning and teachers' pedagogical decisions.

Formative assessment rests on a constructivist theory of learning which promotes the active role of students in their learning, a stark contrast to the traditional model of behaviourism. The ideals of a constructivist approach have been discussed by Boghossian (2006), Sener (1997) and Hackmann (2004). In addition, the constructivist model has not been without its critics, particularly as moving away from this approach is not without its problems. While there is a lack of consensus regarding the constructivist view, there are several shared ideas, largely based around the need of the learner to construct knowledge and meaning from their prior experiences (Hackmann, 2004).

Reference has been made to prolific researchers within the assessment literature, indeed, Black and Wiliam's (1998) seminal piece was considered to be a real turning point in the literature at the time of publication. Despite this, their work has subsequently been subject to much critique, with Dunn and Mulvenon (2009) highlighting methodological, measurement and statistical issues in the studies reviewed by Black and Wiliam (1998). Furthermore, their claims (Black and Wiliam, 1998) that formative assessment can lead to improved student performance have not been well received by some researchers (Dunn and Mulvenon, 2009).

The literature also reveals several researchers who advocate the importance of self-assessment to promote successful formative assessment (Black *et al,* 2003; Boud, 2000). While definitions and

therefore agreement over self-assessment abound, there are some commonalities, namely in relation to its purpose (Blatchford, 1997; Montgomery, 2000). Despite critique, self-assessment has also been aligned with improved student performance (McDonald and Boud, 2003; Sharma *et al,* 2016). Black and Wiliam (1998) also promote teacher questioning as a key formative assessment tool. However, research has shown that questioning cannot always be viewed in this way (Jiang, 2014). Furthermore, questions asked in classrooms are often low-cognitive in nature (Tan, 2007). While there is some evidence in the literature to promote the positive impact of formative assessment on learning, the lack of consensus regarding definitions can in turn affect both the implementation of and effects of formative assessment.

3 A Focus on Summative Assessment

In this chapter I will examine literature sources on summative assessment, more specifically the role and purpose of high stakes testing and the impact of its use as an accountability tool. Alongside this, the effects of high stakes testing on classroom assessment practice will be considered to examine how such external summative assessments impact teachers' application of assessment on a day-to-day basis. The concepts of reliability and validity in relation to assessment will also be explored. The extent to which assessments are reliable and/or valid coupled with their ability to produce information about students which is precise and accurate is of the upmost importance given both the emphasis placed upon external high stakes summative assessments and their use as an accountability tool. I will then turn to consider to what degree teachers work within a value-practice gap. While teachers attach value to elements of their practice, for example developing skills of self-criticality and independence amongst their students, their ability to implement such values into day-to-day practice is often hindered by the need to meet internal and external expectations driven by high stakes summative assessments. This chapter will conclude by re-introducing my research questions and emphasising links to the preceding literature discussion as a precursor to the methodology chapter.

3.1 Summative Assessment

High stakes testing has multiple purposes, one key purpose relates to its role as an accountability tool. "The results of tests (and examinations) are intended to provide a means of making schools accountable for the education they provide to different stakeholders" (West, 2010: 23). While summative assessment can be used both internally and externally, for example internally to report to parents or grading for record-keeping purposes, much emphasis in the English education system is placed upon the use of summative assessment for external purposes including certification by examination bodies, access to employment or higher education and monitoring and accountability. Since 1992, examination results have been released into the public domain and form a component of league tables, a key part of school accountability (Leckie and Goldstein, 2017). Examinations taken at the age of 16 and 18 can significantly affect the future of students and their teachers as

well as schools (Leckie and Goldstein, 2017; West, 2010). When assessment information is used to inform decisions relating not only to the student but also to teachers and the school itself, the results become high stakes from which pressures and constraints often ensue.

League tables often serve to assist parents in the selection of a secondary school for their child, thus, schools are in competition with each other for pupils. Since 1992 the percentage of students achieving 5 or more GCSEs at grade C or higher has been published annually. However, measures of school progress have been revised on a few occasions "from "value-added" (2002-2005) to "contextual-value added" (2006-2010) to "expected progress" (2011-2015) to "progress 8" (2016)" (Leckie and Goldstein, 2017: 193). Furthermore, use of such data to inform parental choice can raise difficulties as the reporting of the percentage of 5 GCSEs at grade C or above can be mistaken as a means to assess the quality of schools. According to Leckie and Goldstein (2017), a school with a higher 5A*-C percentage is often deemed to be more effective and ultimately more attractive to parents than a school with a lower percentage. School rankings based on value-added are now increasingly being utilised to enhance accountability. Although Schiltz, Sestito, Agasisti and De Witte (2018) cite issues with the use of value-added as the measure requires predictions to be made (with value-added scores being the difference between actual and predicted performance). Therefore, it could be stated that direct comparisons of this nature are unreliable as "no attempt has been made to allow or adjust for [. . .] confounding effects" (Leckie and Goldstein, 2017: 194), such as the make-up of the intake from one year to the next. In contrast, progress measures are deemed to represent a more reliable means of comparison as they "implicitly attempt to adjust for what are often substantial differences in the composition of pupils' prior attainments and other characteristics between schools at intake" (Leckie and Goldstein, 2017: 194). Despite this, it can be argued that school league tables should be considered alongside other sources of school information to provide a full and detailed picture. Owing to high stakes tests teachers' beliefs regarding what they deem best for their students can often fall behind the need to prioritise the demands of school policies and parental expectations (Black, 1998: 120). Further pressures can come in the form of Ofsted. Poor examination results and how schools perform in league tables can

also lead to increased scrutiny and in some cases with schools being placed in “special measures” by Ofsted, which can in turn threaten their longevity (Leckie and Goldstein, 2017). There is evidence to suggest that high stakes assessment can adversely affect both teacher and student behaviour, although this is not the only contributing factor, other examples include “gender, participation in class, self-concept and social competence, independence, classroom behaviour, and work habits” (Madaus and Kellaghan, 1992: 126).

Types of assessment affect what and how teachers teach. In the research literature, high stakes testing is often met with distaste as researchers cite an array of unintended negative consequences. It has been suggested that high-stakes testing can narrow the curriculum, lead to a growth of teacher-centred pedagogy, and promote teaching to the test (Au, 2007; Copp, 2018). Harlen (2007) argues that simply focusing teaching on what is assessed serves to narrow opportunities for students to learn. One such argument is related to teachers’ loss of creative input into their teaching, teaching styles are changing to incorporate more teacher-centred pedagogy, where teachers lecture, and students are simply passive recipients of knowledge in the learning process. According to Smith (2004), following research into maths education during a period of modularisation of exams, “far too much time is devoted to examinations and preparing for examinations- “teaching to the test”- and that this is at the expense of the understanding of the subject itself” (93). In their review of exam reform, Baird *et al* (2019) argue that modular exams do not provide opportunities for deep learning as teachers and students are constantly working towards regular examinations. Furthermore, modularisation can be associated with the interruption of regular teaching and a greater focus on exam preparation. Although this is not always the case as students are not required to undertake all modular exam series (Baird *et al,* 2019). Baird *et al* (2019) contend that the stakes attached to linear examinations, which were introduced in 2015, are arguably higher. Their work, based on examination reform, found that an increased focus on school assessments, for example in the form of mock exams, comparable to modular exams, and the teaching of revision and exam skills were associated with the transition to linear (Baird, *et al,* 2019). Moreover, they found that owing to more challenging content which must

be taught to students, there were fewer opportunities for the promotion of deep learning than anticipated, instead there was a need to develop memorisation skills (Baird *et al,* 2019).

Au (2007) argues that knowledge has become increasingly fragmented through the teaching of small, isolated test-sized pieces and teachers teaching to the test at the expense of other subject knowledge (262). In addition, many of the more complex skills associated with teaching are rendered less and less acceptable relative to high stakes standardised testing. Harlen (2005) contends that in such circumstances "teachers make little use of assessment formatively to help the learning process" (209). While such a focus on passing tests often contributes to test score increases, it is possible that such a rise in test scores is not synonymous with a rise in achievement as teachers are simply "training" students to pass a test. According to Linn (2000) the increases in test scores can be attributed to familiarity with the test content as opposed to increased achievement. Despite this, data from Ofqual (2019) suggest that A Level results in England in 2019 were lower than the previous year at Grade A and above, with 25.2% in 2019 and 26.2% in 2018, thus suggesting that achievement, certainly at the higher grade, is in decline. However, this could be attributed to changes in A Level cohort and subject choice. Furthermore, 2019 A Level results in all curriculum areas showed that the award of grades at a C or above has fallen from 76.8% in 2018 to 75.5% (Nye and Thomson 2020). As depicted in Figure 3.1 below, this figure is at its lowest since 2015. Moreover, the award of A/A* grades has also decreased from 26.2% in 2018 to 25.2% (Nye and Thomson, 2020). However, Nye and Thomson (2020) suggest that this could be attributed to a 1.2% decline in A Level entries in 2019, as illustrated in Figure 3.2, rather than a decrease in performance. West (2010) suggests that competition between exam boards for students to sit their examination papers constitutes a reason for the growth in high grades. Furthermore, schools are keen to increase their standing in league tables and as such may choose examination boards that they anticipate will boost the achievements of their students (West, 2010). In addition, "high-stakes tests are inevitably designed to be as "objective" as possible since there is a premium on reliable marking in the interests of fairness" (Harlen, 2005: 209). Therefore, what can be assessed reliably is greatly reduced, including important skills such as problem-solving and critical thinking. In their

study which investigated the extent to which students from low income, and ethnic minority backgrounds can acquire cultural capital, Hong and Youngs (2008) found that high stakes testing not only contributed to the narrowing of the curriculum but also made it more difficult for pupils to develop higher order thinking and problem-solving skills. While Dewitt *et al* (2013) acknowledge that test standards require pupils to be "provided with active, meaningful criteria" (409), their work has shown that high stakes tests do not assess such criteria, namely higher-order cognitive skills.

Figure 3.1 - A Level Grades in all Curriculum Areas

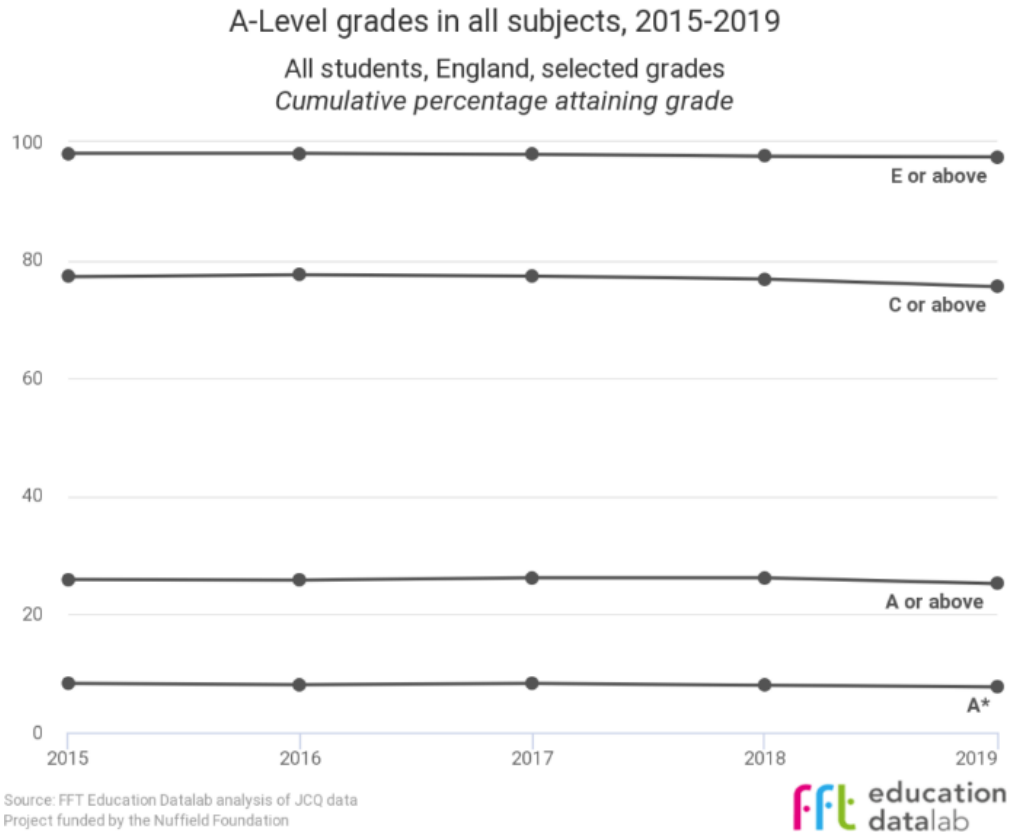
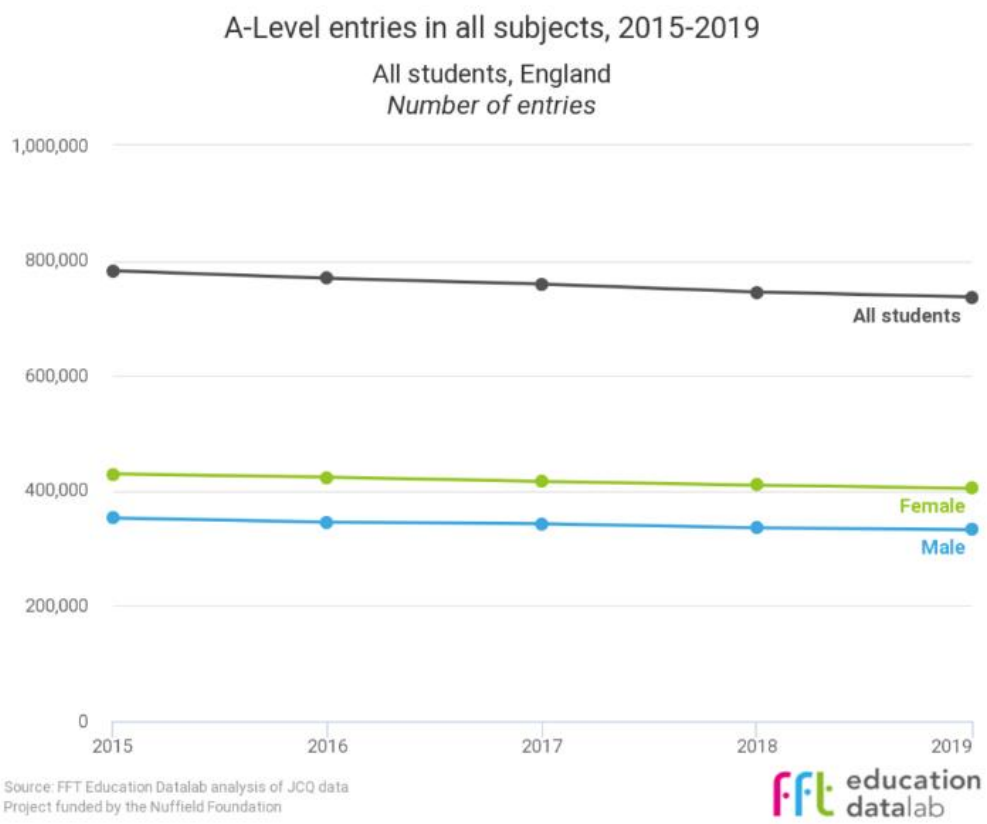


Figure 3.2 - A Level Entries in all Curriculum Areas



While much work in high stakes testing is synonymous with negative connotations and unintended consequences, there is evidence in the literature to counter such critiques. Within education, high stakes tests are largely resisted due to their links with accountability. Tests act an as accountability tool as "policy-makers realise that they cannot directly regulate instruction in classrooms, but they can indirectly influence instruction by attaching rewards or sanctions to the results of mandated tests" (Madaus and Russell, 2010/11: 21). Cizek (2001) discusses several positive consequences, which he argues can be attributed to high stakes testing. For example, Cizek (2001) argues that professional development for teachers is becoming more focused on the curriculum and results while knowledge about testing is becoming increasingly prevalent amongst educators which in turn trickles down to the classroom level. Indeed, Leahy and Wiliams (2012) cite the importance of professional development for teachers, particularly given the fast-moving nature of their working environment. According to Madaus and Russell (2010/11), the production of information relating to student performance alongside the provision of information relating to the quality of schools serves to inform decisions about school choice and "open doors of opportunity to those previously

shut out by holding teachers and schools accountable for student achievement and helping them to focus attention on students who were poorly served in the past" (26).

3.2 The Reliability and Validity of Assessment

According to Halliday (2010) assessment is deemed to be valid when it tests what it intends to test. However, Lambert and Lines (2000) argue that this is an oversimplification of the concept. Regardless, it is paramount that tests can generate information about learners which is precise, what Black (1998: 37) terms "confidence in the result". This is of particular importance when assessments are high-stakes and results are interpreted for the purpose of accessing further study or employment. "They are measures which deeply influence their "life chances", what university and professions they will enter, how rewarding and secure their employment will be, how much they will earn and so forth" (Jarvis, Holford and Griffin, 2003: 158). While the ideal remains for assessments to be both fully reliable and valid, it is widely recognised that it is not easy for reliability and validity to co-exist (Davis, 2006). For example, while a written exam tests students' capacity to write under controlled conditions, it can be argued that there is a lack of validity as writing is being tested as well as knowledge.

One issue relates to the differing interpretations of a learners' assessment performance by different markers. Davis (2006) suggests that "close agreement between assessors can only be obtained by a tight marking scheme and/or a very prescriptive set of exemplar responses" (11). Consequently, this does not promote the exploration and discussion of rich knowledge. Assessing rich knowledge requires a more open and creative response from students, if this is not the case, there is a risk that the assessment becomes nothing more than a test of memory. As such Davis (2006) argues that for an assessment of rich knowledge to be valid, reliability is compromised as a result. However, Curren (2006), maintains that it is better to assess rich knowledge, "albeit imperfectly than not to attempt to assess such knowledge at all" (27).

Unitary validity relates to the inferences which can be made in light of tests. In other words, regardless of how valid a test is based on its construction, the results need to be used appropriately,

otherwise it cannot be deemed to be valid. Indeed, Stobart and Gipps (1997) argue that "it is the use of test results that determines validity" (42). Therefore, validity is associated with professional responsibility, it is not simply the domain of test developers. From this perspective of unitary validity, it is therefore possible to question "the assumptions lying behind the heavy reliance on data from external examinations" (Lambert and Lines, 2007: 10). Lambert and Lines (2000) argue that examinations are becoming less valid as their purpose becomes broader and the results are used in more varied ways, as a means to evaluate student performance as well as provide a mechanism to judge schools and local authorities. As such, the purpose of each use becomes increasingly unclear, ultimately creating a "confusion of purpose" (Lambert and Lines, 2007: 10). Messick (1996) argues that validity relates to the interpretation of test scores. It can be contended that coaching or teaching to the test, an unintended consequence of high stakes testing (Au, 2007), could adversely affect "the validity of the interpretation and use of the coached scores" (Messick, 1996: 6). Conversely, some means of test preparation, for example anxiety reduction can serve to increase validity.

According to Frederiksen and Collins (1989) assessments which are deemed to be authentic and direct are "systematically valid" as they promote "changes that foster the development of the cognitive skills that the test is designed to measure" (27). Systematic validity could be shown through improvements in cognitive skills becoming apparent once the test has been in existence within an educational system for a length of time. It is therefore important to consider whether there are general characteristics of a system of testing which add to or diminish the systemic validity of a test. In addition, Frederiksen and Collins (1989) argue that the extent to which a test is considered to contribute to educational improvement is predicated on the "directness of a cognitive assessment" and "the degree of subjectivity or judgement required in assigning a score to represent the cognitive skills" (28). Other authors, for example, Morrow (1986) cite the concept of "washback validity", this relates to the extent a test is valid being contingent upon its impact on teaching. Ideally, positive washback should be associated with a smooth transition from learning to test exercises, activities should be interchangeable. However, the influence on teaching can be

problematic insofar as “a poor test may be associated with positive effects and a good test with negative effects” (Messick, 1996: 2). Moreover, a test could be seen to affect teaching content but not practice as well as teacher behaviour, for example, whether a teacher employs formative or summative assessment strategies. Consequently, washback is a product of testing which impacts on validity in the event that it can be viewed as an effect of the test itself rather than other forces at play on the educational arena (Messick, 1996). Despite this, authentic and direct assessments do not always generate positive washback owing to the nature of the assessment itself and the education system, particularly the teaching and assessment setting. Firstly, “the ideal forms of authenticity and directness rarely if ever exist” (Messick, 1996: 244). Furthermore, construct under-representation and construct-irrelevant variance are often in existence. As such, the test cannot be seen as a fully representative example of criterion behaviours (Messick, 1996). This can be attributed to test performances not being “scored and interpreted in ways that are unlikely fully or faithfully to capture the criterion domain processes (Messick, 1996: 245). Alderson and Wall (1993) suggest that washback cannot be relied upon to determine test validity.

As claimed by Messick (1996) there are a number of threats to validity, namely construct underrepresentation, which is a threat to authenticity and construct-irrelevant variance, which is a threat to directness. In the case of the former, the test is too narrow and “fails to include important dimensions or facets of focal constructs” (Messick, 1996: 4). However, in the case of the latter, the assessment is too extensive. While it is difficult to fully eradicate such threats to validity, the impacts should be minimised. The assessment literature emphasises the importance of assessment being reliable as well as valid. Assessment is considered to be reliable when it “consistently produces the same results irrespective of extraneous factors” (Halliday, 2010: 370). Nevertheless, this is deemed problematic insofar as assessments are often taken by large groups of people in different places and are marked by multiple examiners. As such the extent to which assessment criteria are interpreted differently by different examiners can be questioned. It can be argued that the concepts of validity and reliability are linked. Many exams now incorporate different styles of exam paper to strike a balance between reliability and validity.

3.3 The Value-Practice Gap

Evidence from the literature suggests that teachers want their pupils to be independent and self-critical learners, good problem solvers and effective communicators (Harlen, 2007: 42). Despite this, teachers often feel that they must "prioritise meeting targets for test results, ahead of encouraging wider learning that they actually value" (Harlen, 2007: 42). Research conducted by James and Pedder (2006), has shown that teachers work within a value-practice gap. In other words, teachers are often obliged to side-line their aspirations across much of their teaching practice. As shown in James and Pedders' (2006) study, teachers often attach greater value to certain aspects of their practice, however in reality, these do not always match those elements of practice which must be pursued in order to meet both internal and external expectations.

James and Pedder (2006) identified three dimensions of assessment: "making learning explicit", "promoting learning autonomy" and "performance orientation" (129) which they argue forms the basis of assessment in the classroom. Their results suggest that while teachers highly regarded the first and second dimensions of assessment, "making learning explicit" and "promoting learning autonomy", thus deeming them important to meet the needs of their students, significantly less value was attributed to "performance orientation" (130). Therefore, the pressures and constraints of external high stakes tests can make it difficult for teachers to pursue their values in their practice. According to James and Pedder (2006), teachers in their study "felt impelled to invoke a set of performance-orientated principles and practices at odds with their professional judgement" (132). While teachers tend to value the concept of formative assessment, many acknowledge that there is a clear mismatch between their personal beliefs and values and the reality of the situation. James and Pedder (2006) based their research on a survey of 558 teachers in 32 schools. While their work serves to support the existence of a value-practice gap, there are some noteworthy points. Namely, two of the three dimensions of assessment identified by James and Pedder (2006), could be deemed to be contradictory in nature as "making learning explicit" and "promoting learning autonomy" are related to assessment for learning whereas the third dimension, "performance orientation" is synonymous with assessment of learning. However, due to the emphasis on

performativity, “teachers are willing to acknowledge its legitimacy in the policy context” (James and Pedder, 2006: 129). Despite this, it is arguable that assessment regimes can always be considered as being constructed to enable formative and summative assessment to co-exist, on the proviso that they are neither confused nor conflated. Therefore, it could be said that the value-practice gap is simply a manifestation of this complexity.

A study conducted by Winterbottom, Brindley, Taber, Fisher, Finney, and Riga (2008) sought to build on the work of James and Pedder (2006) through an examination of trainee teachers’ values in relation to assessment practices and the extent to which their teaching mirrored these values. Winterbottom *et al* (2008) identified the greatest value-practice gaps were in relation to promoting learning autonomy and performance orientation. Winterbottom *et al* (2008) compared their findings with those of James and Pedder (2006) who focused on qualified teachers. Winterbottom *et al* (2008) based their study on a sample of 220 secondary PGCE trainees from one UK Higher Education Institution. It is possible to argue that while this sample incorporated a range of schools whose approaches to assessment policy differed greatly, it does not provide a true representation of UK schools from a statistical perspective. In addition, it is pertinent to question whether the extent of value-practice gaps reported in this study are fully represented as there is a possibility that “trainees would give less authentic responses because they have stronger perceptions of the “correct’ answers”” (Winterbottom *et al,* 2008: 196).

Tiknaz and Sutton (2006) examined and investigated teachers’ thinking regarding planning for formative assessment in the medium term. This study also identified differences between teacher values and teacher practice insofar as while teachers considered it to be important to develop assessment tasks to show students how they are progressing in their learning, in practice they found it difficult to execute. In addition, teachers utilised assessment checklists with interim assessments, which meant that less time was required to make written comments. These checklists comprised a pre-defined list of subject specific and generic targets, therefore, it is possible to question whether this approach enables individual students to close the gap in their learning. Further work conducted by Smith, Hill, Cowie, and Gilmore (2014) examined assessment beliefs of

trainee teachers. Smith *et al* (2014) argued that trainee teachers' beliefs regarding assessment were influenced by their own experiences of assessment as opposed to their experiences of learning about assessment theories as part of their teacher training.

3.4 Concluding Thoughts:

This chapter has sought to explore the role and purpose of summative assessment alongside a consideration of its use as an accountability tool. Although summative assessment can be used internally by teachers in their classrooms, much emphasis is placed upon its external use. High stakes tests inform accountability measures such as league tables and are often viewed negatively as a result. Furthermore, the results of high stakes tests are often used to determine the future of both students and their teachers (West, 2010). The resultant pressures and constraints can result in a number of unintended negative consequences being associated with high stakes testing, including the narrowing of the curriculum, teacher centred pedagogy and teaching to the test (Au, 2007; Harlen, 2007). Moreover, some argue that this can be at the detriment to formative assessment (Harlen, 2005). Black and Wiliam (1998) argue against the inclusion of grades or summative marks on student work as they suggest that this can discourage students from using teacher feedback. Butler (1988) also argues that the greatest learning gains can be attributed to comment-only marking. Furthermore, the provision of feedback to students by teachers enables the identification of any areas of difficulty, and the implementation of any remedial action.

In addition, this chapter has also examined the extent to which teachers work within a value-practice gap. Harlen (2007) discusses the skills which teachers want to develop in their students, but these are often overlooked due to the focus on high stakes testing with meeting targets taking prescience. As such, James and Pedder (2006) argue that teachers work within a value-practice gap. Teachers attach much value to particular aspects of their practice, for example "promoting learning autonomy" but in reality, they are often forced to side-line such aspects with "performance orientation" driving their focus (James and Pedder, 2006).

Having examined the key concepts within the assessment literature, it can be argued that there are several complex issues which intersect and ultimately serve to shape the assessment landscape. Reviewing literature on formative and summative assessment in chapters 2 and 3 has provided a theoretical underpinning to my research and sought to highlight issues within the wider assessment literature which has informed the development of my research. This has contributed to the generation of the following research questions. Links to the preceding literature discussion and the research questions to which they pertain are also identified below.

1. How do teachers use formative and summative assessment strategies with their students and why do they use them?

Formative assessment rests upon a constructivist theory of learning. However, there is no consensus regarding an agreed upon definition of formative assessment. In the research literature there are a diversity of definitions and the varying levels of implementation of formative assessment contributes to variation in effects (Bennett, 2011). It is also pertinent to consider the blurring of the boundaries between formative and summative assessment. Formative assessment can increase learning gains (Black and Wiliam, 1998) and have a positive impact upon achievement (Wininger, 2005).

2. How does high stakes testing affect teachers' approaches to assessment?

There are several purposes of high stakes testing, for example its use as an accountability tool. High stakes testing can lead to teachers feeling compelled to prioritise demands of school policies/parental expectations. In addition, high stakes testing is often associated with unintended negative consequences (Au, 2007; Copp, 2008). It is also important to examine the extent to which assessments are reliable and/or valid. The ability of assessments to generate information about learners which is precise is of importance against a backdrop of high stakes testing. "Washback validity" (Morrow, 1986) relates to the extent to which a test is valid being contingent upon its impacts on teaching and learning.

3. How do teachers use formative assessment to prepare their students for the summative?

Formative assessment is based upon a constructivist framework emphasising the importance of students' role in the learning process. The transition to constructivism builds upon the work of Piaget and Vygotsky. Constructivism promotes a focus on in-depth rather than superficial learning. However, the boundaries between formative and summative assessment can often become blurred. There are several key elements of formative assessment which are considered to be successful, for example self-assessment (which builds on the concept of constructivism) and teacher questioning (Jiang, 2014).

4. What are teachers' beliefs about assessment strategies? Is there a relationship between teacher values and day-to-day practice?

It is pertinent to examine the provision and nature of feedback in the form of marks, grades and/or comments. Teachers' use of formative assessment can be associated with weak practice (Crooks, 1988; Black, 1993a) which could be attributed to emphasis on external summative testing. Some researchers (Boud and Soler, 2016) argue that "sustainable assessment" can contribute to learning in the long term, beyond a programme of study. Teachers want their students to develop skills of independence and self-criticality (Harlen, 2007), although meeting targets often takes priority. Teachers work within a value-practice gap (James and Pedder, 2006; Winterbottom *et al,* 2008). The values teachers attach to elements of their practice often do not play out in the classroom owing to the need to meet internal and external expectations.

In the following chapter, I will now turn the focus to the methodological framework and methods I have used in order to undertake my research and generate answers to my research questions, derived from both my own experiences and engagement with the research literature. A case study approach required a narrow focus in order to enable me to explore how these broader themes actually played out in the classroom, with the focus being on a case study approach within one institution, termed Mitford College in the North East of England.

4 Methodology

The purpose of this chapter is to examine the theoretical underpinnings and rationale for the research methodology. The research setting will be explored alongside a consideration of the ethical issues which are pertinent to my study. The research process itself will be discussed and strengths and limitations of the data collection techniques outlined. The chapter concludes by examining the means by which the data has been analysed and applied within the context of the study.

The overarching framework for this research project takes the form of a case study approach. Within this framework two major sources of data will be utilised, observation and focus group interviews, in order to examine how teachers operationalise assessment strategies in their classrooms with students in Years 12 and 13. There are a number of reasons for the adoption of a case study approach. This study has a limited time frame for completion and is also of a small scale. However, the use of a case study allows the framing of a bounded unit, for example an institution and provides the opportunity for rich, detailed data to be collected "during an intensive but short period of time" (Hamilton *et al,* 2012: 11). As Merriam (1998) states "I can "fence in" what I am going to study" (27). In order to facilitate the generation of rich data, multiple data tools were employed to ensure that "the issue is not explored through one lens, but rather a variety of lenses which allows for multiple facets of the phenomenon to be revealed and understood" (Baxter and Jack, 2008: 544).

4.1 A Case Study Approach - Background

Case studies are often used in educational research to "enhance our understanding of contexts, communities and individuals" (Hamilton, Corbett-Whittier, and Fowler, 2012: 3) and this approach constitutes a widely used qualitative research methodology in this field. According to Yin (2014) case study is popular to "contribute to our knowledge of individual, group, organisational, social, political and related phenomena" (4). Despite this it can be argued that case study is still lacking in

status somewhat in social science research as it does not have “well-defined and well-structured protocols” (Yazan, 2015: 134).

Case studies began to grow in popularity in both the UK and USA in the 1970s largely as a response to the positivist model with its overwhelming focus on quantitative research and measurement. Oancea and Pring (2008) in their work, cite growing policy critique of educational research in the 1990’s and early twenty-first century with its narrow focus, based around “what works” and the superiority granted to experimental designs and randomised control trials. Arguably, such approaches advocate a one size fits all model which “ignores the complexity of education settings and the significance of the diverse individuals and organisations that enhance that complexity” (Hamilton *et al*, 2012: 5).

Considering this, case study enables the observation of effects in real contexts as opposed to developing understanding based upon decontextualized “evidence” (Hamilton *et al*, 2012: 6). Case study “allows investigators to focus on “a case” and retain a holistic and real-world perspective” (Yin, 2014: 4). Despite the popularity of case study, it is often an approach which is traditionally deemed to lack rigour. Notwithstanding, case study has the potential to “offer insights that might not be achieved with other approaches” (Rowley, 2002: 16). As will be discussed later in Section 4.3, case study can be used to generate answers to “how?” and “why?” questions and from this standpoint can be used for exploratory, descriptive, or explanatory work (Rowley, 2002). It can be argued that a case study approach promotes an examination of the world around us and enables an investigation into a particular phenomenon to be conducted within its own context, “it is not necessary to replicate the phenomenon in a laboratory or experimental setting in order to better understand the phenomena” (Rowley, 2002: 18).

Three of the key authorities in the domain of case study research are Yin (2014), Merriam (1998) and Stake (1995) who each have differing standpoints vis-à-vis the use of case study in educational research. While Yin (2002) does not overtly discuss his epistemological perspective, his approach to case study indicates a positivist perspective. Yin’s (2014) four conditions pertaining to design quality- construct validity, internal validity, external validity, and reliability align closely to concepts

of objectivity, validity, and generalisability in positivist research (Crotty, 1998). Stake (1995) bases his philosophical underpinnings on a constructivist paradigm as "most contemporary qualitative researchers hold that knowledge is constructed rather than discovered" (Stake, 1995: 99). Stake (1995) sees those carrying out case study research as "interpreters" who must ultimately report "their rendition or construction of the constructed reality or knowledge that they gather through their investigation" (Yazan, 2015: 137). Further knowledge construction occurs on the part of the reader. Merriam (1998) shares this epistemological perspective- "the key philosophical assumption upon which all types of qualitative research are based is the view that reality is constructed by individuals interacting with their social worlds" (Merriam, 1998: 6). I consider my epistemological stance to align more closely with a constructivist paradigm as I recognise knowledge as a social construction where truth is relative and is "dependent on one's perspective" (Baxter and Jack, 2008: 545). Just in the same way that formative assessment rests upon a constructivist theory of learning which is built around the premise that learners construct their own knowledge (Sener, 1997), the constructivist paradigm "recognises the importance of the subjective human creation of meaning" (Miller and Crabtree, 1999: 10).

4.2 Choosing a Case Study

At the outset of my research project, I was employed as a full-time teacher working in an 11-16 mainstream school with access to groups of students and a staff body to partake in my research. I therefore intended to examine assessment practices at two different stages in the school in which I was working for comparative purposes, Years 7 and 11, and take on the role of a teacher-researcher.

As the research project progressed, personal circumstances led me to leave the teaching profession and take up a full-time position working outside the classroom. I was consequently faced with the difficulties other outside researchers face in terms of gaining access to a site to study and negotiating with gatekeepers. As such, finding and gaining access to research sites can be problematic. Due to the changes in my professional situation and the resultant logistical difficulties of accessing a population to study I subsequently changed the year groups upon which the research

is based. Despite this, the overall focus of my research remained, I have continued to examine how teachers operationalise assessment practice with their students as they navigate an education system with an end goal built around high-stakes summative assessment. Gaining access to this establishment enabled me to examine assessment practices utilised by teachers of Sixth Form students as they worked towards A Level and BTEC qualifications.

Yin (2003) describes a number of scenarios which lend themselves to a case study approach, examples include inability to shape the behaviour of those participating in the study, the desire to examine contextual conditions as the researcher believes that they are important to the phenomenon being studied and unclear boundaries between the phenomenon and the context. In my research the case is defined within a specific institution as teachers' use of assessment, the unit of analysis relates to a process, how teachers operationalise assessment. However, this could not be examined without the context, in this instance the institution or even more specifically, the classroom environment, the setting in which assessment takes place. Both Yin (2003) and Stake (1995) discuss the importance of placing boundaries on a case to ensure that the study does not become too broad. Here established boundaries need to take the form of a concise definition of formative and summative assessment and an indication of where assessment is being utilised, in Year 12 and 13 BTEC and A Level classes in one institution. Therefore, boundaries are both institutional and discursive. The case study for my research is based around one Sixth Form institution in the North East of England in which I used to work. Therefore, the case study is bounded to the confines of one institution. While it would be useful to produce a comparative case study of more than one institution, the limited timescale for this study is a prohibiting factor. Within the institution itself the case study is bounded through a focus on teachers' use of formative and summative assessment with students in Years 12 and 13. Bounding of this nature also enables comparisons to be made between the use of assessment in Year 12 and 13 respectively, with students in Year 13 sitting high stakes summative assessments. The case study is further bounded through the use of data collection techniques, both classroom observation which was not open ended and had a specific focus and focus groups which were based around an interview schedule.

It is also important to establish the type of case study to be carried out. Within the research literature there are a number of categorisations of case study put forward, for example, Yin (2003) describes case studies as explanatory, exploratory, or descriptive whereas Stake (1995) discusses intrinsic, instrumental, or collective case studies. For the purposes of my study the type of case study takes the form of intrinsic, as the case is of interest in its own right, it is intrinsically interesting. An intrinsic case study is concerned with capturing "the case in its entirety and the purpose of the research is to understand more fully the person, department or institution that makes up the case" (Hamilton *et al*, 2012: 11-12).

4.3 Defining a Case Study

Definitions of case study in the literature abound, however there is some disagreement as to the nature of these definitions. According to Yazan (2015) this has become "a contested terrain" (134). While Nisbit and Watt (1984) suggest that case study relates to a particular occurrence with the aim of generating a more wide-ranging principle, Cresswell (1994) describes a case study as "a single instance of a bounded system" (12). Furthermore, Tight (2010) goes on to suggest that a case study is concerned with looking in depth at a small sample. Interestingly, Chong and Graham (2013) use the analogy of a "Russian Doll" to present their understanding of a case study, in order to glean an understanding of the case at the micro-level understanding of the meso- and macro-contextual levels in which it sits is required (24). Regarding my study, teacher use of formative and summative assessment in the classroom sits within a macro-level context of high stakes testing.

As noted in Section 4.1 Yin (2002), Merriam (1998) and Stake (1995) are three of the key authorities in case study research. It is therefore pertinent to examine how they diverge in the definitions they present in their work. According to Yin (2002) a case is defined as "a contemporary phenomenon within its real-life context, especially when the boundaries between a phenomenon and context are not clear and the researcher has little control over the phenomenon and context" (13). Therefore, Yin (2014) associates a case study with answering "how?" or "why?" questions relating to the phenomenon in question. Yin (2014) also discusses definitional shortcomings, for example, early social science texts referred to case study as an approach to be used in the preliminary or

exploratory stages of research only. Over time, a case study approach has evolved. Other shortcomings relate to the confusion of case study research with "doing "fieldwork" as in ethnography or participant observation" (Yin, 2014:15). In contrast, Stake (1995), supports the view of Smith (1978), that a case should be seen as a "bounded system" (2) and is "a specific, a complex, functioning thing" (2). In addition, Stake (1995) suggests that a case study lends itself to the study of programs and people as opposed to events and processes.

Merriam (1988) in contrast to both Yin (2002) and Stake (1995) has altered her definition of case study over time with a change in focus from the outcome to the case itself. Merriam's later definition supports the views of Stake (1995) and Smith (1978), suggesting that the case is a bounded unit. The case "is a thing, a single entity, a unit around which there are boundaries. I can "fence in" what I am going to study" (Merriam, 1998: 27). However, the existence of multiple definitions of case study, relating to it either as a process, the case or bounded unit or the outcome, all serve to distort Merriam's (1988) transition in focus from outcome to case. Merriam (1998) suggests that a case could constitute a person, a programme, a group, a process, an institution, therefore presenting a wider definition than both Yin (2002) and Stake (1995).

Other researchers have also promoted the use of case studies, for example, Stenhouse (1978) argues that case study research provides a means to examine the complex nature of education, but it needs to be verifiable, although according to "quite different principles from those governing verification and cumulation in experimental sciences" (21). Flyvbjerg (2006) argues that case study is often associated with a bias towards verification, "a tendency to confirm the researchers preconceived notions" (234). Certainly, it can be argued that bias towards verification is associated with all methods and it is falsification which typifies case study as opposed to verification. Indeed, Campbell (1975) contends that "even in a single qualitative case study, the conscientious social scientist often finds no explanation that seems satisfactory" (181-182). Therefore, the emphasis upon falsification can be seen to characterise a case study, Stenhouse (1978, 1979) discusses the key divergence between case study and quantitative approaches which previously formed the mainstay of educational research, he illustrates this through the "study of cases versus the study of

samples". Stenhouse (1978, 1979) argues that what is required to enable the verification of case studies is quite different to that associated with the scientific model. Stenhouse (1978) suggests that verification has historically occurred through the critique of evidence "which is available on the same terms to all scholars" (22). On this basis Stenhouse (1978) argues for archives of case records, upon which case studies are based.

Case study has emerged as an approach to research which "can capture rich data giving an in-depth picture of a bounded unit or an aspect of that unit" (Hamilton *et al*, 2012: 10). Nonetheless, there is some confusion in the research literature relating to the nature of case study and whether it is indeed a method, methodology or research design. Hamilton *et al* (2012) argue that case study should be viewed as an approach to research whereas Elliott and Lukes (2008) suggest that it should be seen as a genre. They argue that a case study in this respect enables the examination of complex relationships and attitudes in a bounded unit through the utilisation of a range of data collection tools. Regardless, there is still much debate about how case study is characterised alongside the associated difficulties of defining a "case".

4.3.1 Critiques

Despite their growth in popularity, some researchers argue that a case study approach, particularly a single case study approach, lacks value and should therefore be used only as a pilot method. Case study is often synonymous with the examination of a single phenomenon at the early stages of an investigation, owing to the provision of a hypothesis which can then be more widely tested within the context of a bigger number of cases (Abercrombie, Hill and Turner, 1984). This is reflective of the traditional, hierarchical view whereby surveys and histories are deemed "appropriate for the descriptive phase, and that experiments are the only way of pursuing explanatory or causal inquiries" (Yin, 2014: 7).

