From speech to script? Can collaborative peer assessment help students to improve extended writing in the GCSE RE classroom?

Richard Hackett

A Research & Development Project

Submitted for the MSc Learning & Teaching 2020

DEPOSIT AND CONSULTATION OF THESIS

One copy of your dissertation will be deposited in the Department of Education Library via the WebLearn site where it is intended to be available for consultation by all Library users. In order to facilitate this, the following form should be completed which will be inserted in the library copy your dissertation.

Note that some graphs/tables may be removed in order to comply with copyright restrictions.

Surname	Hackett
First Name	Richard
Faculty Board	Education
Title of Dissertation	From speech to script?' Can collaborative peer assessment help students to improve extended writing in the GCSE RE classroom?

Declaration by the candidate as author of the dissertation

1. I understand that I am the owner of this dissertation and that the copyright rests with me unless I specifically transfer it to another person.

2. I understand that the Department requires that I shall deposit a copy of my dissertation in the Department of Education Library via the WebLearn site where it shall be available for consultation, and that reproductions of it may be made for other Libraries so that it can be available to those who to consult it elsewhere. I understand that the Library, before allowing my dissertation to be consulted either in the original or in reproduced form, will require each person wishing to consult it to sign a declaration that he or she recognises that the copyright of this thesis belongs to me. I permit limited copying of my dissertation by individuals (no more than 5% or one chapter) for personal research use. No quotation from it and no information derived from it may be published without my prior written consent and I undertake to supply a current address to the Library so this consent can be sought.

3. I agree that my dissertation shall be available for consultation in accordance with paragraph 2 above.

Introduction 5
Responding to ‘more challenging GCSEs’ 7
The impact of the new specification on GCSE results at Academy X. 8
Possible benefits of peer assessment. 10
Guiding questions for a review of the literature. 10

How might a focus on extended writing help pupils meet the new syllabus’ demands? 12

What are the potential benefits of peer assessment in helping develop extended writing skills? 14
Defining peer assessment. 14
Is there evidence that formative assessment improves outcomes? 15
How teachers and pupils benefit from formative assessment. 15
Challenges of implementation 17
Overcoming challenges of implementation. 19
Feedback as the heart of formative assessment. 21
Formative assessment in RE. 23
What does peer assessment offer than self assessment cannot? 24
Pupil perspectives on peer assessment. 27
The benefits of peer assessment for differing ability levels 28

What might happen when pupils work together in a group peer assessment model? 30
Sociocultural insights into how peer assessment can help develop extended writing? 30
The Zone of Proximal Development as an aid to concept formation. 31
Spontaneous and scientific concepts in GCSE religious education 33
Applying models of concept development in RE 34
Sociocultural theory and the development of writing. 35
Risks of adopting a sociocultural model of group peer assessment 37
Questions arising from the Literature 38

Methodology 39
Overview of the intervention 39
Preparing the ground for Group Peer Assessment 40
Designing the group peer assessment model 43
What will be assessed, how and by whom? 43
Stage 1 - devising the assessment criteria 43
Stage 2 - gathering assessable material 43
Step 3 - constructing groups 44
Feedback pro-formas 45
Structured observation of group peer assessment in action. 46

Listening in 46
Data collection 49
Potential limitations of the GPA model. 50
Post intervention questionnaires and interviews 51
Participants and their consent 53
Further Ethical Concerns 54
The impact of the coronavirus school shutdown and subsequent alterations to the study 56

Tables and data 59
Selected pupil questionnaire results. 59
Teacher questionnaire results 63

Findings and discussion 68
Research Question 1: Does peer assessment in groups help pupils to produce better extended writing? 68
Teacher and pupil views on the benefits and drawbacks of assessment for learning 72
Research Question 2: Does Structuring feedback help the development of extended writing skills? 75
Potential of feedback proformas 77
Research Question 3: Can working in mixed ability groups help the development of pupil's extended writing skills? 78

Conclusion 88

Bibliography 92

Appendix 102
Appendix 1 a 102
Appendix 1 b: 103
Appendix 2 104
adjusted level descriptors 104
Appendix 3 a 106
Appendix 4a - questionnaire 109
Appendix 4b - completed questionnaire sample 111
Appendix 5 - GPA feedback pro-forma 115

Introduction

I have recently completed my eleventh year as a teacher of Religious Education. During this time there have been significant changes to the nature and quantity of content which must be covered in the subject at GCSE, as well as to how it is examined. While the majority of pupils in my classes have not been adversely affected by these changes, and attainment remains high for most, this is not true for all.

The nature of the subject matter in Religious Education invites opinion and comment from pupils, often leading to discussion and debate. Indeed, this dialogic approach is encouraged (Weisse, 2011) as a way to develop pupil understanding (e.g. Jackson, 1999; Myatt, 2018). That the subject is (at my school) optional, usually means that pupils have chosen and enjoy it, although they may also have been attracted by the perception of relative easiness (Geoggiou and Wright, 2018) . Whatever the reason for their choice, over the last decade I have taught dozens of apparent 'fans' of RE, perfectly capable of articulating credible *oral* answers to religious, philosophical or ethical questions, offering opinions and and defending their responses to challenges, but who ultimately fail to translate these words into extended *written* answers. Their insights, interest and knowledge fail to produce a 'good' GCSE grade.

The majority of my timetable consists of Philosophy, Religion and Ethics (PRE) in Key Stage (KS) 3 and Religious Studies (KS 4 and 5). I have also taught English at KS 3 and 4, GCSE Sociology and am co-ordinator of PSHE (personal, social and health education). Teaching GCSE across three subjects has given me insight into the greater demands placed upon pupils by 2016's GCSE reforms and the efforts made by departments to help candidates prepare. My

observation is that my current school has no whole-school policy or recommended set of strategies for helping pupils to improve their extended writing skills. For those pupils who consistently struggle to transfer their thoughts or words into writing, there is no method deployed consistently across the curriculum to give guidance about how to produce better quality extended writing. My concern is not that pupils lack understanding or knowledge (although some may), but rather they cannot or will not write enough, perhaps not even knowing what 'enough' is. It is as if they have been asked to assemble a piece of flat pack furniture, for which they *may* have *most* of the pieces, but they have been given neither the instructions nor the tools. The end result may vaguely resemble the intended design, but will inevitably fall short of expectations.

'Academy X', is a coeducational academy-convertor of 1700 pupils. Located in Hertfordshire, intake consists of pupils from Southwest Hertfordshire and Northwest boroughs of London. It is partially selective, admission administered through the Southwest Hertfordshire consortium. Consistently rated as 'good' by Ofsted, this status was confirmed in September 2019. Data evidences the school's success, with both attainment 8 and progress 8 scores improving in recent years, as outlined below (a progress 8 score of 0.5 is considered to be well above expected levels of progress).

(2018) Attainment 8: 5.66 = C	Progress 8 0.25
(2019) Attainment 8: 5.89 = B	Progress 8 0.39

71% of pupils achieved five strong passes (grade 5 and above) including Maths and English, and 88% achieved this standard in Maths and English alone. This 'success'

excludes the 29% not achieving 5 strong passes. The most recent OFSTED report highlights that while most pupils do well, those with SEND and other needs are sometimes left behind. While I will not focus on SEND pupils specifically, this indicates that there is more help needed for those who struggle to do what is required for success, which includes extended writing.

Responding to ‘more challenging GCSEs’

Although attainment at Academy X has remained largely consistent for the last decade, with 70-79% of pupils achieving at least 5 GCSE passes, the government decision to increase challenge when developing new GCSE specifications, introduced in 2016, disrupted this. Publicity (https://newgcses.campaign.gov.uk) issued prior to the launch of the new specifications states that changes were intended to make attainment of the highest grades at GCSE more difficult. This was to prevent perceived grade inflation, defined as a rise in average grade awarded without a concomitant rise in achievement (Ayyapan, 2018). The new qualifications are twice referred to in the publicity as ‘more challenging’. The head of OFQUAL, the quasi-non-governmental agency responsible for guaranteeing standards in qualifications called, prior to the changes, for “qualifications that are more stretching for the most able students, using assessments that really test knowledge, understanding and skills" (Stacey, 2013).

A direct comparison between the pre and post reform specifications issued by Pearson Edexcel shows that increased challenge takes the form of questions which demand

more specific knowledge of religious beliefs and teachings, coupled with a requirement to write in an evaluative way which uses religious knowledge to justify conclusions. This is perhaps justifiable given the perception that in GCSE RS, pupils could 'achieve very high grades… without.. deep knowledge and understanding of religious beliefs and teachings' (Georgiou and Wright, 2018, p102). Increased challenge is primarily manifested in the eight 'd-type' questions which students must tackle. Previously, these questions required approximately four reasons for religious or personal views on a topic. The new specification requires critical evaluation of these reasons, and a conclusion which draws together a line of argument. (See appendix 1a for examples of questions types and associated mark schemes. Legacy papers are no longer published by the exam board, so source for direct comparison is not possible).

The impact of the new specification on GCSE results at Academy X.

The impact of the new specification was an immediate drop in attainment. A 'residual' of 1 denotes a candidate achieving a grade higher than their 'calculated' target grade. Calculated target grades are based on data from the Fischer Family Trust (FFT), drawing on key stage 2 data and, according to the now defunct website of the National College of Teaching, 'different indices of social disadvantage'. FFT's website (https://fft.org.uk/fft/target-setting/) states criteria such as English being an additional language, qualifying for free school meals, ethnicity and additional educational needs.

Academy X's calculated grades are measured by FFT 20, suggesting 'high prior attainment' (FFT Aspire, 2018). The final cohort to sit the legacy examination achieved residual grades of 0.71. This dropped to 0.41 in 2018, the first year of the new exam. Although 2019 saw the residual rise to 0.74, this coincided with a large increase in students failing to achieve a grade 4, considered equivalent to the previous 'good pass' grade of C. Less able pupils appear to have been most adversely affected by the increase in challenge.