However, Flyvbjerg (2006) suggests that this is an oversimplification, to the extent that it can be misleading, and he also discusses what he deems to be several misconceptions associated with case study research. Some researchers, for example Yin (2009) argue that it is not possible to generalise

from a single case, therefore contribution to scientific research cannot be made. Flyvbjerg (2006) cites the example of Galileo who rejected Aristotle's law of gravity based upon the results of a single experiment. Case study also played an important role in the work of Darwin, Marx, and Freud. According to Flyvbjerg (2006) "formal generalisation, whether on the basis of large samples or single cases, is considerably overrated as the main source of scientific progress" (226). In addition, Flyvbjerg (2006) argues that formal generalisation "is only one of many ways by which people gain and accumulate knowledge" (227). Furthermore, Kuhn (1987) emphasised the importance of researchers having a wide range of skills for undertaking scientific work, with generalisation forming only one such skill. In addition, Ruddin (2006) suggests that generalisability in relation to case studies is not important as it is symptomatic of positivism in a non-positivistic form of research.

Some researchers, for example Merriam (1988) and Stake (1995) view the focus of case study as being a single entity or case whereas others deem this to be a limiting factor and instead advocate a comparative approach, for example Yin (1989). According to Ragin (1992), it is misguided to suggest that the use of multiple case studies is superior to single case studies as single case studies are "multiple in most research efforts because ideas and evidence may be linked in many different ways" (225). This is particularly important as even within a single case study the use of multiple methods provides the potential for the linking of evidence from a range of sources. Qualitative methods more generally, as well as case study, it can be argued provide opportunities for researcher subjectivity, issues of subjectivity are also pertinent in relation to quantitative methods. Although such methods have traditionally been considered to be less rigorous than quantitative methods. According to Pearson, Albon and Hubball (2015) case study methodology can be considered flexible in relation to the research questions which can be explored and the data collection methods which can be undertaken. In other words, they are not confined by "methodological traditions" (3). They also contend that case study is credible and generalisable on the proviso that work is carried out within the parameters of the case context and adheres to ethics regarding both participants and data management. Pearson, Albon and Hubball (2015) contend that "a case is usually representative in some way, either through its typicality or atypicality, of a broader

group of cases (3). Although caution is needed on a case-by-case basis, regarding how relevant findings are to other contexts (Pearson, Albon and Hubball, 2015). Some researchers such as Campbell (1975) state that case study is rigorous albeit differently to quantitative methods. Notions of rigour are equally problematic in quantitative studies. Case studies are often synonymous with "a substantial element of narrative" (237) and as such do not fit neatly into scientific formulae. However, for the case study researcher a "thick" narrative is a positive.

4.4 Research Methods

One key feature of case study research is the use of multiple data sources which some researchers suggest enhances the robustness of the data as this is a form of triangulation (Patton, 1990: Yin, 2003). "Each data source is one piece of the "puzzle" with each piece contributing to the researcher's understanding of the whole phenomenon" (Baxter and Jack, 2008: 554). In my study I have employed multiple data sources as a means of triangulation, with each data source constituting a piece of the puzzle, to generate rich data sources to feed into the analysis. These data sources and the rationale for their use is shown in Table 4.1 below.

Table 4-1 - Data Sources Used in this Study

Data Source	Rationale
Classroom Observation –non-participant.	To examine teachers' use of formative and summative assessment in the moment. The use of an observation schedule provided a focus for the observation (Appendix 3).
Field Notes	Used to record and report observations and reflections. They can also serve to trigger a memory of an observed event.
Focus Groups	To discuss the rationale behind teachers' use of formative and summative assessment strategies and explore teachers' professional attributes and values in relation to their practice.
Document Analysis	Several documents have been drawn upon to provide context and additional richness to the data. Examples of documents teachers used in the classroom have been included to enrich discussions (such as worksheets) as well as assessment criteria and exam board documents from their respective websites. These have been used to illustrate the parameters within which teachers must work when it comes to high stakes summative assessments.

The subsequent convergence and analysis of the aforementioned data sources works to produce a rich body of data which in turn increases knowledge and understanding of the case. There are several variables within a single case, as such "to catch the implications of these variables usually requires more than one tool for data collection" (Cohen *et al* 2017: 376). In addition, Rowley (2002) discusses the importance of drawing upon evidence from several sources, "the richness of the case study evidence base derives largely from this multi-faceted perspective yielded by using different sources of evidence" (23).

4.5 The Data Collection Process

In this section I will begin by discussing each of my research methods in turn, considering how I utilised each one in my study. The rationale for choosing and the merits of each method will then be discussed in greater depth later in the chapter, in Sections 4.5.1 to 4.5.4 below.

In the initial stages three data collection methods were planned for in the study: semi-structured interviews, classroom observations and questionnaires. Questionnaires were included with the aim of not only investigating teachers' rationale for their approaches to assessment but also eliciting interesting responses to be followed up at interview. An online survey was generated via Online Surveys and the link to access the survey was sent to my contact at the research site who distributed it to participants (see Appendix 14). In total, 14 people completed the survey. Upon reflection, I deemed this to be too few participants to provide meaningful results, therefore, the questionnaire was omitted from the study. Originally, I intended to conduct semi-structured interviews with teachers in the institution under study. However, due to time constraints, affecting both myself and teaching staff, I decided to carry out focus group interviews with multiple participants instead.

As previously discussed in Section 1.4, I carried out my research in a mainstream, mixed Sixth Form College in the North East of England which I have given the pseudonym of Mitford College. There are approximately 1,500 students enrolled at Mitford College and 85 members of teaching staff. Students can choose to study from an extensive range of courses, including over thirty A Levels and over ten BTEC subjects.

During the data collection phase of my research, I undertook two focus groups interviews with participants from 11 subject areas, as shown in Table 4.3. I also carried out 28 classroom observations, each of 30 minutes in duration, Table 4.2 shows the subjects which were observed. Both focus group participants and subjects observed were selected by my contact within the institution to enable me to encounter a variety of subject areas and staff members and fit with the needs of the institution. In addition, I collated and analysed 25 documents which were sourced from both online and the classroom observations themselves. Table 4.4 shows the origin of these documents.

For the purposes of classroom observations, I selected a structured non-participant approach as this enabled me to adopt a "passive, non-intrusive role" (Cohen, Lawrence, and Marion, 2017: 545) in the classroom. As a non-participant observer, I observed activities taking place in the classroom, with the knowledge of participants, without taking part in the activities themselves. Classroom observation constitutes an important data collection tool in my enquiry as it provided a means to record behaviour in the moment. The use of an observation schedule, as shown in Appendix 3, enabled me to maintain consistency across the observations, while my research questions shaped the content of the schedule. There were specific things which I needed to look for, namely examples of formative and summative assessment strategies and whether they were teacher or student led. This was particularly important given the time limited nature of this study to ensure that a manageable amount of data was collected. Table 4.2 below shows the range of subjects which were observed as part of my study.

Table 4-2 - Subjects Observed

	Subject Observed
Spring Term	Accounting (A Level), Biology (A Level), Chemistry (A Level), Classics (A Level), Core Maths (A Level), Criminology (A Level), English Literature (A Level), Film Studies (A Level), Further Maths (A Level), Geography (A Level), History (A Level), Law (BTEC), Media (A Level), PE (A Level), Religious Studies (A Level), Sociology (A Level).
Summer Term	Biology (A Level), Business Studies (A Level), Criminology (A Level), English Language (A Level), Fine Art (A Level), French (A Level), Health and Social Care (BTEC), History (A Level), ICT (BTEC), Maths (A Level), Physics (A Level).

During each observation I also completed field notes, an example of which is shown in Appendix 5, to act as Swain (2006) describes, as an aide memoire and to supplement the observation schedules with more detail. During each lesson I noted down key events and occurrences, which ultimately related to my research questions, including detail of formative and summative assessment strategies, any challenges which arose and examples of exam preparation. I typed up the field notes as soon as possible after each observation while the detail remained at the forefront of my mind. I also reviewed the material and developed and applied a series of codes to the data.

In my study, classroom observations and field notes were supplemented by focus group interviews to provide an opportunity to uncover the rationale behind teacher use of assessment. Indeed, the focus groups allowed for the interrogation of what I had uncovered through the classroom observations and enabled the compilation of data from different participants at the same time. This was a pertinent factor in my study given the time pressures not only on myself as the researcher but also the teachers who participated in the focus groups. Participation in the focus groups in my study gave teachers the opportunity to discuss their perspectives on assessment, an issue which impacts on their day-to-day practice, and many feel strongly about. The focus groups also provided an arena for discussion, an opportunity which teachers often lack owing to pressures and demands on their time.

At the beginning of each focus group, I set the ground rules, including the right to withdraw, one person to speak at a time and the "facilitation of discussion" (Gibbs, 2012: 188). While I had prepared an interview guide, as shown in Appendix 4, I was aware that new areas of discussion could arise. For example, during one focus group, discussion turned to research carried out by one member of staff on factors impacting upon student progress within Science. As the moderator I sought further information in relation to this to make the most of this new information before returning the discussion to the original focus. Table 4.3 below shows the subject each focus group participant taught, there were seven participants in focus group 1 and four in focus group 2, each representing a different curriculum area within the college. Moreover, participants held a variety

of job roles within the college, ranging from subject teacher, and head of department to assistant headteacher and had varying levels of teaching experience, from newly qualified to a significant number of years of experience in the profession.

According to Merriam and Tisdell, (2015) data analysis should take place alongside data collection as this enables the researcher to reflect upon their research so far and begin to identify tentative themes. “Data that have been analysed while being collected are both parsimonious and illuminating” (Merriam and Tisdell, 2015: p197). As part of my data analysis, the coding of field notes and focus group transcripts involved assigning a combination of single words and phrases to the data (a coded transcript is shown in Appendix 6). This enabled me to identify the emergence of key and recurrent themes in the lesson observations which were cross-referenced with the focus group transcripts. Miles and Huberman (1994) use the term “descriptive coding” to refer to such codes which offer a summary of the data. Furthermore, this approach to coding is of particular use when the researcher is required to assign codes to data from different sources, as in my study. Bryman and Burgess (2002) suggest that coding provides the means to organise large quantities of data. It can be argued that coding and the generation of concepts is closely aligned (Miles and Huberman, 1984). In my study, codes represent “the building blocks for emergent rather than pre-specified concepts” (Bryman and Burgess, 2002 p.7).

Table 4-3 – Subject Taught by Each Focus Group Participant

	Subject
Focus Group 1	Biology (A Level), Business Studies (A Level), Health and Social Care (BTEC), Law (BTEC), Media Studies (A Level), Physics (A Level), Sociology (A Level).
Focus Group 2	Criminology (A Level), English (A Level), Film Studies (A Level), PE (A Level).

4.5.1 Classroom Observation

Classroom observation plays an important role not only in classroom research, but also in teachers’ professional development (Hopkins, 2008: 75). As teachers we are “continually monitoring what is going on in classrooms and constantly questioning what is happening” (Wilson, 2009: 83) to inform

and improve practice further. While case studies are associated with the use of multiple data sources and methods of analysis, observation is often a central feature of many case studies. This can take many forms from "natural" (such as ethnography) to artificial, from structured (for example structured non-participant observations) to unstructured (such as ethnographic observations). According to Bryman, (2008), structured observation is concerned with adherence to a specific set of rules with an observation schedule being used to systematically record participants' behaviour. As a researcher it is possible to adopt one of several positions when carrying out observations, from full participant to merely a spectator. However, it is important to discuss the issue of reactivity effects, that is to say the potential for my presence as an observer to have an impact on the course of events. Merriam and Tisdell (2015) contend that it is important for the researcher to "identify those effects and account for them in interpreting the data" (147).

Simply gaining access to a population to study can prove to be a stumbling block to observational research and is "not a matter to be taken lightly but one that involves some combination of strategic planning, hard work and dumb luck" (Von Maaen and Kolb, 1985: 11). Simply gaining entry to an organisation does not automatically mean access to a population to study. Indeed, once access to an organisation has been granted, problems may still arise. For example, staff within an organisation may prove to be suspicious of researchers and see them as a means to check up on them. In addition, staff may also be concerned that their comments will go back to management (Bryman, 2008: 403). Although this is an issue which affects all fieldwork, not simply observation.

Observation has the potential to provide insights which may not come to light through other means such as interviewing as individuals may be unaware of certain behaviours they display or may not wish to openly discuss certain issues with the researcher. According to Bryman (2008), the use of an observation schedule presents the possibility of "imposing a potentially inappropriate or irrelevant framework on the setting being observed" (269).

While observation can be a time-consuming data collection tool, this approach has the potential to provide access to social interactions and the opportunity to collate detailed data, perhaps it could be argued, more detailed than data gleaned from other sources (Simpson and Tuson, 1995).

Interviews, for example, may preclude people from opening up about all that is relevant and regular events may simply be overlooked. “Individuals may never have become aware of them in a conscious fashion and are therefore unable to talk about them in an open and articulate way” (Simpson and Tuson, 1995: 16). Observation can work well in conjunction with other methods which focus on the meanings behind observed behaviours. Observation can enrich data collected via other means and can add to the understanding of what is happening. This approach has been adopted by other researchers, for example Delamont (1976), who employed semi-structured interviews alongside observation.

Classrooms are inherently complex places with numerous interactions and events taking place at any one time, therefore the observational focus needs to be clear and concise. While the presence of another adult in the room is not an alien concept (as it can be argued that children frequently encounter other members of staff conducting lesson observations or learning walks), it is important to be aware of reactivity effects. It is possible to argue, therefore, that simply entering the field involves participation through one’s presence in the social space under investigation.

As previously discussed in Section 1.5, leaving the teaching profession presented me with the challenge of gaining access to a sample population. However, I used my contacts in a school in which I previously taught to set up my fieldwork for this study by initially writing to the Head teacher. As I will discuss in Section 4.6 below, I was later contacted by a former colleague, an Assistant Head teacher responsible for Teaching and Learning, who became my contact and ultimately the gatekeeper who oversaw my access to the institution. While the gatekeeper was extremely accommodating of my research, the lessons which I observed were arranged by him and I received a timetable for each of my visits showing the lessons to observe that day (as shown in Appendix 15). The gatekeeper also arranged and selected participants for the focus group interviews. Thus, the gatekeeper facilitated my access to teaching staff who participated in both the lesson observations and focus group interviews. Despite this, I was given the opportunity to observe a range of subjects across the institution, both A Level and BTEC, as shown in Table 4.2. Furthermore, no teacher was observed on more than one occasion, but I did observe several subjects (Biology,

Business, Criminology, History, Maths) twice. In addition, as shown in Table 4.3 above, the focus group participants represented a range of subject areas, although more teachers of A Level as opposed to BETC subjects participated. Also, as shown in Table 4.2 above, significantly more A Level subjects were observed as part of this study. Both of these instances are reflective of the curriculum offering at Mitford College which is based around a higher proportion of A Level subjects.

4.5.2 Field Notes

Hammersely and Atkinson (2007) describe field notes as "the traditional means in ethnography for recording observational and interview data" (141) whereas Hopkins, (2008) describes field notes as "a way of reporting observations, reflections and reactions to classroom problems" (104). While field notes can be subjective, it is possible to argue that this problem could be mitigated by the fact that I am not conducting this enquiry in my place of work, I am a "scheduled visitor" as opposed to a "participatory resident" (Shimahara, 1988: 86). Despite this, other methods were integrated into the research process with the aim of building up a full and detailed picture. In addition, it is important to recognise their limitations insofar as "they cannot provide a comprehensive record of the research setting" (Hamersley and Atkinson, 2007: 156). Consequently, the use of other methods is fundamental.

Despite their merits, some researchers, for example Walford (2000) argue that field notes are no longer a widely used approach due to the time commitment necessary to write them up effectively. Swain, (2006), on the other hand, suggests that field notes are highly personal to the researcher and act as an aide memoire as the research progresses. "It is important that they are seen as part of a reflexive process, that is subject to re-appraisal as conceptual understandings deepen" (Swain, 2006: 202). In addition, Swain, (2006) argues that field notes do not provide a comprehensive record of all we observe, instead "impressions and unrecorded recollections based on more unreliable fragments of memory will also, inevitably intrude into the construction of the overall picture" (202). Therefore, it is important to review and revisit field notes over time. Particularly as the character of field notes may differ vis à vis focus and detail as the research progresses.

In some research settings, the taking of copious notes may be deemed to be disruptive. However, in an educational environment, while note-taking is somewhat "normal" and expected, disruption must still be avoided. For example, Olesen and Whittaker, (1968) discuss their experiences of taking field notes where writing when the students were not seemed to "attract [the tutor's] attention and on a few occasions, she seems to falter in what she is saying" (28). Therefore, it is important to consider at what point during the observation is most appropriate to make field notes. In the event that not all observable actions are recorded during the observation, it is of the upmost importance to add this to field notes as soon as possible thereafter.

As mentioned in Section 1.5, I used field notes in my study not only to act as an aide memoire but also to provide additional detail relating to teachers' use of assessment in the classroom, alongside the observation schedule. During the observations themselves, I deemed it important to consider my positionality and the potential impact my presence could have on events in the classroom. Therefore, I took the decision to only add to my field notes and complete the observation schedule when students were working independently rather than when the teacher was talking to the class so as not to cause a distraction.

According to Hammersley and Atkinson (2007) the quality of notes "diminishes rapidly with the passage of time, the detail is quickly lost and whole episodes can be forgotten or become irreparably muddled" (142). During an observation, even brief notes can be beneficial insofar as they can serve to trigger a memory of an observed event. According to Schatzman and Strauss (1973) "a single word, even one merely descriptive of the dress of a person, or a particular word uttered by someone is usually enough to "trip off" a string of images that afford substantial reconstruction of the observed scene" (95). While it is time-consuming, time must be allocated for writing up field notes.

4.5.3 Focus Groups

Focus groups are a data collection method used in qualitative research to elicit responses from a group of individuals who are knowledgeable on a topic. It can be argued that focus groups are

constructivist in nature owing to the way in which data are "socially constructed within the interaction of the group" (Merriam and Tisdell, 2015: 114). Hennink (2014) contends that one distinct feature of focus groups relates to their ability to generate "a different type of data not accessible through individual interviews" (2-3).

Focus groups can be used as a means to triangulate data from other sources. Denzin (1970) suggests that "the combination of methodologies in the study of the same phenomenon" (291) has the potential to reduce bias. Furthermore, focus groups can be utilised in order to draw on the knowledge and experience of participants and can be used to examine "not only what people think but how they think and why they think that way" (Kitzinger, 1995: 299). This is of importance within the context of my study as focus group interviews provided a means to explore the rationale behind teachers' approaches to assessment. The group dynamic of a focus group can "help people to explore and clarify their views in ways that would be less easily accessible in a one-to-one interview" (Kitzinger, 1995: 299).

Adopting a focus group approach can be beneficial insofar as the group dynamic can encourage communication between research participants and those who may not like to participate in a one-to-one interview may be more compelled to take part. Moreover, there is also the potential for active participation by those who initially feel they have nothing to contribute. Focus groups allow the collection of data from multiple participants at the same time. Focus groups also enable participants to discuss their thoughts on a topic they deem to be important and express their views to others.

According to Halcomb *et al* (2007) one weakness of focus group research can relate to the inability of moderators to effectively facilitate group interactions. A further critique of focus groups concerns their analysis. Gibbs (2012) argues that "the unit of analysis is the collective perspective" (189), in other words, data should be presented in such a way as to demonstrate interactions between participants.

Hamilton *et al* (2012) discuss the need for criminal records checks to be undertaken for those wishing to conduct research in locations whereby they may encounter young people or vulnerable groups. Although I no longer work as an educator in a school, my current position of employment requires me to have an enhanced DBS certificate. Therefore, this removed the need for such checks to be undertaken prior to me beginning my research. Permission that was granted to me by gatekeepers at the institution enabled me to move freely around the site, from one classroom to another during the research phase as they were happy with my DBS clearance. In my study I carried out two focus groups with teachers from Mitford College.

Focus groups can be difficult to arrange, particularly from the position of an outside researcher. However, for the purposes of my study, participants in both focus groups (shown in Table 4.3 above) were selected by my contact (the gatekeeper) in the college. Those who took part had not only volunteered but were also available on the dates and times which I was able to attend the study site to carry out my research around the commitments of a full-time job. It is pertinent to note that all teachers who took part in the focus group interviews also participated in the classroom observations (as shown in Table 4.2 above), thus enabling me to discuss with teachers the rationale behind their use of assessment in the classroom. Furthermore, the use of two data sources in conjunction with one another serves as a means of triangulation to not only enhance the robustness of the data and but also generate rich data to inform the analysis.

4.5.4 Document Analysis:

Bowen (2009) describes document analysis as "a systematic procedure for reviewing or evaluating documents" (27). Just like other forms of data, documents need to be examined to extract meaning from them and develop our understanding of phenomenon. Documents can consist of both textual and visual materials which have not been influenced by the researcher. Furthermore, document analysis is often utilised alongside other qualitative methods as part of triangulation. Eisner (1991) suggests that the triangulation of data provides "a confluence of evidence that breeds credibility" (110). Merriam (1988) suggests that much value can be gained from using documents as they can "help the researcher uncover meaning, develop understanding, and discover insights relevant to

the research problem" (118). There are several advantages which can be attributed to document analysis. Firstly, a large proportion of documents can easily be accessed via the internet, in the case of my research, exam board documents have been accessed this way. In addition, documents can be described as stable as there is no need to account for researcher effects (Merriam, 1988). However, it can be argued that documents are generated for reasons other than research and as such may not be detailed enough for the purpose of research (Bowen, 2009: 31). Table 4.4 below shows the type of documents I analysed as part of the research process and the subjects they are associated with. I elected to analyse documents published by examination boards, these consisted of both A Level Subject Specifications and exemplar and past examination questions accessed from exam board websites. This enabled me to exemplify how teachers were using exam board documents to inform their teaching and provide a rationale for tasks undertaken in class. I also analysed resources which I collected as part of the classroom observations themselves, such resources provided a means to demonstrate how teachers were using assessment tools as part of their day-to-day practice.

Table 4-4 - Source of Documents

Type of Document	Subjects documents collected from
Practice exam question from past paper	Geography- Figure 5.1
Practice exam question from past paper	Classics- Figure 5.2
Self-audit of exam performance	Business Studies- Figure 5.3
Exam skills audit	Business Studies- Figure 5.4
Kahoot revision tool	Health & Social Care- Figure 5.5
Blockbusters revision game	Health & Social Care- Figure 5,6
Connect 4 revision game	Health & Social Care- Figure 5.7
Subject grade boundaries 2018 examinations	Physics- Figure 5.8
A Level Specification	Media Studies- Figure 5.9
Context sheet	Film Studies- Figure 5.10
A Level Specification	Film Studies- Figure 5.11
The modular exam process	Gov.uk- Figure 5.12
The linear exam process	What Uni? - Figure 5.13
6-mark exam question	Physics- Figure 6.1
Exam questions	Business Studies- Figure 6.2
Levelled (banded) exam question	Geography- Figure 6.3
Assessment criteria	Criminology- Figure 6.4
A Level Specification	History- Figure 6.5
A Level Specification	Geography- Figure 6.6
Mark scheme for Assessment Objective 1	Religious Studies- Figure 6.7
A Level Specification	Religious Studies- Figure 6.8
Institution entry requirements	Institution website- Figure 6.9
Mark scheme for 5-mark exam question	Geography- Figure 6.10
BTEC Authentication Form	Pearson website- Figure 6.11
BTEC Applied Law Specification- Making valid assessment decisions	Pearson website- Figure 6.12

4.5.5 Coding the Data Set

To identify key and recurrent themes in the data sets I collected I carried out a thematic analysis of both my field notes and focus group transcripts. Thematic analysis involves the identification of meaningful categories to make sense of often large quantities of data (Maguire and Delahunt, 2017). This also enabled me to generate links between the data I collected in the field and the theory in the assessment literature. Braun and Clarke (2006) offer a six-phased approach to thematic analysis which I utilised to inform my own thematic analysis. The first phase is concerned with becoming familiar with the data which involved reading and re-reading the transcripts and making some initial notes based on my first impressions of the data. Following this I began to generate some codes. This enabled me to identify ideas which came up on several occasions throughout the transcripts which were relevant to my research questions. This then led me to

search for themes within the data through the grouping of codes which fit together. Following this I began to pull together data which linked to each theme and reviewed whether each piece of data supported the theme before defining the themes. This involved what Braun and Clarke (2006) term as identifying "the "essence" of what each theme is about" (92) and searching for subthemes and links between themes. Several themes and codes which I identified as part of this process have subsequently been used in Chapter 5 as section headings. Table 4.5 below shows the codes and subsequent themes which I identified through the process of thematic analysis.

As part of the process of developing codes to apply to my data I drew upon both my reading of the existing assessment literature and my experiences of professional practice. As discussed in Sections 2.2.7 and 2.2.8 I have examined literature on formative assessment strategies. Through an examination of the data in the transcripts I identified both self and peer assessment as reoccurring ideas and as such they became codes which in turn were grouped together under the theme of formative assessment strategies. Furthermore, both my professional practice and reading of the assessment literature also informed the types of assessment (as shown in Table 5.1 below) which featured on my observation schedule (as shown in Appendix 3). For example, I placed 'verbal feedback' into the 'other' category; from my professional experience, verbal feedback can be a timely activity, owing to larger class sizes, and a full lesson can often be set aside to undertake such an activity. Therefore, I considered that I was less likely to observe multiple instances of verbal feedback taking place in lessons which were providing a snapshot of teachers use of assessment strategies, but the 'other' category still provided a means of recording should it be required.

Table 4-5- Themes and Codes

Theme: Formative Assessment Strategies	**Theme:** Summative Assessment Strategies	**Theme:** Challenges	**Theme:** Best Practice	**Theme:** Frequency	**Theme:** Student action/ response	**Theme:** Exam Preparation
Codes: Self-Assessment Peer Assessment Feedback - (comments) Questioning Model answers	**Codes:** Feedback – (grades) Mock exams Past exam questions	**Codes:** Timing Marking Changing exam structure Changing student mindsets	**Codes:** Upskilling students Reviewing & revisiting information	**Codes:** Cycles Every lesson	**Codes:** Response to feedback Independent learning Changing student mindsets	**Codes:** Mock exams Assessment criteria

4.6 Ethics:

Ethical behaviours are now identified by Governments as well as professional and educational associations. Resnik (2010) defines ethics as the "norms of conduct that distinguish between acceptable and unacceptable behaviour" (1). The following educational associations, BERA (British Educational Research Association), AERA (American Educational Research Association) and SERA (Scottish Educational Research Association) have two guiding principles- respect and responsibility. It is important that the researcher respects those taking part in the study, taking account of knowledge, democratic values, and academic freedom. Furthermore, researchers have a responsibility to not only participants but also sponsors of research and the community of educational researchers (BERA, 2004: 5). Therefore, these regulatory frameworks ensure that as researchers we carefully consider any ethical implications of our research. A number of steps and procedures were followed to ensure that this study adhered to ethical guidelines. An application was submitted (in September 2017) to the School of Education Ethics Sub-Committee at Durham University (as shown in Appendix 10), ethics approval was subsequently obtained.

Guillemin and Gillam (2004) define ethics in practice as "the everyday ethical issues that arise in the doing of research" (263). Brooks, te Riele and Maguire (2014) argue that ethical practice should be concerned with not only the relationships developed with participants during the research process but also the end product, the knowledge generated "from the initial design, through data

collection and analysis to dissemination" (5). According to Hughes (2005) "ethical practice is an ongoing interaction of values in shifting contexts and relationships, rather than something delivered by a signed consent form or adherence to a static set of principles" (231). It is therefore important to consider my positionality as a researcher. Indeed, the potential impact of power dynamics is not always evident. One key issue concerns that of informed consent. Hammersley and Atkinson (2007) contend that this centres on the idea that "people must consent to being researched in an unconstrained way, making their decision on the basis of comprehensive and accurate information about it; and that they should be free to withdraw at any time" (210). However, this could be complicated by the fact that the researcher may not always tell those under study everything about the research. This could be attributed to not knowing all that will be involved at the initial stages when negotiating access with gatekeepers, a lack of interest from participants or alternatively opting to limit the amount of information given to those taking part as some information could influence participants' behaviours thus rendering findings invalid.

As part of the process of negotiating access to the study site I utilised contacts from previous schools in which I have worked. I therefore wrote to the Head teacher of one institution, informing her of my aims, objectives, and methodology, a copy of the letter is included in Appendix 2. Following this I received an email from a colleague with whom I used to work who is now responsible for Teaching and Learning within the school, my letter to the Head teacher had been forwarded to him. I arranged to meet with him face-to-face to discuss the nature of my research in more detail. Following this meeting, it was agreed that I could conduct my research within the establishment and upon his return to work after the Christmas break. Dates for conducting research were agreed via email (Appendix 8). As I work full-time it was necessary to arrange dates which enabled me to take time away from my job.

Negotiating access to an educational establishment does not start and end with the Head teacher, particularly as other staff may be suspicious of researchers' actions and the purpose of their study. According to Hammersley and Atkinson (2007, "it is far more than the granting or withholding of permission for research to be conducted" (43). Gatekeepers often consider how the organisation

will be portrayed by the researcher and as such may insist on “blocking off certain lines of inquiry or shepherding the fieldworker in one direction or another” (Hammersley and Atkinson, 2007: 51).

A number of stages have been employed to ensure that ethical guidelines are reflected in the undertaking of focus groups. From an ethical perspective, issues of confidentiality are associated with focus groups as all participants are privy to the discussion. However, at the outset of each focus group participants were encouraged to view the focus group as a public meeting so that they could be mindful about their contributions. Focus groups were audio recorded, this is discussed in more detail below. However, from an ethical perspective all respondents were informed at the start of the focus group the reasons for the recording taking place, how the recordings were to be utilised and stored and that they were to be destroyed following the completion of the transcription process. Audio recordings were made on a voice recorder app on my iPhone (‘Voice Memo’s). Storing such data on this device is secure thanks to the face ID function, as such my iPhone can only be unlocked through recognition of my face. Once audio recordings were transcribed, they were stored on my laptop which is password protected. To conceal the identities of participants, only fictional names, or pseudonyms were applied to the transcript. At the end of the focus group a debriefing was held to enable those who had taken part to mention any concerns they may have had and ensure that my contact details were made available. It was also pertinent to inform participants that at any point they can request that their comments are removed from the transcripts. In addition, confidentiality, and anonymity, which are both ethical issues, must be upheld across all data sources, for example it is necessary to remove details, from both transcripts and the final write up, with the aim of protecting the identity of participants and the research site. However, it is pertinent to note that given the limited number of Sixth Form Colleges in the North East of England, full anonymity might not be possible.

BERA, through the publication of their ethical guidelines, aim to promote ethical values in educational research. Adherence to and consideration of such ethical values is exemplified in my study in a number of ways. Prior to conducting my research, I provided copies of both my ethics and research proposals via email to ensure that participants were fully aware of what my research

involved. Therefore, I used my positionality as a researcher to ensure that participants were fully briefed on the aims of my research and what was required of them during the process. Moreover, before both the classroom observations and focus groups participants were informed of their right to withdraw at any point. Participants need "enough information about the nature and purpose of the research for them to be able to make an informed choice about whether or not to take part" (Rapley, 2007: 25). Participants were also offered the opportunity to view any documentation produced by myself as part of the research process. The transcription process also offered an opportunity for the involvement of participants in the research process, insofar as providing a copy of the transcript for approval prior to the research progressing further. This enabled participants to express any concerns they had regarding their portrayal in the research. Some researchers suggest that people should be able to control information about them and should be able to give their permission for researchers to use it. "By assigning such ownership rights to people they can be protected from the consequences of information they regard as confidential being disclosed publicly by the researcher" (Hammersely and Atkinson, 2007: 213). However, others argue this has the potential for evidence to become distorted.

4.7 Research Quality

Evaluating qualitative research has arisen out of both internal needs and external challenges. Despite this, there are several problems associated with the evaluation of research quality. Indeed, it could be argued that this helps to explain why general criteria is yet to be developed (Flick, 2018). Evaluation of quantitative research on the other hand, is closely aligned with standardisation and cannot easily be applied to qualitative research. "Standardisation is counter-productive for research situations that are derived from a relatively flexible use of methods" (Flick, 2018: 80). Furthermore, qualitative research is often called into question as it is "so judgement dependent" (Patton, 2015: 653), whereas quantitative research adheres to rules. Patton (2015) also suggests that the credibility of qualitative research can be questioned as findings can be subject to researcher bias and qualitative methods can be deemed to be weaker than quantitative. However, the methods could be deemed different, as opposed to weaker. It is also possible to question whether a one-size

fits all approach in relation to qualitative research is appropriate given the alternating viewpoints in the literature. Elliot, Fischer, and Rennie (1999) argue for a generalist approach, which derives from a quantitative background, incorporating quality criteria for evaluating qualitative research, whereas Reicher (2000) instead favours approach-specific criteria. Nonetheless, "a general discussion about what good qualitative research is and is not sometimes misses the differences in the approach and aims of different sorts of qualitative research" (Flick, 2018: 8). Ruddin (2006) contends that generalisability for qualitative methods such as case study, is not necessary as generalisability is linked to positivism and case study does not constitute positivist research. Alternatively, Yin (2009) argues that it is the type of generalisability which is important vis à vis case study research with "analytic" as opposed to "statistical" generalisability being the focus. Indeed, statistical generalisation is not an appropriate means to generalise from a single case study, owing to the case not being a "sampling" unit (Yin, 2014). This "can help researchers to understand other similar cases, phenomena or situations…a logical rather than statistical connection between the case and the wider theory" (Yin, 2009: 380). Pring (2015) also argues that unlike generalisability from a scientific perspective, case studies have the potential to draw attention to similarities in alternative contexts.

Cases are not usually chosen at random, particularly as the issue linked to the selection of a case is that of access. Certainly, the selection of a case in my thesis was closely aligned with issues of accessibility. Regarding case study generalisation, Yin (2014) cites the concept of theoretical or analytical generalisation. "Case studies, like experiments, are generalisable to theoretical propositions and not to populations or universes. The case study, like the experiment, does not represent a "sample" and in doing a case study, your goal will be to expand and generalise theories (analytic generalisation) and not to enumerate frequencies (statistical generalisation)" (21). In other words, findings from a single case cannot be viewed as part of a larger sample, instead concern should be based around how the findings inform the development of theory. The aim is to "analyse the situation and to arrive at certain concepts, propositions or hypotheses that might

explain what is happening and why in the particular setting that has been investigated" (Denscombe, 2017: 63).

It can therefore be contended that there is clearly a lack of agreement regarding how to determine the quality of qualitative research. According to Flick (2018) outlining criteria for qualitative research, particularly criteria which are deemed appropriate amongst researchers is not easy to achieve. As previously stated, applying the traditional criteria of reliability, validity and objectivity is arguably not appropriate for qualitative data. While in quantitative research repeated data collections are synonymous with data reliability and stability (Flick, 2018), in qualitative research "identical repetition of a narrative in repeated narrative interviews is more a sign of a "constructed" version than of the reliability of what has been said" (Flick, 2018: 27). Maxwell (1992) has attempted to align validity into the domain of qualitative research. Indeed, there is agreement amongst several researchers (Gibb, 2007; Silverman, 2011) that the requirement for verification should be enacted (Denscombe, 2017). Maxwell (1992) discusses a typology of five validities: descriptive, interpretive theoretical, generalisability and evaluative and contends that the first three types of validity are concerned with the researchers' analysis and the process undertaken to interpret the material. Maxwell (1992) likens generalisability to external validity in other contexts and this, along with evaluative validity is deemed to be less important for qualitative research. A further issue regarding the generalisability of qualitative research relates to its "attachment to contexts" (Flick, 2009: 407), in other words, findings and analysis often come out of the context of a specific case and are "based on analysis of relations, conditions, processed etc., in them" (Flick, 2009: 407).

Crotty (1998) discusses the associations objectivity, validity and generalisability have with positivist research. Given the qualitative nature of my thesis I deem it to be appropriate to reject the positivist notion of validity, although it is important to account for quality within my research. One strategy to increase quality in qualitative research relates to the concept of member consensus or validation. Lincoln and Guba (1985) deem this to be "the most crucial technique for establishing credibility" (314). Indeed, Groeben (1990) utilised this approach to seek member consensus regarding statements made in an interview, in such instances consent is given for the individual case as

opposed to the subsequent interpretation. However, it could be argued that a condensation of what was discussed, rather than the full transcripts, could be seen as a form of interpretation (Flick, 2018).

Creswell and Miller (2000) discuss nine strategies of "validity procedures" which can be applied to qualitative research, the following strategies can be applied to my thesis to determine validity. To demonstrate that data can be accurate and appropriate, Lincoln and Guba (1985) use the term "credibility", which encompasses the use of triangulation and respondent validation. I have employed triangulation within my thesis, incorporating classroom observation, alongside focus groups and document analysis as a means to enhance the robustness of the data as well as generate rich data to feed into the analysis. This use of multiple data sources is termed construct validity by Yin (2014). Triangulation has the potential to show that data from different sources essentially produces the same results. Alternatively, different results can be brought to the fore. With regard to my thesis, reviewing data from three different sources served to substantiate my findings, data from focus groups supported the observations I had made in classrooms and the analysis of documents (for example, exam board specifications and assessment criteria) affirmed what teachers had both utilised in their lessons and discussed in focus groups. Thus, within the context of qualitative research, triangulation is "a strategy of managing diversity in a field under study by applying several methodological approaches covering different perspectives in that field" (Flick, 2018: 87). Simply relying on one method can increase vulnerability to mistakes associated with that method. Moreover, "different types of data provide cross-data consistency checks" (Patton, 2015: 661) and can serve to increase confidence in conclusions, thus increasing credibility. Krueger (1993) cites ten factors relating to quality control in focus groups, one of which relates to careful data handling. Problems can be associated with the transcription process which can in turn impact research quality. One such issue relates to the clarity of the recording, however, to mitigate such potential problems, I used two recording devices during the focus group interviews, a Dictaphone and a recording app on my iPhone. Furthermore, it is the role of the researcher to make judgements about what to include. In my research each respondent was given a pseudonym alongside some

information to identify them as individuals, while maintaining their anonymity, more specifically the subject they taught in the institution under study. I made verbatim transcriptions of the recordings made during the focus groups. As Poland (2002) states the aim is to account for what occurred. Denscombe (2017) argues that the inclusion of extracts from transcripts can lack context and become less meaningful as the words are no longer connected to the text which came before and after. In my thesis, data were included and incorporated to emphasise themes and ideas which were made by participants.

Furthermore, Creswell and Miller (2000) identify three lenses through which to examine the quality of a study through: the researcher, the participants and the external (readers). They then go onto attach the nine strategies to one of these lenses as well as a paradigm in qualitative research (the post positivist or systematic, the constructivist, the critical) which they argue affects how qualitative researchers approach quality (validity) issues. According to Creswell and Miller (2000) triangulation is viewed through the lens of the researcher and as a post positivist paradigm while member consensus is perceived through the lens of the participants. From the perspective of the researcher only lens, triangulation is carried out by researchers to identify within the data, commonalities such as themes/categories by discounting areas which cross over (Creswell and Miller, 2000).

Table 4-6 - Validity Procedures within Qualitative Lens and Paradigm Assumptions (Cresswell and Miller, 2000: 126)

Paradigm assumption/Lens	**Postpositivist or Systematic Paradigm**	**Constructivist Paradigm**	**Critical Paradigm**
Lens of the Researcher	Triangulation	Disconfirming evidence	Researcher reflexivity
Lens of Study Participants	Member checking	Prolonged engagement in the field	Collaboration
Lens of People External to the Study (Reviewers/Readers)	The audit trail	Thick, rich description	Peer debriefing

Despite this approach, Flick (2018) argues that this classification can be problematised insofar as it comes up against the same issue as other concepts of validity in qualitative research, “they do not come with benchmarks defining when they are fulfilled and when they are not” (92).

Generalisability in quantitative research is often obtained by use of statistical sampling procedures and cannot be applied to qualitative research. Thomas and Myers (2015) argue that case study "offers little in the way of generalisable information" (30). Moreover, Lincoln and Guba (1985) state that in relation to this issue "the only generalisation is: there is no generalisation" (314). Denscombe (2017) argues that undertaking observation as part of qualitative research can lead to the researcher becoming, to some extent, "an integral part of the data collecting technique" (327). It is therefore possible to question the extent to which another researcher carrying out the same observations would reach the same conclusions. While it is not possible to be certain of this, there are means to approach this which Lincoln and Guba (1985) term "dependability" and Yin (2014), "reliability". Thus, it is necessary to provide "a fully reflexive account of procedures and methods, showing the readers in as much detail as possible the lines of enquiry that led to particular conclusions" (Seale, 1999: 157).

4.8 Concluding Thoughts

This chapter has sought to explore the research methodology which underpins my doctoral thesis. The chosen data collection techniques have been examined alongside a consideration of the rationale for their selection. The importance and application of ethical issues has also been discussed. In this chapter I have also considered factors affecting the determination of quality of qualitative research and the means I have employed to mitigate any quality related issues. The subsequent two chapters will now begin to outline the key findings of the research following the analysis of the data which has been collected.

5 Analysis Chapter: Part 1- Using Formative and Summative Assessment in the Classroom

5.1 Introduction

The aim of this chapter is to present and begin to draw together the findings from both the focus group interviews, lesson observations and document analysis. Key and recurrent themes relating to how teachers operationalise formative and summative assessment with their students in Years 12 and 13 and their associated challenges will be explored. This discussion will subsequently enable me to provide a response to my first research question (Section 7.1 below): How do teachers use formative and summative assessment strategies with their students and why do they use them? There will also be a consideration of the actions and responses of students in relation to assessment practice as well as the impacts the changing examination structure has had on teachers' approaches to assessment and whether high stakes testing affects their approaches to assessment. This will feed into my discussion of my second research question (Section 7.2): How does high stakes testing affect teachers' approaches to assessment? Links will also be made to the body of literature, previously discussed in Chapters 2 and 3, to sit the findings of this study within the wider context of the assessment for learning literature as a whole.

Teachers who participated in the first focus group represented diverse subjects: Business Studies, Sociology, Physics, Biology, Media Studies, Health and Social Care and Law. How these subjects are assessed also differs with A Level qualifications having linear examinations and BTEC courses, such as Law and Health and Social Care, centred predominantly on coursework. Both A Level and BTEC programmes of study are included in my research as they are representative of the curriculum offered at Mitford College. Distinct aspects relating to how they are assessed make them worthy of separate treatment. Teachers from PE, Criminology, English, and Media Studies populated the second focus group. The sample has been fully discussed and theorised in the methodology chapter.