Departmental data shows that pupils who are calculated to achieve grades of 6+ historically do so, often significantly outperforming their target (i.e. residual grades of more than +1). It is those with calculated grades of 3 or 4, usually boys, who don't meet/surpass targets. Work produced by the current year 11 cohort shows that those with lower calculated targets are again struggling to write enough, in the right way to fulfil the assessment criteria, while their more able peers of both sexes outperform their targets consistently. Most of these pupils are able to verbalise knowledge about the course content, but this is rarely reflected in what is committed to paper, and barely ever in a way that meets the examination criteria. This suggests that they may benefit from more regular assessment for learning, particularly collaborative peer assessment. Teacher-led instruction of the assessment criteria, and assessment itself, has not enabled this group to translate their knowledge into good quality extended writing, so perhaps more active involvement in assessment could increase their understanding of what is required?

I am about to embark on a new role as head of department at a newly founded, academically non-selective school, and hope that this research will help me support a cohort of pupils who may have lower prior attainment than those I have taught until now, as well as help my current school develop a method for improving pupils' extended writing.

Possible benefits of peer assessment.

Sociocultural theory suggests that pupils learn best through interaction, either with peers or with teachers (Vygotsky, 1978; Gielen et al, 2010; Bennet, 2011). Peer assessment is a variant of formative assessment of which there is no prescribed method of delivery. This suggests that there is room for collaborative models to be tested and evaluated, and shared if successful. It has been suggested that peer assessment can lead to increased motivation to perform well in assessments, as well as revealing 'the mechanism of assessment' (Gielen, 2010) which improved clarity and internalisation of learning goals. By listening to conversations taking place, either live or recorded, I may be able to gather evidence of this taking place, while potentially developing a useful model to share with colleagues. If evidence can be found that the process is one in which pupils engage, if their learning experiences are positive, and if extended writing skills improve, this might suggest that peer assessment can be used by teachers more regularly, even for extended written work.

Guiding questions for a review of the literature.

My literature review is guided by the search for a form of collaborative peer assessment which will require pupils to be able to identify the extent to which their own written work, and that of their peers, meets the criteria of the GCSE mark scheme. This requires the pupils to have a breadth and depth of subject knowledge and the ability to use this knowledge to structure evaluative essays which reach a supported conclusion (as per the mark scheme, appendix 1b) It is guided by the following questions:

1. How might peer assessment of extended writing help pupils meet the new syllabus' demands?
2. What are the potential benefits of peer assessment in helping develop extended writing skills?
3. What might happen when pupils work together to assess each other's work?

Literature Review

How might peer assessment of extended writing help pupils meet the new syllabus' demands?

It was allegedly the pursuit of a 'knowledge rich curriculum' which prompted the changes in exams leading to this intervention. For Micheal Young, (1999) schools have a responsibility to furnish young people with 'powerful knowledge'. A key element of this is found in 'relations' between different forms of knowledge, including 'theoretical knowledge', (Ibid, 1999) allowing for generalisations. However Mary Myatt (2018) suggests that RE has been dumbed down, and lacks deep engagement with subject matter and challenge in task. Perhaps exposure via peer assessment to the evaluative style of writing required by the new specification will increase the pupil assessors' ability to make Youngian generalisations, grounded in a deeper understanding of subject matter. Richard Kueh (2018) has argued that powerful knowledge in RE can be realised through the evaluation of beliefs against concepts such as truth, evidence and proof. Facts alone are valuable, but the power of knowledge in RE lies in the way it allows pupils to understand and respond evaluatively to the way these beliefs manifest; knowledge of both dogma, and of how dogma can be critiqued. Peer assessing the kind of extended writing required by the new specification demands exposure to and attention/awareness of this kind of relation making, evaluation of and generalising about religions.

Some scholars argue that RE should allow pupils to construct meaning. The importance of meaning making was first stated, note Conroy et al (2012), in OFSTED's 2010 report into the subject. For Conroy, RE fails to allow for 'meaning making' on the personal level, so pupils are unable 'to attend to the meaning (religious claims) have for adherents' (2012). He also claims that 'specifically religious content (and) intellectual challenge' (Conroy et al, 2012) are particular concerns for the subject. It is possible to assert that the requirements of the new specification have moved the subject forward in these areas, with demands for evaluation and justification of arguments replacing what Conroy et al call the 'your own opinion can't be wrong' (2012) culture within the subject. Peer assessing work may allow pupils to see how their peers have achieved this.

Collaboratively peer assessing extended writing may be a useful tool in allowing conversations between pupils about how belief systems affect the views of adherents, and where in/consistencies lie. This could help achieve the 'knowledge and critical awareness of a range of perspectives involved in the study of religion(s) and worldview(s)' called for by Freathy and Davis (2019). The teacher educator Michael Richardson, coming from an American perspective, has also called for 'open, honest and respectful' (Richardson, 2017) communication within religious education in state schools; again, peer assessment creates a forum for this. In his essay 'Is RE still not working' (2018), David Lundie argues that 'increasingly knowledge-focused' curricula are detaching religion from an understanding of lived experience. His concern is that pupils learn only definitions or idealised versions of religions. I would argue that this

intervention, by exposing pupils to evaluative writing through collaborative peer assessment, may make space for pupils to develop their understanding of how to write about religions' theological consistencies and inconsistencies. This removes the idealising of religions.

We can say that the significant amount of knowledge that pupils are required to learn equates to what Micheal Grimmitt (1987) termed 'learning about religion' - having a grasp of the 'facts' particular to a faith and its practise. Peer assessment, intended to improve each others' critical writing, may equate to 'learning from religion', which Grimmitt intended to include reflection upon what has been learned (Fancourt, 2010). If we slightly adjust Grimmit's 'learning from' idea to include the evaluation of a religion's internal consistencies and inconsistencies, we can see how time dedicated to being able to gain exposure to how others have communicated them in their writing, may help pupils meet the demands of the syllabus.

What are the potential benefits of peer assessment in helping develop extended writing skills?

Defining peer assessment.

Peer assessment is formative rather than summative. Dylan William (2018) traces the use of the term ‘formative assessment’ to the academic philosopher Micheal Scriven, and to the work of Benjamin Bloom, both active in the 1960s. Bennet, (2011) tells us that Bloom saw formative assessment as a tool for providing ‘feedback and correctives at each stage in the teaching-learning process’. Black and William (1998), define formative assessment as ‘all .. activities undertaken by teachers and/or their students, which provide information...to modify teaching and learning’. It has been argued (e.g. Deluca, 2018) that the definition is still under review, as contemporary research formulates guidance on implementation. Formative assessment includes self assessment (SA) and peer assessment (PA), together rebranded as assessment for learning (AfL) to avoid jargon overload (Croussard cited in Baird et al, 2017).

Is there evidence that formative assessment improves outcomes?

Numerous studies (e.g Black and William, 1998; 2009: William, 2018; Bennet, 2011, DeLuca, 2017) suggest that formative assessment has ‘a substantial impact on learning’ (William, 2018) though Baird et al. (2017) warn that there persists an absence of large scale quantitative studies on the subject. It has been suggested that ‘systematic use of assessment to improve learning remains the exception rather than the rule’ (William,

2018), and that there is not yet a well defined set of practices which identify formative assessment (Bennet, 2011; William, 2018).

How teachers and pupils benefit from formative assessment.

One motivation of this intervention was to enable more frequent assessment without a corresponding rise in teacher workload, a motivation shared with teachers studied by Andersson and Palm (2018) who were highly motivated to implement successful 'AfL' models partly for this reason. Anrdrade and Valcheva (2009) explicitly mention lack of teacher time as a contributing factor in interest in SA. They also highlight that SA is a formative activity. It is a process of reflection by pupils on the extent to which their work meets a given criteria and allows for revision and improvement (2009). The teacher remains engaged, monitoring the process and deciding grades.

The process of reflection can increase independence. Boud and Brew (1995) point to the 'emancipatory knowledge' students can gain, a concept attributable to the philosopher Jurgen Habermas (1987), who differentiates between three types of knowledge interest: *work knowledge* (in the GCSE RE classroom: knowledge of religious facts and teachings, alternatively termed 'substantive knowledge' by Scwhab (1978); *practical* (the ability to apply knowledge to a question, termed 'syntactic' by Schwab) and emancipatory. For Boud and Brew, formative assessment is emancipatory knowledge because it allows pupils to reflect upon their progress towards meeting

assessment criteria - teachers are involved only in-so-far as they share the assessment criteria and cannot 'arbitrate the process entirely (Brooks et al, 2011), with the pupil deciding, during this process, about their progress.

Formative assessment can enable pupils to understand how assessment criteria are met. Terry (1997) developed Habermas' ideas, and asserts that emancipatory knowledge equates to '*knowing why*' a piece of work archives a particular mark, rather than '*knowing how*' to achieve a mark or '*knowing that*', which consists of knowing facts. This helps us to see how peer assessment can emancipate a pupil from dependence on her teacher. '*Knowing why*' (Terry, 1997) allows her to comprehend how a given piece of work stands in relation to a mark scheme. The immediate question this raises, however, is whether pupils of differing abilities can apply any or all of these forms of knowledge. Whether they are only available to high prior attainers, and if so, how they can benefit their lower prior attaining peers in a peer assessment exercise is explored in '*the benefits of group peer assessment for differing ability levels*', at the end of this section.

It might also be wise to acknowledge that Boud and Brew's exciting assertion that 'students do not simply accept the standards .... they engage in a critique .. (which) informs .. judgements about their work' (1995) might distract us from the fact that in GCSE exams the criteria are, at least for the period of the examination, fixed. The critique in which students engage can only be on how well their work meets the criteria.