5.1.1 What was Observed: Initial Findings

As part of this study twenty-eight lessons were observed, each for thirty minutes in duration, Appendix 7 shows the range of subjects, a combination of A Levels and BTECs and their associated observation number. A label has been created and assigned to each subject, these are referred to in Appendix 1 to identify the subjects in which different types of formative and summative assessment strategies were observed. I selected focus groups to follow up key themes which I observed during lessons and discuss teachers' rationale for their use of assessment strategies. Moreover, this approach generated rich data which fed into the subsequent transcription and analysis which has informed the content of this chapter.

As shown in Table 5.1 formative assessment strategies were more commonly observed than summative assessment strategies, questioning was observed in all lessons. It is important to note that some subjects (Biology, Business, Criminology, History, Maths and Media) were observed more than once. In addition, strategies to gauge understanding and recall were observed in 10 lessons, as shown in Table 5.1 below. However, other formative assessment strategies, namely graphic organisers and peer assessment were observed less often. Formative assessment strategies classified as other included verbal feedback (incorporating advice on how to achieve higher marks), exam skills and self-audit, key term bingo and traffic lights to review progress in a lesson. Regarding summative assessment strategies, exam style questions were most frequently observed, taking place in 9 lessons. Students were provided with marks and/or grades in 5 lessons. However, no instances of multiple-choice tests or unseen examinations were observed. The latter could be attributed to students having recently sat mock examinations in each of their subjects.

In addition, it is possible to identify several common factors evident in subjects, which were observed on more than one occasion with a different teacher. For example, in History, lesson objectives were displayed on the whiteboard and referred to at the start of both observed lessons (Observations 15 and 19) to gauge student understanding at the outset. Learning objectives were then returned to later in the lesson to enable students to reflect on their progress so far, in relation to those objectives. Furthermore, students were asked to self-assess their work, with the teacher

of one Maths class asking students to correct any wrong answers with a purple pen (in order to clearly show anyone who looked at the students' book that the work had been self-assessed) (Observations 6, 14 and 28). Whereas in Geography, following the completion of a 5-mark exam style question, students were asked to self-assess their answer against the mark scheme which was displayed on the whiteboard (Observation 5). Students were also encouraged to use a different coloured pen when carrying out peer assessment in Criminology (Observation 1). Finally, in Biology, lesson objectives were also referred to and displayed at the start of each lesson, with one teacher identifying when an objective had been met through the use of different coloured fonts (Observations 9 and 18).

Table 5-1 - Frequency of Assessment Types Observed in Lessons.

Type of assessment	Frequency	Subject
Formative Assessment		
Sharing learning objectives	7	HISx2, BIOx2, LIT, SOC, FILM,
Questioning	28	FR, ENG, COMP, HIS, ACC, CLA, PE, BIO, PHY, HSC, GEO, BUS, CRI, CHEM, MED, LIT, RS, LAW, SOC, MA, FILM, ART,
Peer Assessment	2	BIO, CRI,
Self-Assessment	6	FR, MA, BIO, GEO, BUSx2,
Assessment strategies to gauge understanding	10	HIS, CLA, PHY, Ma x2, LIT, BIO x2, FILM, HSC
Modelling	5	ENG, ACC, ART, CHEM, CRI,
Graphic Organisers	3	ACC, HSC, CRI,
Other	12	FR, ACC, PE, BIO, ART, HSC, BUS, CRI, CHEM, RS, LAW, MA
Summative Assessment		
Exam style questions	9	ACC, CLA, BIO, GEO, MA, CHEM, CRI, SOC,
Written essay/report	1	FR,
Presentation	1	LAW,
Performance task (to assess a set of skills)	1	ENG,
Feedback- marks/grades	6	ACC, MA, CLA, BUS, CRI, RS,
Other	3	CHEM, MED, RS,

5.2 Assessment Frequency

Despite this variance in subject and final assessment style there was much agreement amongst staff relating to the frequency of classroom assessments and their approach to reviewing feedback and student progress. Paul (Business Studies Teacher, Focus Group 1) discussed the requirements

regarding assessment frequency which are documented in the institutions' assessment policy to be followed by all staff. It is therefore expected that two pieces of student work are quality marked (that is the provision of detailed feedback) by subject teachers over the course of a half term. Such detailed feedback requires teachers to comment on the positive aspects of a students' work before identifying areas for improvement, often in relation to the assessment criteria. Students are then able to use this information to improve the quality of their work and ultimately close the gap in their learning. However, teachers discussed carrying out assessments and providing feedback even more regularly than this.

Teachers described using formative assessment "almost constantly" and "probably most lessons". Examples included use of mini whiteboards to promote recall of information at the start of a lesson, verbal feedback, particularly when doing coursework or controlled assessment and questioning. As previously stated in Section 5.1.1, questioning was identified in all lesson observations. Formative assessment was deemed by participants to be part of best practice, embedded within teaching and learning with the aim of generating opportunities for teachers to acquire feedback from their students in order to enable effective teaching to occur (Black, 1998: Cowie and Bell, 1999). Although, it is possible to question what is meant by the term "best practice" here as such a term can have different meanings and be interpreted differently by individuals.

> [Formative assessment should be] embedded in teaching and learning… [to evaluate] where they are in terms of the learning. Have they grasped what you are teaching them? And it's always informing where you go next with your lesson.
>
> (David, PE Teacher, Focus Group 2).

Teachers discussed undertaking assessments with their students, approximately every 2-3 weeks, taking the form of an essay or past exam style question. In Biology and Physics teachers stated that following such assessments they provide formative feedback for their students to act upon and encourage them to identify areas they need to improve.

> They improve them using set protocols in green pen, traffic light mark schemes, red, amber, green, and then they do some independent work on the areas they need to improve (Louise, Biology Teacher, Focus Group 1).

There is much work in the literature to suggest that feedback is a key component of formative assessment. However, Ramprasad (1983) argues that it is only feedback when it is used to close the gap in students' learning. In addition, Sadler (1989) cites the importance of the role of the student in closing this gap, thus reflecting a constructivist approach to learning. As stated by Louise (Biology Teacher, Focus Group 1), students play an active role in the feedback process.

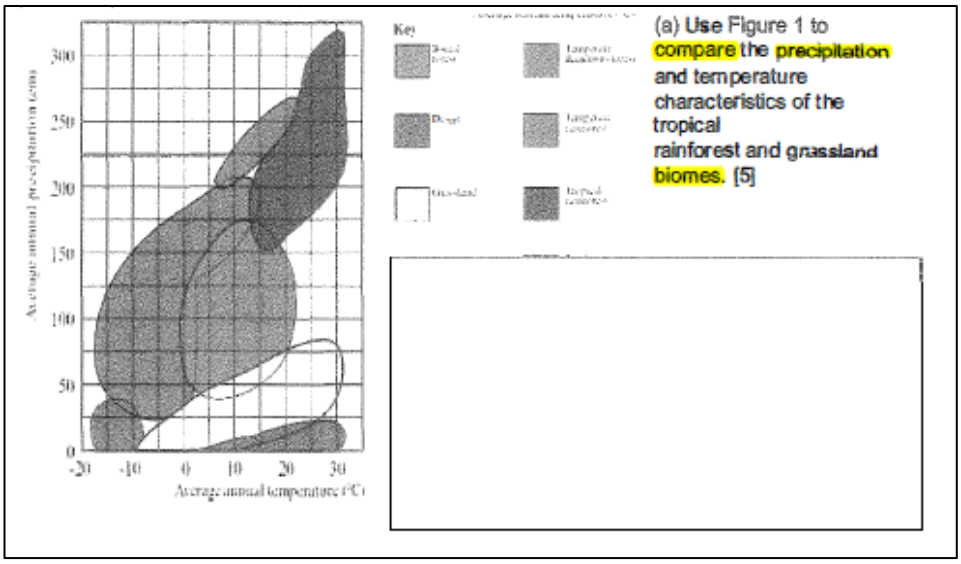
A similar process was described by Lesley,

> I would say that what I do in A Level Sociology and Health and Social Care is very similar, we give them the question and the time that they would get in the exam (Sociology Teacher, Focus Group 1).

Timed assessments were observed in several lessons, Criminology, Geography and Classics. In Criminology students were given 5 minutes to complete a 5-mark exam question. Students were not given any pointers from the teacher as to how much time they had remaining due to changes in exam rules whereby students can no longer have watches/be told timings in exams. This was then followed by peer assessment while the teacher asked students a series of questions to build up a model answer on the board. Students were reminded of the need to tick their partner's work/add in any missing points (Field Notes, Observation 1) to ensure the answer in their partner's book was fully complete. The teacher also tried to encourage more students to contribute to the class discussion by requesting that "someone who hasn't spoken" answer a question each time (Field Notes, Observation 1). Thus, it can be argued that the teacher was trying to increase the reliability of questioning as a method of formative assessment by encouraging everyone to take part. According to Walsh and Sattes (2005) allowing some students to become non-responders promotes "inequitable learning opportunities" in classrooms (77).

In Geography students were also given 5 minutes to complete a 5-mark question, as shown in Figure 5.1 below, however, prior to students answering the question the teacher used questioning to check students understood the command words and were given some pointers relating to the resource the question was based on. “When asked to compare, what are we doing? What is a comparison?” (Field Notes, Observation 5). This style of examination question is particularly significant as students are required to complete not only short answer and longer essay style questions but also data interpretation questions based upon an unseen resource such as the one depicted in Figure 5.1 as part of their external summative assessment.

Figure 5.1 - Practice Exam Question and Resource (used in Observation 5)



In the above Figure the command words are highlighted. The teacher used questioning to elicit from students an understanding of what the question was asking them to do. The following pointers were discussed: what is shown by both the horizontal and vertical axes as well as what the colours on the graph represent. In addition, students were reminded of the need to only focus on tropical rainforests and grassland biomes, as per the question.

As in Criminology students were not informed of how long they had left to complete the question. Upon completion the mark scheme was displayed on the board and students were asked to self-mark their work. In Classics students were given a 10-mark question and the teacher utilised questioning to check students’ understanding of key terms and concepts (highlighted in Figure 5.2).

"What does a good answer to this look like? Is there a technical word we need to clarify?" (Field Notes, Observation 11). Students were informed that they would have 13 minutes to complete the question as per the exam and the start and end time was written on the board. While students worked on the question the teacher moved around the room and provided some verbal feedback.

Figure 5.2 - Practice Exam Question and Resource (used in Observation 11)

1. Explain how this statue creates an image of Augustus as Imperator (10) c. 13mins

2. Evaluate which was more effective in conveying Augustus' image as imperator to the citizens of Rome, literature or visual/material culture. (30) c.42mins

The teacher used questioning to review students' understanding of key terminology in the question. Examples of technical language, in other words, language which is specific to the subject under study are highlighted.

In contrast to both Geography and Criminology, students were informed of the time remaining. The question was set aside to be teacher marked. Thus, teachers were drawing upon both summative and formative assessment strategies in this regard to review progress and aid student improvement. In the observations discussed above the teachers utilised questioning to elicit student responses either relating to their understanding of the question or their answer. Therefore, onus was placed upon the students and their actions and responses. In addition, teachers in a range of subjects were observed encouraging students to expand and embellish their answers, for example in Sociology and History use of specialist terminology was encouraged. In French questions

were asked and answers were expected in French. While in English, Media and History student responses were often met with “Can anyone add to that?” “Can we be more precise about that?” “Tell me more please” and “How do we know that?” Similarly, in Criminology the teacher sought further information from students: “give me an example”, “imagine I’m the examiner marking this paper, what else do I need to hear about?” (Field Notes, Observation 27).

5.3 The Cycle of Assessment

Approaches employed to review work and exam papers in class were discussed in depth in both focus group interviews. Several examples of this were seen during the classroom observations. This process was described as cyclical by one member of staff and was said to be an intentionally timely process to promote the recall and revisiting of information. Several examples of pre-planned formative assessment were observed in lessons (Cowie and Bell, 1999). For example, in Religious Studies students were using the mark scheme along with comments from their teacher to make improvements to their work and ultimately close the gap in their learning. Also, in Business Studies students were required to review their test papers and carry out a self-audit. Students were asking questions of the teacher while they worked, for example, “I would have got the mark if I had…” They were reflecting on where they went wrong (Field Notes, Observation 23).

Figure 5.3 - Self-Audit (used in Observation 23)

Score

Q	Focus	Score	Further Revision Required
1	YED		
2	%Change Market Share		
3	GPM		
4	Break Even		
5	Current Ratio		
6	Profit		
7	Sales Revenue		
8	Corporation Tax		
9	Acid Test Ratio (Difference)		
10	% Growth & Difference		
11	GPM		
12	CPA		
13	Cap. Utilisation (production volume)		
14	Profit Variance		
15	Exchange Rate		
16	YED		
17	PED		
18	S&D		
19	S&D		
20	Crowdfunding (Written)		
21	S&D		
22	S&D		
23	ROCE		
24	Decision Trees		
25	Stock Chart		
26	CPA		
27	Labour Productivity		
28	OPM		
29	% Change		
30	% Change		
31	PFYM		
32	Product Differentiation (written)		
33	Impact of Cap. Uti. (written)		
34	S&D		
35	S&D		
36	20 Marker Investment Appraisal		

Undertaking a self-audit encouraged students to reflect upon their performance in the recent mock exam and identify areas of strength and areas requiring further improvement.

Furthermore, students completed an exam skills audit to review the skills they need to utilise to answer particular types of question.

Figure 5.4 - Exam Skills Audit (used in Observation 23)

Question Type	What skill(s)/ techniques should you use Or How should you approach these type of question
4 Mark Calculation Q	
4 Mark Explanation Q	
8 Mark	
10 Mark	
12 Mark	
20 Mark	

The exam skills audit encouraged students to examine the differing skill requirements of a range of examination style questions, from a 4-mark calculation question to a 20-mark essay question, representing a greater focus on skill acquisition.

Examples were also discussed in the focus group interviews, in Physics and Business, staff referred to carrying out reviews over several weeks to promote the revisiting of information to aid retention.

> They are upgrading question by question over the next couple of weeks with some guidance from me.
>
> (Alan, Physics Teacher, Focus Group 1)
>
> I break it up and do a question per week for the next 4 weeks.
>
> (Paul Business Teacher, Focus Group 1).

According to Hattie and Jaeger (1998) feedback should be "polymorphous, referring to subsequent information aimed at assisting the learner in meeting the goals of the learning process" (113), thus suggesting that the provision of feedback is an on-going process and not simply a one off. This

process was also linked to staff CPD based around the concept of learn, forget, learn. Several staff members championed this approach with one stating “it seems to work better, just that kind of constantly revisiting stuff” (Paul, Business Teacher, Focus Group 1). Black (1998) argues that there is a need for teachers to partake in CPD to successfully support their students with formative assessment.

Cowie and Bell (1999) also cite interactive formative assessment, which can occur at any time during student-teacher interactions. This form of formative assessment often occurs on an ad hoc basis, out of a learning activity. While teachers may plan to undertake interactive formative assessment, they are unable to “plan for or predict what exactly they and the students would be doing or when it would occur” (Cowie and Bell, 1999: 107). Examples were also observed in lessons, for example in Criminology, the building up of a model answer on the whiteboard based on student responses following teacher questioning.

Simon (Criminology Teacher, Focus Group 2) discussed the approach employed when setting end of topic assessments. Instead of teaching a topic and assessing students summatively straight after he said he had elected to set the assessment mid-way through the teaching of the next topic to mirror the exam.

> It’s trying to gear them more towards that exam where you are going to learn something in January and get a question on it in May/June time (Simon, Criminology Teacher, Focus Group 2).

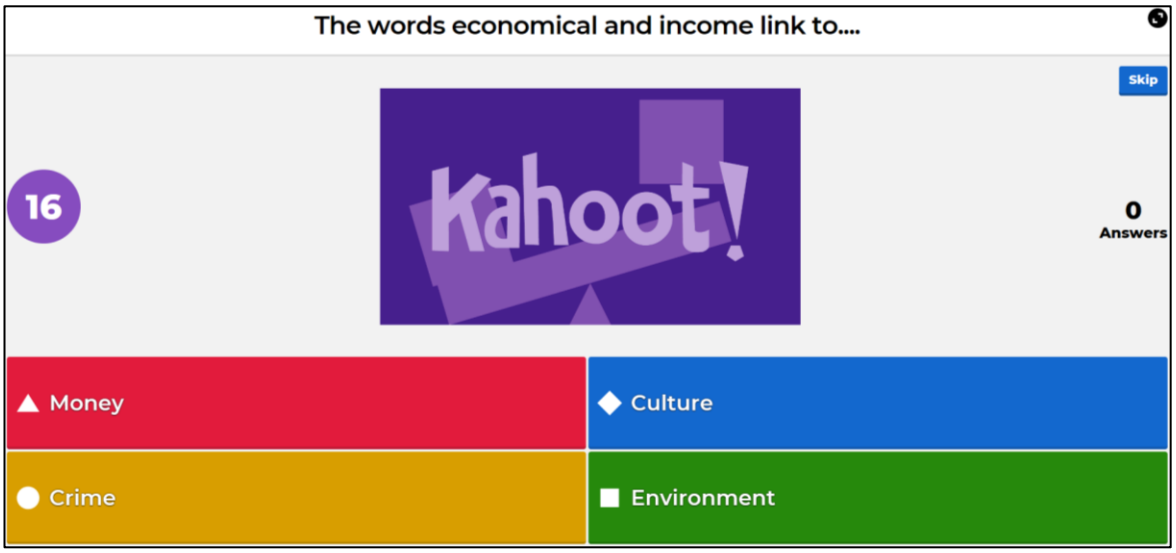
Completing a test at the end of a topic was deemed as giving a false impression of how much they know. Setting a summative assessment on a previous topic while part way through the teaching of the next topic necessitates students to have to recall and revise information to check their understanding rather than simply being told something and subsequently tested on it. In other words, they are being given the opportunity to “forget” material, therefore providing a more realistic insight to what the final examination will be like.

Additional formative assessment strategies were referred to by Amy (Health and Social Care Teacher, Focus Group 1), examples included the use of interactive resources such as Kahoot to review what students have learnt during a lesson or as part of a wider topic, (as shown in Figure 5.5 below). Further examples were observed in lessons, for example Bingo, Connect4, and Blockbusters, which were described by Amy (Health and Social Care Teacher, Focus Group 1) as

> just a really good way of rounding up a topic or a lesson as students can all participate by using their phones.

This was utilised as a revision strategy to assess students' knowledge of Unit 1 topics and content as they had an exam coming up in two weeks.

Figure 5.5 - Kahoot Revision Tool



Kahoot is a freely available, online game-based learning platform. Teachers can quickly and easily generate games to support learning or revision. Students are then able to participate by using their phones. Students are required to freely download the Kahoot app, before entering a game pin to join each game. They then choose their answers to a series of revision questions and submit their answer via their phone.

Figure 5.6 - Online Interactive Revision Game- Blockbusters (Teachers direct, 2005)

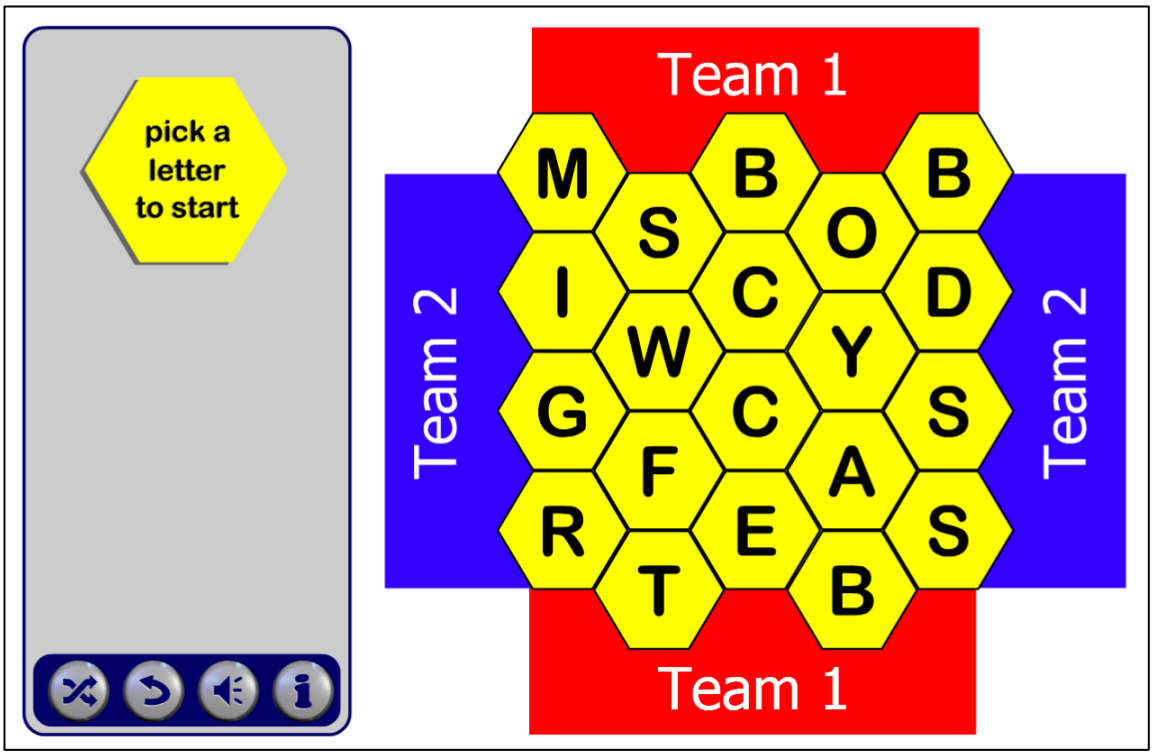
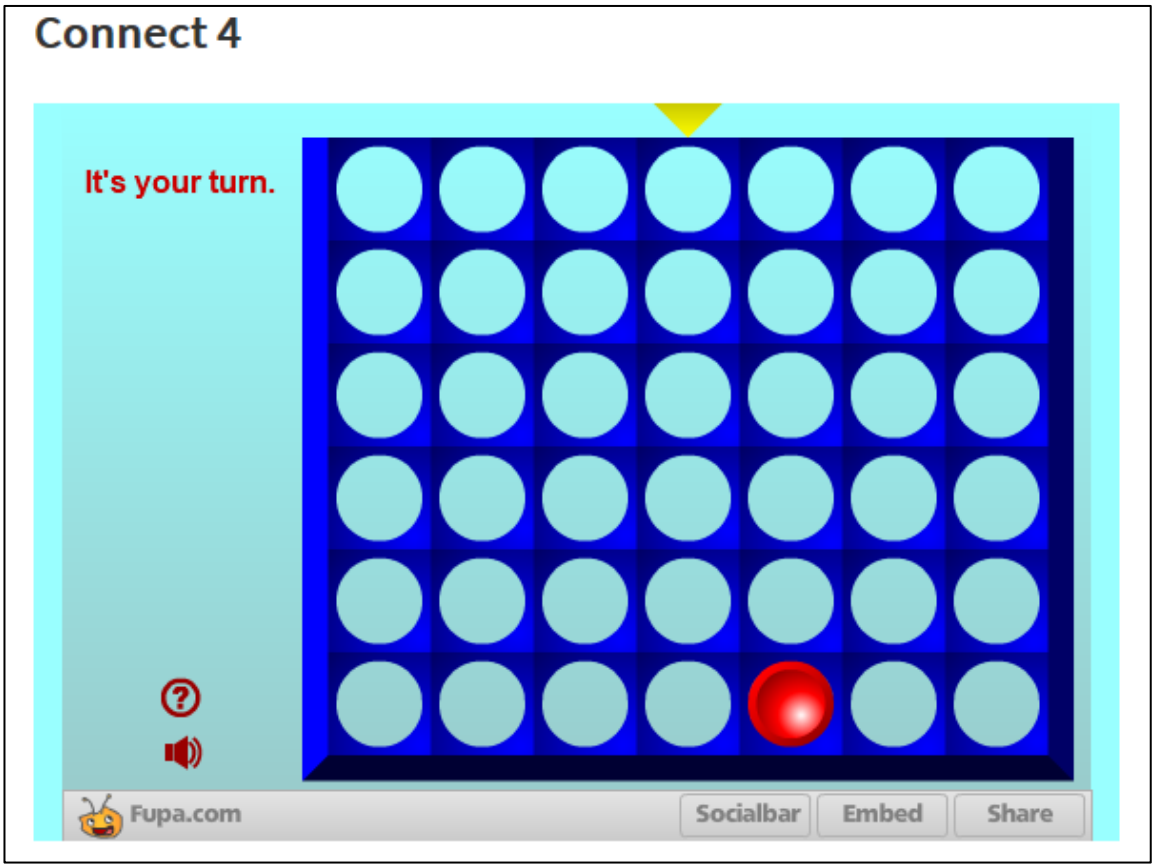


Figure 5.7 - Online interactive revision game- Connect 4 (Revision World, 2019)



Interactive online resources were not referred to by teachers of A Level subjects. Despite this, one student was observed making use of Kerboodle during a chemistry lesson which could be suggestive of not only independent learning but also a need to clarify understanding of a certain phenomenon (Observation 2). Kerboodle enables students to access a range of resources in a digital format to support their learning in class, examples include auto-marked tests, videos, animations, and podcasts. Online Bingo was observed as part of a Health and Social Care lesson to aid with the revision of all Unit 1 topics in preparation for the forthcoming examination.

5.4 Challenges to Formative Assessment

While much emphasis and value were placed upon formative assessment, several challenges were raised relating to its successful implementation. One such challenge was that of time.

> For example, we do a lot of assessment in these two-three-week cycles, you are talking about a 30, 40, 50-mark bit of work which you mark with short feedback, even then it still takes a bit of time (Alan, Physics Teacher, Focus Group 1).

As such some teachers discussed strategies they have implemented to mitigate some of the time related difficulties. In Science, teachers rewrite the mark scheme to produce a student friendly version, which acts as a supporting crib sheet alongside their individually marked work. Students are directed to this resource when they require feedback relating to common misconceptions. This was deemed to be a successful strategy.

> I've found that it really helps because it stops me as I'm about to write something on the paper, adding that it's almost like just a kind of little prompt, look at this bit on that, so it's like a connecting resource (Louise, Biology Teacher, Focus Group 1).

Participants in Focus Group 2 discussed engaging in Continuing Professional Development (CPD) to look at marking strategies and the extent to which it is necessary to mark a piece of work in its entirety in order to provide the same level of feedback to students. While whole school staff CPD was concerned with assessment, individual members of staff were focusing on specific assessment strategies and testing them out as part of their day-to-day practice. Simon (Criminology Teacher,

Focus Group 2) discussed his experiences of using this approach with his classes. Other identified challenges related to the ability of the teacher to assess the whole class. For example, in Criminology, where class sizes can be up to 30, the use of questioning only enables a certain number of students to be asked.

> I gave comments on the first question, the second question was similar, didn't give comments on it. I'm going to compare it with questions in the mock exam, I'm going to see if they have made any major improvements from that final end of topic assessment with the mock.
>
> It is very difficult in a lesson that is an hour long to teach them all something and understand whether they have all actually got something out of it (Simon, Criminology Teacher, Focus Group 2).

Without getting all students to produce written work which is teacher marked, it is difficult to see whether all students have made progress. Therefore, alternative strategies, such as the use of mini whiteboards were discussed to enable teachers to assess the progress of their students quickly and easily at both the start and end of a lesson. Thus, students can write their response on the whiteboard and hold it up for the teacher to see. While students could see the answers of their peers, any attempts to alter their own whiteboard response would be noted, not only by the teacher but also those around them. Furthermore, it could be argued that this is compounded by the quick-fire nature of the task. In Health and Social Care (Observation 26) whiteboards were used to enable students to test their knowledge of key concepts through playing Bingo. Whereas in Maths whiteboards were utilised as part of the plenary for students to note down their answers to quizzes to review their learning in the lesson (Observations 6, 14 and 28). This was also the case in Biology (Observations 9 and 18). In Chemistry students were asked to use their whiteboards to sketch two graphs (the concentration of iodine versus time and the concentration of manganese versus time) to demonstrate their knowledge (Observation 2). Despite this, in all observed lessons teachers utilised student names when asking questions. While they are unable to ask each student in the class, they are therefore able to check the knowledge of specific students as opposed to the

same students repeatedly putting their hands up. In a Criminology lesson the teacher was observed moving around the room to direct her questioning to those sitting at the back (Observation 1).

Black and Wiliam (2010) argue that formative assessment can improve the performance of low achievers, however, time constraints can hinder the formative process. Following the completion of mock examinations, a lack of time prohibited the provision of detailed feedback to students.

> What you need in a perfect world is one-to-one time to say, "this is where you've gone wrong on this question", but you obviously do not have the time to do it (Ashley, English Teacher, Focus Group 2).

5.5 Student Actions and Response

As previously stated in Section 5.2, staff placed much emphasis on student action and response. While student response to teacher feedback was deemed to be important, teachers also stated that they aimed to up-skill their students, developing their independent thinking, reflecting the constructivist model of learning that Assessment for Learning/formative assessment rests on.

Evidence of changing student perceptions towards assessments was provided by Alan (Physics Teacher, Focus Group 1). Alan used the term "up-skill" to describe the process of students getting to know how assessment systems work and how they need to respond to them. Staff highlighted the particular importance of this in light of changes to examination structure and the movement to linear A Levels as students need to become resilient as at first, they may not achieve academic success at the level they have been accustomed to. For example, in Science students need to understand that they will not fully comprehend everything they study and should aim for an understanding of circa 40% of subject content, which Alan attributed to the grade boundaries being set so low.

> I've started to think that if they up-skill themselves, you know, over the course of the year, year, and a half, I'm back to my primary task of making them understand the knowledge that they can then apply themselves […] That's quite hard, if you are giving them a test and

they do not understand any of it, then you fail, it's quite demoralising (Alan, Physics Teacher, Focus Group 1).

Figure 5.8 - Subject Grade Boundaries for Physics Examinations 2018

AQA

Subject grade boundaries – June 2018 exams

A-level – reformed linear

Subject Code	Subject Title	Maximum Mark	Grade Boundaries A*	A	B	C	D	E
7357	MATHEMATICS ADV	300	230	181	158	135	112	90
7272	MUSIC ADV	300	242	209	179	149	120	91
7582	PHYSICAL EDUCATION ADV	300	211	194	170	146	122	99
7408A	PHYSICS ADV (ASTROPHYSICS)	250	183	150	125	100	75	51
7408E	PHYSICS ADV (ELECTRONICS)	250	186	153	128	103	78	53
7408C	PHYSICS ADV (ENGINEERING)	250	184	152	127	102	77	52
7408B	PHYSICS ADV (MEDICAL)	250	181	147	122	98	74	50
7408D	PHYSICS ADV (TURNING POINTS)	250	181	149	124	99	74	50
7182	PSYCHOLOGY ADV	288	214	190	161	132	104	76
7062A	RELIGIOUS STUDIES ADV OPTION A	200	168	140	114	88	62	36
7062B	RELIGIOUS STUDIES ADV OPTION B	200	168	140	114	88	62	36
7062C	RELIGIOUS STUDIES ADV OPTION C	200	168	140	114	88	62	36
7062D	RELIGIOUS STUDIES ADV OPTION D	200	168	140	114	88	62	36
7062E	RELIGIOUS STUDIES ADV OPTION E	200	168	140	114	88	62	36
7192	SOCIOLOGY ADV	240	182	164	142	121	100	79

In order to attain a pass (grade E) on the Physics examination, students are required achieve circa 20% of the marks on each paper they sit (in other words 50 to 53 (depending upon the paper) out of 250 possible marks). Whereas a grade B requires a student to achieve approximately 50-51% of the marks (between 125 and 128 out of 250 marks,) dependent upon which paper they sit. Therefore, this reinforces the point made by Alan (Physics Teacher, Focus Group 1) who states that students do not need to fully understand everything they are taught in order to pass the exam as the grade boundaries are low (AQA, 2018).

Changing attitudes towards assessments were also deemed to play an important role in day-to-day teaching. For example, the importance of getting students out of the habit of using their notes to answer a question was discussed. However, an alternative response was noted by Ashley (English Teacher, Focus Group 2) who said that students were often able to use their notes as a crutch for

the first assessment in Year 12. In Classics students were observed completing a timed examination question. Prior to beginning this task, students were informed by the teacher that they could not have access to their notes for the duration of the task. In other words, the teacher was attempting to imitate exam conditions. It could be argued that this constitutes an important part of the learning process, particularly as students now sit examinations at the end of a two-year programme of study, rather than at regular intervals over the course of the two years. As such, emulating exam conditions as far as possible in class helps students to prepare for what the real, final examination will be like. It is possible to argue that the validity of such assessments could be affected should the assessment being completed be teacher as opposed to exam board generated, an issue which has arisen following the transition to linear examinations and the subsequent dearth of exam board resources.

> Our challenge has been the lack of resources that's out there on the new spec. So, although we spend a lot of time marking these mocks, we are spending a lot of time creating them. And we're creating model answers and standard answers based on what we perceive as being how they are going to be marked…So we are a bit in the dark still in terms of what we are asking students to do (David, PE Teacher, Focus Group 2).

With assessment material being generated by teachers, validity could be called into question as teachers are not fully aware of exam board requirements. Conversely, Stobart and Gipps (1997) contend that validity is not simply the domain of the test developer, there is also a role for professional responsibility in terms of how the results are used.

In addition, it was suggested by Robert that the attitudes and habits of students at GCSE level can affect their approaches to assessment in Sixth Form. For example, at GCSE a bright student will be able to do very little revision and still do ok. Thus, value was placed upon changing student perceptions of assessment. If

> "they came here with that attitude, it's getting that out of them and it's making them think in a different way" (Robert, Media Studies Teacher, Focus Group 1).

Paul (Business Studies Teacher, Focus Group 1) discussed the changes students encounter in their learning careers as they embark upon their A Level courses, therefore, further reinforcing the importance of changing student perceptions vis-à-vis assessment. Changing student perceptions from assessment at GCSE was also discussed in relation to how students approach an exam question. Different attitudes to assessment at GCSE level were also acknowledged by Alan (Physics Teacher, Focus Group 1). Furthermore, teachers referred to the attitudes and perceptions they encounter from some students regarding the organisation of their course documents.

> So, students are with us at the moment embarking linear courses but probably come from modular GCSEs, so they're not used to that structure of learning really (Paul, Business Teacher, Focus Group 1).

> It's getting away from the GCSE of write 16 lines, "I'm bound to get something" […] One of my Year 13s, I'll never forget, it was on consultation evening and he said that he used to be one of these lads that would, you know, 2 weeks before the exam, "I'll buy the revision guide, I'll just go through it and I'll walk in and do it" (Alan, Physics Teacher, Focus Group 1).

> I'm hammering on about files and they go "at GCSE I just got the revision guide", I say "that won't be happening. There will be a file check every week!" (Robert, Media Studies Teacher, Focus Group 1).

As such much emphasis and indeed value was placed upon developing students' abilities to be independent learners and acknowledge the importance of working on and building their assessment skills over the course of the two-year period of study.

Positive student perceptions of assessment were observed in a few lessons. For example, in Religious Studies students were asked to review their essay feedback and improve their work by responding to teacher comments. The teacher circulated around the room and spoke to students on a one-to-one basis. Several students asked questions concerning what they need to do to get to the next grade boundary. Similarly, in Chemistry students were observed asking the teacher

questions to clarify their knowledge and understanding of what they would be required to do in the exam.

Student responses to the provision of feedback in Criminology were discussed by Simon (Focus Group 2). Students are provided with written feedback in the form of numbered improvement points with the aim of getting them to think about how they could improve their own work as opposed to relying on what the teacher tells them to do differently. Thus, the onus is placed onto the student which overall, they were said to respond well to as it makes clear what is expected of them.

5.6 Challenges to Summative Assessment

Challenges were also identified in relation to the successful implementation of summative assessment in the classroom. As with formative assessment, time was discussed as a key challenge related to summative assessment, particularly the volume of marking required to turn around mock exams. There was not deemed to be sufficient time to provide students with individual formative feedback. Teachers placed value on the allocation of time to whole-class feedback, for example to discuss the identification of common misconceptions as well as areas of strength. Therefore, it can be argued that while mock exams provide a means of "rehearsing" for the summative, useful feedback can still be given to students, albeit on a whole class as opposed to individual level.

Alongside this a further benefit was recognised.

> It was amazing how many of them were shaking their hands and getting fatigue in their hands, as they haven't written solidly for 2 hours for a while (David, PE Teacher, Focus Group 2).

Moreover, Simon (Criminology Teacher, Focus Group 2) discussed the benefits of students sitting a full mock exam in each subject.

> I think getting over that kind of anxiety and maybe reducing the stress levels from having practiced it and them having succeeded at it must have some benefit as well.

Alan (Physics Teacher, Focus Group 1) and Ashley (English Teacher, Focus Group 2) did not advocate the giving of grades to students early in the academic year. However, in the lead up to their final exams, following the completion of a mock exam, it was deemed appropriate to provide a grade, particularly as there is lack of time to mark with comments. For example, at the time of the second focus group the Criminology exam was 19 days away and the recently completed mocks were to be teacher marked and fed back to the class, leaving only 2-3 lessons between the mock and final exam for revision and work on areas requiring improvement. Thus, the up-skilling of students over the 2-year course really comes into its own here as:

> it is really reliant on them being able to identify the areas where they can improve based on their papers and their marks (Simon, Criminology Teacher, Focus Group 2).

Regarding up-skilling their students, teachers referred to developing the ability to be an independent thinker and set personal challenge (James and Pedder,2006). Furthermore, the importance of increasing resilience amongst students was also discussed to enable them to confidently tackle anything that comes up on the paper and changing attitudes towards assessment, particularly considering the move to linear examinations.

> The really powerful bit is getting them to look at assessment in a different way. Once they start to see the value of it being protracted and looking at mark schemes and independently thinking and challenging themselves, it certainly has benefits (Alan, Physics Teacher, Focus Group 1).

While feedback can be given on a whole class basis, for example relating to common misconceptions and improving answers, there is a lack of time to discuss performance on a one-to-one basis, particularly with those students who are less well placed to review their own progress.

> What the ones that haven't done well need is someone to sit down with them and say "this is where you have gone wrong, this is what you should do to improve it" but again you haven't got the time to do that (Ashley, English Teacher, Focus Group 2).

In Media Studies two students had recently got an E and a U grade in their mock exam but were now able to learn from their mistakes; “yeah I’m not going to do what I’ve done in the mock again”. The increasing resilience of students and their improving work ethic was attributed to the cyclical approach to assessment by Robert. As previously stated in Section 5.5, students were observed, in Religious Studies asking the teacher how they could progress to the next grade boundary.

5.7 From Modular to Linear

There was much agreement that in a modular system with regular examination periods, much time was spent on exam preparation (Baird *et al*, 2019). A modular examination system involves students taking exams in stages over a two-year period with opportunities for resits in the next exam series. The transition to a linear system, across all subjects represented in this study, now sees students sit final exam(s) at the end of a two-year period. Consequently, students no longer have the ability to re-sit a poorly performed exam in January or the summer exam series. One teacher emphasised the value of deep learning as a means of up-skilling students and discussed how his classroom had become “an arena for discussion and debate”.

> We were constantly teaching to, we have to get this, and you will be tested on that, especially in the days when we had January exams […] When it comes to those end of year exams learning is so deeply embedded that they cannot help answer questions in a certain way (Robert, Media Studies Teacher, Focus Group 1).

In other words, adopting a deep approach to learning could be seen as a means to encourage a movement away from a surface learning approach, to better prepare students for examinations. It is the process of cognition, not the learning, which differs between approaches, with a surface approach to learning not promoting an in-depth exploration of subject material. Baird *et al*, (2019) argue that modular exam series do not promote deep learning. Robert’s (Media Studies Teacher, Focus Group 1) emphasis on deep learning is therefore in contrast to those surface learning approaches which can be adopted in the wake of preparing for high stakes summative assessments.

Instances of discussion and debate were observed in Media Studies whereby group discussion was based around the appeal of computer games for their audience.

Figure 5.9 - A Level Media Studies Specification

Audiences	**Media Forms**
How audiences are **grouped** and **categorised** by media industries, including by age, gender and social class, as well as by lifestyle and taste	Advertising Newspapers Radio Video games
How media producers **target**, **attract**, **reach**, **address** and potentially **construct** audiences	Advertising Newspapers Radio Video games
How media industries target audiences through the **content** and **appeal** of media products and through the ways in which they are marketed, distributed and circulated	Advertising Newspapers Radio Video games

The highlighted section on the document above signifies the link between the exam board specification and the task undertaken in the observed lesson. As such, students are required to have knowledge of how media producers attract audiences in relation to video games (Eduqas, 2019).

While in History students also worked in groups to discuss questions relating to the Cold War. Similarly, in Film Studies students discussed their knowledge of the Spanish Civil War and Franco in relation to the way in which social, cultural, and historical factors can influence film context with their peers

Figure 5.10 - Context Sheet (used in Observation 13)

Context and film - Looking at time and place

Year / Decade the film is set	
What is the significance of this time period?	
Year / decade the film was made	
How might this affect the way the film depicts events? (Does it reflect current situations?)	
Are any major events featured/ referenced in the film? If so what?	

Students were observed using a context sheet to explore the background to a film they were studying.

The teacher emphasised the importance of the task by stating that students needed to be able to discuss the significance of the historical period in which the film was set and apply this knowledge to compare two films in the final exam. Students' responses were then reviewed with the teacher using questioning to explore their ideas. Any additional information gleaned from the class discussion was recorded on the context sheet. Thus, it can be argued that in this context students were self-assessing their own work and adding any further information they deemed to be important or which they had omitted and ultimately taking ownership of their own work.

Figure 5.11 - A Level Film Studies Specification

Component 01: Film History		
Topic	**Key Idea**	**Learners should have studied**
Contexts of filmmaking	Social, cultural, political, historical, institutional	• the social, cultural, political, historical and institutional contexts in which the films studied are made.
Film movements and stylistic developments	Critical approaches to film narrative	• film narrative, including the formalist and structuralist conceptions of film narrative.
	Critical approaches associated with film	• the claims of naturalism and realism as against the expressive.

The highlighted section on the document above signifies the link between the exam board specification and the task undertaken in the observed lesson. Therefore, students need to have studied factors, which affect the context of films under study (OCR, 2018).

In all three instances the group discussion was followed by a wider class discussion.

Nonetheless, it was suggested that in some subjects there was still not enough time for this.

> I do not think I've got the beauty of that because I feel like I'm hammering through at a pace because there's still so much to get through (Paul, Business Studies Teacher, Focus Group 1).