Challenges of implementation

Implementing more peer to peer engagement may pose challenges and face opposition from pupils, especially in a time critical examined subject. Moreover, as peer assessment derives from AfL, an as yet unproven pedagogy (De Luca, 2018), it is important to be aware that pupils may be reluctant to participate. At least one study (Kluger and DeNisi, in Bennet, 2011 p278) found that 38% of formative assessment activities had negative effects on attainment. Boud (1999) identifies a trio of issues raised by implementation of self and peer assessment models. One is perception; pupils' believe that they are involved in a solution to a problem of assessment rather than learning. Hence, little value is placed on understanding the process. It is necessary for teachers to ensure that the value of being able to self or peer assess is understood (Goodrich, 1996). Emphasis could, for example, be placed on peer assessment's potential to develop emancipatory knowledge (Habermas, 1987), manifesting as the ability to 'know why' (Terry, 1997) work achieves a particular mark. Secondly: SA/PA may be deployed on an ad hoc basis (Boud, 1999), appearing tokenistic. Lack of engagement with the assessment criteria is the third problem, and it seems reasonable to suggest that this is a consequence of the second problem. If pupils are not regularly practised in peer formative assessment, then they will be unable to engage in it effectively, with each episode being reduced to a re-learning of the criteria.

With the above in mind, it is important to ensure that pupils' possess familiarity with and confidence in using the GCSE RE mark scheme. Andrade and Valcheva (2009) discovered that practise makes perfect in three identifiable ways. Firstly, positivity towards SA increases with experience; secondly, clarity of teacher expectations was vital in enabling effective SA; finally SA was a task carried out 'mindlessly' (2009) until realisation dawned that done properly it could improve work and grades. These findings offer a response to Boud's (1999) concerns about pupil scepticism. They offer a justification for pursuing a model with SA at its foundation, because they show that initial pupil scepticism can be overcome once they realise it 'emancipates' them from a reliance on their teacher. With properly developed SA skills, a pupil should be able to apply SA to a piece of work prior to submission, be able to accurately predict the mark they will achieve and, should their prediction be proven wrong, be able to identify why.

Overcoming challenges of implementation.

Using formative assessment implies my belief that at least some pupils could become consistently reliable and accurate assessors of both their own and their peers' work, and that 'immaturity is not a..barrier (Brooks et al, 2011).

Moulding pupils into effective self/peer-assessors requires a number of important conditions to be met. For Goodrich (1996) an 'awareness of the value of self-assessment', which has been dealt with above, is one of a number of criteria. He also recommends access to clear criteria on which to base the assessment, such as a

GCSE mark scheme. A specific task or performance to assess is needed, in this case pupils' own written answers. Models of assessment should be made available, such as examples of marked work, including modelling/scaffolding. Further required is direct instruction in and assistance with self-assessment, for example the modelling of answers which has taken place throughout the course. Finally, practice is recommended, which for my intervention, would be the introduction of regular peer assessment itself.

To Goodrich's recommendations, Ross (2006) suggests the addition of revision of work post assessment. Andrada and Valcheva (2009) have gone as far as to say that these conditions can 'ensure .. *all* students get ... feedback they need'[1]. The question of whether these measures will truly offer the benefits of SA to the whole range of abilities and motivation types is explored in the section 'the benefits of peer assessment for differing ability levels'. Ross (2006) cites studies from Bergee and McEnery & Blanchard which found that self-assessed marks agreed with peer-assessed judgements rather than with teachers'. He suggests that this could be explained by differences between pupil and teacher interpretation of criteria. We cannot discount the notion that novice assessors may lack the confidence and separation from those being assessed which allows teachers to make corrections. Ross (2006) also suggests that pupils may focus on superficial features of a piece of work, rather than the detail. For example, if the mark scheme asks for three

[1] Emphasis mine

developed reasons, then the novice self assessor might focus on quantity and ignore quality.

The first key to developing effective models of AfL appears to be allocating time to allow pupils to develop confidence through practise. Topping (2006) argues that a novice assessor's lack of skill can be compensated for by the relative abundance of time they have. The second key is teacher intervention (Goodrich, 1996: Ross, 2006). The emphasis placed on teacher monitoring and intervention during AfL may equate to the standardisation of professional examiners, with experts regularly reminding the relative novice of the standard.

The process of training pupils to become competent self-assessors has been found to instil self efficacy beliefs. Bandura (cited in Ross, 2006) found that pupils who perceived themselves to have been successful in a task will adopt a more positive view of their abilities in later tasks, although this invites us to ask what happens to those who perceive failure? Training pupils to self-assess can also transmit the positive message that teachers trust them and value pupil judgements (Usher and Panjeres, 2005). It will be necessary, therefore, for me to gather baseline data which measures pupil confidence as self-assessors before allowing them to undertake group peer assessment, and for me to ensure that my monitoring of the group peer assessment is regular, to avoid the ‘grade inflation’ noted above.

Feedback as the heart of formative assessment.

One key element of formative assessment found in a variety of descriptions (Bloom, 1969; Black and William, 1998, 2009; 2011; Hattie, 2009; DeLuca, 2018) is feedback. Kluger and DeNisi (1996) define feedback as 'information from an external agent regarding some aspects of one's performance' and Black and William (1998) assert feedback's 'centrality to formative assessment'. In the classroom one would assume that feedback therefore proceeds from teacher to pupil. Burke et al (2010) have noted discrepancies between the developmental usefulness of feedback as perceived by tutors and pupils, with the latter often struggling to know what to do with it. Hattie (2009) asserts that feedback is most effective when it comes *from* pupil *to* teacher. His meta-analysis of 134 studies on feedback found a positive impact of .73 (as of publication in 2009, reduced to .70 in 2015) which is well above the point at which positive difference becomes 'visible', and within the 'zone of desired effects (2009). However, of the studies Hattie cites only the most effective contained student-to-teacher feedback. If we keep in mind Goodrich (1996) and Ross (2006), and their call for teacher monitoring of self assessment, we can see that this is where pupil to teacher feedback can be located. For Hattie too, feedback needs to relate to specific learning goals and must be acted upon through revision, requirements echoing those suggested by Goodrich (1996) and Ross (2006). Importantly, Hattie suggests that feedback is most effective when 'at, or just above, the level where the student is working' (2009, p177). This highlights a potential difficulty of feedback in peer assessment. It will take a

certain level of expertise for peers providing feedback to avoid the trap of being too personal: i.e. ‘you are great’ (Hattie, 2009). While peer assessment might benefit from allowing greater peer interaction (Gielen et al, 2010), it risks increasing only the interest in and attitude towards a task, rather than performance in it (Deci and Ryan, 1994). A method will need to be designed to moderate the feedback peers give each other.

There is a need to be mindful that in seeking a Herbarmassian ‘emancipation’ of self and peer assessors, and in demonstrating that the transfer of trust (Usher and Panjeres, 1998) from teachers to pupils is real, we do not create a situation where confusion reigns over the chance to develop a real understanding of the assessment criteria and how to meet it. For example, this could occur if there is disagreement over how to carry out assessment, with the disagreement overshadowing any development in understanding. This could be achieved by following the advice of Nichol (2006) and using tightly self regulating models of peer-to-peer feedback to keep conversations focussed. Kollar and Fischer (2010) agree with this notion, and identify feedback provision and reception as distinct stages of the peer assessment procedure which, if interactive, guard against the recipient’s potential passivity.

Formative assessment in RE.

The literature on self and peer assessment tends to be generic in nature with only a limited amount specific to RE. This is probably because formative assessment in RE is

not unique, but rather can take advantage of ‘generic stock’ to grow a more ‘subject specific sub-species’ (Brooks et al, 2012). A study into formative self assessment in RE carried out by Nigel Fancourt noted that one element of RE as defined by policy makers, requires pupils to learn *from* RE. Fancourt points out that this requires engagement in ‘evaluation of beliefs and values’ (2010). Fancourt identifies a similar theme in Robert Jackson’s reflexive approach to RE. This prescribes a constructive critique of material along with a running critique which is self critical and methodological (Jackson, 2004, cited in Fancourt, 2010). It is arguable that this is precisely what is required of the students in order to gain top marks at GCSE. They must be able to state how a given reason, belief or practise fits within the broader teachings of a faith; this requires the constructive critique called for above. Fancourt terms this ‘reflexive self-assessment’, (2010) and states that an optimal model of this in RE will require pupils to assess not only their knowledge, but also (amongst other elements) ‘how effective they are at evaluating the views and values of others’ (2010). Thus, while the process of self and peer assessment may be largely generic across subjects, we can see that it allows pupils to engage in what have been deemed important aspects of engagement within RE as a discrete subject.

What does peer assessment offer than self assessment cannot?

I am keen to encourage increased interaction and collaboration between pupils so that they are exposed to a wider range of examples of extended writing. Thus peer assessment will take primacy over self assessment in my intervention. There are

multiple advantages in terms of feedback and competency afforded by peer assessment. For Topping (2009), the key benefit of peer assessment is the quantity and immediacy of feedback it affords. Peer assessment and its feedback ‘is plentiful’ (Topping, 2006; Gielen, 2011), because it can happen more regularly than teacher assessment - peer assessors making up for in time available what they lack in assessing skill. This process affords ‘multiple cognitive gains’ (2009) including better error spotting and questioning. This, argues Topping, provides students with the skills to identify what high quality work looks like (2009).

A decade earlier than Topping, however, Boud et al (1999) had concerns about group vs individual focus in peer assessment, asking whether it can be shown that the experience of reflection which peer assessment attempts to produce carries into individual practice, or if the method is instead tokenistic, risking individual achievement? Their overall judgement of peer learning was positive despite concerns about whether the process or the end product is the focus of endeavours. They highlighted numerous reasons to instill peer and group assessment and learning, with a focus on increased opportunity for reflection and communication within the subject area, leading to what Zimmerman (2002) has termed ‘competent disciplinary practice’. When pupils work collaboratively to assess each other’s work, they gain access to a wider variety of examples (Reinzhold, 2016), allowing identification of work at different levels. Reinzhold termed these beneficial aspects of peer assessment ‘awareness, performance awareness and gap closure’ (2016). In practice, a pupil struggling to consistently

perform at a higher level can see what performance at that level consists of. This process also ‘promotes self assessment by making otherwise invisible assessment processes more transparent’ (Ibid, 2016)). Here we see how closely the two systems of formative assessment are related.