Consequently, in some subjects, teachers do not feel that they have the time to integrate and allocate so much time to discussion and debate in their teaching, as they are concerned that they will not have time to teach all of the necessary content which could appear on the final exam. It is therefore possible to argue that this poses a significant threat to examination validity. Owing to the pressures of high stakes testing, it could be said that that teachers are obliged to prepare students for exams in a particular way, for example through the giving of practice tests and coaching students to be able to answer questions, "rather than in using and applying their understanding more widely" (Harlen, 2007: 21). This often results in "teaching to the test" (Au, 2007). However, according to

Sweiry *et al* (2002) for exams to be valid "we should be assessing conceptual knowledge and competence in a particular subject" (2). It can be contended that emphasis on practice tests and examination technique to ensure students are fully prepared to pass their exams means that conceptual knowledge is often lost along the way. Therefore, exam validity could be under threat here as there is a risk that not all the required content will have been taught before students sit the exam. This in turn has the potential to adversely affect student exam performance, therefore, it is possible to argue that teachers need to focus their teaching on what the exams set out to test. Assessment of rich, conceptual knowledge requires a more creative response from students, but for such assessment to be valid, reliability is compromised as a result (Davis, 2006).

While time to teach the required content remains problematic in some subjects, Paul (Business Studies Teacher) did note that this was not to the same extent as previously under the modular system. In addition, Paul also acknowledged that he did not feel like his teaching style differed greatly in the linear system, with the exception of "constantly going back to what I did at the start to keep that drip feeding through", thus emphasising the importance of revisiting knowledge. Examples of teachers revisiting knowledge with their students was evident in some lesson observations. In Sociology, as a starter activity, students were asked to note down 10 things relating to Social Action Theory, thus revisiting prior learning. This approach was advocated by David (PE Teacher, Focus Group 2) who stated that he regularly used whiteboards, for example through quick fire questions or quizzes, to promote recall of information at the start of a lesson.

In other subjects, for example Media Studies and PE staff discussed finishing teaching subject content approximately one month before the final exam to focus on revision. Despite this, several teachers noted that this seemed to have had an adverse effect upon attendance.

> Their mind-set is "right I've got my notes, I've got everything I need, yes, it is revision but if I miss that lesson then potentially it's not as bad as if I was to miss a lesson where he is delivering new content" (David, PE Teacher, Focus Group 2).

Simon (Criminology Teacher, Focus Group 2) discussed how students have turned up to lessons and asked if they can do their own revision; “they just think that because you’ve finished the teaching, that’s it”. Similarly, Natalie (Film Studies Teacher, Focus Group 2) outlined the difficulties relating to the attendance of Year 13 resit students at revision classes.

Changes in examination structure were also identified as a challenge for staff as many students they work with have sat modular GCSEs but are embarking upon linear A Level courses. Paul (Business Studies Teacher, Focus Group 1) stated that “we are operating in a very different world now compared to what we used to with students”. According to Ashley (English Teacher, Focus Group 2), for the most part the changes have been viewed positively by Year 12 students as they perceive that they have a lot more time to prepare for their exams. While current Year 13 students would have sat modular GCSEs and are therefore used to this system, there was much agreement that the next cohort of Year 12 students should be “more used to having to wait to be assessed, as it won’t be as much of a shock” (Natalie, Film Studies Teacher, Focus Group 2). However, David (PE Teacher, Focus Group 2) discussed how students no longer have the security they once did when going into their final exams. For example, under the modular system (whereby exams could be taken in January as well as June) PE students were aware prior to the final exam how many marks they had accrued up to this point. Therefore, they were not wholly reliant upon their performance in their final exam for the grades they were awarded.

Both Paul (Business Studies Teacher, Focus Group 1) and Lesley (Sociology Teacher, Focus Group 1) reported that this was compounded by difficulties they are experiencing in accessing exemplar materials from exam boards to support their teaching which in turn affects their ability to provide formative feedback. Thus, it can be argued that high stakes testing is distorting curriculum processes. Teachers are struggling to obtain sample materials and exam style questions from exam boards to give to their students to complete. In addition, a lack of sample mark schemes also makes it difficult for teachers to provide useful formative feedback to their students, as they do not fully know how questions will be marked and what examiners are looking for.

> The main problem I've got with the formative assessment is getting a clear line from the exam boards about what the style of the questions is (Lesley, Teacher of Sociology, Focus Group 1).

A lack of information from exam boards was also deemed to be problematic in BTEC Health and Social Care. While the textbooks suggested that the format of questions would reflect the allocation of 2,4,6 and 8 marks, in the final examination the question expected to be worth 6 marks was actually worth 4. In other words, the exam board approved textbook suggested that there would be a combination of 2,4,6 and 8-mark questions on the final exam but there was instead a combination of 2, 6- and 8-mark questions. This further compounds the difficulties teachers are facing in the light of changes to the examination structure which in turn affects their ability to prepare their students. This issue has arisen following the transition from modular to linear A Level exams, which occurred in September 2015, the first results from linear examinations were published in 2017. Prior to this, A Level exams were modular, with two exam windows each academic year, one in January and one in May/June. However, the January window has now been removed, meaning that students have fewer opportunities to retake exams as a result. This was accompanied by a review of A Level content, thus resulting in the revision and republication of syllabuses. Furthermore, both AS and A Levels are now formerly assessed at the end of their respective programmes of study, with AS Levels no longer contributing to an A Level qualification, unlike previously (Ofqual, no date). The process can be illustrated as follows:

Figure 5.12 - The Modular Exam Process (Gov.UK, 2018)

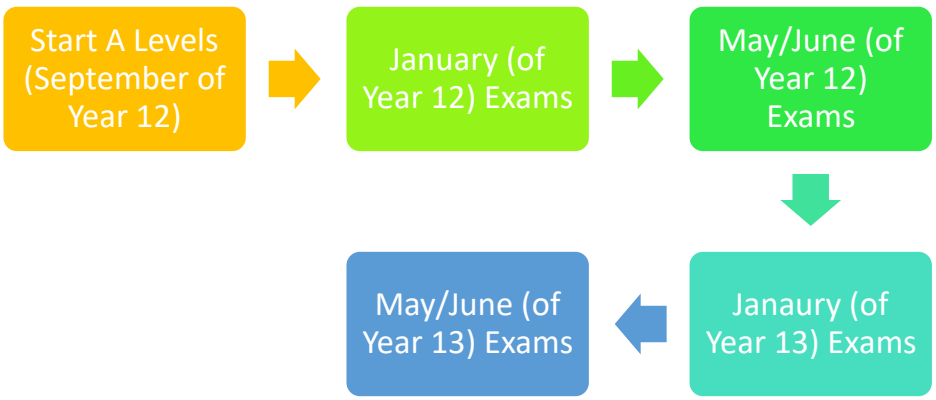
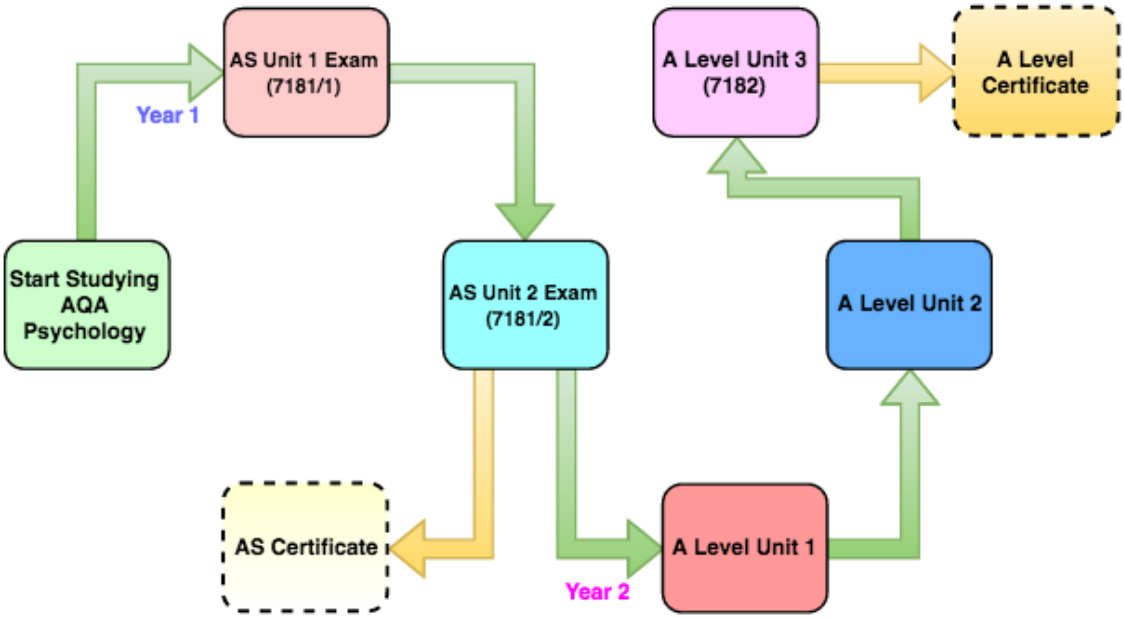


Figure 5.13 - The Linear Exam Process (What Uni? 2019)



Such instances could be deemed to be problematic as it has the potential to skew exam preparation and expectations of what the paper will be like. As such this affects the ability of teachers to generate sample exam questions which are representative of the final paper. Therefore, building students up to be resilient to anything that appears on the exam paper was deemed to be important. This was also cited as problematic by David (PE Teacher, Focus Group 2). He too reflected on the lack of available resources to support teaching and stated that staff were spending a considerable amount of time on creating resources, mock papers, and model answers. Despite this, there is a lack of confidence relating to their accuracy.

> We're not clear on how they are going to be assessed and how these long answer questions are going to be marked (David, PE Teacher, Focus Group 2).

Robert (Media Studies Teacher, Focus Group 1) reinforced the importance of formative feedback and encouraging students to take ownership of knowledge in light of changes to the examination structure. In terms of pedagogy, this involves moving students away from being passive recipients of knowledge to take an active role in their learning so that they can use formative feedback to improve their work and feed forward in the assessment cycle. Robert, (Media Studies Teacher, Focus Group 1) discussed the tendency to move away from plying students with the knowledge

they need to know to answer a particular question so that they can "approach questions creatively and demonstrate the scope of knowledge".

Despite this, I identified a contradiction when observing a lesson in Media Studies, taught by the same teacher. Students were informed when they need to write something down by stating "take this down" (Field Notes, Observation 4), thus the teacher was emphasising what they need to know for their examinations. Therefore, such actions could be regarded as coaching or perhaps even spoon feeding. A similar instance was observed in a Criminology lesson with the teacher informing students "you need to write this down" (Field Notes, Observation 27) when discussing specialist terminology. While it is possible to argue that these are examples of teaching to the test, an unintended consequence of high stakes testing (Au, 2007), asserting the importance of key terminology could be perceived as a fundamental element of the curriculum, regardless of specialism. The importance of knowledge was also emphasised by Alan (Physics Teacher, Focus Group 1), if students up-skill themselves "in my day-to-day teaching I focus on the knowledge bit". In a Biology lesson on DNA students were made aware of some information that is not explicitly listed on their exam specification, thus reinforcing the importance of a broad knowledge base. This contrasts with Au (2007) who argues that teaching to the test occurs at the expense of other subject knowledge.

5.8 Self and Peer Assessment

Both peer and self-assessment are examples of formative assessment (Black and Wiliam, 1998; Boud, 2000; Boud and Soler, 2016). In the research literature it is suggested that not all teachers welcome such approaches, as it requires a move away from the traditional view that pupils are recipients of knowledge. Some researchers argue that successful formative assessment should incorporate self and peer assessment as "student's learning can be enriched by marking their own or one another's work" (Black *et al,* 2003: 51). Discussion during the focus groups pointed towards a lack of popularity for these forms of assessment amongst some members of staff. This represents a further example of how the rhetoric found in the literature can differ in practice. Peer and self-assessment were described by Robert (Media Studies Teacher, Focus Group 1) as "a waste of time".

In contrast however, Paul (Business Studies Teacher, Focus Group 1) stated that such strategies were employed on occasion to enable students to see the work of others, although marks were not recorded by the teacher in his mark book. In addition, peer and self- assessment was said to be utilised in BTEC Law, due to the course being assessed entirely through coursework, as a means to "keep pace and keep them engaged" (Laura BTEC Law Teacher, Focus Group 1). In addition, Laura stated that she would not use these strategies for an A Level course. Conversely, some values of peer and self-assessment were discussed by Natalie (Film Studies Teacher, Focus Group 2).

> It forces them to use the mark scheme and be aware of all the different components of it.

In a recent mock exam in Business Studies Paul stated that students had completed a 2-hour paper comprising 2 ten-mark questions and 2 twelve-mark questions which were all quality marked (the provision of detailed written feedback) by subject teachers taking "an absolute eternity", thus reflecting the challenge of time. In Sociology, Lesley advocated the use of peer assessment to mitigate some of the demands placed on teachers vis-à-vis marking. It is possible to argue that this use of peer assessment reflects a means to reduce teacher workload, a somewhat pragmatic response. Alternatively, it is possible to question whether this form of assessment was employed to enable students to learn more about the process and ultimately become more independent learners. Furthermore, this approach could be deemed to be a means to prepare students for the summative. Perhaps such forms of assessment could serve multiple purposes.

> There's only so many questions they can ask so in order to get the sheer volume they literally do an essay every week and obviously I can't keep up with the marking (Lesley, Sociology Teacher, Focus Group 2).

However, it can be argued that students should be properly trained to use both peer and self-assessment effectively (McDonald and Boud, 2003). Despite this perspective, self and peer assessment can be seen as a means to save staff time. It can therefore be argued that training students to effectively utilise peer and self-assessment is a timely process, one which does not fit well into a system focused on high stakes summative assessments. Indeed, Falchikov (1986) in her

study reported that students found peer and self-evaluation schemes to be “time consuming, hard, challenging but also helpful and beneficial” (157). However, Lesley goes on to cite the benefits of this process for students. Nonetheless Lesley does emphasise the importance of training students to undertake this form of assessment. In this instance therefore students can access a wide variety of examination style questions in the run up to the exam.

> When they go into the exam, they’ve seen that question before […] if I train them well to be examiners, then train them well in Year 12, by the time Year 13 comes I can literally just give them, here’s another one, here’s another one. (Lesley, Sociology Teacher, Focus Group 1).

Having seen exam questions before could arguably be construed as coaching students to pass an exam or even rote learning. Ross (2006) argues that agreement between student and teacher assessments tends to be higher when students have been taught the means by which to assess their work. Furthermore, Black and Wiliam (1998) argue that training for students can lead to improvements in school performance. For Lesley it is simply about making this process of self and peer assessment reliable enough for her to use as part of her professional practice.

> I find it really, valuable if the students are on board and trained well. It’s just a good way of getting them to do all the questions you’ve got (Lesley, Sociology Teacher, Focus Group 1).

Despite this view, issues were still identified, namely relating to students’ tendency to be too generous when marking the work of their peers. As outlined in Section 5.2, following the completion of an exam style question in Classics it was set aside to be teacher marked. It is possible to infer that this could be due to students struggling to apply the mark scheme realistically when self or peer marking, particularly for a question with a considerable number of marks attached.

> I just think they do not want to mark in the way that teachers would mark so they are far too lenient, in my experience they do not really mark it properly (Ashley, English Teacher, Focus Group 2).

> I think the issues I've had in the past relate to where they are just a bit too generous! And they will just give their friends a good mark because they do not want to fall out or whatever (Ashley, English Teacher, Focus Group 2).

This is also discussed by McDonald and Boud (2003) who argue that students do not always know what a good piece of work looks like. Consequently, if students do not know what a good piece of work looks like, they are unable to construct meaning and make productive improvements (Chanock, 2000; Hyland, 2000). Black and Wiliam (1998) advocate peer and self-assessment as a key element of successful formative assessment. It is therefore possible to question the extent to which students have been properly trained to undertake such forms of assessment and the value placed on it by the teacher concerned. This is particularly important as several teachers, during the focus group interviews, discussed the difficulties they encounter in relation to such forms of assessment.

A further difficulty associated with peer assessment was discussed by David (PE Teacher, Focus Group 2). If students are marking the work of their peers and the teacher is not marking it, then the teacher is unaware of any gaps in students' learning. David suggested that this is the reason why they no longer utilise peer assessment in PE, instead advocating the use of self-assessment following teacher marking having taken place. Furthermore, Ashley stated that while students will do peer or self-assessment if asked by their teachers, she suggested that they do not value it in same way and will "invariably ask if it can be properly marked".

> Once I've marked it there's then scope for them to do a lot of self-assessment in terms of reflecting on what they have produced in comparison to the mark scheme (David. PE Teacher, Focus Group 2).

Some instances of both peer and self-assessment were observed in lessons, for example, peer assessment took place in Criminology following the completion of an exam question. Students were asked to swap answers with a partner and write the words "peer assessment" under the answer so that when people look at student files (for example members of the Senior Leadership Team), they

can see a variety of marking styles. Self-assessment in Geography was observed following the completion of an exam style question whereas in other subjects, such as Classics, as previously stated, emphasis was placed upon teacher feedback. In both French and Biology students were asked to review and check their answers after having completed an exam style question, although the term "self-assessment" was not explicitly used. Such instances serve to get students used to the assessment process and develop their "assessment literacy".

There was some agreement that when using peer assessment students struggle to apply the mark scheme realistically. This perspective is also reflected in the research literature as it can take time to embed the necessary skills which may not be realistic given the focus on high stakes testing and student performance (Wylie and Lyon, 2015). According to Nitko (1995) some students "lack the metacognitive skills needed to evaluate the quality of their own progress" (327). Despite this, some researchers, for example Boud (2000) argue that improving self-assessment skills plays an important role, particularly as teacher feedback has become so mainstream to the point that it is often overlooked and becomes under-conceptualised. Instead, the promotion of self-assessment encourages pupils to develop skills of independent learning. In addition, there are some who align the use of self-assessment with improved student performance (McDonald and Boud, 2003).

5.9 Concluding Thoughts

This chapter has sought to analyse and review formative and summative approaches to assessment and their associated challenges from the perspective of teachers who participated in this study. Questioning has been identified as a key formative assessment strategy, with other approaches, such as peer assessment, being employed to a much lesser extent. While a number of challenges associated with both formative and summative assessment were identified, namely time-related, teachers discussed strategies they have employed to mitigate such issues. For example, the provision of formative feedback at the whole class level, as opposed to the individual level following the completion of a mock exam. Assessment, both formative and summative, was described as cyclical and part of an ongoing process with regular assessment, incorporating formative feedback and opportunities for feed forward taking place every 2-3 weeks. Furthermore, formative

assessment should be embedded into day-to-day practice and complemented with regular summative assessment. There has also been a consideration of the actions and responses of students in relation to assessment practice. Much emphasis was placed upon students themselves taking an active role in the feedback process to identify areas they need to work on and ultimately close the gap in their learning. As such, teachers attached value to developing within their students skills of independent thinking and learning, particularly following the transition to linear examinations.

In this chapter, the impacts of a changing A Level examination structure have also been examined. Teachers described problems relating to accessing exemplar materials from exam boards which adversely impacts their ability to prepare for the summative. This in turn affects their provision of formative feedback to students as teachers have a lack of knowledge relating to what examiners are looking for. Therefore, it can be argued that the application of Assessment of Learning in the classroom is influenced by external factors. The following chapter will explore how high stakes testing affects teachers' approaches to assessment alongside a consideration of the extent to which teachers use formative assessment to prepare students for the summative. Finally, teachers' beliefs about assessment strategies and their professional attributes will be examined.

6 Analysis Chapter: Part 2- Preparing for the Summative

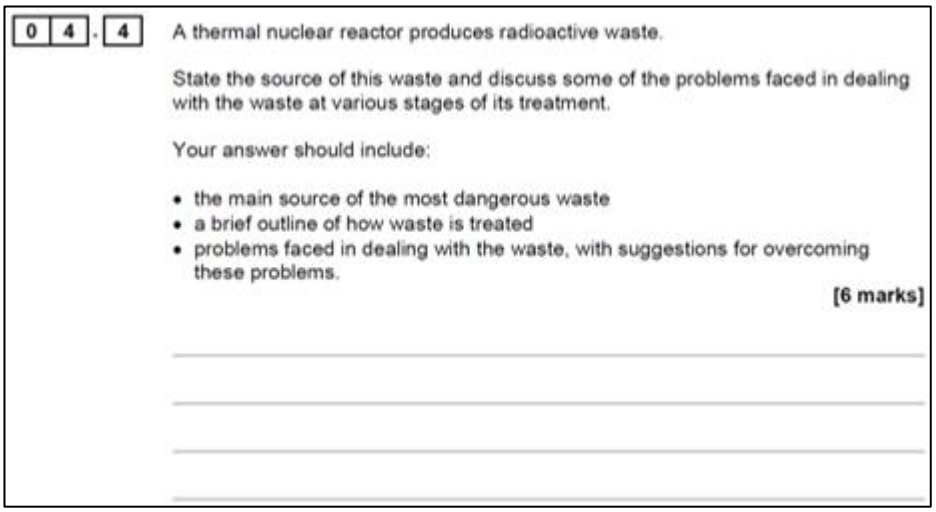
This chapter aims to build on the analysis of findings from Chapter 5 through a discussion of additional key and recurrent themes. To explore the extent to which teachers operationalise formative and summative assessment strategies with their students in Years 12 and 13, this chapter will examine strategies and approaches to assessment employed by teachers to prepare students for external examinations. This will also enable me to explore whether high stakes testing is deemed to affect teachers approaches to assessment and whether teachers use the formative to prepare students for the summative. The analysis of such findings will contribute to my third research question (Section 7.3): How do teachers use formative assessment to prepare their students for the summative? Finally, teachers' beliefs about assessment strategies will be considered alongside the extent to which there is a relationship between teacher values and day-to-day practice. This will form the basis of my discussion of my fourth research question (Section 7.4): What are teachers' beliefs about assessment strategies? Is there a relationship between teacher values and day-to-day practice?

6.1 Preparing for External Examinations

Regarding question style and marks attributed to them there is a degree of subject individuality. For example, an A Level Physics paper has one levelled question worth a total of 6 marks, with the rest of the marks accrued from a combination of 1, 2, 3- and 4-mark questions. Whereas Sociology and Business Studies have questions worth 21 and 25 marks, respectively. As such it is possible that this factor may influence teachers' approaches to assessment and exam preparation. The lack of longer essay-style questions in Physics, unlike other subjects, has influenced Alan's approach to preparing students for such examinations. Alan (Teacher of Physics, Focus Group 1) stated that in his day-to-day teaching the main focus was on the transmission of knowledge with him having taken a step back from

> the skill of question analysis, looking for clues in the question, structure, and calculations, all that kind of stuff.

Figure 6.1 – 6-Mark A Level Physics Examination Question



0 4 . 4 A thermal nuclear reactor produces radioactive waste.

State the source of this waste and discuss some of the problems faced in dealing with the waste at various stages of its treatment.

Your answer should include:

- the main source of the most dangerous waste
- a brief outline of how waste is treated
- problems faced in dealing with the waste, with suggestions for overcoming these problems.

[6 marks]

This levelled question represents an example of the highest scoring question on a Physics exam paper. All other questions on the paper are worth 1, 2, 3, or 4 marks (AQA, 2017).

Conversely, in Business Studies students are required to answer questions worth substantially more marks. As such, Paul (Business Studies Teacher, Focus Group 1) discussed his approach to preparing students for these summative assessments. Paul discussed providing students with a clear structure of what an answer to a 12 or 10-mark question should look like. Moreover, students command higher marks in Business examinations through the development of a clear and logical argument as opposed to the presentation of a vast amount of knowledge. Therefore, the skills of substantiating and justifying an argument are deemed more important. This contrasts with the approach discussed by Alan (Physics Teacher, Focus Group 1) who emphasised the importance of focusing on knowledge acquisition.

Figure 6.2 - A Level Business Examination Questions

1 (a) Explain how Sony could be disadvantaged by operating in a dynamic market.

(4)

(c) Assess the possible benefits to manufacturers of video games consoles of the liberalisation of the Chinese market.

(10)

(d) Assess the likely impact of the proposed expansion of ASEAN on a business such as Sony.

(12)

(e) Evaluate the importance of branding in helping businesses such as Sony to enter the Chinese games console market.

(20)

These represent examples of the range of questions students are required to answer on a Business exam paper and the maximum marks available for each question (Pearson, 2014).

Figure 6.3 - An Example of a Levelled (Banded) Exam Question from Geography

1 a (ii) With reference to *Figure 1*, suggest why there is likely to be an increasing need for shoreline management.	AO1	AO2.1a	AO2.1b	AO2.1c	AO31	AO32	**Total**
			6				**6**

Indicative content

- Increasing need may be because of (causality) increasing storms, coastal flooding, sea level rise, together with more human activities along coasts e.g. building at Wells-next-the-sea, tourism. Increasing costs arise if coastal flooding were to occur in the areas shown in Figure 1 (rebuilding, insurance, emotional, impact on infrastructure such as the A149 and the B1105).

Marking guidance

Near the upper end, answers may, through introducing the idea of Shoreline Management Plans (SMP), show applied understanding by suggesting that piecemeal coastal protection schemes may not always be compatible with coastline needs shown in Figure 1 and elsewhere within the same sediment cell i.e. the concept of 'integrated strategies' (systems concept).

Near the lower end there will be limited application of the increased need for SMPs.

Credit other valid approaches.

Levelled (or banded) questions attach a specific number of marks to an expected outcome. For example, for an answer to the above question to be awarded 1-2 marks (band 1) limited or fragmented knowledge is demonstrated compared to band 3 (5-6 marks), where a much wider range of knowledge is applied. A student must demonstrate that their answer successfully meets these outcomes to be awarded the marks. This reinforces the need, expressed by participants of the focus group interviews, for students to become acquainted with assessment criteria prior to sitting their external summative assessments (Eduqas, 2017).

Whereas in other subjects, for example Sociology, value was placed upon ensuring students had as many opportunities to encounter exam questions as possible, facilitated using model answers and peer assessment. Model answers were deemed to be particularly valuable, by some teachers, as exemplar material from exam boards is lacking.

> When I get these model answers I make like a jigsaw of them and sometimes, there's 12 questions that would come up in an exam, here's all the model answers cut up that we have done throughout the year and they have to spend the full hour putting the answers together (Lesley, Sociology Teacher, Focus Group 1).

Inferences were also made to different examples of teaching to the test (Black, 1998), for instance Alan (Physics Teacher, Focus Group 1) stated that his wife is a History teacher as well as an examiner, as such "she could go into a school and get the grades up by a grade in half an hour". It could be argued that as an examiner she would know exactly what was required of students to succeed in the exam and could utilise this knowledge to help prepare them for the exam, although this could be seen as coaching or teaching to the test.

Robert (Media Studies Teacher, Focus Group 1) stressed the importance of getting students to understand assessment objective criteria, which could be deemed to be an example of up-skilling students. It could be contended that this represents not only an example of Assessment for Learning, but also up-skilling as students were being encouraged to take responsibility for their own learning, which in turn serves to build their resilience in the face of examinations. This was exemplified through students using teacher feedback alongside assessment criteria to improve their work and close the gap in their learning. References were made to approaches students must adopt to move through the levelled assessment criteria. For example

> you can't end up in level 2 or 3 if you have only addressed half of the content. And if you've not done the advantages and disadvantages, you're stuck in level 2 and then a really good answer you move up to level 3 (Robert, Media Studies Teacher, Focus Group 1).

In both Geography and Biology, teachers were observed talking their students through the assessment criteria, following the completion of an exam question, to show them how to progress to the next level, which is an important aspect of assessment for learning. According to Robert (Media Studies Teacher, Focus Group 1) if students do not know how to access level 4 of AO1 (Assessment Objective 1) "then we haven't done our job really" (in other words if students do not know how to obtain top marks for each objective they are being assessed against) and reflected on the importance of reassuring students if initially they do badly that that it is okay. In three observed lessons teachers made specific reference to the exam board specification when setting tasks for students to complete. In Criminology, Assessment Criteria from the exam board specification were used to frame a revision task which required students to generate spider diagrams for each agency

of social control (Prisons, CPS, Probation Service) they had studied to date. The Assessment Criteria shown in Figure 6.4 were displayed on the whiteboard as a guide to students as to what they needed to include in their spider diagrams.

Figure 6.4 - Assessment Criteria for Criminology Unit 4 Crime and Punishment

Unit 4 learning outcomes	**Assessment criteria**	**Content**	**Amplification**
The learner will:	*The learner can:*		

LO3 Understand measures used in social control	**AC3.1** Explain the **role** of **agencies** in social control	**Role**
		• aims and objectives
		• funding
		• philosophy
		• working practices
		  ○ types of criminality
		  ○ types of offenders
		  ○ reach (local, national)
		
		Agencies
		• government-sponsored agencies
		  ○ police
		  ○ CPS
		  ○ judiciary
		  ○ prisons
		  ○ probation
		• charities
		• pressure groups
			Candidates should be able to identify agencies involved with social control and explain their role in achieving social control.
			
			Synoptic links:
			Learners can apply their understanding from Unit 3 to this criterion.

	AC3.2 Describe the **contribution** of agencies to achieving social control	**Contribution**
		• tactics and measures used by agencies
		  ○ environmental
		  ▪ design
		  ▪ gated lanes
		  ○ behavioural
		  ▪ ASBO
		  ▪ token economy
		  ○ institutional
		  ○ disciplinary procedures
		  ▪ rule making
		  ▪ staged/phased
		• gaps in state provision
			Candidates should understand the range of techniques used by the agencies and be able to examine their contribution.
			
			Synoptic links:
			
			Learners should apply their understanding of:
			• *policy and campaigns from Unit 1*
			• *criminological theories from Unit 2*
			• *the processes used to bring an accused to justice in Unit 3*
			to the role of the different agencies.

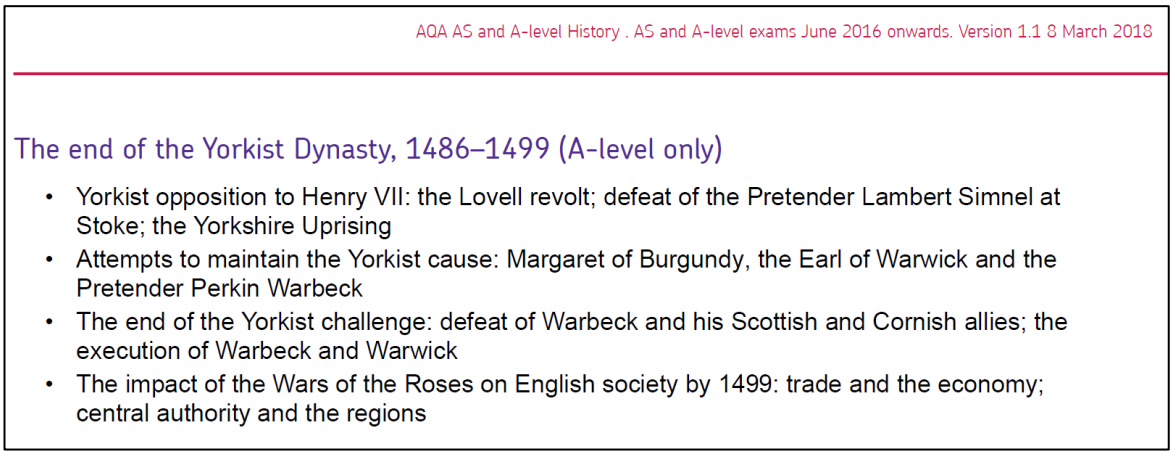
Unit 4 learning outcomes	**Assessment criteria**	**Content**	**Amplification**
The learner will:	*The learner can:*		
	AC3.3 Examine the **limitations** of agencies in achieving social control	**Limitations** • repeat offenders/recidivism • civil liberties and legal barriers • access to resources and support • finance • local and national policies • environment • crime committed by those with moral imperatives	Learners should understand the limitations of social control agencies and able to examine the implications of these limitations. *Synoptic links:* *Learners should apply their understanding of criminological theories from Unit 2 in their examination of the limitations. Learners will also draw on their understanding of policy and campaigns for change in examination of the limitations of agencies.*

	AC3.4 Evaluate the effectiveness of **agencies** in achieving social control	**Agencies**
		• government sponsored agencies
		  ○ police
		  ○ CPS
		  ○ judiciary
		  ○ prisons
		  ○ probation
		• charities
		• pressure groups
			Learners should be able to draw together their learning to evaluate the success or failure of agencies in achieving social control.
			
			Synoptic links:
			
			Learners should apply the skill they developed in Unit 3 to evaluate information in terms of:
			• *bias*
			• *opinion*
			• *circumstances*
			• *currency*
			• *accuracy*
			
			The types of evidence, as set out in Unit 3, include:
			• *evidence*
			• *trial transcripts*
			• *media reports*
			• *judgements*
			• *Law Reports*

The highlighted sections on the above Figure denote the Agencies of Social Control students were required to include in their spider diagrams. Students needed to include the role of each agency, any associated limitations, and their effectiveness. Thus, this represents an example of exam board materials being shared directly with students (WJEC, 2017).

Similarly, in both Geography and History the lesson titles reflected the exam board specification with subsequent tasks directly related to the described content, shown in Figures 6.5 and 6.6, respectively. For example, in Geography students were observed researching the management of the impacts which resulted from Storm Desmond whereas in History students reviewed source material pertaining to Historians views of the impact of the War of the Roses on society.

Figure 6.5 - A Level History Specification of Content



AQA AS and A-level History . AS and A-level exams June 2016 onwards. Version 1.1 8 March 2018

The end of the Yorkist Dynasty, 1486–1499 (A-level only)

- Yorkist opposition to Henry VII: the Lovell revolt; defeat of the Pretender Lambert Simnel at Stoke; the Yorkshire Uprising
- Attempts to maintain the Yorkist cause: Margaret of Burgundy, the Earl of Warwick and the Pretender Perkin Warbeck
- The end of the Yorkist challenge: defeat of Warbeck and his Scottish and Cornish allies; the execution of Warbeck and Warwick
- The impact of the Wars of the Roses on English society by 1499: trade and the economy; central authority and the regions

The above Figure shows the content students are required to know for their examination in this subject. At the start of the lesson this was displayed on the whiteboard to show students how their lesson links to their final examination and where it sits within a wider context. As such this is a further example of sharing exam board materials with students (AQA, 2019).

Figure 6.6 - A Level Geography Specification of Content

3.5.5 Impacts and management of climatic hazards	• Impacts of hazards associated with low-pressure systems on the environment and human activity • Impacts of hazards associated with high-pressure systems on the environment and human activity • Strategies to manage climatic hazards

This was displayed on the whiteboard during the lesson (Observation 5) to show students how the Case Study they were working on linked to the exam board specification. Students were required to examine perspectives on Storm Desmond, relating to the impacts and management of this climatic hazard. This represents a further example of exam board materials being made available to students (Eduqas, 2019).

Upon the completion of a 10-mark exam question in Classics, students reviewed the mark scheme with their teacher and discussed what was required to score highly in such a question. Although this question was teacher marked, this enabled students to see what their teacher would be looking for in their work and where the marks could be accrued. In Religious Studies students were observed responding to teacher comments and using the mark scheme (shown below in Figure 6.7) to help them to make improvements to their work. This enabled students to see what they needed to do to achieve a higher mark.

Figure 6.7 - Religious Studies Mark Scheme for Assessment Objective 1

Band	**Assessment Objective AO1 – Part (a) questions** 25 marks *Demonstrate knowledge and understanding of religion and belief, including:* - *religious, philosophical and/or ethical thought and teaching* - *influence of beliefs, teachings and practices on individuals, communities and societies* - *cause and significance of similarities and differences in belief, teaching and practice* - *approaches to the study of religion and belief.*
5	**21-25 marks** • Thorough, accurate and relevant knowledge and understanding of religion and belief. • An extensive and relevant response which answers the specific demands of the question set. • The response demonstrates extensive depth and/or breadth. Excellent use of evidence and examples. • Thorough and accurate reference made to sacred texts and sources of wisdom, where appropriate. • Thorough and accurate use of specialist language and vocabulary in context.
4	**16-20 marks** • Accurate and relevant knowledge and understanding of religion and belief. • A detailed, relevant response which answers the specific demands of the question set. • The response demonstrates depth and/or breadth. Good use of evidence and examples. • Accurate reference made to sacred texts and sources of wisdom, where appropriate. • Accurate use of specialist language and vocabulary in context.
3	**11-15 marks** • Mainly accurate and relevant knowledge and understanding of religion and belief. • A satisfactory response, which generally answers the main demands of the question set. • The response demonstrates depth and/or breadth in some areas. Satisfactory use of evidence and examples. • Mainly accurate reference made to sacred texts and sources of wisdom, where appropriate. • Mainly accurate use of specialist language and vocabulary in context.
2	**6-10 marks** • Limited knowledge and understanding of religion and belief. Basic level of accuracy and relevance. • A basic response, addressing some of the demands of the question set. • The response demonstrates limited depth and/or breadth, including limited use of evidence and examples. • Some accurate reference made to sacred texts and sources of wisdom, where appropriate. • Some accurate use of specialist language and vocabulary in context.
1	**1-5 marks** • Very limited knowledge and understanding of religion and belief. Low level of accuracy and relevance. • A very limited response, with little attempt to address the question. • The response demonstrates very limited depth and/or breadth. Very limited use of evidence and examples. • Little or no reference made to sacred texts and sources of wisdom, where appropriate. • Some grasp of basic specialist language and vocabulary. **N.B. A maximum of 2 marks should be awarded for a response that only demonstrates 'knowledge in isolation'**
0	• No relevant information.

As previously stated in Section 5.3, students had completed a 20-mark exam style question, which had been teacher marked. They were observed using this feedback along with the mark scheme shown above to improve their answers.

The question students had completed was an example of a 20-mark Assessment Objective One question. In other words, students were required to demonstrate their knowledge and understanding of religion and belief (Eduqas, 2019).

Examples were also noted of teachers getting students used to the terminology employed by exam boards. In the same Religious Studies lesson, when introducing the next topic to be studied, "Two Views of Jesus", the teacher made explicit not only where the new topic fit in but also related the topic to sample questions they could be asked in the final exam. Students were informed that the two scholars cited in the specification, as shown in Figure 6.8 below, (Crosson and Wright) could feature in a question asking them to compare and contrast their standpoints and beliefs (Observation 8).

Figure 6.8 - A Level Religious Studies Specification of Content

Option A: A Study of Christianity

Theme 1: Religious figures and sacred texts	
Knowledge and understanding of religion and belief	
D.	**The Bible as a source of wisdom and authority:** How the Christian biblical canon was established. Diverse views on the Bible as the word of God: different understandings of inspiration (the objective view of inspiration; the subjective view of inspiration; John Calvin's doctrine of accommodation).
E.	**The early church (in Acts of the Apostles):** Its message and format: the kerygmata as presented by C. H. Dodd, with reference to Acts 2:14-39; 3:12-26. The challenges to the kerygmata (with reference to the historical value of the speeches in Acts and the work of Rudolf Bultmann). The adapting of the Christian message to suit the audience.
F.	**Two views of Jesus:** A comparison of the work of two key scholars, including their views of Jesus with reference to their different methods of studying Jesus: John Dominic Crossan and N. T. Wright. Crossan: Jesus the social revolutionary; using apocryphal gospels; seeing Jesus as a product of his time; what the words of Jesus would have meant in Jesus' time. Wright: Jesus the true Messiah; critical realism; texts as 'the articulation of worldviews'; seeks to find the best explanation for the traditions found in the Gospels.
Issues for analysis and evaluation will be drawn from any aspect of the content above, such as: • The extent to which the Bible can be regarded as the inspired word of God. • Whether the Christian biblical canonical orders are inspired, as opposed to just the texts they contain. • The extent to which the kerygmata (within the areas of Acts studied) are of any value for Christians today. • Whether the speeches in Acts have any historical value. • The validity of using critical realism to understand Jesus. • The validity of using apocryphal gospels to understand Jesus.	

6.2 Grades versus Comments

There is much reference and debate in the literature regarding the provision of grades and/or comments on student work (Black and William, 1998; Black *et al,* 2003; Butler, 1988). Alan (Physics Teacher, Focus Group 1) stated that giving students grades on assessments was a practice which stopped around 2 years ago, this is supportive of the work of a number of researchers (Black and William, 1998; Stobart, 2006). In addition, Alan explained, "I've kind of almost moved away from the headline percentage as well". Instead, choosing to break results down by topic so that students can easily identify areas they need to work on. By contrast, in Criminology, students are given a numerical mark, not a grade.

> We break it down into the topics and say you've got 70% of those topic questions right so it's less about "I got 30% overall". It's more a case of that did me in (Alan, Physics Teacher, Focus Group 1).

> We do not feel it is appropriate to grade it A, B, C, D because it's not a big enough sample to say that's an A Level piece of work (Simon, Criminology Teacher, Focus Group 2).

For example, students may have only answered one exam style question. Simply answering one exam question does not justify the giving of a grade as this is not representative of completing a full paper. This was echoed by Ashley (English Teacher, Focus Group 2) who stated that the first 6 months of Year 12 is not A Level standard, notably as the first Year 12 assessment is open book whereas ultimately the final summative assessment will be closed book. It could be argued that the start of Year 12 is a transitional phase, particularly as students who attend Mitford College arrive from a whole host of secondary schools. As Ashley (English Teacher, Focus Group 2) states, students who have just completed their GCSE examinations are not prepared nor equipped to produce work of A Level standard in the initial stages. Regardless of this, students still ask for a grade. She explained that this is exacerbated by the provision of target grades, which is the reason

> behind them having that expectation of wanting to know how they are doing in relation to their target grade (Ashley, English Teacher, Focus Group 2).

This was attributed to students receiving their target grades at the start of Year 12. At GCSE level students often become accustomed to having their work graded. Although this is not the case at A Level. Thus, a change in attitude towards assessment is required and this can become a steep learning curve for students in Year 12. In some cases, therefore the way in which the discourse plays out in practice differs somewhat from the rhetoric found in the literature. However other teachers' practice did not support marks versus grades. For example, in Sociology

> if I'm training them, this is how you do a 12-mark question, with Sixth Formers and I do not give them a mark out of 12 there will be hell on! (Lesley, Sociology Teacher, Focus Group 1).

Emphasis was placed upon the importance of student action as they were informed that it is not the number on the front of the paper, which is important, instead they need to learn from the experience. Paul (Business Studies Teacher, Focus Group 1) said that he planned to provide students with the opportunity to improve each question as they go through the paper as a class.

> It is a lot about getting them to do something with it rather than just being passive in the process.

Evidence from the literature suggests that high stakes testing can produce unintended consequences, one such example being students becoming passive recipients of knowledge in the learning process (Au, 2007). However, the approach championed by Paul, goes against this argument in the literature as Paul's approach is concerned with students actively participating in the learning process to close the gap in their learning. Indeed, it could be argued that while students are not passive recipients of knowledge, Paul is using this formative approach to prepare students for the summative.

As indicated earlier in Section 5.7, difficulties were discussed relating to accessing support from exam boards in the wake of the new examination structure. This lack of input also affects teachers' abilities to provide students with overall grades as there is still much confusion over how marks convert to grades. Notwithstanding this difficulty, teachers spoke of using exam board assessment

criteria to mark work. As such it is easy for students to see their level of working even if teachers are unable to equate the marks to a specific grade.

Several instances were observed of students asking their teachers for verbal feedback, for example, in Core Maths students were keen to check that they were completing tasks correctly. Equally in Chemistry students asked the teacher questions to clarify their understanding. Similarly, teachers provided their students with verbal feedback on their work in five of the observed lessons, Accounting, PE, Core Maths, BTEC Law and Chemistry.

Despite the issues teachers discussed relating to examination changes, strategies to reduce the demands on teacher time were referred to, such as the collation of the best student answers to an essay question to create a bank of model answers. While this was deemed to be a successful strategy in Sociology, albeit by two teachers, not all teachers agreed. Robert (Media Studies Teacher, Focus Group 1) described the model answer as the final thing you go to and stated that "for me it's a case of not a model answer as in you should trot it out". Instead, he discussed the importance of getting students to look at their own answers critically and upgrade them using teacher feedback. Students were observed using teacher feedback to upgrade their work in both Business Studies and Religious Studies lessons, in both cases after having received a mark and grade alongside comments. Consequently, he suggested that "by the time you get to model answers, they almost stop thinking of it as a take way" (Robert, Media Studies Teacher, Focus Group 1).

This links to ideas discussed previously relating to changing student attitudes towards assessments. One possible reason for the differences in approach could be the style of exam question with Media questions asking for students own creative input and therefore providing less opportunity to "trot out" a model answer response. This can be exemplified through the following sample questions. An example Media Studies question: "With reference to your own detailed examples, explore how texts are constructed to target different audiences" (WJEC, 2017), compared to a Sociology essay; "Outline and explain two problems of using the functionalist perspective to study today's society" (AQA, 2018). As previously stated in Section 6.1, subject individuality, relating to question style and the number of marks attributed to a question can influence teacher approaches to assessment.

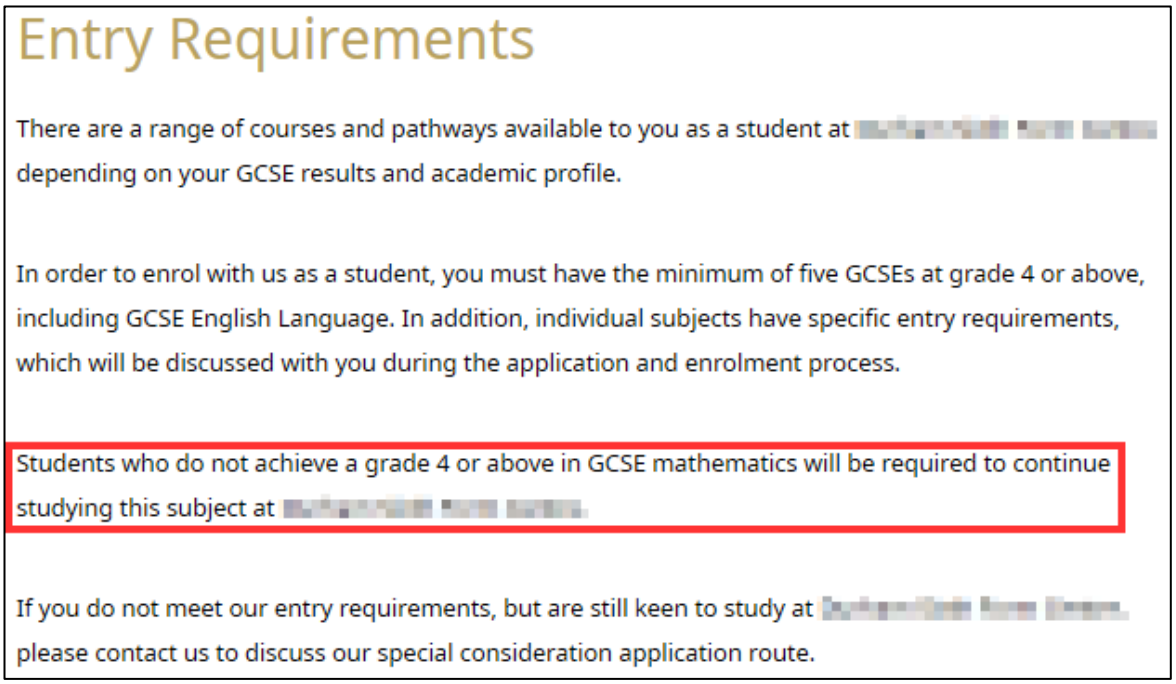
The demands and assessment criteria of other subjects led to some teachers describing a completely different approach. For example, in Business Studies more marks are accrued for application and context as opposed to knowledge. In other words, simply stating what they know about a topic commands less marks than applying knowledge to a particular context or example. As such Paul (Business Studies Teacher, Focus Group 1) stated that his approach consists of providing students with a structure, "so a 12 mark needs to look like this, a 10 mark needs to look like that". Furthermore, Paul (Business Studies Teacher, Focus Group 1) explained that students who have attended different schools to Mitford College have not necessarily been trained in this way and "can only get to a certain point because they're only kind of knocking out knowledge".

However, this difficulty was not apparent in all subjects. In Science, teachers felt that while approaches to marking exam papers have changed the questions themselves remain the same, thus enabling them to draw on resources from years previously.

> I've got the benefit of the questions haven't really changed; the Physics is still the same from what it always has been. So, I can still use papers from 10-15 years ago. I have to look at the mark scheme to check that the level of what they are expecting is still consistent but the question itself, there's nothing fundamentally wrong with the question (Alan, Physics Teacher, Focus Group 1).

Louise (Biology Teacher, Focus Group 1) discussed some research she had conducted into factors affecting student progress in A Level Science subjects, including background, subject choices, A Level Curriculum and GCSE attainment. A comprehensive understanding of Maths at GCSE level was deemed to positively impact, and support students' progress in A Level Science subjects, whereas the study of Double or Triple Award Science at GCSE level was found to have a much lesser impact. As such GCSE level Maths is a fundamental precursor to the study of Science at A Level. This is reinforced through the reference made to GCSE Maths as part of the entry requirements to study at this institution. Figure 6:9 displays the generic entry requirements for students to study at this institution. Equally students who do not possess the desired grade in GCSE Maths are required to continue their study to achieve this qualification.

Figure 6.9 - Institution Entry Requirements



Entry Requirements

There are a range of courses and pathways available to you as a student at [redacted] depending on your GCSE results and academic profile.

In order to enrol with us as a student, you must have the minimum of five GCSEs at grade 4 or above, including GCSE English Language. In addition, individual subjects have specific entry requirements, which will be discussed with you during the application and enrolment process.

Students who do not achieve a grade 4 or above in GCSE mathematics will be required to continue studying this subject at [redacted]

If you do not meet our entry requirements, but are still keen to study at [redacted] please contact us to discuss our special consideration application route.

While individual subjects may have specific entry requirements the above highlights the importance of achieving a desired grade in both GCSE Maths and English (Mitford College, no date).

A subsequent student questionnaire, conducted by Louise (Biology Teacher, Focus Group 1) found that students attributed their progress in Science to assessment and feedback they received from their teachers and the quality of the resources they were given to support their learning. She remarked that

> the two things that we've been doing much more of and kind of in a much more deliberate way and kind of thought-out way, it seemed that they found them as useful as we thought they were.

6.3 Formative Assessment, Summative Assessment or Both?

With regard to forms of assessment, there was unanimous agreement that teachers placed value on a combination of both formative and summative assessment. They were considered by David (PE Teacher, Focus Group 2) to "work hand in hand". While data has a significant role to play, teachers acknowledge the importance of feedback and improvement.

> The numbers and things are meaningful, and we like them, but you also need to know how to make it better and that feedback is absolutely crucial (Paul, Business Studies Teacher, Focus Group 1).

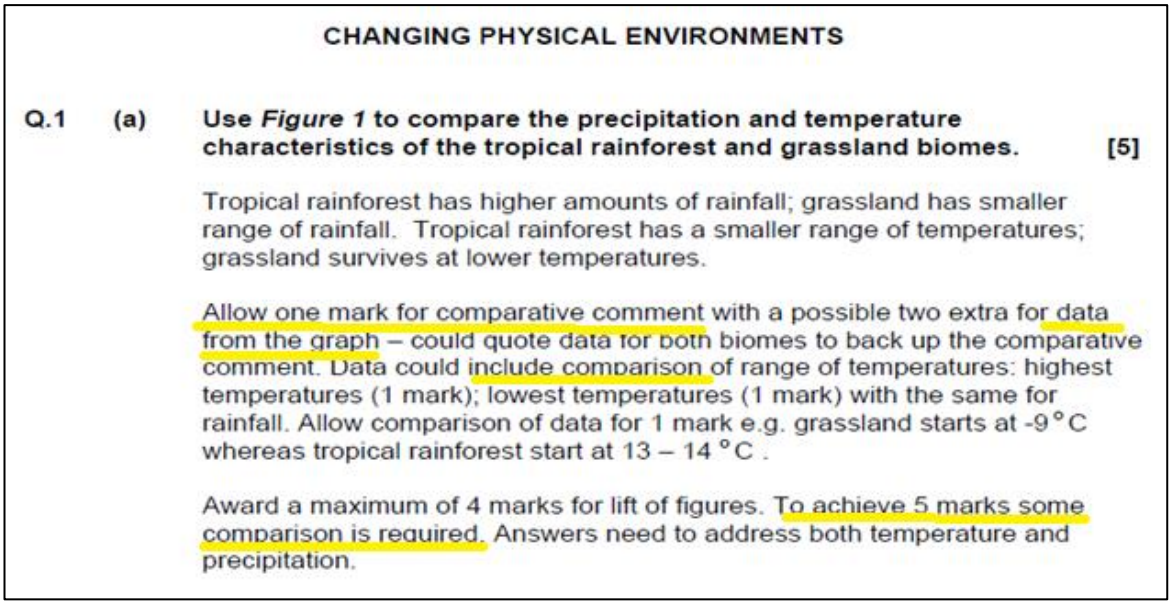
Conversely some researchers, for example Simpson (1990) and Harlen (2005) favour a distinction between formative and summative forms of assessment and as such do not view them as connected entities. David (PE Teacher, Focus Group 2) described formative assessment as a form of best practice "embedded into your good lesson delivery". However, Simon (Criminology Teacher, Focus Group 2) emphasised the importance of supporting formative assessment with regular summative assessment. Without summative assessment

> you wouldn't have as much information as to how they were doing and lose, kind of, the impact of following lessons where you look at correcting any misconceptions from the end of unit test (Simon, Criminology Teacher, Focus Group 2).

While more examples of formative assessment were observed in lessons, there were a range of instances whereby formative and summative assessment strategies were used in combination, thus reflecting the views upheld in the focus groups. In Geography a summative assessment was employed to gauge student knowledge, students were given 5 minutes to complete an exam style question. Prior to the completion of the question, the teacher used questioning to check student understanding of command and subject specific vocabulary. The same task was then used for formative purposes insofar as students self-assessed their work and upgraded their answers in light of this. The mark scheme (shown in Figure 6.10) was displayed on the whiteboard and the teacher

talked through it, pointing out what students needed to do in order to pick up marks, for example "1 mark for a comparative comment, 1 mark for data" (Observation 5,).

Figure 6.10 - Mark scheme for 5-Mark Exam Question (used in Observation 5)



CHANGING PHYSICAL ENVIRONMENTS

Q.1 (a) Use *Figure 1* to compare the precipitation and temperature characteristics of the tropical rainforest and grassland biomes. [5]

Tropical rainforest has higher amounts of rainfall; grassland has smaller range of rainfall. Tropical rainforest has a smaller range of temperatures; grassland survives at lower temperatures.

Allow one mark for comparative comment with a possible two extra for data from the graph – could quote data for both biomes to back up the comparative comment. Data could include comparison of range of temperatures: highest temperatures (1 mark); lowest temperatures (1 mark) with the same for rainfall. Allow comparison of data for 1 mark e.g. grassland starts at -9 °C whereas tropical rainforest start at 13 – 14 °C .

Award a maximum of 4 marks for lift of figures. To achieve 5 marks some comparison is required. Answers need to address both temperature and precipitation.

The underlines in the Figure above represent the key points the teacher pulled out of the mark scheme when talking through it with students.

Similarly, in Criminology students completed a summative assessment in the form of an exam style question to establish their level of knowledge. This task was then used to elicit feedback from students through teacher questioning to generate a model answer on the whiteboard. Students were also asked to peer assess. In Accounting students worked through a series of exam style questions and received verbal feedback from their teacher. While in both Business Studies and Religious Studies students received teacher marked work, which incorporated both marks and grades alongside comments which they used to make improvements and upgrades to their work.

A further example of successfully combining formative and summative assessment was discussed by Alan (Physics Teacher, Focus Group 1) whereby students are encouraged to analyse their own performance in a mock exam. This strategy, of drawing a graph to represent individual performance, is aimed at encouraging a change in attitudes towards assessment.

> It got them away from saying "I got 30 out of 70" because my point was "well great, but you lost three quarters of your marks on these two questions" (Alan, Physics Teacher, Focus Group 1).

It can be argued that this approach encourages student to focus on the areas they need to work on and serves to move them away from looking purely at the grade, mark, or percentage on the front of the paper.

> It's a lot about getting them to do something with it rather than just being passive in the process (Paul, Business Studies Teacher, Focus Group 1).

> If it's a 4-mark calculation, but you've done this and this, but you've actually only ended up with 1, for me they are less bothered about the fact that "oh well what does that mean? Does that mean I'll get a C in the end after 2 years?" It's more about "oh well I knew how to do that, where did I lose the 3 marks on it?" (Alan, Physics Teacher, Focus Group 1).

Conversely, this contrasts with the unintended consequences cited in the research literature (Harlen, 2007), which deems students to be passive recipients of knowledge in the learning process. Summative assessment was also thought to play an important role in developing knowledge and understanding and identifying gaps in learning. This was exhibited by Paul (Business Studies Teacher, Focus Group 1) who discussed a recent calculations test as students who scored less than 60% were attending after school intervention.

> It was a low-stakes, little test in the classroom but that's given me some meaningful data that I can then use to put some intervention in place and hopefully improve their understanding of those calculations (Paul, Business Studies Teacher, Focus Group 1).

However, in discussion with teachers of BTEC subjects, big differences were identified compared to A Level subjects. One key challenge for the staff of such subjects is their inability to provide detailed feedback to students which is the result of parameters put in place by the exam board. This is shown in Figure 6.11. Instead, students are required to refer back to a checklist of what should be included

in their work. While this checklist was teacher generated, it was based upon the guidance provided by the examination board.

Figure 6.11 - BTEC Applied Law Authentication Form

Authentication form

This form must be signed by both the learner and the teacher.

Qualification (delete as appropriate):
BTEC Certificate/ BTEC Extended Certificate

Unit number		**Unit name**	
Centre number		**Centre name**	
Registration number		**Learner name**	

Notice to Learners

The work you submit for assessment must be your own. You may be disqualified from at least the subject concerned if you copy from someone else, allow another learner to copy from you, or if you cheat in any other way.

Declaration of Authentication

I have produced the attached work without assistance, other than that which my teacher has explained is acceptable within the specification.

Signed

(Learner).....

Date.....

Declaration by the Teacher

I confirm that the **learner's** work was conducted under the conditions laid out by the specification. I have authenticated the **learner's** work and am satisfied that to the best of my knowledge the work produced is solely that of the **learner**.

Signed (Teacher) .....

Name (Block capitals)

.....

Prior to submission of their work students are required to sign a declaration stating that they have not received help with their work. (Pearson, 2019).

This was deemed to be frustrating as teachers wanted to give their students more help. In addition, students' response to this can be somewhat negative as they ask

"why haven't you told us this?" "What do I need to do?" It's trying to get them to be more independent (Laura, BTEC Law Teacher, Focus Group 1).

Figure 6.12 - BTEC Applied Law Specification- Making Valid Assessment Decisions

Making valid assessment decisions

Authenticity of learner work

Once an assessment has begun, learners must not be given feedback on progress towards fulfilling the targeted criteria.

An assessor must assess only learner work that is authentic, i.e. learners' own independent work. Learners must authenticate the evidence that they provide for assessment through signing a declaration stating that it is their own work.

Assessors must ensure that evidence is authentic to a learner through setting valid assignments and supervising them during the assessment period. Assessors must take care not to provide direct input, instructions or specific feedback that may compromise authenticity.

As shown above, the exam specification document states that staff are unable to provide feedback to students once an assessment has commenced which has the potential to compromise the authenticity of the students' work (Pearson, 2017).

Despite this staff can provide brief verbal feedback which was seen during the classroom observations. When work is handed in it is marked as either a pass, merit, or distinction. Laura (BTEC Law Teacher, Focus Group 1) described the process as frustrating.

> It is really hard because you want to tell them, and you want to like give them a little more help, but you can't.

Nonetheless, teachers do not always encounter positive student responses in such situations as "they think you're not doing your job properly" (Laura, BTEC Law Teacher, Focus Group 1) and continue to ask what they need to do. This was also observed in A Level subjects with students in Core Maths and Chemistry asking their teachers for clarification, for example relating to what they could be asked to do in the exam (Field Notes, Observation 2).

As previously stated in Section 5.5, A Level teachers discussed the importance of up-skilling their students to become more independent learners, this is also deemed to be important in BTEC subjects.

> It's trying to get them to be more independent and think "well actually I've got to read through my work and work out which areas need improving and what have I missed off" (Amy, Health and Social Care Teacher, Focus Group 1).

Several instances were observed of the need for students to show a greater level of ownership, for example in Core Maths 5 students had not completed homework with the teacher reminding them that they only had 3 months until their exam and "I can't wait a week for homework to turn it back around" (Field Notes, Observation 6). One question on the homework was worth 15-marks, which represented a quarter of their exam. Therefore, it could be contended that this further exemplifies the pressures teachers face when preparing their students for the summative. In addition, one student did not understand some of the work as they had missed the lesson on this. Students were reminded of the need to catch up if they miss a lesson. Teachers attempting to collect missing homework were also observed in History and Physics. Despite this, some instances of student independent work were observed, in Chemistry one student was challenged over the use of his phone, although he was using this to access Kerboodle, an online learning resource. Furthermore, Louise (Biology Teacher, Focus Group 1) and Alan (Physics Teacher, Focus Group 1) discussed their approaches to assessment feedback which incorporates students completing independent work on areas they need to improve. Thus, in some contexts, students were seen to be taking ownership of their learning.

In the research literature much discussion relates to the unintended consequences which ensue from high stakes testing including teacher-centred pedagogy and teaching to the test (Harlen, 2007). During the classroom observations very few instances of the former were identified. In two Biology lessons teachers adopted a "lecture style" approach to transmit information to students prior to them completing a task. While it could be argued that in this instance the students were passive recipients of knowledge, they were then expected to actively utilise and apply this

information to complete a series of tasks. In the focus group interviews teachers discussed the importance of up-skilling students and developing their ability to be independent particularly considering the new examination structure. However, Lattimore (2001) argues that high stakes tests incite negative attitudes amongst students towards tested material. Moon *et al* (2007) go further to discuss the impact of high stakes testing on Gifted and Talented students. They contend that while some Gifted and Talented students respond well to the challenge schools offer them, others can become disinterested and demotivated. Furthermore, Moon *et al* (2007) found that a proportion of Gifted and Talented students feel "frustration and resentment at the slow pace of learning" (xiii) while the impacts of high stakes testing are particularly profound for disadvantaged pupils. Arguably, this is in contrast to findings of my study whereby in the lessons I observed several instances of independent student working were identified, for example in Chemistry students were using Kerboodle to research additional information, and in Accounting students were observed independently working through a series of exam style questions. In addition, in the focus groups teachers discussed instances whereby students were receptive of the assessment process, for example the system of numbered improvement points adopted in Criminology. Indeed, Polesel, Rice and Dulfer (2014) suggest that high stakes tests can serve to motivate pupils as they focus them on the end goal, ultimately to pass the test.

6.4 External Summative Assessments Methods and Teachers Classroom Practice

As discussed in Chapter 4, document analysis of exam board documentation has been undertaken to demonstrate the parameters within which teachers must operate when it comes to high stakes summative assessments. Following this, I have compared the external summative assessments for each subject observed and cross-referenced this with the types of assessments (both formative and summative) which were observed in the classroom. This is shown in Appendix 13.

The use of exam style questions was observed in nine lessons, all of which were different subjects. Of those nine subjects, seven are summatively assessed by examinations which account for 100% of the course mark.

Table 6-1 – Use of Exam Style Questions in Subjects Externally Summatively Assessed by Examination

Subject	**External form of summative assessment**	**Type of formative assessment observed**	**Type of summative assessment observed**
Accounting	100% examination	Questioning, modelling, graphic organizers, other	Exam style questions, feedback-marks/grades
Biology	100% examination	Sharing learning objectives, Questioning, peer assessment, self-assessment, gauge understanding, other	Exam style questions
Chemistry	100% examination	Questioning, modelling, other	Exam style questions, other
Classics	100% examination	Questioning, gauge understanding	Exam style questions, feedback-marks/grades, using assessment criteria
Criminology	Controlled assessment & exam	Questioning, peer assessment, modelling, graphic organizers, other	Exam style questions, feedback-marks/grades, using assessment criteria
Geography	80% exam, 20% individual investigation	Questioning, self-assessment	Exam style questions, using assessment criteria
Maths	100% examination	Questioning, self-assessment, gauge understanding, other	Exam styles questions, feedback-marks/grades, using assessment criteria
Further Maths	100% examination	Questioning	Exam style questions, using assessment criteria
Sociology	100% examination	Sharing learning objectives, Questioning	Exam style questions

The provision of marks and/or grades was observed in six subjects (Table 6.2), five of these subjects are fully assessed by examinations at the end of the two-year programme of study.

Table 6-2 – Provision of Marks/Grades in Subjects Externally Summatively Assessed by Examination

Subject	**External form of summative assessment**	**Type of formative assessment observed**	**Type of summative assessment observed**
Accounting	100% examination	Questioning, modelling, graphic organizers, other	Exam style questions, feedback-marks/grades
Business Studies	100% examination	Questioning, self-assessment, other	Feedback-marks/grades
Classics	100% examination	Questioning, gauge understanding	Exam style questions, feedback-marks/grades, using assessment criteria
Criminology	Controlled assessment & exam	Questioning, peer assessment, modelling, graphic organizers, other	Exam style questions, feedback-marks/grades, using assessment criteria
Maths	100% examination	Questioning, self-assessment, gauge understanding, other	Exam styles questions, feedback-marks/grades, using assessment criteria
Religious Studies	100% examination	Questioning, other	Feedback-marks/grades, other, using assessment criteria

Similarly, assessment criteria were referred to in five subjects, three of which are assessed by 100% examination, as depicted in Table 6.3. Therefore, it can be argued that in subjects which are externally assessed by examination teachers employ specific summative assessment strategies to prepare their students for such high stakes tests.

Table 6-3 - Use of Assessment Criteria in Subjects Externally Summatively Assessed by Examination

Subject	External form of summative assessment	Type of formative assessment observed	Type of summative assessment observed
Classics	100% examination	Questioning, gauge understanding	Exam style questions, feedback-marks/grades, using assessment criteria
Criminology	Controlled assessment & exam	Questioning, peer assessment, modelling, graphic organizers, other	Exam style questions, feedback-marks/grades, using assessment criteria
Geography	80% exam, 20% individual investigation	Questioning, self-assessment	Exam style questions, using assessment criteria
Further Maths	100% examination	Questioning	Exam style questions, using assessment criteria
Religious Studies	100% examination	Questioning, other	Feedback-marks/grades, other, using assessment criteria

However, in all nine subjects which culminate in assessment via examinations, a plethora of formative assessment strategies were evident. As previously stated in Section 5.1.1, questioning was observed in all lessons, while self and peer-assessment were seen in eight lessons. Additionally, teachers in ten lessons employed strategies to gauge student understanding of a task or topic. This can be seen to substantiate evidence from the focus group interviews in which participants discussed not only the importance but the value of utilising a combination of formative and summative assessment strategies. It is of particular note that of the eleven subjects which are externally assessed by examination only, teachers in all but one of those lessons were observed using a combination of both formative and summative assessment strategies as part of their day-to-day teaching practice. Nine observed subjects are externally assessed via a combination of examinations and coursework, although in all cases examination is weighted more highly than coursework, further reflecting the emphasis placed on external summative assessments. In five of these lessons, no summative assessment strategy was evident, whereas formative assessment strategies were observed in all lessons. While in these subjects, the ratio of coursework to examination has a greater weighting towards exams, with the exception of Health and Social Care,

the absence of summative assessment could be attributed to students having recently completed mock exams in their subjects. Evidence therefore suggests that teachers are using formative assessment strategies to prepare their students for the summative, particularly as their use is prominent in subjects which are wholly or for the most part assessed by examination. As such, this finding could be seen to contradict the work of both Harlen (2005) and Man Sze Lau (2016). Indeed, Harlen (2005) argues that in the wake of the pressures of high stakes testing teachers do not utilise formative assessment as part of the learning process. Furthermore, Man Sze Lau (2016) contends that emphasis on accountability and certification associated with high stakes summative assessment has pushed formative assessment "out of its place in the relationship originally intended by Scriven and Bloom" (512). As Health and Social Care is a BTEC which provides students with an A Level equivalent work-based qualification, there is a greater weighting towards internally assessed coursework.

Table 6-4 - Frequency of Assessment Types Observed in relation to Form of Summative Assessment in Each Subject

Formative Assessment Type	Observed Frequency	Number summatively assessed by 100% examination	Subjects	Number summatively assessed by exam & coursework	Subjects
Exam style question	9	6	ACC, BIO, CHEM, CLA, Ma x2, SOC.	2	CRI, GEO.
Provision of marks/grades	6	5	ACC, BUS, CLA, MA, RS.	1	CRI.
Use of assessment criteria	5	3	BIO, MA, RS.	2	CRI, GEO.
Peer Assessment	2	1	BIO.	1	CRI.
Self-Assessment	6	5	BIO, BUS (x2), FR, MA.	1	GEO.
Gauge student understanding	10	4	BIO x2, CLA, MA x2, PHY.	4	FILM, HSC, HIS, LIT.
Questioning	19	10	ACC, BIO, BUS, CHEM, CLA, FR, MA, PHY, RS, SOC.	9	CRI, ENG, FILM, GEO, HIS, HSC, LIT, MED, PE.
Sharing learning objectives	5	2	BIO. SOC.	3	FILM, HIS, LIT.
Modelling	4	2	ACC, CHEM.	2	CRI, ENG.

It is also useful to examine the extent to which teachers' use of formative and summative assessment strategies in the classroom is affected by the type of questions students are required to complete in their final summative examinations for each subject they study. I have reviewed the structure of questions which make up the examination papers in all subjects which are externally summatively assessed by examination (as shown in Appendix 9) and compared this to the types of formative and summative assessment observed (Appendix 13).

Three subjects I observed (Fine Art, ICT and Law) have no form of external summative examination, this reflects a degree of subject individuality, with Fine Art requiring students to complete a portfolio of practical work thus requiring a much more personalised response from the student. It can therefore be argued that while the assessment brief is the same for all students undertaking the subject, it is interpreted differently by different students, thus leading to the production of more personal assessment material. The remaining two subjects (ICT and Law) require the completion of internally assessed coursework units to enable students to achieve a BTEC qualification. As previously stated in Section 6.2, data that I have analysed suggest a degree of subject individuality in relation to the style of examination questions, with Media Studies questions, for example, commanding a greater degree of creative input from students, unlike Sociology.

Of the twenty subjects which have summative assessments in the form of examinations at the end of the two-year programme of study, fifteen require students to complete essay-based or extended answer questions as shown in Table 6.5 below.

Table 6-5 - Subjects which Use Essay-Based/Extended Response Questions

Subject	Structure of Questions	Type of Summative Assessment Observed
Accounting	Multiple choice, short answer & extended answer (25 marks)	Exam style questions, feedback- marks / grades,
Biology	Short answer and extended answer (25 marks)	Exam style questions
Business Studies	Short answer and extended answer questions (20 marks)	feedback- marks / grades
Classics	Short answer questions. Essay based (20 marks)	Exam style questions, feedback- marks / grades, using assessment criteria
English Language	Short answer and essay based (20, 30 marks)	Performance task
English Literature	Extended answer questions (25 marks)	
Film Studies	Essay based (40 marks)	
French	Short answer questions. 2 x 40-mark questions	Written essay, report
Geography	Short answer and essay based	Exam style questions, using assessment criteria
Health and Social Care	Short and extended answer questions	
History	Essay based (1x 30 marks, 2 x 25 marks)	
Media Studies	Short answer and extended answer questions. (10, 12, 15 and 30 marks)	Other
Physical Education	Multiple choice, short answer & extended answer (up to 15 marks)	
Religious Studies	Essay based (20 and 30 marks)	feedback- marks / grades, other, using assessment criteria
Sociology	Short answer (4,6,10 marks) and essay based (20, 30 marks)	Exam style questions

It can therefore be argued that students require opportunities to practise answering such long questions. The use of exam style questions was observed in five of these fifteen subjects, as shown by the yellow colour coding in the table above. Furthermore, the use of assessment criteria was observed in three of these subjects (highlighted in green above). However, it is pertinent to note that at the time of the classroom observations students had recently completed mock examinations in all their subjects.

In the fifteen subjects which incorporate essay based or extended answer questions in their final exam (depicted using the pink colour coding above), a range of formative assessment strategies were observed. There were eleven subjects in which up to two formative assessment strategies were used by teachers in each lesson and twelve in which three or more were used, as shown in Table 6.6.

Table 6-6 - Number of Formative Assessment Strategies Observed in Subjects with Extended Answer Questions in their External Summative Assessments

Up to two formative assessment strategies observed	**Three or more formative assessment strategies observed**
CLA	ACC
ENG	ART
GEO	BIO
ICT	BUS
LAW	CHEM
MA	CRI
MED	LIT
PE	FILM
PHY	FR
RS	HIS
SOC	HSC
	MA

This therefore substantiates the argument that teachers are using the formative as part of their day-to-day practice to prepare their students for the summative.

Table 6-7 - Frequency of Formative and Summative Assessment Strategies Observed in Years 12 and 13

Subject	Year of Study	Frequency of formative assessment observed	Frequency of summative assessment observed
Accounting	Year 12	4	2
Biology	Year 12 x2	6	1
Business Studies	Year 12 x2	3	1
Chemistry	Year 13	3	2
Classics	Year 12	2	3
Criminology	Year 12, Year 13	5	3
English Language	Year 12	2	1
English Literature	Year 13	3	
Film Studies	Year 13	3	
Fine Art	Year 13	3	
French	Year 12	3	1
Geography	Year 12	2	2
Health and Social Care	Year 12	3	
History	Year 13 x2	3	
ICT	Year 12	1	
Law	Year 12	2	1
Maths	Year 12, Year 13	4	3
Further Maths	Year 13	1	2
Media Studies	Year 12	1	1
Physical Education	Year 13	2	
Physics	Year 12	2	
Religious Studies	Year 13	2	3
Sociology	Year 12	2	1

As previously stated, more instances of teachers using formative assessment strategies were observed overall. However, it is useful to note that use of formative assessment strategies as part of teachers day-to-day practice was a frequent occurrence in lessons being taught in both Years 12 and 13. It can therefore be argued that teachers are using formative assessment as a means to prepare their students for the summative at the outset in Year 12 as well as in Year 13. Furthermore, more instances of summative assessment were observed in subjects which are externally summatively assessed by 100% examination and were a feature of lessons in both Years 12 and 13, apart from Criminology.

6.5 Professional Attributes

My analysis of the focus group data uncovered several professional attributes which teachers deemed to be significant and they attached value to. Much emphasis was placed upon promoting

the independence and up-skilling of students, which, for the most part, was attributed to the changing examination structure. According to Harlen (2007: 42) teachers want their students to be independent and self-critical, good problem solvers and effective communicators and ultimately place value on these attributes. However, it is argued in the literature that teachers often feel that they must "prioritise meeting targets for test results, ahead of encouraging wider learning that they actually value" (Harlen, 2007: 42). In addition, James and Pedder (2006) suggest that teachers work within a value-practice gap. In other words, they contend that there is a deficit between the ideals and values which teachers hold and what they can realistically achieve on a day-to-day basis with their students. It could be contended that the demands placed on teachers in light of high stakes testing often serves to constrain them from being able to pursue their aspirations as part of their day-to-day teaching practice, instead having to focus on meeting targets.

In contrast to these arguments the importance of promoting the skills of independent working were discussed in the focus groups with reference made to actively encouraging students to become more independent in their learning. In addition, Louise (Biology Teacher, Focus Group 1) referred to the independent work students complete both in class and for homework upon receiving teacher feedback for their assessment work.

> It's trying to get them to be more independent and think "well actually I've got to read through my work and work out which areas need improving and what have I missed off" (Amy, Health and Social Care Teacher, Focus Group 1).

> They take that home, review it and bring in their independent work exam questions (Louise, Biology Teacher, Focus Group 1).

Alan (Physics Teacher, Focus Group 1) discussed the impact of this approach on students.

> The really powerful bit is getting them to look at assessment in a different way. They start to see the value of it being protracted and looking at mark schemes and kind of independently thinking and challenging themselves and going back and revisiting. They've been really positive about it and it's actually spurred a lot of them to engage with it.

Furthermore, examples of independent student working were observed during lessons. In Chemistry students were working independently to compile a summary of homogenous catalysts, whilst in Religious Studies students were reviewing teacher feedback on an essay before working on an improved response. Similarly, in Accounting students were completing a cross section of exam style questions on an individual basis.

The importance of students being able to review their feedback and work independently, particularly in the run up to the final exam, was emphasised by Simon (Criminology Teacher, Focus Group 2).

> We won't have that time really for any kind of one-on-one teacher student contact time so it is really reliant on them being able to identify the areas where they can improve based on their papers and their marks.

Despite the importance teachers attached to the promotion of independent working and up-skilling, one difficulty was identified. This was observed in a Business Studies lesson where students were asked to make improvements to their test papers, although, one student completed this task ahead of everyone else.

> You might think we could spend half an hour improving that very usefully but some of them, despite how many times you go back to them, will feel like they're done in like 5 minutes (Ashley, English Teacher, Focus Group 2).

According to Harlen (2005) teachers make "little use of assessment formatively to help the learning process" (209). It can therefore be argued that the emphasis upon formative assessment acts as a precursor to ultimately prepare students for high stakes summative assessments.

In both focus groups, there was a consensus amongst teachers that formative assessment played a key role in their day-to-day practice. However, it can be argued that the aforementioned teachers are working against a backdrop of high stakes testing and are using formative assessment to get their students ready for the final, external examination. This was exhibited by Paul who discussed having used the results from a recent summative assessment to improve students learning.

> It was a low stakes little test in the classroom but that's given me some meaningful data that I can then use to put intervention in place and hopefully improve their understanding of those calculations (Paul, Business Studies Teacher, Focus Group 1).

Simon (Teacher of Criminology, Focus Group 2) discussed incorporating strategies into his teaching of Year 12 students with the aim of preparing them for linear exams, even though they do not sit their exams until the end of Year 13. It can therefore be contended that the development of students' exam preparedness skills has trickled down.

6.6 Concluding Thoughts

This chapter has sought to explore and analyse the following themes, identified from focus group interviews, lesson observations and document analysis: the approaches to assessment teachers use to prepare students for high stakes summative assessment, formative assessment as a means to prepare for the summative, and teachers' professional attributes.

Teachers approaches to preparing students for external examinations were in some cases affected by a degree of subject individuality, for example the essay-based nature of the Sociology exam led the teacher to place emphasis on the use of modal answers. Whereas in Physics the focus was on the transmission of knowledge. Emphasis was also placed on the use of exam board assessment criteria to show students how to progress to the next level in several subjects. For example, taking the form of an exam style question, alongside formative feedback, and opportunities to feed forward.

Teachers deemed there to be a need to change student perceptions vis à vis assessment, and develop their resilience, particularly considering the transition to linear examinations. In other words, teachers need their students to be able to tackle whatever comes up on the final summative examination as the opportunity for multiple resits which used to exist under the modular system is no longer available. As examinations are now sat at the end of a two-year period of study, there is increasingly more pressure on students to perform in the final exam.

As I established in Chapter 5, fewer instances of summative assessment were observed. However, formative assessment was frequently used to prepare students for or as a response to the summative. Questioning was frequently used in this way, for example, to check students' understanding of command words in an exam question before embarking on the task. The combination of formative and summative assessment was also exemplified through the completion of an essay or exam style question which was subsequently teacher marked. Such feedback was later utilised by students to close the gap in their learning. This has implications for Assessment for Learning insofar as formative assessment is not simply being used to improve students' learning and develop their academic skills but to improve their performance in readiness for high stakes tests. In other words, its use could be seen as having an ulterior motive, it is not being used as a teaching strategy in its own right, implementation is fuelled by the demands and pressures of high stakes tests.

Teachers attached value to several professional attributes, including the promotion of independence amongst students, particularly following the transition to linear examinations. This was exemplified through students using teacher feedback to close the gap in their learning. Many teachers also attached value to the nature of the feedback they gave to their students, with much emphasis being placed upon the comment marking as opposed to marks and grades.

Considering the emphasis placed upon high stakes testing in the English education system, it can be argued that teachers' approaches to assessment have implications for Assessment for Learning. Having presented and analysed the data and discussed it in light of the wider research literature, I will now turn the focus in the next chapter to pull the analysis together to answer my research questions.

7 Research Questions: Reflecting on my Findings

The previous two chapters have used data from the focus group interviews, classroom observations and document analysis to draw together key findings and recurrent themes identified through the process of data analysis. This chapter aims to examine the four research questions upon which this study is based and provide an answer to each question in turn, subsequently building on the analysis presented in Chapters 5 and 6.

As discussed in Chapter 3, there is much work in the literature documenting the use of both formative and summative assessment in the English education system with assessment being seen to serve a range of purposes. The contested nature of this area of research alongside my findings I feel makes my research questions particularly interesting. The discussion of each of my four research questions in this chapter enables me to reflect upon the findings I have identified over the course of my research and make links back to the existing research literature to relate my findings to the work of others in this field. Areas of agreement as well as any disparate findings will be identified and discussed. A consideration of my research questions will enable me to reach conclusions relating to how teachers operationalise formative and summative assessment with their students at Mitford College and their rationale for doing so. I will also be able to explore the extent to which high stakes testing affects teachers' assessment practices and consider how formative assessment is employed by teachers as a means to prepare their students for the external summative assessment at the end of Year 13. Finally, I will examine whether a value-practice gap is in existence.

7.1 How do teachers use formative and summative assessment strategies with their students and why do they use them?

While summative assessment strategies were observed during lessons, this was to a lesser extent than formative assessment strategies which it was argued should be embedded into day-to-day teaching and lesson delivery. This was reflected by at least one example of formative assessment being observed in each lesson as part of this study.

There was a consensus amongst focus group participants that formative assessment was synonymous with a continuous cycle of assessment, which involves regularly reviewing and revisiting information to aid retention. This was demonstrated in Biology with students being encouraged to act on feedback they receive from their teachers to close the gap in their learning. In Business Studies students were reviewing their feedback from a recent mock examination over a number of weeks. Questioning was considered to play an important part in day-to-day teaching, this was observed in all twenty-eight lessons. However, for questioning to be described as an assessment tool it should serve to enhance students' understanding and their responses should feed into teacher decision-making which in turn contributes to student progress (Jiang, 2014).

Other examples of formative assessment strategies included the use of model answers and the sharing of learning objectives. Model answers were observed in Criminology with the teacher using questioning following the completion of an exam style question to build up a model answer on the whiteboard. Furthermore, the benefits of using model answers was also identified in Sociology, particularly in the run up to final examinations, thus demonstrating the ways in which a formative assessment strategy might deliberately look to the summative. Learning objectives were shared with students at the start of both History and Biology lessons and were later returned to as a means to enable students to reflect on their progress in relation to the objectives. Alternatively, in Health and Social Care online interactive resources were utilised to engage students in a process of formative assessment. While instances of peer assessment were observed in lessons this was to a much lesser extent than other formative assessment strategies. As discussed in Chapter 5 peer assessment elicited different viewpoints amongst teachers. In addition, it was contended that students perceived peer assessment as lacking value, instead preferring their teachers to mark their work properly.

A key principle of Assessment for Learning relates to knowledge of the assessment criteria, published by exam boards such as AQA and Eduqas. From a pedagogical perspective, teachers can encourage students to engage with such criteria through the use of self and peer assessment which has the benefit of enabling them to become more familiar with what they are required to do. On

the other hand, it is possible to question the extent to which peer assessment can be trusted as a means of reliable feedback. Indeed, marks from such forms of assessment can lack reliability and training should be undertaken as a means of mitigation (Boud, 1989). As such teachers perceive there to be greater value in getting students used to the assessment criteria as opposed to a vehicle for meaningful feedback. Despite differing opinions, some teachers did cite using peer assessment as a tool to enable students to see the work of their fellow students. However, in such instances, teachers stated that the marks which were allocated as part of the peer assessment process were not recorded for any purposes. It can therefore be argued that no stakes were attached to teachers using peer assessment in this way. Although higher stakes could be associated with peer assessment through a students' positionality within the class. Similarly, peer assessment was employed in BTEC to maintain student engagement. Conversely, as discussed in Chapter 5, there were some advocates for the use of peer assessment, particularly when students are trained well.

Peer assessment was observed in Biology, with students completing a quiz to review their knowledge of the lesson which was subsequently peer assessed. Students were asked to read out answers with the teacher using questioning to extend responses, "what is missing?" what would the third mark be for?" (Field Notes, Observation 18). In French students were asked to mark and check their own work following the completion of a practice exam question while the teacher asked (in French) for the phrases which supported their answer. However, the term self-assessment was not explicitly used in the French class, instead, the term 'review' was utilised. Indeed, the use of self-assessment has been aligned with improved student performance (Black and Wiliam, 1998 and McDonald and Boud, 2003).