It may also be the case that peer assessment’s advantages are owed to the removal of the teacher from the process. Gielen (2011), found that peer assessment benefits secondary pupils because they share a ‘wavelength’. Gielen argues that the gap in intellect between peers is less great than between pupils and their teacher and thus explanation of the process is more accessible to the recipient, though quality control must presuppose use of a method of regulating feedback, as called for by Nicholl (2006). It is also possible to suggest that peer assessment may succeed by offering individuals of different learning dispositions the opportunity to benefit from each other. For example, those lacking dispositional resilience, possessing instead ‘brittleness’ (Carr and Glaxton, 2004) might benefit from a more reciprocal environment, where peers become involved in their learning via the formative feedback peer assessment affords. Carr and Claxton point to evidence that learning can flourish when carried out in reciprocal relationships with others (ibid, p111). Neil Mercer’s large scale study on the use of dialogue between pupils in lessons concluded that ‘being drawn into more extended explanations and discussions’ provides a suitable opportunity for practice of using language as a tool for thinking (Mercer and Littleton, 2007).

Peer assessment is not a ready made panaceia, and alongside the potential *challenges to implementation* discussed above, Aoun et al (2016) warn that it may be necessary to structure such discussion and reflection in order to make it truly valuable. This echoes calls from Nichol (2006) and Goodrich (1996) for structured methods of feedback, especially important as scholars such as Hogg (2017) state that her students were likely to be unfamiliar with peer assessment strategies. That this is a concern as recently as 2017 reminds of warnings (Bennet, 2011; Black, 2018) that an agreed and widely understood model of peer assessment is not yet formulated.

Pupil perspectives on peer assessment.

How an assessment model is perceived by pupils is likely to impact its effectiveness in improving outcomes. Topping (2009) quotes findings from a study by Sadler and Good in which as many as 70% of peer assessed grades agreed with those of teacher assessors. This reliability is not always reflected in the views of pupils. Deluca (2018), in a study of German secondary schools, found that peer feedback was viewed as less helpful than teacher feedback in helping progress. Ngar-fun's Hong Kong based study made similar findings, where lack of peer expertise was perceived. Topping (2017) argues that once such concerns are overcome, the benefits of peer assessment are likely to be high. Gielen et al (2010) found that as few as 23% of participants felt they benefited from peer assessment. Li et al (2019) attempted a meta-analysis of studies on peer assessment and found that progress was hindered where trust in peer expertise

was low. They recommend thorough training in the process in order to get the kinds of gains Topping claims are possible. One model suggested in the work of DeLuca (2018) is to focus feedback on process, with suggestions about how improvements can be made, rather than grades. This would echo the ideas of Boud (1999) and Reinzhold (2016) that the reflection and discussion afforded by peer assessment is one of its key benefits. All of this points to a need to clearly communicate the aims of the strategy to my pupils, and to ensure that modelling and monitoring of the process takes place.

The benefits of peer assessment for differing ability levels

Peer assessment will only help to improve pupils' extended writing skills if it is engaged with. Otherwise, it may only assist those already engaged and succeeding while compounding the struggles of those who require the most assistance to improve their extended writing. In their review of the research, Williams et al (2019) found that both 'high' and 'otherwise achieving' students showed a preference for and enjoyment of collaborative learning. They report no difference between differently achieving groups in sensitivity to the 'free rider effect' (a term attributed to Orbell and Dawes, 1981). Gielen et al (2010) suggest rating peers' contributions can lessen opt-out. Williams et al (2019) found that higher achieving pupils showed no preference for the ability levels of peers in their group, although there was often a preference for friendships to be recognised. Research carried out by Chichekian and Shaw (2017) found that more 'gifted' peers enjoyed the conflict that can be produced in collaborative work, and enjoyed 'holding

firm' when disagreement arose. This suggests that I might hope to find pupils willing to commit to the uncomfortable act of challenging a peer's misconceptions. Cautionarily, though, Cera Guy at al. (2019) found that fewer high-achieving students had confidence that their 'overwise-achieving' peers would contribute their share of work and complete tasks. As with the process and quality of feedback, I will need to monitor the groups' performance. The same study also found that less able pupils are likely to have lower expectations of the social interactions they will experience, and of the amount of work expected of them in group work situations. Given the nature of my investigation, I will need to monitor the contributions of lower achieving pupils.

In cases where a pupil is yet to demonstrate the extended writing skills required by the GCSE, can collaborative peer assessment strategies aid progress? Gielen et al (2011) found that performance on tasks which will be peer assessed is high because of increased external motivation to do well, where the desire to appear competent in front of peers might trump a lack of internal motivation. Other studies, including work from Cole (1991) and Gibbs (1999) support the idea that pupils will care more about the quality of work they present to their peers. Teacher assessment allows for discretion, so that poor performance is only seen by the teacher and poor feedback only seen by the pupil. Relationships are also important, and the possibility of greater familiarity between peers than between teacher and pupil is a reason to expect both that pupils will work harder to impress their peers, (Cole, 1991) but also that they are more likely to take feedback on board. Gibbs (2006) argues that a major problem with teacher to pupil feedback is that engagement with it is low. Maclellan (2001) studied a broad range of

tutor to pupil feedback, finding it rarely prompted discussion, reflection or understanding. Others have found evidence that pupils will try to develop strategies to avoid having to share work due to ‘fear of competition’ (Fancourt, 2013).

What might happen when pupils work together to assess each others' work?

Sociocultural insights into how peer assessment can help develop extended writing?

Peer assessment can be scrutinised via Vygotskian sociocultural theory, which places emphasis on learning as a social activity, Vygostsky arguing that, 'human learning presupposes .. a social nature ..a process by which children grow into the intellectual life of those around them' (1978, p88). By encouraging dialogue between pupils, GPA may encourage such growth.

Vygotsky sees development as 'the transition from interpsychic to intrapsychic..i.e., from the social, collective activity..to..more individualised activity (1962, p133). Interpsychic is communication from one person to another; intrapsychic is when the individual has an orienting 'conscious understanding' (Ibidm 1962)) of a concept, and is able to work with it independently. Importantly, as a justification for attempting peer assessment, though the presence of a demanding task does not in itself cause the development required to solve it, Vygotsky is clear that the absence of challenge will at least delay development of thinking at its 'highest stages' (1962, p59). Before proposing the inter to intrapsychological model, however, Vygotsky offers a detailed description of how concept formation takes place.

The Zone of Proximal Development as an aid to concept formation.

Vygotsky proposed the Zone of Proximal Development (ZPD) as the location of the development of concept formation. In ‘Mind and society’ (1978) the ‘ZPD’ is defined as ‘the distance between the actual developmental level ..determined by independent problem solving and the level of potential development as determined..under adult supervision or in collaboration with more capable peers’. The key is that a learner will be able to do something with assistance which they cannot do on their own, but the potential for independent work will be there. Reviewing work with a more capable (or perhaps a better prepared pupil) would, it is theorised, likely highlight errors and prompt the pupil to take her writing in a more accurate direction, one which was unavailable to her when working independently. Vygostsky talks in terms of the ‘buds’ or ‘flowers’ of development, which by definition precede the ‘fruits’ of independent functioning (1978, p86). The bud shown in the example outlined above is that the pupil is dealing with the culturally alien concept of reincarnation, which, although not a Buddhist concept, at least shows the ‘bud’ of awareness of alternative beliefs about life after death.

The different functions given to the two external agents in this first depiction of the ZPD are interesting. That the *adult* ‘supervises’ while the *more capable peers* are ‘collaborators’, matches roles allocated in the literature on self and peer assessment previously discussed. The Vygotskian scholar Seth Chaiklin (2015) points out that Vygotsky’s particular interest was in how the development of external functions of

cultural development (of which writing is one) takes place. Chaiklin says that Vygotsky conceptualised development as progressing from one configuration of psychological functions to another. This reminds us of Cole and Gibbs' (cited in Gielen, 2010) suggestion that I might expect feedback from peers to be viewed differently to the instructions that a teacher may give, if peers are indeed 'on the same wavelength' (Gielen, 2010), working towards the same goal. Those who can produce extended writing remain closer in the chronology of their development to those who cannot, and so are perhaps well placed to assist their peers make 'appropriate content based generalisations (Chaikiln, 2015) required to be able to write evaluatively about religious attitudes to theological, ontological and societal claims.

How the ZPD actually enables development is helpfully clarified by James Wertsch (1979) who explains that development is about self-regulation of ego-centric speech (Wertsch, 1979, p3). For Vygotsky, egocentric speech involves 'no interest in his interlocutor, ... to communicate', (1962, p15). Perhaps those yet to produce extended writing are involved in ego-centric thinking, because they don't interact with the questions. To develop the thought which must precede developed extended writing requires interaction. This brings us back to the location of higher mental function's development, which Wertsch summarises (1979) as occurring first on the social, 'interpsychological plane' before becoming internalised on the individual 'intrapsychological' level. This transition takes place via the kind of imitation and practise that my intervention aims to provide. Involvement in peer assessment might encourage this imitation. Neil Mercer suggests that collaborative learning between

peers might allow development via co-construction of ideas, whereby individuals within a group benefit from exposure to the ideas and strategies of peers (Mercer, 2013). This can lead to ‘transformation’ of their subsequent thinking, so they ‘begin to reason intramentally in a more dialogic way (ibid, 2013). The process of collective reasoning forms a ‘template’ which allows individual reasoning to develop with a greater degree of self-regulation, as suggested by Wertsch (1979).