Summative assessment was utilised by teachers to not only test student knowledge and progress but also to replicate timed exam conditions. It could be contended that this represents an example of a formative aspect being added to a summative assessment. In other words, getting students ready for the summative could be seen to incorporate the interpretation and answering of exam questions under exam conditions but could also be extended to include practising the format of examinations, for example the replication of timings. As such, practice taking place in class will feed

forward to the summative. In addition, teachers used summative assessment to enable students to practise not only the physical skill of handwriting for a prolonged period of time but also applying their knowledge to specific questions. Great value was attached to carrying out summative assessments, particularly in the lead up to the final examination as it enables teachers to see what needs to be taught differently to enhance understanding as well as what students have done well and what they need to improve on. Teachers also suggested that performing well in the mock exam had the potential to increase student confidence going into the final exam.

As I discussed in Chapter 5, the following summative assessment strategies were observed in lessons: exam style questions, timed essays, presentations, performance tasks and the provision of feedback in the form of marks and or grades. However, teachers unanimously agreed that formative assessment needs to be implemented alongside regular summative assessment. In other words, the two complement each other to provide a rounded, cyclical approach to assessment. The importance of combining formative and summative assessment was exhibited through the use of numbered improvement points in Criminology to provide formative feedback to students on summative assessment material.

Examination style questions were the most frequently observed summative assessment method which took the form of students providing written answers to several questions taken from past papers and/or exam board exemplar materials. For example, this was undertaken as a starter activity in both Geography and Criminology, whereas in Classics it formed the main task of the lesson. Furthermore, in several lessons, this method was paired with a formative assessment strategy. In Criminology the completion of an exam-style question was followed by modelling and peer assessment while in Geography students self-assessed their work. Mock examinations were discussed as part of the focus group interviews to enable students to learn from the process prior to the final examination. For instance, in Business Studies students had recently completed a 2-hour paper in exam conditions which incorporated two 10-mark questions, two 12-mark questions and two 20-mark questions along with a number of smaller marked answers. The long answers were subsequently quality marked (which involves the provision of detailed feedback on both the

positive aspects of a students' work and areas for improvement which students can use to enhance their work) and feedback was given to students who used this feedback, along with the mark scheme to improve each question in turn. This could be seen to contradict the argument that assessments which are built around low-level aims, namely the recall of information, do not serve to inform teachers about students' progress (Stiggins *et al*, 1989). However, summative assessment strategies at Mitford College for example in the form of exam-style questions, fed into a cycle of assessment. The timing of summative assessments was also deemed to be important with some teachers setting end of unit tests part way through the teaching of the next topic in order to mirror the exam which could be seen as a formative function.

As discussed in Chapter 2, expressing a clear distinction between formative and summative assessment as exemplified through use of terminology, can be seen to overlook the role which summative assessment can play in the support of learning (Bennett, 2011). It can be argued that in Mitford College summative assessment is utilised as a means to support learning by way of complementing formative assessment practice as part of the cycle of assessment.

7.2 How does high stakes testing affect teachers' approaches to assessment?

My research demonstrated that high stakes testing affects teachers' approaches to assessment in Mitford College in a number of ways. The first impact of note is that of the provision of feedback to students. Black and Wiliam (1998) argue that feedback should be comment only, the practice and views of several teachers at Mitford College were centred on the provision of comment only marking. Indeed, Butler (1988) attributes the greatest learning gains to comment only marking.

In Religious Studies, students were observed receiving teacher feedback following the completion of an essay question. Students were required to respond to the comments their teacher had provided to improve their work. The teacher moved around the room to speak to students and answer their questions, questioning on a one-to-one basis (Field Notes, Observation 8). In Science, teachers discussed having moved away from both giving grades and the headline percentage on assessments, instead favouring a break down topic by topic alongside comments. This approach

has enabled students to refocus their attention on the areas they need to work on to improve their exam performance as opposed to focusing on the overall grade. Furthermore, it has enabled the teacher to focus on students' knowledge acquisition and application. As mentioned earlier in Section 5.5, the benefits of numbered improvement points were discussed to feedback to students.

As such, much emphasis is placed upon the action of the student to take responsibility for and ownership of their learning and ultimately become learners that are more independent. This can be exemplified through students using Kerboodle in a Chemistry lesson to support their own learning without direction from the teacher. Moreover, in Religious Studies, students were observed independently engaging in the cycle of assessment by using written feedback from their teacher alongside the assessment criteria to improve their answer to an exam style question. This was deemed to be of particular importance following the transition to linear examinations. Therefore, students need to become more independent and resilient in the face of external assessments taking place only at the end of Year 13. This is reinforced by the lack of guidance teachers have received on how exam papers will be marked and graded. As discussed earlier in Section 5.7, in Health and Social Care, the format of questions and associated mark allocations differed in the examination from what teachers had previously been told by the exam board themselves.

The push towards independent learning was observed in several lessons. For example, in Film Studies, students were given an exam style question and were asked to explore their understanding by generating a series of questions to invoke their thought process. That is to say how they interpreted the question and what they deemed to be important to consider when answering the question which could be seen as developing assessment literacy. This has the capacity to develop independent learners. In addition, as students encounter more exam style questions, they have a greater awareness of what to expect in the final exam. However, as will be discussed later, this could be seen as coaching students to pass exams (Au,2007). Similarly, in Accounting students were observed working independently on a cross section of exam style questions, which became progressively more difficult. While students were able to receive verbal feedback from the teacher,

they were also able to take ownership and check their answers against those on the model answer sheet. This also prevented students from having to wait for the teacher to check their work and could ultimately be promoting their development as independent learners.

While teachers agreed that providing grades to students is not always feasible, several teachers favoured the provision of marks to students. For example, in Sociology the teacher deemed it beneficial to add marks to student work given the nature of preparing students to sit examinations which have certain mark allocations attached. Providing marks to students was also considered to be important in Science to enable students to see where they have dropped or accrued marks when completing assessments. Despite this, there were instances when the giving of grades was deemed to constitute the most appropriate form of feedback. For instance, following the mock exams, in the run up to the final exams, a lack of time prohibited feedback in the form of comments. It can be argued that this is attributable to the shift from modular to linear examinations, as students must sit their A-Level exams at the end of Year 13, thus removing the opportunity for multiple resits.

Although teachers placed much emphasis upon the use of formative assessment in their teaching, time pressures and constraints were identified as a key factor, which in some subjects affected the means teachers employed to provide their students with formative feedback. In the focus group interviews, teachers of Science discussed timesaving strategies they have implemented in this regard. As previously stated in Section 5.4 teachers carry out assessments with their students in 2-3-week cycles, which generates a lot of marking as a result. In Science, teachers have rewritten the mark scheme, turning it into an explanation of what students need to include in their answer. This was acknowledged as being time consuming initially, however the teachers concerned stated that it saved a significant amount of time in the long-term.

Recent changes in A-Level examinations, incorporating the transition from a modular to a linear examination system, have created several difficulties for teachers and this in turn has affected their approaches to assessment. In the focus group interviews several teachers discussed the problems they have encountered in relation to the availability of resources from exam boards to support their teaching and assessment. This has not only led to a great expenditure of time to create new

resources for use with students, but also impacted teachers' ability to carry out formative assessment owing to a dearth of information on the style and structure of questions and answers. Further challenges in relation to switching the lesson focus from undertaking controlled assessment during lessons to preparing for exams were also discussed. As mentioned in Section 6.1, model answers were promoted as a valuable tool in Sociology, although this was not the case in all subjects, for example Media Studies where they were deemed to be of little use. This was attributed to the requirements for a degree of student creative input. Given this, it is possible to question whether other teachers of Sociology would share this view.

In addition, clear differences in approaches to assessment were identified and discussed in focus group interviews, from one subject to the next, thus reflecting a degree of subject individuality. Where the same subjects were observed on more than one occasion with a different teacher, a number of similarities were identified. For example, in both Biology and History lessons, the teacher shared learning objectives with their students at the start of the lesson and subsequently returned to them later to assess student progress. In Classics, students were observed completing a 10-mark exam style question. Prior to this, the teacher utilised questioning to prepare students for the upcoming assessment, for instance through the clarification of technical language. Later in the lesson, the teacher moved onto a 30-mark question and students were asked to work independently to clarify their understanding of the question. In Classics, student answers were set aside to be teacher marked but, in both Criminology, and Geography students were given a role to play in the assessment of their answers. This could be attributed to the length of the question to be marked.

Changes in both examination structure and specifications has also led some teachers to alter the means they employ to prepare students for assessments. Examples include the promotion of discussion and debate to engage students in learning. Following the transition to linear exams, teachers discussed the need to change how content is delivered to students to ensure that learning is deeply embedded in preparation for exams. This contrasts with the modular exam system which involved students sitting exams in both January and June. Owing to the greater frequency of exams,

teachers stated that the volume of content they were required to cover from one exam to the next, did not enable them to build in time for discussion and debate. In History, debate was centred on the Cold War and the concerns of countries such as the United States at the threat of communism spreading. Although not all teachers felt that they were able to allocate so much time to discussion and debate which could be linked to the amount of content which must be taught in advance of the summer examination period. It could be argued that this is attributable to the demands of the individual subject and the required content to be covered. It can be contended that in the wake of high stakes testing teachers spend too much time preparing students to sit examinations to the detriment of knowledge of the subject (Smith, 2004). Furthermore, as previously stated in Section 5.7 teachers in the focus group interviews expressed concern and uncertainly in relation to the organisation and structure of the new linear examinations, citing a lack of specimen materials from the examination boards as a significant issue.

7.3 How do teachers use formative assessment as a means to prepare their students for the summative?

Given the role of external summative assessment in the English education system and the associated pressures for certification and accountability, it can be argued that formative assessment has become side-lined (Black, 1998; Guskey, 2005; Man Sze Lau, 2016; Shavelson *et al*, 2004). However, it is possible to question the extent to which such claims are still relevant today, over twenty years since the publication of Black and Wiliam's (1998) seminal piece. As previously mentioned in Section 2.1, while there are several definitions of formative assessment, there is largely a broad consensus over what formative assessment constitutes. In addition, definitions have changed over time (Vlachou, 2015). In the research literature formative assessment is portrayed as a key element of effective teaching and learning. As a process, it is structured, interactive and involves both teachers and learners working together to elicit, interpret and act on information about student learning (Vlachou, 2015). Nonetheless, in practice there have been a number of issues. Despite formative assessment constituting a key part of educational strategy, there is evidence to suggest that its implementation is somewhat flawed (Vlachou, 2015). Arguably, it is

possible to question whether anything has changed and the extent to which summative assessment remains in the driving seat. During my research, many examples of teachers utilising formative assessment to prepare their students for the summative were identified. Therefore, this challenges the idea that formative assessment is almost side-lined in the wake of external summative assessments.

As previously discussed, teachers were observed both referring to and utilising mark schemes and assessment criteria documentation with their students. For example, in Religious Studies, while responding to teachers' comments, students were also encouraged to refer to not only the mark scheme but also grade boundary descriptors. Despite this, several students asked their teacher what they needed to do to reach the next grade boundary. It can therefore be contended that students valued the establishment of a clear dialogue between themselves and their teacher to improve their performance. This contrasts with the argument put forward by Nichol and Macfarlane-Dick (2006) that students often struggle to decipher the meaning of feedback from their teachers. Furthermore, Evans (2013) suggests that feedback can be received in contrasting ways by learners. Indeed, providing effective feedback can be seen as challenging, particularly as research has shown that feedback from teachers often omits specific guidance on how to progress (Higgins, Hartley, and Skelton, 2001) or is difficult to follow up (Poulos and Mahony, 2008). Whereas, in my study students were instead observed turning to their teachers to have their questions answered. In Science, teachers spoke of rewriting the mark schemes to support learning and increase their accessibility to students. Similarly, in a Year 12 Geography lesson, students were required to utilise the mark scheme, after having completed a practice exam question to self-mark their work. Equally, in Core Maths, when going through answers to exam style questions on the whiteboard, the grade boundaries were also displayed to enable students to see how many marks they needed to accrue in order to reach a particular grade. In addition, teachers in both Geography and History lessons made links to the exam board specification by way of utilising content listed in the specification for their lesson titles.

Further examples of teachers using formative assessment to prepare students for the summative were seen in Accounting and both Core and Further Maths lessons. In both Maths classes, teachers gave students exam style questions to complete for homework. However, not all students had completed this with the teacher emphasising the importance of them taking ownership and taking the time to complete such assignments, although the importance was emphasised after the deadline had passed. In contrast, in Accounting, students were required to complete any work not done in class at home. As students work at a different pace in class, students were informed that home working is necessary to ensure that all students attempt the more difficult questions as part of their exam preparation. Upon completion, students received feedback from their teacher. In the aforementioned examples, teachers were not only encouraging students to become independent learners and take ownership of their work, but also highlighting the rationale for undertaking tasks of this nature. In seven lessons teachers provided students with a clear rationale for the tasks they were required to complete, thus emphasising their importance in relation to the final summative assessment. For example, in Film Studies, the teacher provided students with a rationale for the task they were completing by emphasising the knowledge requirements of the final exam.

Teacher questioning also played a key role in a Media Studies lesson. The teacher relayed to students what they need to be able to do in the final exam, that is not portraying the narrative but the reasons why it is so appealing to audiences. After having watched a video on Assassins Creed, the students were asked to engage in discussion with their peers, as the teacher circulated around the room. Following this, the teacher led a whole-class discussion and students were encouraged to develop their answers and extend and embellish their responses. Conversely, research has shown that low-cognitive questions are more frequently used as part of teachers day-to-day practice (Jiang, 2014; Tan, 2007). As discussed in Chapter 6, the style of Media Studies exam questions requires students to be able to provide their own creative input. Therefore, through this task it can be suggested that the teacher was encouraging students to develop this skill in relation to the Assassins Creed video.

The use of peer and self-assessment strategies could be a means to prepare students for the summative. It could be argued that this is attributable to a degree of subject individuality, for example in Sociology students are required to complete essay questions worth both 10 and 30 marks, respectively. As mentioned in Chapter 5, the use of peer assessment in Sociology enables the teacher to use a formative strategy to prepare students for the summative. Alternatively, it could be viewed more pragmatically that this represents a way for the teacher to reduce their marking workload.

There was a consensus amongst teachers that formative assessment was synonymous with best practice and should form part of teachers' day-to-day practice. Moreover, the regularity of formative assessment was also deemed to be important with such assessments being undertaken every 2-3 weeks. Following the marking of summative assessments, the dissemination and review of feedback formed part of a cycle of assessment, which lasts several weeks. This was considered beneficial insofar as students had often forgotten information after a few weeks and the constant revisiting was said to keep it fresh in their minds. This process was initiated by teachers as part of their day-to-day practice. Despite this, not all teachers initially saw the benefit of this approach. In such instances, it can be argued that teachers are using the formative feedback they provide as a means to prepare students for the summative. Students can review the feedback they receive and use it to make improvements to their work and subsequently learn from any mistakes prior to the final exam. The importance of the cycle of assessment (involving the regular reviewing and revisiting of information to aid retention) was exemplified in Sociology. Students were encouraged to draw on prior learning by recalling and writing down 10 facts about social action theory. The teacher used questioning to elicit student responses, however, some students were unable to recall this theory, despite it having been a topic on a recent exam. Therefore, this example reinforces the importance of teachers revisiting information with their students to prepare them for the summative.

Teachers in several lessons set their students in class assessments in the form of exam style questions, which were timed, this was noted in Classics, Criminology and Geography. As such, it

may be said that this replication of timed conditions from external summative assessments represented a means of practising for the shape or format of the summative which constitutes a part of assessment reliability. In both Criminology and Geography, students' answers did not receive feedback from the teacher. In Criminology, peer assessment was utilised, while in Geography students self-assessed their answers. Meanwhile in Classics, answers were teacher marked. This could be attributed to the number of marks attached to the questions in Classics, 10 and 30, respectively. Thus, it could be argued that this necessitated teacher feedback to enable students to clearly see their level of progress. This further substantiates earlier arguments relating to the use of peer assessment, with its implementation being seen as time-consuming (Black *et al,* 2003) and requiring students to fully understand the targets of their learning (Wylie and Lyon, 2015), which is not always the case (Nikto, 1995). However, this contradicts the approach advocated in Sociology with the teacher promoting use of self and peer assessment for the marking of 10 and 30-mark essay questions.

The above-mentioned examples all demonstrate how teachers in this study utilised formative assessment to prepare their students for the summative. It can therefore be argued that, based on this evidence, formative assessment is not being side-lined in the wake of high stakes summative assessments. Instead, it forms a key part of teachers day-to-day practice as a means to prepare students for the summative.

7.4 What are teachers' beliefs about assessment strategies? Is there a relationship between teacher values and day-to-day practice?

Teachers expressed several beliefs and values relating to assessment strategies. Much emphasis was placed upon the importance of a cycle of assessment, the changing of student mind-sets (with the encouragement of students to look at assessment in a different way being a key part of this) and the role of students themselves in the assessment process. My study has identified a value-practice gap in relation to one area of teachers' day-to-day practice. At the time of the focus group interviews, teachers had recently completed mock examinations with their students as part of their

preparation for the final examinations. Teachers attached value to the idea of allocating time to provide one-to-one feedback to students, this would be particularly useful for those students who performed less well. Despite constraints on time hindering the provision of individual feedback, mock exams can still be seen as a means to "rehearse" for the summative. Notwithstanding useful feedback can still be given to students at the whole-class level. While teachers attached value to the provision of one-to-one feedback following a mock exam, in practice this was not deemed to be feasible owing to time constraints, thus resulting in a value-practice gap.

In practice teachers provide individual feedback to students in the form of a grade and feedback to the whole class relating to common misconceptions and how particular questions could be answered better to accrue more marks. Therefore, the values which teachers expressed relating to the provision of feedback on a one-to-one basis were not translated into practice. However, other factors, which teachers discussed and attached value to, were evidenced in their day-to-day practice as part of the lesson observations. Teachers attached considerable value to the promotion of independent learning and the development of resilience amongst students, particularly following the transition to linear examinations. It can be argued that teachers want their students to become independent learners who exhibit good problem-solving and communication skills (Harlen, 2007). Furthermore, assessments have the potential to not only fulfil formative and/or summative functions, but also develop skills of sustainable assessment that will be applicable to future learning situations (Boud and Soler, 2016).

Examples of the promotion of independent learning were observed in lessons. In Business Studies students were asked to complete an audit of exam skills they deemed necessary to answer a range of question types. Meanwhile in Fine Art students were observed working independently on their portfolio of work. Similarly, in Accounting students were working independently on a series of practice exam questions, which became progressively more difficult. In Classics, students were given a 30-mark examination question and required to review the question and generate initial ideas. Whilst in Film Studies, students were asked to come up with questions in response to an unseen exam question to explore their thinking. It is possible that such approaches encourage

students to become more independent in the face of exam style questions, particularly as they become more used to examining questions they have never seen before, which in turn can serve to increase their resilience.

There was much agreement amongst the teachers at Mitford College that peer assessment is not a valuable assessment tool. Despite this, both Paul (Business Studies Teacher, Focus Group 1) and Laura (BTEC Law Teacher, Focus Group 1) deemed peer assessment to be seen more usefully as a process to get students used to the assessment criteria. As mentioned earlier in Chapter 2, both peer and self-assessment are not always favoured by teachers, as they are required to let go of the traditional idea that students are simply the recipients of knowledge. However, teachers in my study attributed their caution and trepidation with such forms of assessment to the value and accuracy of peer and self-assessment in practice. It can be argued that embedding the necessary skills can be an extremely time-consuming process, one which may not be possible owing to the demands and pressures of high stakes summative assessments (Black *et al*, 2003). Furthermore, students' responses when assessing their own work may not always be honest (Ross, 2006). Other concerns were also raised relating to the accuracy of the marks students give to their peers, with leniency being cited as a particular problem. For example, in Biology students were required to use the mark scheme to give their peers a mark out of six. Despite this, a couple of teachers stated that they do use peer assessment, for example in BTEC Law to keep students engaged and in Business Studies to enable students to see the work of others. Peer assessment was observed in a Biology lesson whereby students were required to peer mark answers to a quiz, which tested their knowledge of the lesson. Furthermore, peer assessment can acquaint students with the criteria against which their work is marked. Arguably, students' learning can be enhanced by partaking in this process (Black *et al*, 2003).

As previously mentioned in Section 7.1 instances of self-assessment were observed in some lessons, although the activity was not explicitly referred to as self-assessment. In a Biology class, students assessed their own progress during the lesson by placing post-it notes on the appropriate traffic light colour as they left the room. It can be contended that students are capable of such

assessments on the proviso that they are prepared appropriately and are given a rationale for undertaking the assessment (Mowl and Pain, 1995). In Sociology, the use of self-assessment was advocated but this was contingent upon students being trained well. It can be suggested that training students in such practices can positively impact their performance.

Teachers discussed the importance of changing student mind sets from the approach they had become accustomed to at GCSE level. Many teachers supported the view that giving a grade to students for each piece of assessed work was counter-productive, the exception to this was following the final mock exam. Overall, they did not attach value to the giving of grades. This was attributed to students receiving their target grades at the outset of their course. As previously stated, self-assessment as a part of formative assessment, enables students to work with and become familiar with assessment criteria. The use of the assessment criteria to show students what they need to do to reach the next grade could be a distortion of self-assessment. While this could enable students to take ownership of the process in terms of their self-assessment and progress, attaching grades could be considered to be going against formative assessment. Furthermore, at the outset of a course, assessments are not anything like A Level standard, thus further reinforcing the argument to not provide students with a grade.

Teachers attached much greater value to students responding to feedback in an active manner as opposed to simply being passive recipients of feedback, for example, in Religious Studies students were observed responding to such feedback to improve their essay answer. While students in this subject were provided with a grade and/or mark alongside teacher comments, notwithstanding, they were engaging in the feedback cycle and using the comments they received formatively, contradicting the argument of Black *et al* (2003). It can also be contended that in the wake of high stakes testing, teachers are required to produce summative data to monitor student progress for school leaders and managers (Tiknaz and Sutton, 2006). Emphasis was placed on ensuring that students were familiar with and understood assessment objective criteria. Mock exams were also deemed to be an important part of the learning process for students as they provide a means for them to learn from any mistakes they make. Although carrying out mock exams is time-consuming

from a marking perspective, teachers emphasised the value in being able to identify gaps in their student's knowledge prior to the final exam.

After having observed twenty-eight lessons across a range of subjects and conducted two focus group interviews, several links have been identified between the values teachers hold and their day-to-day practice. Indeed, my research has shown that teachers at Mitford College attached value to a process of cyclical assessment, the changing of student mindsets and students themselves playing an active role in the assessment process, examples of which were observed as part of my study. For example, through students actively responding to feedback as part of a feedback cycle. Further examples included the promotion of independent learning and the development of resilience, particularly following the transition to linear exams. However, I have also uncovered evidence to support the existence of a value-practice gap. Teachers attached value to the provision of one-to-one feedback following the completion of mock examinations but in reality time pressures, arguably exerted in the wake of high stakes summative assessments, prevented this from taking place, therefore producing a value-practice gap. Despite this, undertaking mock exams was still deemed to be a valuable process. There also were aspects of teaching practice which teachers did not attach value to, namely the use of peer assessment, attributable to a lack of marking accuracy by students and time constraints as well as the giving of grades, with the exception being following a mock exam. Therefore, it can be contended that my study has uncovered evidence supporting the existence of a relationship between teachers' values and day-to-day practice alongside the existence of a value-practice-gap. The following chapter aims to summarise and draw together the key findings from this study. In addition, several limitations will be explored before recommendations for future research are made.

8 Conclusion

8.1 The Main Research Findings

The aim of my study was to investigate how teachers operationalise formative and summative assessment with students in Years 12 and 13, who are working towards high stakes summative assessments, in the form of A Levels and BTECs, at the end of a two-year programme of study. This chapter draws conclusions relating to teachers use of formative and summative assessment, the impact of high stakes testing on their assessment practice and finally, teachers' beliefs and values in this regard.

My study has identified that formative assessment is a commonly used approach to classroom assessment, employed by teachers in their day-to-day practice, with questioning being particularly prevalent. Teacher use of questioning observed as part of my research placed onus upon students to embellish and extend their answers. It can therefore be contended that students were being actively encouraged to play an active role in the feedback process. These findings run counter to the work of Jiang (2014), who argues that low cognitive questions are more widely used in teachers' day-to-day practice. Tans (2007) findings are also comparable with those of Jiang (2014). However, in my study teachers were observed using questioning to develop students' understanding and elicit additional information from them as a means to facilitate teaching and learning. It could therefore be contended that the use of questioning contributed positively to the effectiveness of teaching and learning, which in turn supports the work of Black *et al* (2003) Ruiz-Primo and Furtak (2006;2007) and Chen, Hand, and Norton-Meier (2017). Moreover, this could be deemed to be a further example of teachers using the formative to prepare students for the summative.

My findings show that formative assessment should be supported by regular summative assessment. In class, summative assessment in the form of exam style questions was the most frequently used summative assessment strategy. My research has also shown that classroom assessment to review student progress should be frequent and cyclical in nature to aid retention of knowledge. This supports the work of Sun, Przbylski and Johnson (2016) who discussed the

emphasis placed by teachers on short, informal instances of formative assessment to provide a "pulse check", in other words a means to determine where students are up to as part of their day-to-day practice and enabled them to adapt their practice accordingly. Regular formative assessment should also be embedded as part of best practice and built into teaching plans as opposed to simply being bolted onto existing schemes of work. Moreover, the provision of teacher feedback to students should enable them to use such feedback to close the gap in their learning. It can be argued that this goes some way to align with the work of Black and Wiliam (2010) who suggest that feedback only becomes formative when teacher instruction is adapted to meet student learning needs. In addition, work by Looney, Cumming, van Der Kleij and Harris (2018) has centred on the importance of teachers presenting good quality assessment and interpreting evidence to close the gap in student learning.

It can be argued that some of my findings contradict the work of other researchers who contend that teachers' use of formative assessment is synonymous with weak practice, with emphasis instead being placed upon rote learning with few opportunities for reflection (Crooks, 1988; Black, 1993a; Black and Wiliam, 2012; Gardner, 2006). Indeed, Baird *et al* (2014) argue that teachers need to carry out higher quality classroom assessment with their students. In my study teachers placed emphasis upon the value of comment-only marking. Indeed, the significance of comment-only marking is consistent with previous research as Black *et al* (2003) and Stobart (2006) advocate the use of this form of marking. Butler (1988) also attributes the greatest learning gains to comment-only marking. Although as discussed earlier there is a need to ensure that comments and feedback are administered in a timely manner to enable students to apply insights from teacher comments (Carless *et al,* 2011). Furthermore, my study has identified time as a key challenge associated with both formative and summative assessment, although teachers have made some attempts to mitigate such impacts. While high stakes assessment can be deemed to be synonymous with an array of unintended consequences (Au, 2007; Copp, 2018; Harlen, 2007), arguably, evidence from my study suggests that students are being encouraged to take an active role in the learning process, not simply become passive recipients of knowledge.

As discussed in Chapter 2, peer and self-assessment form a key element of successful formative assessment (Black *et al,* 2003; Boud, 2000; Boud and Soler, 2016). However, findings from my study offer a variability of responses in relation to self and peer assessment, with some teachers not considering it to be a valuable assessment tool. While others utilise such forms as assessment to enable their students to see and learn from the work of their peers, teachers stated that no records of marks are kept. Conversely, these findings also support the work of Nitko (1995) and Ross (2006) who contend that self-assessment is not suitable for all students and may not always elicit honest responses. Despite this, from a pedagogical perspective, the use of such forms of assessment enabled students to engage with the assessment criteria produced by examination boards, which is arguably even more important given the transition from modular to linear examinations. Indeed, Sharma *et al* (2016) argue that engagement in self-assessment can lead to an increase in academic performance.

A key finding of my study relates to teachers use of formative assessment to prepare students for the summative, for example through the use of questioning and feedback on exam style questions. In the research literature, several researchers favour a distinction between formative and summative assessment (Harlen, 2005; Simpson, 1990), however, my research has emphasised the importance of the link between them. It can be contended that this is reinforcing the formative/summative overlap and ability of summative assessment to provide support for learning (Bennett, 2011). This is particularly significant given the emphasis placed upon external high stakes summative assessment in the English education system. In addition, in the assessment literature some argue that summative assessment can be used for formative purposes (Gipps, 1996), while my results have shown that formative assessment can be used as a means to prepare students for the summative. While there is evidence to suggest that formative assessment is being integrated into classroom practice on a day-to-day basis, it can be argued that such integration has an ulterior motive. This can be exemplified through the use of self and peer assessment. While such strategies were not deemed to be favourable with all participants, its use enabled students to engage with exam board materials. Furthermore, this evidence goes some way to contradict arguments in the

literature which contend that high-stakes testing has side-lined formative assessment use (Black, 1998; Guskey, 2005; Man Sze Lau, 2016 and Shavelson *et al,* 2004). However, my findings have shown that, in practice, the lines between formative and summative assessment are often blurred. This can be exemplified through teachers utilising formative assessment as part of their day-to-day practice to prepare their students for the summative.

My investigation has also found that teachers attached value to several professional attributes, namely the promotion of independence and resilience alongside the upskilling of their students, which was attributed to the transition to linear examinations. These findings support the work of Harlen (2007) who contends that teachers wish to develop their students' skillset to encompass independence, self-criticality, the ability to problem-solve and communicate effectively. While it can be argued that much time was spent preparing students for exams, teachers actively promoted and encouraged students to become independent learners, which was evidenced in several classroom observations. Arguably this is particularly important following the removal of modular examinations. My study has also generated evidence to suggest that teachers work within a value-practice gap, which supports the work of James and Pedder (2006) and Winterbottom *et al* (2008).

This study contributes to our understanding of how high stakes testing can affect the day-to-day assessment practices of classroom teachers. Harlen (2005) argues that in the wake of high stakes summative assessments, teachers make little use of formative assessment to support learning. However, my study has generated evidence to the contrary with formative assessment being used as a means to prepare students for the summative. Evidence shows that teachers employed formative assessment strategies with students in both Years 12 and 13, arguably this emphasises the importance of formative assessment, particularly given the need for students to be prepared to tackle anything which comes up in their final exams. It can also be argued that the structure of external summative assessment impacts teachers' uses of assessment strategies with classroom summative assessment being more widely observed in subject assessed by 100% examination at the end of a two- year programme of study. Moreover, a degree of subject individuality was also found to influence how teachers approached assessment with their students. Despite this, some

evidence of “training” students to pass exams has been identified which could be seen as coaching or an unintended consequence of high stakes testing (Au, 2007). The repeated use of exam style questions as part of classroom assessment practice could also be seen as students becoming increasingly familiar with test content (Linn, 2000).

Overall, therefore, it seems that teachers use formative assessment strategies as part of their day-to-day practice with their students as a means to prepare them for the summative at the end of a two-year programme of study. As such, it can be argued that my study has raised important questions relating to the nature of formative assessment and how it is utilised in the classroom. One such question relates to where the formative resides. It is pertinent to consider whether this is in the tool itself or in the use of the tool (Dixson and Worrell, 2016). As previously discussed, an assessment tool can only be described as formative if it is used by the teacher to give feedback to students (Black and Wiliam, 2010). This research has shown not only that teachers use the formative to prepare their students for the demands of external high stakes summative assessment but also to enable students to close the gap in their learning. On this basis it can be argued that the formative resides in the application and use of the tool as opposed to the tool itself. Simply applying a tool which could be described as formative does not automatically lead to the application of the tool with a formative function, in other words, the tool must be used in a particular way in order to close the gap in students’ learning.

8.2 Limitations of my Study

As previously discussed in Chapter 4, I initially intended to use responses from online questionnaires (as shown in Appendix 14) to form the basis of semi-structured interviews. However, due to a low response rate, the questionnaire data was not included in my study. Since my research was limited to one study site, it was not possible to draw comparisons relating to teachers use of assessment strategies across institutions. My ability to conduct this research across multiple sites was compounded by the fact that I work full time.

The lessons and subjects I observed were chosen and timetabled by Mitford College, I was therefore unable to observe the same teacher on more than one occasion for comparative purposes, although some subjects were observed twice. Furthermore, I had no choice over whether the lessons I observed were Year 12 or Year 13. Observations took place on five separate days between January and May 2018, owing to my work commitments. It is possible to question whether it would have been more effective to carry out the observations more intensively, over the period of one week for example.

The observation which took place in January 2018 occurred not long after students across Mitford College had completed mock examinations. Therefore, the extent to which the assessment practices observed on this day were reflective of 'normal practice' can be questioned. Moreover, while an observation took place in each month from January to May, other months were not examined. Therefore, this would be a useful consideration for future research.

8.3 Closing Remarks and Recommendations for Assessment Practice and Research

This study has gone some way towards enhancing our understanding of how teachers operationalise formative and summative assessment with their students in post-16 education. Through carrying out classroom observations, focus group interviews and document analysis, my findings show that teachers frequently use formative assessment in their teaching practice which is cyclical in nature. My analysis has also led me to argue that teachers use formative assessment to prepare their students for the summative and my work has reinforced the importance of a link between formative and summative assessment. The findings of this thesis also go some way to align with the work of other researchers (Black *et al,* 2003; Black and Wiliam, 2010; James and Pedder, 2006; Stobart, 2006). However, my evidence has highlighted some contradictions with works in the assessment literature (Crooks, 1998; Black 1993a; Gardner, 2006). I have also amassed evidence which shows that teachers attach value to professional attributes but work within a value-practice gap.

An examination of how teachers use formative and summative assessment with their students in Years 12 and 13 has highlighted a number of recommendations which can be made for assessment practice. Teachers at Mitford College placed much emphasis upon a cycle of assessment, involving the reviewing and revisiting of material alongside the provision of teacher feedback subsequently used by students to close the gap in their learning. The use of numbered improvement points as a means to provide feedback to students to encourage them to consider how to improve their own work was also discussed in relation to Criminology. Teachers in my study identified time as a challenge pertaining to both formative and summative assessment. A further recommendation for practice refers to a strategy to mitigate time-related difficulties, as discussed in regard to Science. Teachers rewrite the mark scheme to produce a supporting crib sheet which is given to students alongside their individually marked work and enables students to identify feedback relating to common misconceptions. While my study has provided some theoretical insights into teachers use of assessment, my study has also generated insights for practitioners which could be applied to teachers' day-to-day practice.

This study therefore lays the groundwork for future research into teachers' assessment practices. Further research could usefully explore formative and summative assessment practices in pre-16 education. This would allow for comparisons to begin to be made in relation to how teachers utilise classroom assessment strategies against a backdrop of high stakes testing at different stages in the English education system, more specifically GCSEs and A Levels. This would be particularly pertinent following the recent overhaul of the grading system used for GCSEs, the award of grades A* -G to a system of numbering 1-9. Moreover, an alternative methodological approach could be adopted, for example, in the form of an ethnographic study or a longitudinal case study to investigate teachers' use of assessment over a pre-determined period. As previously stated, there is a dearth of empirical evidence to align formative assessment with significant improvements in students' educational performance, adopting a longitudinal approach would enable this to be investigated further. Future research could be based around an examination of how teachers operationalise formative and

summative assessment strategies with their students before reviewing student performance in external high stakes summative assessments.

My research has identified a number of similarities and differences between A Level and BTEC subjects. Key similarities included the lack of information teachers had received from examination boards following the introduction of new specifications and the transition to a linear exam structure and the emphasis teachers of both qualification types placed upon the importance of upskilling their students. In contrast, BTEC subjects such as Law have a greater focus on coursework and parameters put in place by the examination boards prevent teachers from providing their students with detailed feedback and thus partaking in a cycle of assessment. Therefore, such distinctions in assessment practice would merit further investigation. Furthermore, it would also be pertinent to examine the extent to which teachers of BTEC subjects employ different formative assessment strategies compared to their A Level counterparts, with some BTEC teachers in my study advocating greater use of peer assessment and online interactive resources as a means to keep students engaged. Further research could also be undertaken in relation to student perspectives as teachers in my study emphasised the importance of changing student mindsets with regard to assessment, particularly given the transition to linear examinations. Perspectives of both BTEC and A Level students could also be explored with some differences having been identified in my study. My study has also generated evidence to suggest that teachers work within a value-practice gap which relates to their inability to provide detailed, one-to-one feedback following the completion of mock exams. Further research could explore whether this is also the case in pre-16 education prior to students taking their GCSE examinations.

My research has not only generated interesting findings which have made some contributions to our understanding of formative and summative assessment use in a post-16 teaching and learning environment but has also paved the way for future research to be carried out to further investigate these issues.

Bibliography

Abercrombie, N., Hill, S., & Turner, B. S. (1984). *Dictionary of sociology.* Harmondsworth, UK: Penguin.

Alderson, J.C. and Wall, D. (1993). Does washback exist? *Applied Linguistics,* 14, 115-129.

AQA (2019). *AS and A Level History Specifications.* Retrieved from: https://filestore.aqa.org.uk/resources/history/specifications/AQA-7041-7042-SP-2015.PDF

AQA (2018). A Level Sociology: Paper 1 Education with Theory and Methods. Retrieved from: https://filestore.aqa.org.uk/sample-papers-and-mark-schemes/2018/june/AQA-71921-QP-JUN18.PDF

AQA (2018). *Grade boundaries- June 2018 exams. A Level Reformed Linear.* Retrieved from: https://filestore.aqa.org.uk/over/stat_pdf/AQA-A-LEVEL-RL-GDE-BDY-JUN-2018.PDF

AQA (2017). *A Level Physics Examination Paper2.* Retrieved from: https://filestore.aqa.org.uk/sample-papers-and-mark-schemes/2017/june/AQA-74082-QP-JUN17.PDF

Au, W. (2007). High stakes testing and curricular control: A qualitative Metasynthesis. *Educational Researcher*, *36*(5), 258–267.

Baird. J., Hopfenbeck, T., Newton, P., Stobart, G., and Steen-Utheim, A. (2014). State of the field review: Assessment and Learning. Report for the Norwegian Knowledge Centre for Education. Retrieved from http://taloe.up.pt/wpcontent/uploads/2013/11/FINALMASTER2July14Bairdetal2014AssesmentandLearning.pdf

Ball, D.L. (2000). Bridging Practices: Intertwining Content and Pedagogy in Teaching and Learning to Teach. *Journal of Teacher Education,* 51, 241-274.

Barbour, R. (2007). *Doing Focus Groups.* London: SAGE.

Barnes, N., Fives, H. and Dacey, C.M. (2015) Teachers' beliefs about assessment in H. Fives, and M.G. Gill (eds.) *International Handbook of Research on Teachers' Beliefs*. New York and London: Routledge.

Bassey, M. (1984). Pedagogic research: on the relative merits of search for generalization and study of single events. In T.B.J. Bell (ed.) *Coordinating small scale investigation in educational management* (pp.103-122). London: Harper and Row.

Baxter, P and Jack, S. (2008). Qualitative Case Study Methodology: Study Design and Implementation for Novice Researchers. *The Qualitative Report,* 13(4), 544-559.

Bennett, R.E. (2011). Formative assessment: a critical review. *Assessment in Education: Principles, Policy and Practice,* 18(1), 5-25.

Bhattacharjee, J. (2015). Constructivist Approach to Learning: An Effective Approach of Teaching and Learning. *International Research Journal of Interdisciplinary and Multidisciplinary Studies,* 1 (6), 65-74.

Bichelmeyer, B. and Hsu, Y. (1999). Individually-Guided Education and Problem-Based Learning: A Comparison of pedagogical approaches from different epistemological views, in Proceedings of Selected Research and Development Papers Presented at the National Convention of the Association for Educational Communications and Technology (AECT) (21st, Houston, TX, February 10-14, 1999).

Biggs, J. (1998). Assessment and Classroom Learning: a role for summative assessment? *Assessment in Education: Principles, Policy and Practice,* 5(1):103-110.

Black, P.J. (1993a). Assessment policy and public confidence: comments on the BERA Policy Task Group's article 'Assessment and the improvement of education'. *The Curriculum Journal*, 4: 421–427.

Black, P.J. (1993). The Shifting Scenery of the National Curriculum in R. White (Ed). *The National Curriculum- Which Way Now?* London: Paul Chapman.

Black, P. (1998). *Testing, friend or foe? The theory and practice of assessment and testing*, London: Falmer Press.

Black, P.J. and Atkin, J.M. (1996) *Changing the subject: innovations in science, mathematics and technology education* London, Routledge.

Black, P., Harrison, C., Lee, C., Marshall, B., and D. Wiliam. (2003). *Assessment for learning: Putting it into practice*. Maidenhead: Open University Press.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 7–74.

Black, P. and Wiliam, D. (1998). *Inside the Black Box: Raising Standards through Classroom Assessment.* London: King's College.

Black, P. and Wiliam, D. (1998a). Assessment and Classroom Learning. *Assessment in Education,* 5(1), 7-74.

Black, P. and Wiliam, D. (1998b). Inside the Black Box: Raising Standards through Classroom Assessment. *Phi Delta Kappan,* 80 (2): 139-148.

Black, P. and Wiliam, D. (2009). Developing the Theory of Formative Assessment. *Educational Assessment, Evaluation and Accountability,* 21, 5-31.

Black, P., & Wiliam, D. (2010). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, *92*(1), 81–90.

Black, P. and Wiliam, D. (2012) Assessment for Learning in the Classroom in J. Gardner (eds) *Assessment and Learning* (pp.11-33). London: SAGE.

Blatchford, P. (1997). Student' self-assessment of academic attainment: Accuracy and stability from 7-16 years and influence of domain and social comparison group. *Educational Psychology,* 17(3), 345-360.

Bloom, B.S. (1956). Taxonomy of Educational Objectives. New York: Longmans, Green.

Bloom, B.S. (1969). Some theoretical issues relating to educational evaluation. In R.W. Tyler (ed), *Educational Evaluation: New Roles, New Means.* Vol 68, NSSE; distributed by the University of Chicago Press.

Boghossian, P. (2006). Behaviourism, Constructivism and Socratic Pedagogy. *Educational Philosophy and Theory,* 38(6), 713-722.

Bowen, G.A. (2009). Document Analysis as a Qualitative Research Method, *Qualitative Research Journal,* 9(2), 27-40.

Boud, D. (1989). The role of self-assessment in student grading. *Assessment in Higher Education.* 14(1), 20-30.

Boud, D. (2000). Sustainable assessment: rethinking assessment for the learning society. *Studies in Continuing Education,* 22(2), 151-167.