Spontaneous and scientific concepts in GCSE religious education

The developmental transition from interpsychological to intrapsychological occurs to an individual's understanding of nonspontaneous ‘scientific concepts’ (1962, p86) which are introduced via a formal source, such as an educational setting, and thus on the ‘interpsychological plain’ (1962, p133). Scientific concepts have been defined as of a ‘historical, linguistic or practical’ nature (Fancourt, 2014). These differ from what Vygostsky terms ‘everyday’ or ‘spontaneous’ concepts (1962, p106), which he argues lag behind the development of scientific concepts in some cases (Ibid). Vygotsky compares ‘because’ with ‘although’, as examples where the developmental pace differs. The suggestion is that in both cases, the developmental stage reflects frequency of practice ‘in collaboration with the adult’ (1962, p107). ‘Because’ and ‘although’ are the basic form of analytical thought required for successful extended writing in GCSE RE. The pupil needs to be able to take the scientific concepts (e.g. samudaya, resurrection, nirodha, the golden rule), and use them in tandem with their everyday concepts of evaluation. Thus, the everyday concepts of justification, comparison and contrast need

to ‘proceed upward’ into the conscious thought, where they may meet the scientific concepts as they travel in the opposite direction (p107). As this meeting takes place, the egocentric thought of the pupil can operate with these concepts.

Applying models of concept development in RE

Attempts to make use of Vygotskian theory in assisting the development of extended writing is to assume that evidence exists that the explanations offered by Vygostky a century ago have been used effectively with modern learners. Anne Edwards developed a ‘quadrant model’ (2014) whereby learning events can be structured to promote development. ‘Quadrant 1’ of this model requires the introduction of key ideas and the modelling of engagement with them. This would equate to familiarisation with the mark scheme and exam technique. ‘Quadrant 2’, tightly structured tasks requiring engagement with the subject material manifests as the writing of exam answers. ‘Quadrant 3’ introduces the ‘twin terrors of ambiguity and risk’. This equates to the act of peer assessment, requiring individuals to both expose their work, and engage in critical dialogue about the work of others. This is where Mercer (2010, 2014) suggests we look for ‘exploratory talk’ which is indicative of ‘transformation’ on an individual level (2013). Quadrant 4 requires demonstration of learning, which in this case may be the application of what has been discussed, via application of the kind of developing template proposed by Mercer (2013).

There has been further development and innovation based on Vygotsky’s ideas. The notion of ‘insider’ and ‘outsider’ concepts was introduced by Nigel Fancourt (2014) to

distinguish between scientific and everyday concepts within RE. Fancourt proposes a quadrant in which ‘outsiders’ (from a particular faith) need to be brought into contact with the conceptual framework of the insider (to Sikhism, In Fancourt’s model). For Buddhism, an outsider (non-Buddhist) may possess the everyday concept of compassion, which to the insider becomes ‘karuna’, while an equivalent scientific concept may be that of ‘justice’, which could equate to ‘karma’ for the insider. I suggest that we can apply this version of the quadrant model if we consider the ‘insiders’ to be the more capable peers, while the ‘outsiders’ are those who have not yet demonstrated the ability to think as required by the specification in their extended writing. Peer assessment will build on the outsider everyday concepts of ‘because’ and ‘although’ as discussed by Vygotsky (1984), and allow, through repeated peer assessment, opportunities for more able peers to demonstrate how the scientific concepts of the faiths studied can be organised within the everyday concepts of ‘because’ and ‘although’ to enable pupils to develop the kind of extended writing skills required by the specification.

Sociocultural theory and the development of writing.

The development of writing is seen by sociocultural theorists as a critical developmental stage. Vygotsky sees it as marking a ‘critical turning point’ (Vygostsky, 1978, p106) in a child’s cultural development. He sees written language as developing from the codification of spoken words, which themselves codify real entities and relations (p106), towards a state where spoken language is removed and written language ‘directly

symbolises the entities and relations' (p106). Though at this stage focussing on the early stages of writing development, Vygotsky urges that writing be 'relevant to life' and arouse an 'intrinsic need' in developing writers. It is possible to argue that preparation for public examinations provides this dose of relevance and urgency.

Writing, which Vygostsky sees as a form of inner speech, comes after social speech, which is 'spontaneous, involuntary and non-conscious' (Vygostsky, 1986). Thus, written work is a good indication of how developed a pupil's understanding of concepts is. In order to write anything, the writer needs the concept they are writing about to have sunk to the intrapsychic level. Once thoughts about concepts can be written down correctly, abstract thinking has been acquired (Thompson, 2012). Peer assessing written work, by encouraging pupils to engage with each others' work, should serve a dual purpose. Firstly, talking about what constitutes a good written answer, and how examples can be improved, should help pupils to develop their understanding of scientific concepts (religious ideas and knowledge), and everyday concepts (justification and comparison) required through allowing regular use of relevant language in oral speech. This will allow for the 'continual movement back and forth from word to thought' (1962, p124) which allows developmental change in the relation between the two. Secondly, it will provide the mediated activity, in the form of a 'supportive but challenging learning environment' (Thompson, 2012) which encourages writing. All the while, the 'more expert other' (Thompson, 2012) i.e. more advanced peer or teacher, will be present. In their study 'Thinking Together', Mercer and Littleton (2007) found that introducing

opportunities for ‘collective reasoning’ (2007, p52) enabled participants to become better at solitary reasoning. They identify ‘exploratory talk’, where knowledge is shared, ideas are challenged and evidence is evaluated (2007, p62) as an indicative that discussions are productive. It will be possible to listen in to conversations to measure the extent to which this kind of talk dominates over the less productive ‘disputational’ or ‘cumulative’ variety, in which discussions are, at worst, confrontational, or at best, uncritical (2007, p59).

This idea of working and thinking together changes the social roles of the pupils, so they ‘play out’ (Leonti’ev, 2006) the role of assessor rather than remaining passive in their role as learner. Backhurst (2009, 2018) argues that the key element of involvement activities is that they cannot be engaged in unintentionally. While a pupil may be able to remain passive in a lesson, once they are placed into a different role, such as that of assessor, they need to engage in that role consciously. The activity of assessing thus becomes an object of mediation.

Risks of adopting a sociocultural model of group peer assessment

There have been warnings that Vygotsky’s ideas are too often seen as a quick fix to complex problems. Chaiklin points out that the ZPD is a zone of ‘development’ rather than ‘learning’. This point is relevant to my investigation because it is not necessarily learning, in the sense of acquiring key ‘substantive knowledge’ (Schwab, 1978; Thompson, 2014), but rather developing their way of using the knowledge which they

should have acquired to produce exam literate extended writing. It will also be important to be mindful of two other warnings from Chaiklin (2003). Firstly, each individual will have their own subjective ZPD. Because development takes place via imitation, and a 'person can imitate only that which is within her developmental level' (1978, p88) it will be important for me to ensure assessment takes place between peers whose developmental levels are different, but not so different that the less developed cannot benefit. Secondly, Chaiklin points out that Vygostsky 'never assumed that learning (in the) ZPD is..enjoyable' (2003). Thus I will need to remain mindful of the usual obstacles to learning which exist in each individual's everyday life, such as their cultural and material circumstances, relationships and so forth. I should not expect a miracle.

Questions arising from the Literature

My search for a form of collaborative peer assessment which can assist pupils in developing their extended writing will be guided by the following research questions which have arisen from the literature.

Research question 1: Does peer assessment in groups help pupils to produce better extended writing?

Research question 2: Does structuring feedback help the development of extended writing skills?

Research question 3: Can working in *mixed ability* groups help the development of pupil's extended writing skills?

Methodology

Overview of the intervention

In response to my first question, to assess the extent to which peer assessment in groups can help pupils to produce better extended writing, I decided to devise a model of group peer assessment (GPA). This would introduce to lessons regular episodes of group peer assessment focussed on the 12 mark 'd' questions from the Edexcel Religious Studies specification, which require extended writing. A record would be kept of scores achieved before, during and after the intervention to identify any changes. Questionnaires, conducted prior to the intervention, would look out for any obstacles to its success (explained further in the following section, page 41-42).

To answer my second question, the extent to which structuring feedback helps in the development of extended writing skills, feedback forms (appendix 5) would be devised in order to guide the conversations around assessment. This is described in more detail on page 46-7.

To evaluate the impact of mixed ability groups, structured observation and recording of the groups participating in the intervention would help me to identify the contributions of individuals of different abilities. This is explained further on page 47-48.

The intervention was to be carried out in collaboration with the head of department. We are a department of two, and she was fully supportive of using the department's time to explore potential methods of helping pupils develop their extended writing skills because she shared many of the concerns I outlined in the introduction to this paper.

Preparing the ground for Group Peer Assessment

This small scale classroom based intervention was to be conducted with two year 11 classes. The ability profile of both groups was mixed. The calculated grades of pupils in the class, as determined by Fischer Family Trust indices, range from grade 9 down to grade 3. Most pupils are predicted to meet their calculated grades, although those at the lower end of the ability profile, particularly 3 to 4 and 4 to 5 borderline show less consistency in their performance. If the intervention works, it is the extended writing of these pupils who will be improved the most. High prior attainment pupils should benefit by further internalising, via discussion and explanation to their less high attaining peers, the elements of successful extended writing in the context of GCSE RS.

Cognisant of the problems of perception of peer assessment (e.g. Boud, 1999; Gielen et al, 2010; Li et al, 2019), my first step was to canvass all pupils' perceptions via questionnaire (see appendix 4). This was intended to identify any objections or deal with negative perceptions which could have prevented the model from helping pupils produce better extended writing.