Boud, D. and Soler, R. (2016). Sustainable Assessment Revisited. *Assessment and Evaluation in Higher Education,* 41(3), 400-413.

Braun, V., and Clarke, V. (2006). Using thematic analysis in Psychology. *Qualitative Research in Psychology,* 3, 77-101.

Brewer, J.D. (2000). *Ethnography.* Buckingham: Open University Press.

British Educational Research Association (BERA) (2004) 'Revised ethical guidelines for educational research' Retrieved from: https://www.bera.ac.uk/publication/revised-ethical-guidelines-for-educational-research-2004

British Educational Research Association [BERA] (2018) Ethical Guidelines for Educational Research, fourth edition. Retrieved from https://www.bera.ac.uk/researchers-resources/publications/ethical-guidelines-for-educational-research-2018

Broadfoot, P. (2000). *Promoting quality in learning: does England have the answer?* London: Cassell.

Brooks, R. te Riele, K. and Maguire, M. (2014). *Ethics and Education Research.* London: Sage.

Brooks, V. (2002). *Assessment in secondary schools: The new teacher's guide to monitoring, assessment, recording, reporting, and accountability*. United Kingdom: Open University Press.

Brown, A. (1994). The advancement of learning. *Educ. Res.* 23(8), 4-12.

Brown, G.T.L. (2004). Teachers' conceptions of assessment: Implications for policy and professional development. *Assessment in Education: Policy, Principles and Practice*, 11(3): 305–22.

Brown, G.T.L. (2006). Teachers' conceptions of assessment: Validation of an abridged instrument. *Psychological Reports*, 99: 166–70

Bryman, A. (2008). *Social Research Methods.* (3rd Eds.), New York: Oxford University Press.

Bryman, A. and Burgess, R.G. (2002). Developments in qualitative data analysis: an introduction. In A. Bryman and R.G. Burgess (eds.), *Analysing Qualitative Data* (pp1-18). Routledge.

Butler, R. (1988). Enhancing and undermining intrinsic motivation: the effects of task-involving and ego-involving evaluation on interest and performance. *British Journal of Educational Psychology,* 58 (1), 1-14.

Campbell, D. T. (1975). Degrees of freedom and the case study. *Comparative Political Studies,* 8(1), 178-191.

Carless, D. (2006). Differing perceptions in the feedback process, *Studies in Higher Education,* 31(2): 219-33.

Carless, D., Salter, D., Yong, M., and Lam, J. (2011). Developing Sustainable Feedback Practices. *Studies in Higher Education,* 36(4), 395-407.

Chanock, K. (2000). Comments on essays: do students understand what tutors write? *Teaching in Higher Education,* 5, 95-105.

Chen, Y., Hand, B., and Norton-Meier, L. (2017). Teacher Roles of Questioning in Early Elementary Science Classrooms: A Framework Promoting Student Cognitive Complexities in Argumentation. *Research in Science Education,* 47, 373-405.

Cheng, L.Y. and Curtis, A. (2012). Test impact and washback: Implications for teaching and learning. In C. Coombe. P. Davidson, B. O'Sullivan, and S. Stoynoff (eds), Cambridge Gide to Second Language Assessment (pp.89-95). Cambridge: Cambridge University Press.

Chong, P.W. and Graham, L.J. (2013). The 'Russian doll' approach: developing nested case-studies to support international comparative research in education. I*nternational Journal of Research and Method in Education*, 36, (1), 23-32.

Cizek, G.J. More Unintended Consequences of High Stakes Testing, *Educational Measurement Issues and Practice,* 20(4), 19-27.

Cizek, G.J. Fitzgerald, S.M. and Rachor, R.A. (1995). Teachers' Assessment Practices: Preparation, Isolation and the Kitchen Sink. *Educational Assessment, 3(2), 159-179.*

Coe, R. (2002). It's the Effect Size, Stupid: what effect size is and why it is important. Paper presented at the Annual Conference of the British Educational Research Association, University of Exeter, 12-14 September Retrieved from http://www.leeds.ac.uk/educol/documents/00002182.htm

Cohen, J. (1969). *Statistical Power Analysis for the Behavioural Sciences.* New York: Academic Press.

Cohen, L. Manion, L. and Morrison, K. (2017). *Research Methods in Education*. (8th ed). New York: Routledge.

Cooper, P. (1993). Paradigm shifts in designed instruction: From behaviourism to cognitivism to constructivism. *Educ. Technol,* 12-19.

Copp, D.T. (2018). Teaching to the test: a mixed methods study of instructional change from large-scale testing in Canadian schools, *Assessment in Education: Principles, Policy and Practice,* 25(5): 468-487.

Cowie, B. and Bell, B. (1999). A model of formative assessment in science education. *Assessment in Education, Principles, Policy and Practice* 6(1), 101-116.

Creswell, J.W. and Miller, D.L. (2000) Determining validity in qualitative inquiry. *Theory into Practice,* 39 (3), 124-30.

Cresswell, J.W. (1994). *Research design: qualitative and quantitative approaches*. London: Sage.

Crooks, TJ. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research*, 58(4):438–481.

Crotty, M. (1998). *The foundations of social research.* Thousand Oaks, CA: SAGE.

Curren, R. (2006). Connected Learning and the Foundations of Psychometrics: A Rejoinder. *Journal of Philosophy of Education,* 40(1), 17-29.

Daugherty, R., & Ecclestone, K. (2006). Constructing assessment for learning in the UK policy environment. *Assessment and learning*, 149-167.

Davis, A. (1998). *The Limits of Educational Assessment.* Oxford: Blackwell.

Davis, A. (2006). High stakes testing and the structure of the mind: a reply to Randall Curren. *Journal of Philosophy of Education,* 40(1), 1-16.

De Grez, L., Valcke, M., and Roozen, I. (2012). How effective are self-and peer assessment of oral presentation skills compared with teachers' assessments? *Active Learning in Higher Education,* 13(2); 129-142.

Delamont, S. (1976). *Interaction in the Classroom*. London: Methuen.

Denscombe, M. (2017). *The Good Research Guide for Small-Scale Social Research Projects* (6th ed). London: Open University Press.

Denzin, N.K. (1970). *The research act: A theoretical introduction to sociological methods*. New York: Aldine.

DeWitt, S.W., Patterson, N., Blankenship, W., Blevins, B., DiCamillo, L., Gerwin, D., Gradwell, J.M., Gunn, J., Maddox, L., Salinas, C., Saye, J., Stoddard, J., and Sullivan, C.C. (2013). The Lower-Order Expectations of High Stakes Tests: A Four State Analysis of Social Studies Standards and Test Alignment, *Theory and Research in Social Education*, 41(3): 382-427.

DfEE. (1997). *Excellence in schools*, London: HMSO.

Diamond, J. (1996, November 14). The roots of radicalism. *The New York Review of Books,* pp4-6.

Dixson, D.D. and Worrell, F.C. (2016). Formative and Summative Assessment in the Classroom, *Theory into Practice,* 55(2): 153-159.

Drever, E. (1995). *Using Semi-Structured interviews in small-scale research (SCRE publication)*. Edinburgh: Scottish Council for Research in Education.

Dunn, K.E., and Mulvenon, S.W. (2009). A critical review of research on formative assessment: the limited scientific evidence of the impact of formative assessment in education. *Practical Assessment, Research and Evaluation,* 14 (7), 1-11).

Eduqas (2019). *WJEC Eduqas GCE A Level in Geography Specification.* Retrieved from: https://www.eduqas.co.uk/qualifications/geography/as-a-level/a-level-geography-spec-from-2016-eduqas.pdf

Eduqas (2019). *WJEC Eduqas GCE A Level in Media Studies Specification.* Retrieved from: https://www.eduqas.co.uk/qualifications/media-studies/as-a-level/Eduqas%20A-Level%20Media%20Studies%20Specification%20(From%202017).pdf?language_id=1

Eduqas (2019). *WJEC Eduqas GCE A Level in Religious Studies Specification.* Retrieved from: https://www.eduqas.co.uk/qualifications/religious-studies/as-a-level/eduqas-a-level-religious-studies-spec-from-2016-e%20200219.pdf?language-id=1

Eduqas (2017). *Sample Assessment Materials.* Retrieved from: https://www.eduqas.co.uk/qualifications/geography/as-a-level/WJEC-Eduqas-A-Level-Geography-SAMs.pdf

Eisenhart, M. (2017). A matter of scale: multi-scale ethnographic research on education in the United States. *Ethnography and Education*, 12 (2), 134-147.

Eisner, E.W. (1991). The Enlightened Eye. New York: Macmillan.

Elkind, D. (2004). The Problem with Constructivism. *The Educational Forum,* 68(4), 306-312.

Elliot, R. Fischer, C.T. and Rennie, D.L. (1999). Evolving guidelines for publication of qualitative research studies in Psychology and related fields, *British Journal of Clinical Psychology,* 38: 215-29.

Elliott, J. and Lukes, D. (2008). Epistemology as ethics in research and policy: the use of case studies. *Journal of Philosophy and Education,* 4 (S1), 87-119.

Evans, C. (2013). Making sense of assessment feedback in Higher Education. *Review of Educational Research,* 83(1), 70-120.

Eysenck, H. J. (1976). Introduction. In H. J. Eysenck (Ed.), *Case studies in behaviour therapy*. (pp. 1-15). London: Routledge.

Falchikov, N. (1986). Product Comparisons and Process Benefits of Collaborative Peer Group and Self-Assessments. *Assessment and Evaluation in Higher Education,* 11(2) 146-166.

Flick, U. (2009). *An Introduction to Qualitative Research* (4th ed). London: Sage Publications.

Flick, U. (2018). *Managing Quality in Qualitative Research.* London: SAGE Publications.

Flyvbejerg, B. (2006). Five Misunderstandings About Case Study Research. *Qualitative Inquiry, 12*(2),
219-245.

Fontana, D., and Fernandes, M. (1994). Improvements in math performance as a consequence of self-assessment in Portuguese primary school pupils. *British Journal of Educational Psychology,* 64, 407-417.

Foster, P. Gomm, R. and Hammersley, M. (eds). (2000) *Case study method: key issues, key texts*. London: SAGE.

Frederiksen, J.R. and Collins, A. (1989). A systems approach to educational testing. *Educational Researcher,* 18 (9), 27-32.

Fuchs, L. S., and Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis. *Exceptional Children, 53, 199-208.*

Fullan, M., and Watson, N. (2000) School-based management: Reconceptualising to improve learning outcomes. *School Effectiveness and School Improvement,* 11, 453-473.

Gans, H.J. (1968). The Participant-Observer as a Human Being: Observations on the Personal Aspects of Fieldwork. In R.G. Burgess (Ed). *Field Research: A Sourcebook and Field Manual.* (pp 53-61). London: Allen and Unwin.

Gardner, J. (2006). Assessment and Learning. London: Sage.

Geertz, C. (1995). *After the fact: Two countries, four decades, one anthropologist.* Cambridge, MA: Harvard University Press.

Gibbs, A. (2012). Focus Groups and Group Interviews in J. Arthur, M. Waring, R. Coe, and L.V. Hedges (Eds) *Research Methods and Methodologies in Education.* (pp.186-191). Thousand Oaks, CA: Sage.

Gioka, O. (2009). Teacher or examiner? The Tensions between Formative and Summative Assessment in the Case of Science Coursework. *Research in Science Education,* 39, 411-428.

Gipps, C.V. (1996). Assessment for Learning. In A. Little, and A. Wolf (Eds.), *Assessment in Transition: learning, monitoring, and selection in international perspective.* Oxford: Pergamon Press.

Glass, G.V. McGaw, B. and Smith, M.L. (1981). *Meta-Analysis in Social Research.* London: Sage.

Goldstein, H. and Leckie, G. (2008). School league tables: what can they really tell us? *Significance,* 5, 67-69.

Gov.UK (2018). Get the facts: AS and A Level reform. Retrieved from: https://www.gov.uk/government/publications/get-the-facts-gcse-and-a-level-reform/get-the-facts-as-and-a-level-reform

Green, A. (2013). Washback in Language Assessment. *International Journal of English Studies,* 13(2), 39-51.

Groeben, N. (1990). Subjective theories and the explanation of human action in G.R. Semin, and K.J. Gergen. (Eds) *Everyday Understanding: Social and Scientific Implications.* (pp. 19-44). London: Sage, 19-44.

Guba, E.G. and Lincoln, Y.S. (1989). *Fourth Generation Evaluation*. Newbury Park, CA: Sage.

Guillemin, M. and Gillam, L. (2004). Ethics, Reflexivity and "Ethically Important Moments" in Research. *Qualitative Inquiry*, 10 (2), 261-280.

Guskey, T.R. (2005). Formative Classroom Assessment and Benjamin Bloom: Research, Theory, and Implications. The Annual Meeting of the American Educational Research Association, Montreal, April 11-15.

Hackmann, D.G. (2004). Constructivism and block scheduling: Making the connection. *Phi Delta Kappan,* 85, 697-703.

Haggerty, K. (2004). Ethics creep: Governing social science research in the name of ethics. *Qualitative Sociology,* 27(4), 391-414.

Halcomb, E.J., Gholizadeh, L., DiGiacomo, M., Phillips, J., and Davidson, P.M. (2007). Literature Review: Considerations in undertaking focus group research with culturally and linguistically diverse groups. *Journal of Clinical Nursing,* 16, 1000-1011.

Halliday, J. (2010). Educational Assessment in R. Bailey, R. Barrow, D. Carr and C. McCarthy. (Eds) *The SAGE Handbook of Philosophy of Education* (pp.369-385) London: Sage Publications.

Hall, K. Webber, B., Varley, S., Young, V. and Dorman, P. (1997). A study of teacher assessment at Key Stage 1. *Cambridge Journal of Education,* 27, 107-122.

Hamilton, L., Corbett-Whittier, C., and Fowler, Z. (2012). *Using Case Study in Education Research.* London: SAGE.

Hammersley, M. and Atkinson, P. (2007). *Ethnography: Principles in Practice* (3rd ed). New York: Routledge.

Han, C., and Fan, Q. (2020). Using self-assessment as a formative assessment tool in an English Chinese interpreting course: student views and perceptions of its utility, *Perspectives,* 28(1); 109-125.

Hanrahan, S.J. and Isaacs, G. (2001). Assessing self and peer assessment: The students' views. *Higher Education Research and Development,* 20(1), 53-70.

Harlen, W. (2005). Teachers' summative practices and assessment for learning – tensions and synergies. *Curriculum Journal*, *16*(2), 207–223.

Harlen, W. (2007). The quality of learning: Assessment alternatives for primary education. Primary Review Research Survey 3 / 4, Cambridge, Faculty of Education, University of Cambridge.

Harlen, W. (2007). *Assessment of learning*. United Kingdom: Sage Publications.

Hattie, J. and Jaeger, R. (1998). Assessment and classroom learning: A deductive approach. *Assessment in Education: Principles, Policy and Practice,* 5(1), 111-122.

Hembree, R. (1988). Correlates, causes effects and treatment of test anxiety. *Review of Educational Research,* 58(1), 47-77.

Hennink, M.M. (2014). *Focus group discussions.* New York: Oxford University Press.

Heritage, M., and Wylie, C. (2019). Teacher Preparation in Mathematics in H. Andrade., R.E. Bennett and G.J. Cizek (Eds) *Handbook of Formative Assessment in the Disciplines.* Routledge.

Higgins, R., Hartley, P. and Skelton, A. (2001). Getting the message across: The problem of communicating assessment feedback. *Teaching in Higher Education,* 6(2), 269-74.

Hong, W.P., and Youngs, P. (2008). Does high stakes testing increase cultural capital among low income and racial minority students? *Education Policy Analysis Archives,* 16(6).

Hopkins, D. (2008). *A Teacher's Guide to Classroom Research.* McGraw-Hill Education.

Hughes, J. (2005). Ethical Cleansing? The process of gaining "ethical approval" for a new research project exploring performance in place of war. *Research in Drama Education,* 10(2), 229-32.

Hyland, P. (2000). Learning from feedback on assessment. In A. Booth and P. Hyland (Eds). *The Practice of University History Teaching*. Manchester: Manchester University Press.

Illeris, K. (2007). *How we learn: Learning and non-learning in school and beyond.* London and New York: Routledge.

Ingram, J., Elliott, V., Morin, C., Randhawa, A., and Brown, C. (2018). Playing the system: incentives to 'games' and educational ethics in school examination entry policies in England. *Oxford Review of Education,* 44(5), 545-562.

James, M., & Pedder, D. (2006). Beyond method: Assessment and learning practices and values. *Curriculum Journal*, *17*(2), 109–138.

Jarvis, P., Holford, J., and Griffin, C. (2003). *The theory and practice of learning*. London: Kogan Page.

Jiang, Y. (2014) Exploring Teacher Questioning as a Formative Assessment Strategy. *RELC Journal,* 45 (3), 287-304.

Kellaghan, T., Madaus, G.F. & Airasian, P.W. (1982). *The Effects of Standardised Testing*. Boston: Kluwer-Nijhoff.

Kellaghan, T., Madaus, G.F. and Raczek, A. (1996). The use of external examinations to improve student motivation (Washington DC, American Educational Research Association).

Kearney, S. (2013). Improving engagement: the use of 'Authentic self-and peer assessment for learning' to enhance the student learning experience. *Assessment and Evaluation in Higher Education,* 38(7); 875-891.

Kitzinger, J. (1995). Qualitative Research: Introducing focus groups. *British Medical Journal,* 311, 299-302.

Kitzinger, J. and Barbour, R. (1999). *Developing Focus Group Research.* London: SAGE.

Krueger, R. (1993). Quality control in focus group research. In D.L. Morgan (Ed). *Successful focus groups: Advancing the state of the art.* (pp 65-88). London: Sage, 65-88.

Kuhn, T. S. (1987). What are scientific revolutions? In L. Kruger, L. J. Daston, & M. Heidelberger (Eds.), *The probabilistic revolution, Vol. 1: Ideas in history* (pp. 7-22). Cambridge, MA: MIT Press.

Lambert, D. and Lines, D. (2000). *Understanding Assessment: Purposes, Perceptions and Practice.* New York: Routledge Falmer.

Lattimore, R. (2001). The wrath of high-stakes tests. *The Urban Review,* 33(1).

Lee, T.W. (1999). *Using Qualitative Methods in Organizational Research*. London: Thousand Oaks.

Leckie, G. and Goldstein, H. (2009). The limitations of using school league tables to inform school choice. *Journal of the Royal Statistical Society,* 172, Part A, pp,835-851.

Leckie, G., and Goldstein, H. (2017). The evolution of school league tables in England 1992-2016: 'Contextual value-added', 'expected progress', and progress 8'. *British Educational Research Journal,* 43(2), 193-212.

Lincoln, Y.S. and Guba, E.G. (1985). *Naturalistic Inquiry.* London: Sage Publications.

Linn, R.L. (2000). Assessments and accountability. *Educational Researcher,* 29(2), 4-16.

Lomax, R., West, M., Harmon, M., Viator, K., & Madaus, G. (1995). The Impact of Mandated Standardized Testing on Minority Students. *The Journal of Negro Education, 64*(2), 171-185.

Looney, A., Cumming, J., van Der Kleij, F., and Harris, K. (2018). Reconceptualising the role of teachers as assessors: teacher assessment identity, *Assessment in Education: Principles, Policy and Practice,* 25(5), 442-467.

Madaus, F.G., & Kellaghan, T. (1992). Curriculum Evaluation and Assessment. In P.W. Jackson (Ed.), *Handbook of Research on Curriculum.* New York: Macmillan Publishing Company.

Madaus, G. and Russell, M. (2010/2011). Paradoxes of high stakes testing. *Journal of Education,* 190 (1/2), 21-30.

Maguire, M., and Delahunt, B. (2017). Doing a Thematic Analysis: A Practical, Step-by-Step Guide for Learning and Teaching Scholars. *All Ireland Journal of Teaching and Learning in Higher Education.* 3, 3351-33514.

Man Sze Lau, A. (2016). 'Formative good, summative bad?'- A review of the dichotomy in assessment literature. *Journal of Further and Higher Education,* 40 (4), 509-525.

Maxwell, J.A. (1992). Understanding validity in qualitative research. *Harvard Educational Review,* 62: 279-300.

McDonald, B., and Boud, D. (2003). The impact of self-assessment on achievement: The effects of self-assessment training on performance in external examinations. *Assessment in Education,* 10(2), 209-220.

Mehan, H. (1979) 'What time is it Denise?': Asking known information questions in classroom discourse. *Theory into Practice,* (4), 285-294.

Meisels, S. J., Atkins-Burnett, S., Xue, Y., Nicholson, J., Bickel, D. D., and Son, S-H. (2003). Creating a system of accountability: The impact of instructional assessment on elementary children's achievement test scores, *Education Policy Analysis Archives*, 11(9). Retrieved [13/09/2018] from https://epaa.asu.edu/ojs/article/viewFile/237/363

Merriam, S. (1988). *Case Study Research in Education: A Qualitative Approach.* San Francisco: Jossey-Bass.

Merriam, S. (1998). *Qualitative research and case study applications in education*. San Francisco: Jossey-Bass.

Merriam, S. and Tisdell, E.J. (2015). *Qualitative Research: A Guide to Design and Implementation.* (4th Ed). San Francisco: Jossey-Bass.

Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13(3), 241-256.

Miles, M.B. and Huberman, A.M. (1994). *Qualitative Data Analysis: An Expanded Sourcebook*, (2nd ed). Thousand Oaks, CA: SAGE.

Miller, W. and Crabtree, B. (1999). *Doing Qualitative Research* (2nd ed). London: Sage.

Mitford College (no date). Institution Entry Requirements. North East: Mitford College.

Moon, T.R., Brighton, C.M., Jarvis, J.M., and Hall, C.J. (2007). State Standardized Testing Programs: Their Effects on Teachers and Students. National Research Centre on the Gifted and Talented, University of Connecticut.

Morgan, D.L. (1988). *Focus groups as qualitative research*. London: SAGE.

Morrow, K. (1986). The evaluation of tests of communicative performance. In M. Portal (Ed.) *Innovations in language testing.* London: NFER Nelson.

Morse, J.M. (1994). Design in funded qualitative research. In N. Denzin and Y.S. Lincoln (eds) *Handbook of Qualitative Research* (pp 220-35). Thousand Oaks, CA: Sage.

Mowl, G. and Pain, R. (1995). Using self and peer assessment to improve students' essay writing: a case study from Geography. *Innovations in Education and Training International,* 32 (4), 324-335.

Müller, U., Carpendale, J., and Smith, L. (Eds). (2009). *The Cambridge Companion to Piaget.* Cambridge: Cambridge University Press.

Naylor, S. and Keogh, B. (1999). Constructivism in Classroom: Theory into Practice. *Journal of Science Teacher Education,* 10 (2), 93-106.

Newton, P.E. (2007). Clarifying the purpose of educational assessment. *Assessment in Education,* 14(2), 149-170.

Nichol, D.J. and Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education,* 31(2), 199-218.

Nisbit, J. and Watt, J. (1984). Case Study. In J. Bell, T. Bush, A. Fox, J. Goodey and S. Goulding (eds) *Conducting Small Scale Investigations in Educational Management* (pp79-92). London: Harper and Row.

Nitko, A. J. (1995). Curriculum-based continuous assessment: A framework for concepts, procedures and policy. *Assessment in Education: Principles, Policy & Practice*, *2*(3), 321–337.

Nye, P. and Thomson, D. (2019). *A Level results 2019: The main trends in grades and entries*. Retrieved from: https://ffteducationdatalab.org.uk/2019/08/a-level-results-2019-the-main-trends-in-grades-and-entries/

Oancea, A. and Pring, R. (2008). The importance of being thorough: on systematic accumulations of "what works" in education research. *Journal of Philosophy of Education,* 4(s1), 4-15.

OCR (2018). *A Level Specification Film Studies.* Retrieved from: https://www.ocr.org.uk/Images/316666-specification-accredited-a-level-gce-film-studies-h410.pdf

Ofqual (2019). *Guide to AS and A Level results for England 2019.* Retrieved from: https://www.gov.uk/government/news/guide-to-as-and-a-level-results-for-england-2019

Oliver, P. (2010). *Student's Guide to Research Ethics.* (2nd Ed). Open University Press: Berkshire.

Olsen, V. and Whittaker, E. (1968). *The Silent Dialogue: A Study in the Social Psychology of Professional Socialization*. San Francisco, CA: Jossey-Bass.

Panadero, E., Brown, G.L., and Stribos, J.W. (2016a). The future of student self-assessment: a review of known unknowns and potential directions, *Educ. Psychol. Rev.* 28: 803-830.

Patton, M.Q. (1990). *Qualitative Evaluation and Research Methods* (2nd ed). London: Sage.

Patton, M.Q. (2015). Qualitative Research and Evaluation Methods (4th ed). Thousand Oaks, CA: Sage Publications Ltd.

Pearson (2019). *Centre Register and Learner Authentication and Record Form.* Retrieved

from: https://qualifications.pearson.com/en/qualifications/btec-nationals/applied-law-2017.coursematerials.html#filterQuery=pearson-UK:category%2FExternal-assessments

Pearson (2017). *BTEC Level 3 National Certificate in Applied Law Specification.* Retrieved from: https://qualifications.pearson.com/content/dam/pdf/BTEC-Nationals/Applied-Law/2017/Specification/9781446934135_BTEC_L3_NAT_CERT_APPLAW_SPEC_Iss1_200616.pdf

Pearson (2014). *A Level Business Sample Assessment Materials.* Retrieved from: https://qualifications.pearson.com/content/dam/pdf/A%20Level/Business%20Studies/2015/specification-and-sample-assessment-materials/Business_A_SAM_Collation.pdf

Pearson, M.L., Albon, S.P., and Hubball, H. (2015). Case Study Methodology: Flexibility, Rigour and Ethical Considerations for the Scholarship of Teaching and Learning. *The Canadian Journal for the Scholarship of Teaching and Learning,* 6(3).

Perie, M., Marion, S., and Gong, B. (2007). The role of interim assessments in a comprehensive assessment system: A policy brief. Retrieved 12th September 2018 from https://www.nciea.org/library/role-interim-assessments-comprehensive-assessment-system-policy-brief

Perrenoud, P. (1991). Towards a Pragmatic Approach to Formative Evaluation. In P. Weston (Ed.). *Assessment of Pupils' Achievement: Motivation and School Success.* Amsterdam, Swets and Zeitlinger.

Polesel, J., Rice, S., and Dulfer, N. (2014). The impact of high stakes testing on curriculum and pedagogy: a teacher perspective from Australia, *Journal of Education Policy,* 29(5); 640-657.

Pollard, A. (1987). *Children and Their Primary Schools: A New Perspective, Vol 2*. London: Falmer Press.

Poland, B.P. (2002). Transcription quality. In J.F. Gubrium and J.A. Holstein (Eds). *Handbook of Interview Research: Context and Method.* (pp 629-50). Thousand Oaks, CA: Sage.

Poulos, A., and Mahony, M.J. (2008). Effectiveness of feedback: The student's perspective. *Assessment and Evaluation in Higher Education.* 33(2), 143-54.

Prawat, R., and Floden, R. (1994). Philosophical perspectives on constructivist views of

learning. *educ. Psychol.* 29(1), 37-48.

Pring, R. (2015). *Philosophy of educational research.* (3rd Ed). London: Bloomsbury.

Ragin, C. C. (1992). "Casing" and the process of social inquiry. In C. C. Ragin & H. S. Becker (Eds.), *What is a case? Exploring the foundations of social inquiry* (pp. 217-226). Cambridge, UK: Cambridge University Press.

Ramaprasad, A. (1983). On the definition of feedback. *Behavioral Science,* 28(1), 4-13.

Rapley, C. (2007). *Doing conversation. Discourse and document analysis*. London: Sage publications.

Reicher, S. (2000). Against methodolatry: some comments on Elliot, Fischer and Rennie, *British Journal of Clinical Psychology,* 39: 11-26.

Remesal, A. (2007). Educational reform and primary and secondary teachers' conceptions of assessment: The Spanish instance, building upon Black and Wiliam (2005). *The Curriculum Journal*, 18, 27–38.

Resnik, D.B. (2010). What is Ethics in Research and Why is it Important? Research Triangle Park, NC: National Institute of Environmental Health Sciences. Retrieved from https://www.niehs.nih.gov/research/resources/bioethics/whatis/index.cfm

Revision World (2019). *Connect 4.* Retrieved from: https://revisionworld.com/games/connect4

Roach, C.M. (2014). Going native: Aca-fandom and deep participant observation in popular romance studies. *Mosaic,* 47(2), 33-49.

Rodriguez, M.C. (2004). The role of classroom assessment in student performance on TIMSS. *Applied Measurement in Education,* 17 (1), 1-24.

Ross, J. A. (2006). The reliability, validity, and utility of self-assessment. *Practical Assessment, Research and Evaluation,* 11(10), 1-13.

Rowley, J. (2002). Using case studies in research. *Management Research News, 25*(1), 16-27.

Ruddin, L.P. (2006). You can generalise stupid! Social scientists, Bent Flyvbjerg and case study methodology. *Qualitative inquiry,* 12(4), 797-812.

Ruiz-Primo, M.A. and Furtak, E.M. (2006). Informal formative assessment and scientific enquiry: Explaining teachers' practices and student learning. *Educational Assessment,* 11(3-4), 205-235.

Ruiz-Primo, M.A. and Furtak, E.M. (2007). Exploring teachers informal formative assessment practices and students' understanding in the context of scientific inquiry. *Journal of Research in Science Teaching* 44 (1) 57-84.

Rushton, A. (2005). Formative assessment: a key to deep learning? *Medical Teacher,* 27 (6), 509-513.

Sadler, D.R. (1989). Formative assessment and the design of instructional systems. *Instructional Science,* 18, 119-144.

Saldaña, J. (2011). *Fundamentals of Qualitative Research.* New York: Oxford Univesity Press.

Saldaña, J. (2013). *The Coding Manual for Qualitative Researchers.* (2nd Ed). Thousand Oaks, CA: Sage.

Schatzman, L. and Strauss, A. (1973). *Field Research: Strategies for a Natural Sociology*, Englewood Cliffs, NJ: Prentice Hall.

Scheurman, G. (1995). Constructivism, Personal Epistemology and Teacher Education: Toward a Social-Developmental Model of Adult Reasoning. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA.

Schiltz, F., Sestito, P., Agasisti, T., and De Witte, K. (2018). The added value of more accurate predictions for school rankings. *Economics of Education Review,* 67, 207-215.

Scriven, M. (1967). The methodology of evaluation. In R. W. Tyler, R. M. Gagne and M. Scriven (Eds), *Perspectives of curriculum evaluation, Volume 1 (*pp. 39-83). Chicago, IL: Rand MCNally.

Seale, C. (1999). *The Quality of Qualitative Research.* London: Sage.

Searle, J.R. (1995). *The construction of social reality.* Simon and Schuster.

Sebatane, E. M. (1998). Assessment and classroom learning: A response to Black & Wiliam. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 123–130.

Sener, J. (1997). C*onstructivism: Asynchronous Learning Networks*. *ALN Magazine.*

Sharma, R., Amit, J., Gupta, N., Garg, S., Batta, M., and Kant Dhir. (2016). Impact of Self-Assessment by Students on their Learning. *International Journal of Applied and Basic Medical Research,* 6(3), 226-229.

Shavelson, R.J., Black, P.J., Wiliam, D., and Coffey, J. (2004). On Linking Formative and Summative Functions in the Design of Large-scale Assessment Systems. Retrieved from: http://www.stanford.edu/dept/SUSE/SEAL/Reports_Papers/On%20Aligning%20Formative%20and%20Summative%20Functions_Submit.doc

Shepard, L.A. (2000). The Role of Assessment in a Learning Culture. *Educational Researcher,* 29(7) 4-14.

Shimahara, N. (1988). Anthroethnography: A Methodological Consideration in R. R. Sherman, and R.B. Webb (Eds.). *Qualitative Research in Education: Focus and Methods.* London, New York, Philadelphia: The Falmer Press.

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, *78*(1), 153–189.

Silverman, D. (2011). Interpreting Qualitative Data: Methods for Analysing Talk, Text and Interaction. (4th edn). Thousand Oaks, CA: Sage.

Simpson, M. (1990). Why criterion-referenced assessment is unlikely to improve learning. *The Curriculum Journal,* 1, pp.171-183.

Simpson, M. and Tuson, J. (1995). *Using Observations in Small-Scale Research: A Beginners Guide*. SCRE: Glasgow.

Sly, L. (1999). Practice tests as formative assessment improve student performance on computer-managed learning assessments. *Assessment and Evaluation in Higher Education,* 24(3), 339-343.

Smith, A. (2004). Making mathematics count: The report of Professor Adrian Smith's inquiry into post-14 mathematics education. London: The Stationary Office.

Smith, L. (1978). An evolving logic of participant observation, educational ethnography and other case studies. *Review of Research in Education,* 6 (1), 316-377.

Smith, L., Hill, M., Cowie, B., and Gilmore, A. (2014). Preparing Teachers to Use the Enabling Power of Assessment in C. Wyatt-Smith., V. Klenowski and P. Colbert (Eds). *Designing Assessment for Quality Learning. The Enabling Power of Assessment,* Vol 1. Springer, Dordrecht.

Smith, M.L. and Rottenberg, C. (1991). Unintended Consequences of External Testing in Elementary Schools. *Educational Measurement: Issues and Practice,* (10) 4.

SSAT. (2019). SSAT's Embedding Formative Assessment Programme. Retrieved from: https://www.ssatuk.co.uk/cpd/teaching-and-learning/embedding-formative-assessment/

Stake, R. (1995). *The art of case study research.* Thousand Oaks, CA: Sage.

Stenhouse, L. (1978). Case Study and case records: towards a contemporary history of education, *British Educational Research Journal,* (2): 21-39.

Stenhouse, L. (1979). Case study in comparative education: particularity and Generalization, *Comparative Education,* 15, 5-10.

Stiggins, R.J., Miller-Griswold, M. and Reed Wikelund, K. (1989) Measuring Thinking Skills Through Classroom Assessment. *Journal of Educational Measurement* 26(3), 233-246.

Stiggins, R.J. (1991). Assessment Literacy. *Phi Delta Kappa International,* 72 (7), 534-539.

Stiggins, R.J. (1999). Assessment, Student Confidence and School Success. *Phi Delta Kappan,* 81 (3), 191-8.

Stiggins, R. (2006). Assessment for Learning: A Key to Motivation and Achievement. *Edge,* 2(2), 3-19.

Stigler, J.W., & Hiebert, J. (1997). Understanding and Improving Classroom Mathematics Instruction: An Overview of the TIMSS Video Study. *Phi Delta Kappan,* 79 (1), 14-21.

Stobart, G. (2006). Influencing classroom assessment. *Assessment in Education: Principles, Policy & Practice*, *13*(3), 235–238.

Stobart, G. and Gipps, C.V. (1997). *Assessment: a teacher's guide to the issues*. Hodder and Stoughton.

Sun, J., Przybylski, R., and Johnson, B.J. (2016) A review of research on teachers' use of student data: from the perspective of school leadership. *Educational Assessment, Evaluation and Accountability,* 28(1), 5-33.

Swain, J. (2006). An Ethnographic approach to researching children in junior school. *International Journal of Social Research Methodology*, *9*(3), 199–213.

Sweiry, E. Crisp, V. Ahmed, A. and Pollitt, A. (2002). 'Tales of the Expected: The influence of students' expectations on exam validity', *British Educational Research Association Conference,* Exeter, September. Available at: www.cambridgeassessment.org.uk/images/109695-tales-of-the-expected-the-influence-of-students-expectations-on-exam-validity.pdf (Accessed: 20th March 2019).

Tan, Z. (2007). Questioning in Chinese University EL Classrooms: What lies beyond it? *RELC Journal,* 38 (1), 87-103.

Taras, M. (2005). Assessment-Summative and Formative- Some Theoretical Reflections. *British Journal of Educational Studies.* 53(4), 466-478.

Teachers Direct (2005). *Quiz Busters*. Retrieved from: www.teachers-direct.co.uk/resources/quiz-busters/subjects/ks5/health-and-social-care.aspx

Thomas, G. and Myers, K. (2015). *The Anatomy of Case Study*. London: Sage Publications.

Tight, M. (2010). The curious case of case study: A viewpoint. *International Journal of Social Research Methodology*, 13 (4) 329-339.

Tight, M. (2019). *Documentary Research in the Social Sciences*. London: SAGE Publications.

Tiknaz, Y. and Sutton, A. (2006). Exploring the role of assessment tests to promote formative assessment in Key Stage 3 Geography: evidence from twelve teachers. *Assessment in Education: Principles, Policy and Practice,* 13(3), 327-343.

Van Maaen, J. and Kolb, D. (1985). The professional apprentice: observations on fieldwork roles in two organizational settings. In Bachrach, S. and Mitchell, S. (eds) *Research in the sociology of organisations* (pp.1-33) Greenwich, CT: JAI.

Vlachou, M.A. (2015). Does Assessment for Learning Work to Promote Student Learning? The England Paradigm, *The Clearing House: A Journal of Educational Strategies, Issues and Ideas,* 88 (3), 101-107.

Verschuren, P. (2003). Case study as a research strategy: Some ambiguities and opportunities. I*nternational Journal of Social Research Methodology,* 6(2), 121-139.

Vygotsky, L.S. (1978). Mind in Society: The Development of Higher Psychological Process. Cambridge, MA: Harvard University Press.

Walford, G. (2000, 11-12 September). The over-use of interviews in ethnographic research. Paper given at the Ethnography and Education Conference, Oxford.

Walsh, J.A. and Sattes, B. D. (2005). *Quality questioning: research-based practice to engage every learner.* Thousand Oaks, CA: Sage.

Walsh, J. A., & Sattes, B. D. (2005). How can quality questioning transform classrooms? questioning to advance thinking, learning, and achievement. *Quality questioning: research-based practice to engage every learner. Thousand Oaks (CA): Sage Publications*.

West, A. (2010). High stakes testing, accountability, incentives and consequences in English Schools. *Policy and Politics,* 38(1), 23-39.

What Uni? (2019). The New Linear A Levels Explained. Retrieved from: https://www.whatuni.com/advice/guides/the-new-linear-a-levels-explained/64354/

Wiliam, D. (2011). *Embedded Formative Assessment.* Bloomington IN: Solution Tree Press.

Wilson, E. (2009). *School-based Research- A guide for education students.* London: Sage Publications Ltd.

Wininger, R. S. (2005). Using your tests to teach: Formative summative assessment. *Teaching Psychology,* 32(2), 164-166.

Winterbottom, M. Brindley, S. Taber, L. Fisher, Finney, J. and Riga, F. (2008). Conceptions of assessment: trainee teachers' practice and values. *The Curriculum Journal,* 19(3), 193-213.

WJEC (2017). *WJEC Level 3 Applied Certificate and Diploma in Criminology.* Retrieved from: https://www.wjec.co.uk/qualifications/criminology/Criminology%202015%20spec.pdf?language_id=1

WJEC (2017). *Media Studies: Media Representations and Responses Examination Paper.* Retrieved from: http://pastpapers.download.wjec.co.uk/s17-1291-01a.pdf

Wylie, E. C., & Lyon, C. J. (2015). The fidelity of formative assessment implementation: Issues of breadth and quality. *Assessment in Education: Principles, Policy & Practice*, *22*(1), 140–160.

Yazan, B. (2015). Three Approaches to Case Study Methods in Education: Yin, Merriam and Stake. *The Qualitative Report*, 20 (2), 134-152.

Yin, R.K. (1989). *Case Study Research Design and Methods.* London: Newbury Park.

Yin, R. K. (2002). C*ase Study Research: Design and methods*. Thousand Oaks, CA: Sage Publications.

Yin, R. (2003). *Case study research: Design and methods* (3rd ed.). Thousand Oaks, CA: Sage.

Yin, R. (2009). *Case Study Research: Design and Methods* (4th ed). Thousand Oaks, CA: Sage.

Yin, R. (2014). *Case Study Research: Design and Methods* (5th ed). London: SAGE publications.

Appendix 1 - Subject Labels

Subject	**Label**
Accounting	ACC
Fine Art	ART
Biology	BIO
Business Studies	BUS
Chemistry	CHEM
Classics	CLA
Computing	COMP
Criminology	CRI
English Language	ENG
English Literature	LIT
Film Studies	FILM
French	FR
Geography	GEO
Health and Social Care	HSC
History	HIS
Law	LAW
Maths	MA
Media Studies	MED
Physical Education	PE
Physics	PHY
Religious Studies	RS
Sociology	SOC

Jenny Peadon

[redacted]

[redacted]

[redacted]

Dear Mrs [redacted]

My name is Jenny Peadon; I was previously employed in your institution in the Geography Department, having left following [redacted] I have subsequently worked in two other secondary schools, before taking the decision in October 2016 to leave the teaching profession. I now work at Sunderland University as a Disability Adviser. In addition, I am completing a Doctorate in Education on a part time basis at Durham University.

Completion of my Doctorate is centred around the design, implementation and subsequent analysis of a piece of independent research based around an area of interest. I have chosen to focus my study on how teachers operationalise formative and summative assessment with students in contrasting year groups, in the face of a climate of high stakes summative testing in the UK education system.

I intend to draw upon three sources in order to construct my data: classroom observation, questionnaires, and a small number of interviews. I am writing to enquire whether it would be possible to come into [redacted], to speak with staff, some of whom I worked with during my time with yourselves, and to hopefully conduct some research.

Through classroom observation, I aim to examine how teachers utilise assessment in their classrooms and what form this takes. Questionnaires will then be used to investigate teacher values in relation to assessment with the aim of following up some ideas at interview.

Ethics approval for my research project was granted by the university in September 2017. Should you require further information about my research please do not hesitate to contact me

(jennypeadon@yahoo.co.uk / 07805924939). Alternatively, you could contact my supervisor at Durham University, Dr Jonathan Tummons (jonathan.tummons@durham.ac.uk).

I look forward to hearing from you.

Kind regards

Jenny Peadon

Appendix 3 - Classroom Observation Schedule

Type of assessment observed	Date: Time: Duration: Subject	Teacher or student led?	Notes
Formative Assessment			
Sharing learning objectives			
Questioning			
Peer Assessment			
Self-Assessment			
Pre-assessment strategies to gauge understanding			
Modelling			
Graphic organisers			
Other			
Summative Assessment			
Unseen examination			
Exam style questions			
Open book exams			
Multiple choice tests			
Written essay / report			
Presentation			
Performance task (to assess a specific set of skills)			
Feedback- marks/grades			
Other			

How often do you use formative assessment?