Questionnaires were designed to fit on two sides of A4 and to be quick and easy to complete, key recommendations in increasing response rates according to Edwards et al. (in Coe et al. 2017). The use of a combination of closed questions, useful for forming comparisons and generalisations about data (Oppenheim, 1992) with open ended questions, especially for follow up and explanation avoided the problem that closed questions alone may restrict answers (Ibid, 1992) while still allowing reluctant writers to participate, avoiding the pitfall noted by Wilson and Mclean (1994) that open ended questionnaires may discriminate against the less 'articulate'. Questions around perceptions of and familiarity with peer assessment were necessary because it would impact the level of preparation needed pre intervention. If a strong resistance to the method were identified, more justification might help engage pupils. If William (2018) is right that formative assessment is not yet an established practice, then I am at risk of assuming knowledge or viewpoints that the pupils may not have (Cohen, 2018 page 383).

Questions about the confidence of pupils when assessing the work of others, and their views on having their own work assessed would allow me to identify where to focus my preparation, so that the time dedicated to the intervention could be best used. If I were to discover a consistent perception of lack of subject knowledge, for example, then time could be spent revising concepts central to the question about to be peer assessed. If concerns centred on knowledge of the mark scheme, time could be directed towards further familiarisation. This would give the GPA model a better chance of succeeding.

Although no pupils were forced to engage in the questionnaire, participation was encouraged. I decided to not to guarantee anonymity and non-traceability, factors suggested by Cohen et al. (2018) as influencing engagement. This is because I felt it would be useful to be able to search for patterns between confidence in being an assessor, or attitude towards the method and identification of pupils would allow me to place these patterns in the context of other factors such as target and mock grades. Ultimately, this study seeks to understand how a group peer assessment model assists development of individual's extended writing; thus, the individual students are the units of observation (Coe, 2017), and the aim is to benefit them by improving their performance in extended writing. However, the option of anonymity was made available to meet the obligation of non-maleficence (Cohen, 2011, p85). It is possible that students could feel uncomfortable offering their honest opinions underneath their name. Allowing for anonymity would afford the additional benefit of potentially increasing engagement, though it risked the possibility that the answers would not be as seriously considered. This position allowed pupils the option of maintaining their privacy, highlighted as 'primordial value' by Cohen (2011, p90), while also allowing them the option of relinquishing it, which respects the self-determination of the participants.

Designing the group peer assessment model

What will be assessed, how and by whom?

Stage 1 - devising the assessment criteria

In response to the importance placed in the literature (Goodrich, 1996; Ross, 2006) on clarity of criteria and task, the first step was to provide the grade boundaries. The exam board's level descriptors (appendix 1b) were unsuitable for use in group peer assessment with a year 11 class, with a readability score being 17.89[2] (indicating years of formal education required for comprehension). Only a minority of pupils read at this age when last tested. To overcome this, I worked in collaboration with my co-researcher to re-word the document, eventually achieving a readability score of 12 years. This much lower score meant that deciphering the grade descriptors posed less of an obstacle to the task. The pre and post alteration grade descriptors are attached in the appendix (2a and b).

Stage 2 - gathering assessable material

The next step was to decide what each group would assess. Initially, I intended for groups to work on collections of their peers' answers. An ethical concern was that pupils had no way to opt out of this model. It removed agency from the pupils, leaving them unable to deploy avoidance strategies such as those noted in the work of Fancourt (2013) where the subject underperforms to avoid being chosen for peer assessment. This clearly inverts the recommendation that individual dignity is given priority

[2] Checked via Google

consideration over potential knowledge to be gained (SSHRC of Canada, in Cohen, 2018). Given that some pupils had expressed, via initial questionnaires, a lack of enthusiasm around having their work peer assessed, to persist would risk breaking the 'golden rule' of ensuring that participants' situations are not diminished by the research (Cohen, 2018, p85). However, I decided that the potential benefits of pupils sharing their own work identified in the literature (i.e Cole, 1991; Gibbs, 1999; Gielen et al, 2010) were worth the risk that some students may be embarrassed, especially because the aim of the exercise was to enhance their performance rather than diminish their situation. Pupils would assess each others' work.

Step 3 - constructing groups

To literature suggests a range of benefits should arise from the GPA model with mixed ability groups (Research Question 3), including an increased motivation to write well (Gielen et al, 2010; Bandura, 1997) an increase in emancipatory knowledge (Boud and Brew, 1995; Terry, 1997), overall improvement in learning, (Black and William, 1998; 2009: William, 2018; Bennet, 2011, DeLuca, 2017) and increasing complexity of thought and reasoning via a social interaction with more capable peers (Vygotsky, 1962, 1978; Chaiklin, 2003). To help achieve this, it struck me that composition of groups deserved careful planning.

The most informative data available were the FFT target grades and mock exam results of the students. If the peer group assessment model were to become a zone of proximal development (Vygostsky, 1962, 1978), then each group would require a 'more capable

peer' (1978). The FFT calculated/target grades used in conjunction with mean CAT score data and individual teacher knowledge of their class could be used in combination to tailor groups which would allow for potential development. The ideal group would have at least one pupil who had achieved in their mock a high target grade, a second who had failed to meet a high target grade and a third with a lower target grade, whether they had achieved it or not. Initial surveys would probe for concerns pupils had about peer assessment, and any arising could be taken into consideration. It was also important to collaborate with student support to ensure that there were no circumstances preventing individual pupils from working with one another. Such issues might include past incidences of bullying, for example. While I was keen to encourage the kind of conflict that Chicheckian and Shaw (2017) found more able pupils thriving on, I didn't want to promote the negative perceptions that Cera Guy et al. (2019) found that the more able had of their less able peers. Aware via 'student bulletins' which contain need to know information about issues concerning pupil wellbeing, that ability based bullying is not unknown at Academy X, this was a concern. The 'golden rule' of not diminishing the pupils situation remained mindfully present throughout the planning process.

Feedback pro-formas

To investigate the impact of structuring feedback on developing better extended writing (Research Question 2), feedback pro-formas were designed. The literature (e.g Hattie) points to the importance of immediateness, intimacy and comprehensibility of peer

feedback. It also (Nichol, 2006; Kollar and Fischer, 2010) points to the importance of structure. To this end, and in collaboration with my co-researcher, a feedback proforma was designed which would allow pupils to identify which if any of the mark scheme criteria had been met. There was also space for pupils to take notes on key points from discussions. This was designed, in part, to reduce any potential 'free-rider effect' (Orbell and Dawes, 1981), as all pupils would be required to record something on the feedback sheet. This would allow me to have a clearer idea of the level of engagement of each pupil, as it would be impossible to observe every single interaction in every episode of peer group assessment. The pro-forma is attached in appendix 5. Due to the school closures, none were actually completed. The forms adopted a 'plus-minus-interesting' model for recording the positive, negative and developmental aspects of work assessed. This was intended to enable all pupils to engage. Keeping it simple would reduce the risk that the already complicated task of assessing a peer's work could be perceived as adding further challenge. I remained concerned that adding additional paperwork of this sort might risk shifting the exercise away from discussion towards box ticking, and was prepared to intervene to remove the feedback sheets if this appeared to be happening.

Structured observation of group peer assessment in action.

Listening in

Although there would be the potential to measure any difference that peer group assessment model made to extended writing by assessing work produced by pupils

post-intervention and then comparing it with their pre-intervention work, this data alone would not be sufficient evidence of development taking place due to or during group peer assessment. I decided that observation of a range of groups engaged in group peer assessment would enable me to find evidence that could support any claim that pupils' knowledge and understanding is deepening.

I was keen to record sessions of group peer assessment. Being able to listen to and transcribe the pupils' discussion would provide a rich pool of data for analysis. It would be easier to avoid the pitfall of selective attention (Cohen, 2011) if all groups were recorded, which would have been time consuming, but transcribing even only the discussions of selected groups would mean that my attention would be across every individual in the recording. Recording would also prevent me from forgetting what was said and by whom. This was a difficult proposal, however, as when broached with pupils, in the interest of good practice overt observation (Cohen, 2011) and informed consent, many were uncomfortable with the idea. To record audio covertly would have been possible but highly unethical, and detrimental to the principle of informed consent and the privacy and agency of the participants (Cohen, 2011, p471). I have never used recording equipment with the class before, and doing so risked upsetting the 'natural setting' of the classroom (Cohen, 2011, p472). Recording has been noted to induce anxiety in participants (Ibid, 2011, p470), and indeed could produce a Hawthorne effect (Cohen, 2011, p186) with the pupils modifying their behaviour in reaction to be recorded. Were a pupil to remain silent throughout an activity, this would deprive the study of any evidence either way as to its impact, and would instead render the

investigation as one into the effect of recording equipment. I was aware that reactivity can produce ‘demand characteristics’ (Shaughnessy, cited in Cohen, 2011), in which a pupil’s behaviour adapts to fit what they think the observer is looking for (in this case, positive engagement and discussion of extended writing against mark scheme criteria). This effect could have benefited the experiment if otherwise reluctant speakers adapted their behaviour in this way, but I realised that a non-recorded ‘observer-as-participant’ (Gold, 1958: cited in Cohen, 2018) model could also produce this effect. My observer role took on the following dimensions as outlined by Cooper and Schindler (cited in Cohen, 2018): it was direct because my presence (and that of my collaborator in her classroom) was required and it was overt, because the pupils knew I was engaged in observation.

My ‘observer-as-participant’ status acknowledges the fact that the observations were carried out overtly, and that it was likely that the pupils would in some way seek my approval or guidance as I observed. Although the aim would be to remain ‘unobtrusive’, my role as the class teacher would never afford ‘complete detachment’. (Cohen, 2018. p457). The potential pitfalls of this method centre around how to record all relevant data, and the idea that my presence risked counter-transference (Cohen, 2011). For example, would my recording of an pupil’s contributions, or lack thereof, be impacted by my previous perceptions of that pupil? This would be problematic because the group peer assessment model would quite likely be a new scenario for the pupil, and thus their reaction to it deserves to be treated separately from their behaviours elsewhere. Ultimately, my concerns around which model of observation to choose remained

hypothetical because schools closures were enforced before any observation could take place.