What are the most effective strategies? Why? What is the value of formative assessment?

Do you use peer / self-assessment? Why (not)?

Have you done CPD on formative assessment? Do you utilise this in your day to day teaching?

How often do you use summative assessment?

What are the most effective strategies? Why? What is the value of summative assessment?

Would you utilise different approaches to assessment with girls / boys? Why (not)?

What student perceptions of formative / summative assessment do you encounter?

Do students play an active role in the assessment process?

When providing feedback do you include marks or comments or both? Why?

What are the challenges with formative assessment?

What are the challenges with summative assessment? (classroom based and external exams)

What challenges have resulted from the changes in examination structure? Do you use more formative assessment since these changes occurred?

Has this changed your approach / students' approaches to assessment?

How would you like to improve your use of formative assessment in the future?

Do you place more value on formative or summative assessment?

Themes:
Formative assessment strategies
Summative assessment strategies
Student action / response
Frequency / timescales
Best practice
Challenges
Exam preparation

Year 13 Criminology 2:20-2:50pm 02/05

2nd half of a double lesson. 1 teacher and the change. Short break at the start.

3 weeks until exam.

3 minutes in silence to write down everything you know about the role of the police service.

(without notes / mock- do on whiteboard).

Teacher moving around while students work. '30 seconds left'. 'put pens down'.

1 minute to explain this to your partner- in pairs work out who is the oldest, they go first. 1 minute to talk to the person next to you. If your partner repeats, hesitates, deviates, put your hand up.

Mock paper returned last lesson – marks and grades.

Focus for lesson- Prisons.

'Rub white boards.

1 student selected to give example (now whiteboard has been cleared).

Same task repeated for prisons. Youngest person explaining knowledge to partner.

Feedback.

Choosing students by name.

Some questioning based on what students say.

Encouraging students to develop answer- ‘give me an example’ / ‘expand’ / ‘give me some more’ /

‘what else do I need to hear about for a role question?’

HMPS- ‘you need to write this down’.

Imagine I’m the examiner marking this paper- ‘I need something else’.

Revision for AC3.1, 3.3 and 3.4.

Spider diagrams- prisons.

AC3.1 Role

AC3.3 Limitations

AC3.4 Effectiveness (linking to assessment criteria).

Recreate for all agencies covered- CPS, Prisons, Probation Service.

Themes:
Formative assessment strategies
Summative assessment strategies
Student action / response
Frequency / timescales
Best practice
Challenges
Exam preparation

Focus Group 2 Transcription

How often do you use formative assessment in your classroom?

Probably most lessons I would say. You know, reading over students' shoulders and either orally giving them some feedback or making some sort of quick written feedback that they can use and hopefully respond to pretty much immediately.

We have stages throughout the year because we have a big bulk of coursework that we do right at the beginning of the year so every single lesson you are, as Julie said, standing behind them, you know giving them some feedback and they're kind of making improvements all the way along.

I think it's just embedded in good teaching and learning is it not? Your good lesson structure and the design, all the time you are using formative assessment to identify where they are in terms of the learning, have they grasped that what you are teaching them, and it's always informing where you go next within your lesson.

I'd agreed with that. Most of my lessons are, kind of obviously quite, quite a lot of questioning throughout the lesson and lots of different activities that are obviously based around assessment. We have the end of unit and topic assessments as well that most people have. But lots of different types of assessment within, embedded in the lesson really. I think it's almost constant. Certainly with, we do some controlled assessment and then we have the exam as well, it's trying to switch them to kind of different forms of assessment so it's that's the bigger challenge that we face at the start of kind of January. They are in controlled assessment mind where, maybe not so

much constant assessment, there's obviously the exam, getting them to prepare for the exam, it's all the time, every single lesson.

Have you got any strategies which you think are more effective than others that you would default to?

I would always default for starters to whiteboard, I literally fire down recall type points, so you can see there and then where they are in terms of it. But I think a lot of it is questioning, actually through your lesson and whatever little activity you have set up. My personal fall back is just that effective questioning to try and entice out what they know and where their gaps are. And then depending on your activity you supplement things like talking partners and pair work and bouncing ideas and little tests. But I think the two that I tend to fall back on, is my starter to be a whiteboard type activity, recall and then questioning.

Mini whiteboards?

Yeah

Yeah, we use them quite a lot. I think, something I've started to do this year in terms of trying to look at different forms of assessment is when the assessment is done, so instead of the traditional end of topic assessment, try to stagger that to try to give them that chance to potentially maybe forget something so it's not pure this is some information now tell me it back. It's trying to gear them more towards that exam where you are going to learn something in January and get a question on it in May / June time so I've tried that a little bit to see whether that makes any difference in terms of how they learn but also giving me more information and then obviously that informs teaching in the next lesson is there certain areas that they are struggling to grasp.

How do you test them?

So it's the same type of test, it's exactly the same assessment, it's just that I will teach a unit or a topic and then where traditionally, and I know other teachers in the department will do it, they'll do the end of topic test which is obviously a written test which we mark, give full feedback to. I'll

then go straight onto the next topic, teach one topic, won't assess them, so I'll do the normal formative assessment but I won't have kind of an end of topic written test and then say right that's one topic done let's go to the next one, I'll continue teaching the next topic and wait until we are maybe half way through that topic and say right well here's your end of topic test from the previous topic. So, it's trying to kind of mirror what they are going to have to do for the exam really. They are already thinking about their current topic and then they are getting tested on their previous topic. So, it's still recall. It's still checking prior learning but it's giving them a little bit more of an opportunity not to forget it but to mirror proper, maybe exam conditions where the skill they are going to have to do in the exam where they are not thinking about it there and then, not just being told some things. I think sometimes that gives a false picture, false impression of actually how much they know. Another thing recently, I've, certainly with the exam, ours it's quite soon, it's trying to get them out of the habit of always using that crutch of notes, so you ask them a question and their first thought is, or you give an activity and the first thought is I'll get some notes out to answer the question but I think that comes from the fact that they've had controlled assessment where they, it's an open book, so they can do that. With the exam they can't so it's trying to get them out of that mindset and obviously every time they look at their notes, they are kind of cheating themselves really. A lot of my students will always try to go straight to their notes, so I try to push them to answer it without checking. Even if they get it wrong or even if I have to scaffold it for them and give them a little bit, I'd rather they get there without looking at something they have just got written down on a piece of paper.

Do you ever use peer or self-assessment?

The students, I find are quite, they will do it but they do not value it in the same way so they'll do it but then they will invariably ask if it can be 'properly marked'. And obviously just in terms of workload especially at this time of year when they are doing quite a lot of assessments, they can't all be teacher assessed with the best will in the world. And so, they will fall back on it, but they definitely see it as a fallback position.

Do you think there is a value in it?

I think there is to a degree because it makes them aware of the mark scheme. I think it forces them to use the mark scheme and be aware of all of the different components of it. I think the issues I've had in the past is where they are just a bit too generous! And they will just kind of give their friends a good mark because they do not want to fall out or whatever! So, I think it is kind of teaching them those skills to be able to do it and then you know after a period of time they can do quite effectively. But I do not tend to use it an awful lot. The mark really is a by-product, it's the process that's important and they can get something out of but the mark they actually give is

It's much more about identifying areas for improvement really.

I certainly wouldn't use it to, as a yardstick in terms of how well they are doing because like I say, for those exact reasons you said. I just think they do not want to mark in the way that teachers would mark so they are far too lenient in my experience they do not really mark it properly I would say. I think it's a worthwhile experience for understanding the mark scheme and I try to use that quite a lot in terms of getting them to look at mark schemes, this is what it's asking you for, how do you get to that point? So, I think it's great for that, but I just do not think they see it as a valuable tool.

I did a lot of it, about 3 years ago we did a lot of it and our results weren't great and we actually moved away from it. One of the facts is if they were marking it , then I didn't mark it and actually what I found was I wasn't identifying where their gaps were, if there was a common theme coming out where a group of students weren't grasping aspects of a topic then I was missing out on that information. So, I found it more beneficial the more I read of their work the better it was. But once I've marked it there's then scope for them to do a lot of self-assessment in terms of reflecting on what they have produced in comparison to the mark scheme. I think there's a lot out of that and there's a lot out of looking at student's standard answers, and a good answer that they've produced and peer assessing their answers compared to either a standard or your

students best answer in there to identify where they have gone wrong but what type of structure and how they should be writing it themselves.

Do you spend a lot of time looking at mark schemes?

Yes

Have you done very much CPD on formative assessment and does that influence your day-to-day practice?

I've done a little bit. I think at the minute we're doing CPD where there's different strands, so I know mine personally is on assessment, so we are doing a study if you like, on different assessment techniques and testing a couple of them out. I've found certainly that process of looking at different types of assessment techniques and I'm trying to use some of those in my teaching. Obviously whether they are a success is an ongoing process to try and find that out but it's good to try and find different ideas and obviously testing them in your own teaching. One thing I've looked at is marking, that's what my, kind of study is based around, do you have to mark the full piece of work to actually get the same level of feedback?

I remember you saying. So, have you got a conclusion?

Well, I did it with the last end of topic test that I did, and it was difficult because of the type of assessments that we have to do in the sense that there were two different questions, but they were linked so I marked one question and then I didn't mark the other one, so I gave comments on the first question, the second question was similar, didn't give comments on it. I'm going to compare it with the questions in the mock exam, compare it with how they do with that type of assessment. The question that I didn't mark fully I'm going to see if, obviously I know how they have done on that and I've still got that information and I'm going to see if they have made any major improvements from that final end of topic assessment with the mock. But it wasn't ideal in terms of the type of assessments that we were doing. The idea is kind of based on longer questions, essay type questions which is what we did for the controlled assessment. I think it would be better to try it next year with the controlled assessment type questions because they

are much longer. So, it's a bit like English for example essays I think it would be interesting to see how it would work with that. Whereas unfortunately the type, it would have worked fine if it was earlier in the year but the time of year it was the assessments we had left there were kind of exam based questions and it was two different questions so it was very difficult to kind of mark one question and not the other. And obviously the results won't be massively reliable, but it will give an idea I suppose because other classes, I only did it with one class so it will be interesting comparing it to other classes whether it makes any difference, whether they are much weaker or have got the same type of improvements.

We have done a lot CPD wise in terms of formative assessment; it has underpinned a lot of what has been delivered. There might be a focal group like questioning, challenge but a lot of the people that are leading these groups are putting in the research and a lot of the research is around formative assessment and the effective use of it in questioning or in challenge of, or in differentiation whichever the focus was.

I just stated at this school in September but my last school we had a couple of CPDs where we had a sort of market place activity and all the teachers would bring an assessment tool to the forum and it was just, I found it really, really useful because it was an opportunity to just go round and see how different people were using assessment.

When you are doing mock exams/end of unit tests do you mark with grades/percentages?

Ours is, obviously we give a mark so we give a numerical mark out of, we do not give a grade as such because it's obviously one piece of work so we do not feel it is appropriate to grade it A,B, C or D because it's not a big enough sample to say well that's an A Level piece of work. So, we mark it numerically, like these exam type questions we are doing at the minute, marking numerically but then it's obviously based around feedback so the actual written feedback that we give. So we use a system of numbered improvement points, so highlight a point in the work, then give a numbered improvement point, obviously trying to not just say where they have gone wrong but try to make them develop the answer or improve the answer themselves so I try and get them to

put a little bit more challenge in there. So, if it's a good piece of work and they develop it further, if it's not quite the level required, trying to get them to think about how they could improve it as opposed to you saying, 'this is what you need to do differently'. I still try to put it back onto them to work out what they need to improve.

Are they quite receptive to that feedback?

Yeah, they tend to do the feedback quite well. We tend to get; I think obviously because of the numbered improvement points they know quite clearly what's expected of them. They've got maybe 6, 7, 8 points they need to address. Some of them will obviously put more detail in than others and obviously we check it again, give it back to them if they do not do it well enough, but I think the numbered system helps because it does make it quite clear that they have to address every single improvement point.

We do something very similar in terms of obviously we'll mark it, but we do not give a grade even if it's an end of unit test. They find it difficult I think in Year 12, it's a change in mindset from GCSE where everything was graded and that's what they like. Sometimes, they will often go 'what's that Sir?' 'Is that an A?' and trying to shift their mindset and actually all I want them to identify is what they do not know yet. And that's where we do a lot of work in terms of the mark scheme and the feedback process and setting them tasks based on what we've identified as being their gaps in knowledge in the assessment. But I think it's a learning curve for them in year 12. Ok they are used to it by the end of Year 13, but it is a little shift in mindset in terms of actually we are not interested in giving you a grade, just identifying what it is you do not know. This is what I want you to do about it.

I agree with that. My students will always say what grade is this? Well, it wouldn't be a grade because it's obviously out of 10 and the exam is out of 75 so it's almost pointless really giving you a grade on it. But they straight away when you give them a 6/10, 7/10 what grade was that? Like I say it doesn't make a difference, it doesn't matter, what's important is what you haven't put in.

Assessments early on aren't anything like A Level standard assessment so, you know our first Year 12 assessment is, they are allowed to have the book, but the ultimate assessment will be a closed book thing. So, it's obviously completely different. Sometimes, especially for the first assessment as well I might even let them have their notes on the table as a sort of crutch because, you know you want them to do, you want them to be able to write and be successful. What they are doing straight away certainly is not A Level standard. But they will still ask for the grade. But you can't always give them it. I'd say I feel like we can give them a grade as we get to this time in the academic year. And also, their target grade remains as a grade, so you know there is a reason behind them having that expectation of wanting to know how they are doing in relation to their target grade. I'd say for the first 6 months you wouldn't really be able to give them anything very useful.

Do you think any approaches are more effective with girls / boys?

I think, I do not know whether it's just my perception or it's actually true or not but I think certainly with the improvements that we do I tend to find that the girls will put a lot more effort into the improvements than the boys will. I think the boys need much more coercing into actually 'this is not good enough', 'you need to put more detail in'. Whether that's just because the girls take a little more pride in their work and if there's something, they haven't done properly they are desperate to make sure it is done properly or not I do not know. Like I say it might just be my perception, it might not actually be true; I've got no data to back that up. It's just a feeling that I get.

I only teach, well I do not only teach girls, but I must teach 90% girls I would say, and I do not find that there's any differences, you know there's conscientious students and there's non conscientious students, it does not seem to make much difference. The boys do not, all 5 of them in Year 12, they do not all tend to do any one particular thing well or not well I do not think.

I find the boys need more sort of scaffolding you know they need to be prepared more for assessments.

There are always the exceptions. That's quite a big generalisation. I think it is a lot of my perception, but I teach a fair few boys, but I teach mainly girls, maybe it's just the type of cohort that I've got. Some of the boys aren't as strong maybe but yeah there might be a slight difference but probably not a huge amount.

No there's nothing obvious I do not think that jumps out that says girls are better at this, boys are better at that. Yes, you've got your generalisations, girls files are neater, they are better prepared and therefore when they do the response to the feedback, they take more care in it.

That's what we found with our controlled assessment because that is very much a case of, a lot of it is organisation because it's open book but masses and masses of bits of paper and I think girls tend to do better but some of that grade is just because they are more organised and it's not that they're better, it's not that they maybe try harder it's just that they take more pride in maybe their work and their organisation. I think that's, when it comes to improvements, they probably do a little bit more because it's a pride element.

What do you think the main challenges are with formative assessment?

I'd say being able to assess the whole class. Obviously, Matthew was saying about whiteboards, like I say we use whiteboards quite a lot and I think with questioning one of the main concerns with questioning even if you are asking really effective and really good questions you can only ask a certain amount of people. So, you've got a class, I mean I have quite big class sizes, up to like 30 in a class. It's very difficult in a lesson that is an hour long to teach them all something and understand whether they have all actually got something out of it other than doing an official, kind of end of topic test where you've actually got to get them to write something down and mark it which obviously we do not have the time to do that every single lesson that's the biggest challenge that I think I would say that we face is just too hard to get a picture of where everyone's at which is why whiteboard strategies are really good. I mean even then with big class sizes it's still not perfect because you can't go around and look at every single whiteboard, but it gives you a better idea than just asking a question and one person putting their hand up. What I find as well

is no-one, well very few people will volunteer an answer but when you ask them there's a good chance that they will know it. So sometimes you can get a picture of 'well no-one is volunteering an answer, does no-one know it?' and yet when I do actually pick them out and say answer that or they do a written piece of work you get a bit of a shock sometimes as to who the better students are, it's not always the ones that volunteer the answers.

Its peer pressure though is it not?

Certainly in Year 12 which I teach, a lot of them are too embarrassed to get it wrong or too scared to kind of stick their hand up and answer a question so you've got to drag it out of them!

There's a time element to that, of dragging it out because you can go at someone for a while but it could be you've spent 5 minutes on one student and yes you are happy that now they know it but you might have X amount of students in the class so.

That's the challenge, getting the balance between the pace of the lesson and keeping everyone engaged but also not asking questions and them saying they do not know and just moving onto someone else. You do not want that because the next person will just say I do not know, and you're stuck. So, you've got to try and get it out of them but, how long do you spend doing that?

What challenges do you think there are with summative assessment?

Time, yeah, I'd say time. In terms of volume of marking, for it to be effective, look at my Year 13's now, how many weeks? 4, 5 weeks out from the main exam. Do a full mock, which I think is vitally important for them because, and to do a couple, because they've got to, well one from a physical point of view the first time they did it, it was amazing how many of them were shaking their hands and getting fatigue in their hands, as they haven't written solidly for 2 hours for a while, and that training element of it. But that appreciation of time management from our subject is huge. They have got 3 long essays to write in a 2-hour exam which are a fraction of what the paper is and so however they do these mocks but the time element to mark them is huge on my behalf. I want to mark them properly to identify both the knowledge they are missing but also the structure in terms of how they are answering it, their style of English and such like. And to feed

that back is, I think the biggest burden but brings with it a lot of value. And there's a lot both I and they will learn from the process.

And it is that variety in terms of how they will respond to the feedback. Obviously, you'll expect, you might think well we could spend half an hour improving that very usefully but some of them despite how many times you go back to them will feel like they're done in like you know 5 minutes.

We're doing mocks this week and I've already started marking some of them and obviously you do kind of hold that feedback to kind of say this is the full paper, the common misconceptions but I think obviously we're marking these mocks we're not marking them with comments because we wouldn't have time to mark them with comments so it's just a mark. This is where we do give them a grade, A, B,C , D whatever and obviously we'll do the full class assessment, common misconceptions, this is how to answer this question better and put some model answers or some of the better answers or weaker answers and do all of the assessment around that but really what the ones that haven't done well need is someone to sit down with them and say this is where you have gone wrong, this is what you should do to improve it, but again you just haven't go the time to do that so there's a great value in doing the summative assessment, obviously as a teacher to understand the common misconceptions, what needs to be taught differently, what's been taught well, and what they've done well, where obviously you can make improvements and develop them but really what you need in a perfect world is one on one time to say right this is where you've gone wrong on this question, this is where you've gone wrong on this question but you obviously do not have that time to do it.

If you have just put the marks on it, it is a good learning process for them I think to look themselves and see 'oh that's why I've got that many marks', 'that's why I've got that, and I haven't got that'. You know 'I've got 6/10 there and it's because I didn't do X, Y and Z'. But obviously the ones that get the least marks will be the ones that are least able to have a reasonable go at that.

That's the thing. It would be great if you just put 3/6 and they could look at that and say 'right ok I need to put this and this and this in' but those that need that development do not have the ability to do that. That's the conundrum that you face.

And do you feel that you have enough time to get them up to speed then?

Well, I think obviously that, I mean we've got 19 days till what, till the external exam, obviously some do not do the mock until Friday, got to mark them, get them back to them. Some of them will probably only have a week, week, and a half maximum, probably 2 lessons, maybe 3 lessons between the mock, getting the mock back and doing the external exam. So when you consider you probably need a lesson for the kind of general whole class feedback and then obviously maybe a lesson or two to actually do the revision sessions and activities based around the areas they need to develop then we won't have that time really for any kind of one on one teacher student contact time so it is really reliant on them being able to identify the areas where they can improve based on their papers and their marks and obviously on how good the general feedback is to the rest of the class. Looking at those model answers, can they look at model answers and take theirs from what it is to the model answer. And look at a weak one and say that's similar to mine so there probably is not enough time but I do not know if there is ever going to be enough time, even if you have 5/6 weeks, but we do not have that amount of contact time.

I know mine are just struggling with timings. For one of the papers, they've got 45 minutes to get 3 questions answered and I mean one of them they've literally got 11minutes and they've got to get as much as they can down. It goes like that doesn't it?

I mean ours is 75 marks in 90 minutes so it's just over a minute a mark.

Have the changes to exam structure caused you any difficulties?

I think what you were saying really about spacing that's where when they've gone linear, we've, as an institution I think started to think more and more about that haven't we.

I do not teach A Level so I do not have that concern of linear but I'm still trying to put in some of the strategies that we've talked about in terms of improving teaching for that linear because I think there's still some benefit in even normal exams that aren't linear like some of the strategies that are taught would work for kind of traditional teaching and assessment. Our main challenge really is going from controlled assessment so quickly to an exam, getting students out of that mind set. You have to kind of drum it into them this is how you do a controlled assessment and then right ok, forget all that. This is how we do an exam. So that's the challenge.

Our challenge has been the lack of resources that's out there on the new spec. So, although we spend a lot of time marking these mocks, we are spending a lot of time creating them. And we're creating model answers and standard answers based on what we perceive as being how they are going to be marked. That's only on a training course and 2 exam mark schemes that are out there, or example mark scheme. So, we are a bit in the dark still in terms of what we are asking students to do. We're not clear on how they are going to be assessed and how these long answer questions are going to be marked. And so, we are still dubious in terms of that. But there are only 2 mock papers out there and so over a 2-year programme we're just trying to create as many examples as we can which is time consuming but has that caveat that we're not entirely sure how accurate it is.

How do you think students are finding the linear exams?

I think for Year 12 it's a bit of a blessing in disguise, they see it as because they think they've got a lot of time so it's not totally negative certainly their response to it because they feel like they've got time to get prepared.

My cohort, they did an AS exam last year, so we sat it. We haven't necessarily felt that whole linear and I do not think it's dawned on them that they are going to sit the two and obviously lose their AS. I think they haven't struggled with it in terms of that, but you are not going to know that now until they sit the exams at the end of 2 years. And I think in comparison to how it was, how many times they could re-sit it and go back to it and going into their final exam they would have

known that they had banked two thirds of the course, they knew exactly where they were at in terms of that so how many marks they would need in their final exam to get what they need, they just do not have that security any more. And so, they never had it in terms of that from an A Level but from obviously looking at it, comparing cohorts you know that there's previously students who were on that module format were in a stronger position and a safer position going into their final exam season than the current cohort are.

And at least when the current, with the GCSEs changing as well aren't they so this is maybe a bad couple of years in the sense that the Year 13s we've got now still very much had their hands held all through GCSE and they did do Christmas and Summer, Christmas and Summer in Year 10 and Year 11 and so they're used to that. But the next set of Year 12s that should be less the case because their GCSEs would have changed so that they're a bit more used to having to wait to be assessed, it won't be quite as much of a shock.

Do you think you are still working on the content right up until the exam?

We tend to schedule it so that we get finished and we've got at least a month to work on revision.

I'm trying to think Year 12 and Year 13, I've only taught Year 12 today.

We finished just after Easter which gives us plenty of time for revision. Now I do not know whether that's had a massive positive effect because I think it's had a bit of a knock- on effect on their attendance. I think because we are into revision so early that I've noticed a drop off in attendance and I was asking them the other day whether their attendance would have actually been better if I was teaching for longer. Because almost their mind set is right, I've got my notes, I've got everything I need, yes it is revision but if I miss that lesson then potentially it's not as bad as if I was to miss a lesson where he is delivering new content.

To be honest I've noticed a few students that miss lessons and I've said, 'where were you?', 'I wasn't feeling great, but it was only revision, so I didn't come in'.

And they all think they can revise at home.

They turn to lessons and 'Sir Can we just do our own stuff', well 'no it's a lesson so I'll be doing revision activities with you and some class revision' 'but I revise better on my own'. They just think that because you've finished the teaching that, that's it. And that the lessons are just really for them to come and sit and kind of make a few revision materials, have a look at some stuff and I think they maybe lose focus a little bit. I think if you taught it for longer the lessons would obviously have the knock-on effect of some of them would be sitting an exam and answering questions on content you have taught them the week or two weeks prior if you went further towards the exam but I think it would focus them a little bit more right up until that point. And obviously the kind of focus would then be on them revising independently or at home which I mean a lot of them will do anyway but if every lesson is revision lesson, I think maybe you'll lose some of that. Certainly, I've noticed attendance with some of them has been an issue and they've actually said to me 'well I knew it was just revision, so I didn't think it was a major issue if I missed it'.

We've had a bit of an issue with Year 13s who are resitting the Year 12 material because we invite them back once a week and trying to get them all to come in is just, has just been really hard. Some of them are obviously more dedicated will come every week and will be really keen to learn exactly what they need to do. It's a real issue.

Is there any way you would like to improve your use of formative assessment in the future?

Less marking for Mike!

Yeah definitely.

I've got about 90 students so yeah markings a bit of an issue but I just think that really in a kind of ideal world what we've already said really would be more kind of one on one time or smaller group time, you know maybe getting 5 or 6 students that have had the same issue, got the same type of result on any type of assessment, same type of mistakes and kind of sit down with them for half an hour and go through it, kind of mini lessons if you like but obviously that's a timing thing is it not? You're not going to have the opportunity to do that really, but I think that would

certainly in my opinion improve things. But I think everyone would want that wouldn't they? It's just not going to happen.

Do you place more value on formative or summative assessment or do you think it is a combination of the two?

I personally think combination. I wouldn't pick one over the other. I think they both have their uses. I think they are both vitally important. And I think there's a need for both of them, but they work hand in hand and at different times depending upon what you need. I think the scope for summative is an important aspect of it, but I think formative is just embedded into your good lesson delivery. I think that's their day in day out in terms of what's going on in the classroom, but I think that needs to be supported with regular summative assessment.

That's always what the results are going to be, based on is it not? Well not always but the majority.

Summative assessment is kind of focused a little bit more. I think if you took away that summative assessment, end of unit tests and regular tests that are kind of mirroring the exam you would maybe kind of lose a little bit of focus and you wouldn't have as much information as to how they were doing and lose kind of the impact of following lessons where you look at a kind of correcting any misconceptions from that end of unit test. I think they go hand in hand really, you couldn't have one without the other. I think you need them both.

It's like you say at the end of the day that's how they are going to get their mark. And so, they need to practice it. And you can have all the formative assessment and you'd be confident they know everything but one of our difficulties is students know it all but put an exam paper in from of them and they can't answer it. And so, there's definitely a need there I think to do regular questions to develop exam technique, not necessarily to access knowledge but to develop exam technique and exam style which is equally as important as the gaining of knowledge and securing that knowledge.

If you do it regularly and they do well it can increase their confidence so like you say obviously getting them ready for that exam, exam practice, exam technique I imagine the students who do well in this mock exam that we've just done will go into that real exam with a great deal of confidence that actually I do know it, I can do it, I can answer these questions, there's nothing to be too concerned about. I think getting over that kind of anxiety and maybe reducing the stress levels from having practiced it and you know them having succeeded at it must have some benefit as well.

Appendix 7 - Observed Lesson Subject and Observation Number

Observation Number	Subject
1	Criminology
2	Chemistry
3	Sociology
4	Media Studies
5	Geography
6	Maths
7	English Literature
8	Religious Studies
9	Biology
10	BTEC Law
11	Classics
12	PE
13	Film Studies
14	Maths
15	History
16	Accounting
17	Physics
18	Biology
19	History
20	French
21	BTEC ICT
22	English Language
23	Business Studies
24	Business Studies
25	Fine Art
26	Health and Social Care BTEC
27	Criminology
28	Maths

On 7 Feb 2018, at 13:36, [redacted] wrote:

Hi Jenny

Glad it was useful. You're more than welcome to come in on another Tuesday. Do you want to do it from 12pm (like the next one) or 1pm (like the first set of obs - i.e.4 lesson obs or 6? Let me know and I'll get on to organising.

Thanks

[redacted]

Subject	Number of exams	Structure of questions
Accounting	2 exams (50% each)	Multiple choice, short answer & extended answer (25 marks)
Biology	3 exams (100% in total)	Short answer and extended answer (25 marks)
Business Studies	3 exams (2 x 35%, 1 x 30%)	Short answer and extended answer questions (20 marks)
Chemistry	3 exams (2 x 35%, 1 x 30%)	Short answer questions (up to 7 marks)
Classics	3 exams (1 x 40%, 2 x 30%)	Short answer questions. Essay based (20 marks)
Criminology	1 exam (50%)	Short answer questions (up to 8 marks)
English Language	3 exams (2 x 30%, 1 x 20%)	Short answer and essay based (20, 30 marks)
English Literature	2 exams (40% each)	Extended answer questions (25 marks)
Film Studies	2 exams (35% each)	Essay based (40 marks)
French	3 exams (1 x 50%, 1 x 30%, 1 is x 20%)	Short answer questions. 2 x 40-mark questions
Geography	3 exams (80% in total)	Short answer and essay based
Health and Social Care (BTEC)	Exam	Short and extended answer questions
History	2 exams (40% each)	Essay based (1x 30 marks, 2 x 25 marks)
Maths	3 exams (100% in total)	Short answer questions
Further Maths	4 exams (25% each)	Short answer questions
Media Studies	2 exams (35% each)	Short answer and extended answer questions. (10, 12, 15 and 30 marks)
Physical Education	2 exams (35% each)	Multiple choice, short answer & extended answer (up to 15 marks)
Physics	3 exams (2x 34%, 1 x 32%)	Multiple choice questions Short answer questions (up to 6 marks)
Religious Studies	3 exams (100% in total)	Essay based (20 and 30 marks)
Sociology	3 exams (100% in total)	Short answer (4,6,10 marks) and essay based (20, 30 marks)

100% exam

Combination of exam and coursework

Name of applicant	Jenny Peadon
Email address	jennifer.peadon @durham.ac.uk
Category *[choose from list]*	Postgraduate student - Research programme
If "Other" please specify	
Student ID number *[students only]*	dzrl15
Programme *[students only – choose from list]*	EdD
If "Other" please specify	
Name of supervisor *[students only]*	Jonathan Tummons
Title of research project	A comparative case study of teacher's use of formative and summative assessment with contrasting year groups.
Date of start of data collection phase of the research *[must be a future date – no research to be conducted until ethical approval obtained]*	January 2018
Is the research funded *[staff only – choose from list]*	No
Name of funder *[staff only]*	
Name of Co-Is if applicable *[staff only]*	
Is this application subject to external ethical review? *[choose from list]*	No
If "yes" please specify who	

FOR OFFICE USE ONLY – Please do not delete this box

Please can reviewers enter the date and select the outcome from the drop-down outcome list below? To open the drop-down list, please select "click here to choose from list". To enter a comment, please click on the yellow highlighted area below and start typing.

Please note that as the review process is anonymous there is no requirement to include initials or signatures in this section.

REVIEWER RESPONSE	REVIEWER COMMENTS
Date:	
Click here to choose from list	

1) a. Does the proposed research project involve data from human participants (including secondary data)?	Yes
b. Is the research project *only* concerned with the analysis of secondary data (e.g. pre-existing data or information records). If yes, please continue with Q6-13	No
c. Is the work non-empirical (e.g. literature review, opinion piece, systematic literature review) If yes, please complete Q11	No
2) Will you provide your informants – prior to their participation – with a participant information sheet containing information about the following: a. The purpose of your research?	Yes
b. The voluntary nature of their participation?	Yes
c. Their right to withdraw from the study at any time?	Yes
d. What their participation entails?	Yes
e. How anonymity is achieved?	Yes
f. How confidentiality is secured?	Yes
g. Whom to contact in case of questions or concerns? *Please attach a copy of the information sheet (template available at **appendix B**) or provide details of alternative approach at Q13 of this form.*	Yes
3) Will you ask your informants to sign an informed consent form? *Please attach a copy of the consent form (template available at **appendix C**) or provide details of alternative approach at Q13 of this form.*	Yes
4) a. Does your research involve covert surveillance?	No
b. If yes, will you seek signed consent post hoc?	Click here to choose from list
5) a. Will your data collection involve the use of sound or image recording devices?	Yes (if yes, please answer Q5b and Q5c below)
b. If yes, will you seek signed consent?	Yes
c. Please specify the type of recording	Audio

6) Will your research report be available to informants and the general public without restrictions placed by sponsoring authorities?	Yes
7) a. Does the research involve unsupervised access to children or vulnerable adults within an activity that is deemed as regulated and would therefore require DBS clearance?	No
b. If yes, can you confirm that DBS clearance is in place or will be in place prior to commencing your research?	Click here to choose from list

8) How will you guarantee confidentiality and anonymity? The identities of both institutions and participants involved in the research project will be protected through the use of pseudonyms. In addition, details will be removed from both transcripts and the final write up, which could lead to the identification of participants or the location of the research site.
9) What are the implications of your research for your informants? The main implication is the potential infringement on the valuable time of teachers. With this in mind, questionnaires will be available online using the Bristol Online Survey tool; therefore making the completion of the questionnaire as quick and easy as possible. Respondents will also be able to provide their contact details, should they wish to take part in a follow up interview.
10) Are there any other ethical issues arising from your research? Full disclosure is imperative, participants, in this case teachers whose lessons are being observed, need to know that they are taking part in research and are happy to participate. The transcription (of semi structured interviews) will provide another means for participants to become involved in the research process insofar as providing a copy of the transcript for approval prior to the research progressing further.
11) ***For non-empirical projects only***, please provide a brief overview of your project, approx.150 words max. Please include the research aims and objectives and your research approach *(Appendices A to C are not required)*. N/A
12) Will your research either-

- Involve the study of an organisation which is proscribed under the terms of the Terrorism Act, or require accessing materials produced by or in support of such an organisation (see https://www.gov.uk/government/publications/proscribed-terror-groups-or-organisations--2)

  Or

- Involve the study of any other current organisation which, as part of its agreed programme, advocates the use of violence to achieve its aims, or require accessing materials produced by or in support of such an organisation.

If you answer yes to either of the above then please contact ed.ethics@durham.ac.uk for an additional appendix to complete.

For further information please refer to the University policy https://www.dur.ac.uk/resources/research.office/local/policy/SecuritySensitivematerialsFINAL1.0.pdf

No- not applicable.

13) Please provide any additional information relevant to your application

Declaration

I have read the Department's Code of Practice on Research Ethics and believe that my research complies fully with its precepts.

I will not deviate from the methodology or reporting strategy without further permission from the School of Education Ethics Sub-Committee.

I am aware that it is my responsibility to inform the organization in which data collection takes place (e.g., school) that ethical approval from the School of Education Ethics Committee has been given, prior to commencing data collection.

Applicant signature* 	Date 14/09/2017
Proposal discussed and agreed by supervisor *[students only]* Supervisor signature*	Date

**To enable electronic submission of applications, electronic (scanned) signatures will be accepted.*
Please note that typed signatures cannot be accepted.

Pseudonyms:

FOCUS GROUP 1

Paul	Business Studies
Robert	Media Studies
Alan	Physics
Louise	Biology
Lesley	Sociology
Laura	Law (BTEC)
Amy	Health and Social Care

FOCUS GROUP 2

Ashley	English
Simon	Criminology
David	PE
Natalie	Film Studies

Appendix 12 - Completed Observation Schedule

Type of assessment observed	Date: 20/02 Time: 1.50 - 2.20 Duration: 30 mins Subject RS	Teacher or student led?	Notes
Formative Assessment			
Sharing learning objectives			
Questioning	1) to check prior knowledge (this working on essay feedback)	2) new topic where is conclusion?	
Peer Assessment			
Self Assessment			
Pre-assessment strategies to gauge understanding			
Modelling			
Graphic organisers			
Other	Advice on how to get more marks		
Summative Assessment			
Unseen examination			
Exam style questions			
Open book exams			
Multiple choice tests			
Written essay / report			
Presentation			
Performance task (to assess a specific set of skills)			
Feedback- marks / grades	reviewing feedback on essay from teacher, responding to comments (to improve answer)		
Other	(worksheet working just to give more info) Responses to be checked by teacher		

Appendix 13 - External Assessments and Types of Assessments Observed

Subject	External form of summative assessment	Type of formative assessment observed	Type of summative assessment observed
Accounting	100% examination	Questioning, modelling, graphic organizers, other	Exam style questions, feedback- marks/grades,
Biology	100% examination	Sharing learning objectives, Questioning, peer assessment, self-assessment, gauge understanding, other	Exam style questions
Business Studies	100% examination	Questioning, self-assessment, other	feedback- marks/grades
Chemistry	100% examination	Questioning, modelling, other	Exam style questions, other
Classics	100% examination	Questioning, gauge understanding	Exam style questions, feedback- marks/grades, using assessment criteria
Criminology	Controlled assessment & exam	Questioning, peer assessment, modelling, graphic organizers, other	Exam style questions, feedback- marks/grades, using assessment criteria
English Language	80% exam, 20% coursework	Questioning, modelling	Performance task
English Literature	80% exam, 20% coursework	Sharing learning objectives, Questioning, gauge understanding	
Film Studies	70% exam, 30% coursework	Sharing learning objectives, Questioning, gauge understanding	
Fine Art	Practical work (60%) and externally set task (40%)	Questioning, modelling, other	
French	100% examination	Questioning, self-assessment, other	Written essay, report
Geography	80% exam, 20% individual investigation	Questioning, self-assessment	Exam style questions, using assessment criteria
Health and Social Care (BTEC)	Exam and coursework	Questioning, graphic organizers, other	

History	80% exam, 20% historical investigation	Sharing learning objectives, Questioning, gauge understanding	
ICT (BTEC)	Internal and external assessments	Questioning	
Law (BTEC)	Coursework	Questioning, other	presentation
Maths	100% examination	Questioning, self-assessment, gauge understanding, other	Exam style questions, feedback- marks/grades, using assessment criteria
Further Maths	100% examination	Questioning	Exam style questions, using assessment criteria
Media Studies	70% exam, 30% coursework	Questioning	other
Physical Education	70% exam, 30% coursework	Questioning, other	
Physics	100% examination	Questioning, gauge understanding	
Religious Studies	100% examination	Questioning, other	feedback- marks/grades, other, using assessment criteria
Sociology	100% examination	Sharing learning objectives, Questioning	Exam style questions

Jisc bos

Assessment

Response ID	Completion date	
324251-324243-29330298	25 Jan 2018, 18:11 (GMT)	

1	How would you describe formative assessment?	Throughout class, continuous assessment, interim feedback
2	How would you describe summative assessment?	At the end of a topic/test
3	Is formative assessment part of your school policy?	Yes
4	Have you received CPD on the use of formative assessment?	Yes
5	Did you receive CPD on formative assessment while completing your ITT?	Yes
6	How often on average do you use formative assessment?	At least once per lesson
7	How often on average do you use summative assessment?	Once per week
8	Which of the following formative assessment strategies do you use?	• Sharing learning objectives • Questioning • Peer Assessment • Self Assessment • Modelling • Pre-assessment strategies to gauge understanding
8.a	If you selected Other, please specify:	
9	Which of the following summative assessment strategies do you use?	• Unseen examinations • Exam style questions • Written essays / reports • Performance task (to assess a specific set of skills)

9.a	If you selected Other, please specify:	
10	Rate the value of formative assessment as an assessment tool.	5 Very Good
11	How would you rate the value of using summative assessment?	2 Very valuable
12	How do you provide feedback to students?	Marks and grades alongside comments
12.a	If you selected Other, please specify:	
13	How should students be given feedback?	
13.1	Feedback to students should be comment only	Disagree
13.2	Students should be given a mark / grade	Disagree
13.3	Students should be given a mark / grade alongside comments	Strongly Agree
14	How effective is high stakes testing as an assessment tool?	3 Somewhat effective
15	Do you think that high stakes testing affects your use of assessment on a day-to-day basis?	No
16	How has high stakes testing affected your assessment practice? Please respond in relation to the following statements: High stakes testing has led me to change my approach to instruction.	3 Disagree
17	High stakes testing has led to more teaching to the test.	2 Agree
18	High stakes testing has led to a narrower curriculum for students.	2 Agree
19	High stakes testing has NOT resulted in a loss of creative input into my teaching.	3 Disagree
20	High stakes testing has NOT resulted in less time being spent on formative assessment.	4 Strongly Disagree

DURHAM SIXTH FORM CENTRE

Data Collection Methods

Jenny Peadon

Lesson Observations

Time	Subject	Year	Teacher	Room
1.20-1.50pm	Criminology	12	[redacted]	F10
1.50-2.20pm	Chemistry	13	[redacted]	B9
2.20-2.50pm	Sociology	12	[redacted]	B3
2.50-3.20pm	Media Studies	12	[redacted]	A3

Time	Subject	Year	Teacher	Room
12.20-12.50pm	Geography	12	[redacted]	G9
12.50-1.20pm	Core Maths	12	[redacted]	G2
1.20-1.50pm	English Literature	13	[redacted]	G19
1.50-2.20pm	Religious Studies	13	[redacted]	F9
2.20-2.50pm	Biology	12	[redacted]	B11
2.50-3.20pm	Law (BTEC)	12	[redacted]	F16

Time	Subject	Year	Teacher	Room
12.20-12.50pm	Classical Civilisation	12	[redacted]	F14
12.50-1.20pm	PE	13	[redacted]	E1
1.20-1.50pm	Film Studies	13	[redacted]	A3
1.50-2.20pm	Further Maths	13	[redacted]	G6

2.20-2.50pm	History	13	[redacted]	G18
2.50-3.20pm	Accounting	12	[redacted]	F4

Time	Subject	Year	Teacher	Room
12.20-12.50pm	Physics	12	[redacted]	B10
12.50-1.20pm	Biology	12	[redacted]	B7
1.20-1.50pm	History	13	[redacted]	G21
1.50-2.20pm	French	12	[redacted]	11
2.20-2.50pm	Computing (BTEC)	13	[redacted]	G3
2.50-3.20pm	English Language	12	[redacted]	E3

Time	Subject	Year	Teacher	Room
12.20-12.50pm	Business	12	[redacted]	F4
12.50-1.20pm	Business	12	[redacted]	F16
1.20-1.50pm	Fine Art	13	[redacted]	A4
1.50-2.20pm	Health and Social Care (BTEC)	12	[redacted]	B2
2.20-2.50pm	Criminology	13	[redacted]	F1
2.50-3.20pm	Mathematics	13	[redacted]	G6