Data collection

The validity of the data collected would depend upon my ability as an observer, and it would be necessary to carry out multiple observations in order to practise recording data accurately (Cohen, 2018).

I decided to limit each observation window to three minutes, which would begin after the work had been read. The focus would be on the kinds of contributions made by each pupil in the group. In each group, pupils would be labeled as follows:

Pupil A: high target grade (<6), target exceeded in mocks
Pupil B: high target grade (<6), target met in mocks
Pupil C: lower target grade (<5) target not met in mocks.

Information about sex, ethnicity, language status and pupil premium would be included because of the possibility that it could explain behaviour in the group assessment activity, for example a higher ability male from a wealthier economic background might be predicted to be more confident at this kind of activity than a lower achiving pupil premium student. Boys and recipients of Pupil Premium are groups of focus for the school's development plan, so any evidence about their performance in this activity

would be of immediate value to the wider school. However, this investigation is about the effectiveness of group peer assessment in developing extended writing skills, and so any findings related to these demographic factors will be a secondary concern.

My observation form was to record occurrences of particular types of utterance from each pupil. Known as ‘event sampling’ (Cohen, 2018). Every time a pupil spoke, a tick was placed against the relevant category of speech, as follows:

a) Comment about the strength of the work
b) Comment about a weakness of the work
c) A comment about how the work could be improved
d) A comment containing subject specific terminology (substantive knowledge (Scwabb, 1978))
e) A comment containing mark scheme terminology (syntactic knowledge (Schwabb, 1978))
f) A comment suggesting development of reasoning (possibly something the pupil has realised for the first time)

Tracking these comments against the speaker’s position in the group (i.e. are they pupil A, B or C) would help answer my third research question, as dominance of the more capable members might suggest that mixing the ability profiles in the group inhibits the participation of the less capable members.

Potential limitations of the GPA model.

It was important to remain aware throughout that this exercise is exposed to multiple limitations. Cohen (2018) warns that the method is behaviourist, and thus an absence of contribution cannot be interpreted as an absence of development or engagement, as I can infer nothing about the motivations of the participant. Because of the method of group composition, it is unlikely that the pupils will be with friends, and this could influence their participation levels. Moreover, it must be noted that even if utterances indicative of a development of understanding do occur, it represents only a 'freeze-frame' (Cohen, 2018). It would be necessary to follow up the observations with further questionnaires and interviews about pupil views on the intervention's effectiveness, because they may feel they are benefitting from both group peer assessment and working in mixed ability groups.

Post intervention questionnaires and interviews and ethical concerns thereof.

The use of follow up interviews would allow me to explore how the participants had felt about engaging in group peer assessment, and its overall impact on their understanding of what is required when carrying out extended writing in GCSE RS. The focus is thus on the 'lived experience' (Coe, 2017) of the pupils. The interviews can be considered collaborative, because the input of the pupils would affect my conclusions (Coe, 2011), and explore the group peer assessment model from the perspective of the participant,

and give an insight into the process from the participant perspective. Interview participants could be drawn from those who had chosen to identify themselves in their questionnaires, based on whether they liked or disliked peer assessment. This would allow me some level of control over the different biases likely to be expressed (Cohen, 2011). It would also be possible to plan questions which covered a range of variables which could positively or negatively affect responses, as demonstrated by Morrison (cited in Cohen, 2011). For example, if pupils maintained a negative attitude towards group peer assessment, it would be possible to frame questions in such a way that other factors (friendships, lack of confidence, low self esteem) could be explored. This would allow insight into the real motivation for responses. My primary motivation in interviewing was to discover how they found the experience of working in a group to peer assess, how useful the feedback pro-formas had been, and whether the re-worded mark scheme had helped. The interviews would be recorded with the permission of the pupils, and where consent was not given, that interview would not be conducted. This decision was to ensure that transcription could be carried out. There are a range of issues specific to interviewing children and young people, one being the ‘perceived asymmetries of power’ (Cohen, 2011). The main ethical concerns which arose when planning interviews concerned the privacy of group members and the self image of interviewees. Having students discuss a group activity risks the interviewee saying something negative about a peer, or feeling coerced into doing so. It is perhaps quite unlikely that some pupils would truthfully report other group members not participating appropriately. It would be possible to frame questions in such a way to avoid the pupil

feeling that they had offered information. For example, if interviewing a member of a group in which pupil X had been observed remaining silent, the question could address this directly. ‘Tell me who in your group didn’t do the work’ could be replaced with ‘I noticed pupil X didn’t participate. How did this make you feel?/Why do you think this was?’. Secondly, as with the planning of groups, it is a risk that a pupil might self-identify as the ‘less-able’ member of a group. This would, at the very least, take away their autonomy and privacy (if we consider that others might identify their place in the group structure). It would be important to be prepared to deal with any challenges along these lines with a justification based not on ability, but performance data.

Participants and their consent

Consent to carry out this intervention was received from the school’s Headteacher (see appendix 3a). CUREC approval was granted on 28/10/19, reference ED-C1A-20-22. The consent of the pupils was informal, but at least partially meets the criteria of informed consent as defined by Diener and Crandall (cited in Cohen et al, 2019). Voluntarism is evident because the pupils were free to disengage in the study, and could have opted out simply by talking. Full information about the purpose of the study was shared with the pupils, so all were aware that the aim was to improve their extended writing. This contributes to the comprehension and competence aspects of informed consent, although the level to which each of these criteria can be met is proportionate to the developmental level of the pupil, as noted by Fine and Sandstrom

(in Cohen, 2018). A key ethical concern arising early in my planning was that pupils may feel exposed by having to share their work. Those who know their work is of lower quality than their peers may feel targeted or victimised, as in ‘rape research’, (Sikes, 2006, cited in Cohen 2018) which prioritises the aims of the researcher, offering no benefit to participants. This may not apply here, as the aim is to help the participants, but the consideration did lead to a change in my original plans, as discussed below.

Further Ethical Concerns

In engaging the pupils in group peer assessment, an ethical concern arises over their agency. Those pupils who expressed a dislike of group work, or a mistrust of peer assessment are likely to feel that they are being ignored. While Usher and Panjeres (2005) talk of the sense of trust peer assessment can transmit to pupils, it may in this case appear that their views have been ignored. Warwick (1982) argues that researchers need to respect human freedom to make and act on choices. Time will need to be spent explaining to these dissenting pupils why this has not been done in this case. The evidence from my literature review in favour of both peer assessment and the benefits of working with more capable peers ensures that I can claim ‘equipoise’, because evidence suggests that this intervention is highly unlikely to harm the progress of pupils (Coe, 2017). As all pupils will participate in this small scale study, in both year 11 classes, it also meets BERA (2011) guidelines that no one group is advantaged or disadvantaged compared to another. It has been pointed out (e.g. Fancourt, 2017) that researchers can respond to ethical dilemmas during an episode of

research. Were any particular problem to arise from the grouping model, I would be able to change the design and composition of them. According to Hersch (2018) it is because this intervention is carried out as research that requires me to aim to achieve equipoise in this way, to ensure that no individual suffers due to the group they are in. Equipoise is also achieved by the fact that both GCSE groups are involved in the intervention, via collaboration with my departmental colleague. This means that neither are at any disadvantage attributable to the intervention when sitting the same exam.

I was also concerned that pupils might notice the formation of the group, and self-identify as the ‘less able’ or ‘less high achieving’ member. This could potentially contradict my responsibilities to protect children’s well being as per the *Every Child Matters* (DfES, 2004) policy, although the same policy requires teachers to enable children to ‘achieve’ and make ‘positive contributions’, both aims of this intervention.

The impact of the coronavirus school shutdown and subsequent alterations to the study

The Coronavirus lockdown of spring 2020, which caused schools to close to all but a very limited number of children from 20th March 2020, removed any realistic chance to complete the planned intervention. With the cancellation of GCSE examinations, even if schools had reconvened in May, year 11 would only have had one week left until their published leaving date and pupil motivation to engage with group peer assessment would undoubtedly have been severely impacted.

Consequently, of the various elements of intervention planned, it was possible to carry out only some. Preliminary surveys of student opinion were conducted via questionnaire when rumours of an imminent lock down were rife. A significant number of pupils and their families had entered self-isolation. This impacted the number of surveys completed.

The planning and formation of groups that would allow for the benefits discussed in the methodology section was possible, though it remained theoretical, school closures occurring prior to any group work being conducted. There was a value in using a combination of data and personal knowledge of the pupils to devise groups, and I had never previously given quite as much thought to the allocation of groups. This also allowed for collaboration between myself and my departmental colleague, in that we

were able to discuss the optimal composure of each group, and share knowledge on the relationships between different individuals.

A further change took the form of a survey of teaching staff. This focussed on their attitudes towards assessment for learning, of which group peer assessment is a form. It was prompted by the large variation in student questionnaire responses about the frequency of peer assessment activities. I believed that surveying teachers would provide a wider range of data with which to work, and help me understand a more varied range of experiences of conducting peer assessment. This may have been useful to answering RQ3 in particular.

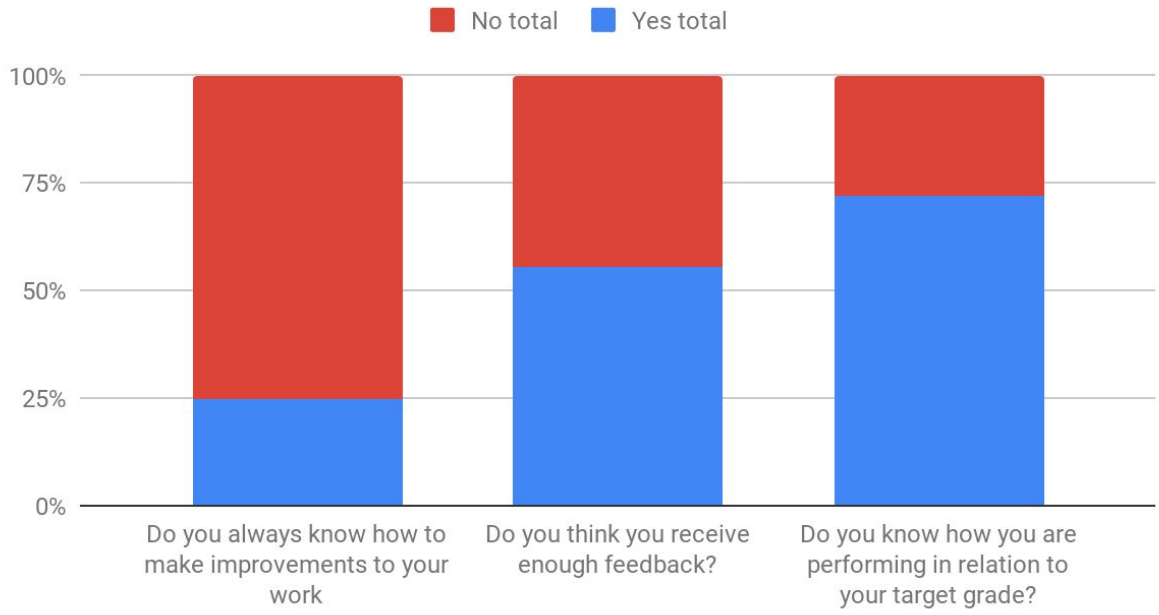
The teacher survey was conducted beginning April 15th, which was the fourth week of school closure, and the second week of Easter holidays. Despite this, the response rate was acceptable, with approximately one quarter of teaching staff responding. The survey was fully anonymised, being conducted using Google Forms. There was no obligation to answer any question, and the function requiring questions to be answered was not used. The main ethical issue arising from this survey is that it exposes a number of teacher traits. For example, when asking about the frequency with which they use AfL, teacher's everyday pedagogical choices are exposed to scrutiny. Similarly, when asked about obstacles they face in conducting AfL oriented activities, options included concerns about pupil enjoyment of AfL. This question was motivated by the literature (e.g. Kluger and DeNisi, in Bennet, 2011; Boud, 1998) but risked suggesting

that teachers are overly guided by their pupils preferences. I felt that this question would be insightful for my third research question on the effectiveness of mixed ability groups, because The guarantee of anonymity allowed teachers to answer truthfully, as evidenced by the answers given. The results of the pupil questionnaires and teacher survey are presented for reference in the following pages.

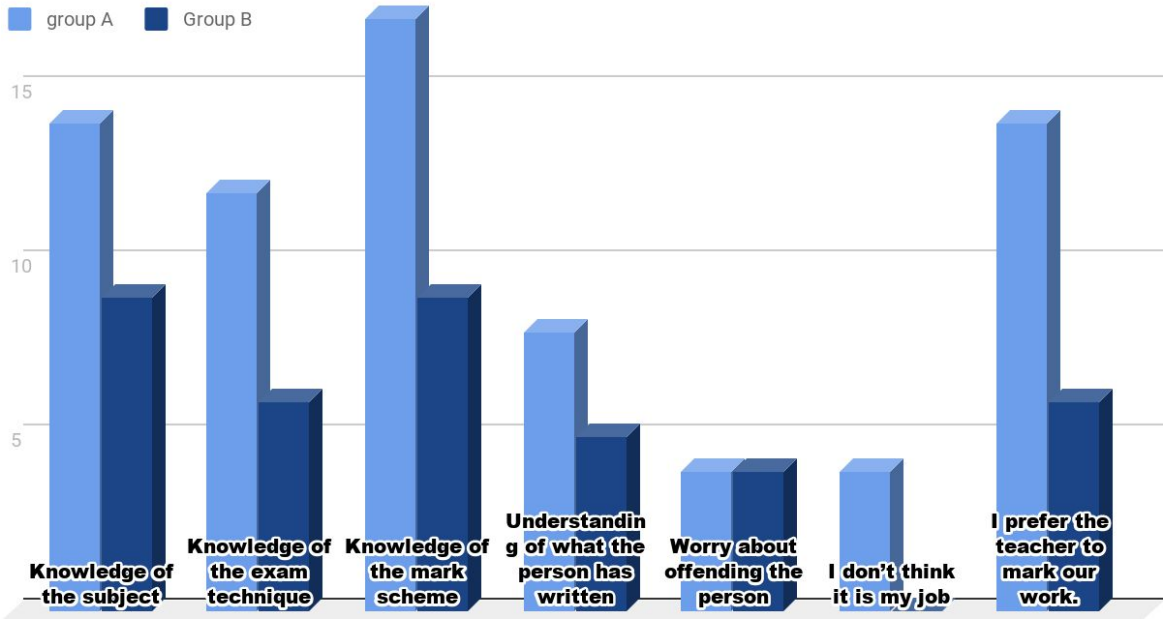
Tables and data

Selected pupil questionnaire results.

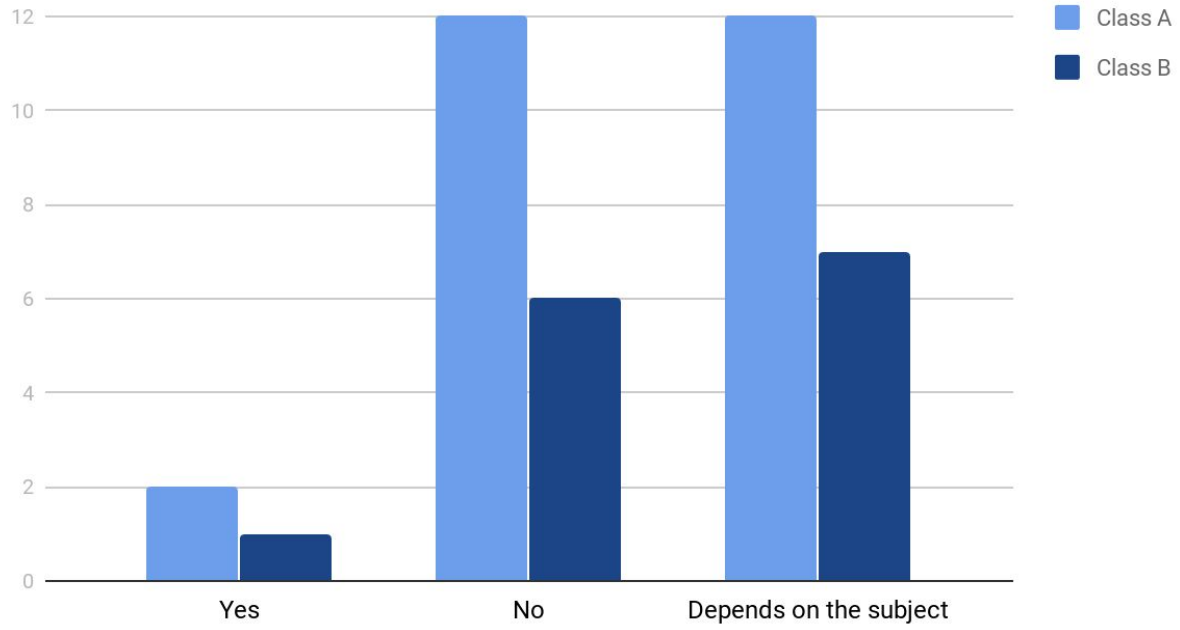
Graph 1: results from group A and group B



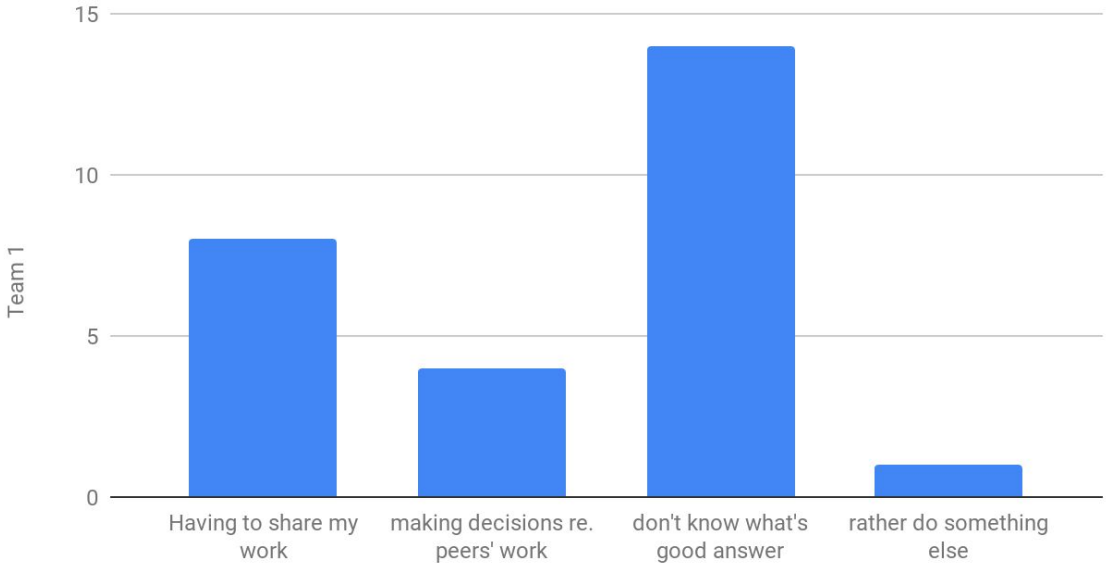
Graph 2: Reasons for answer to question 9 (Do you feel confident when marking somebody else's work?)



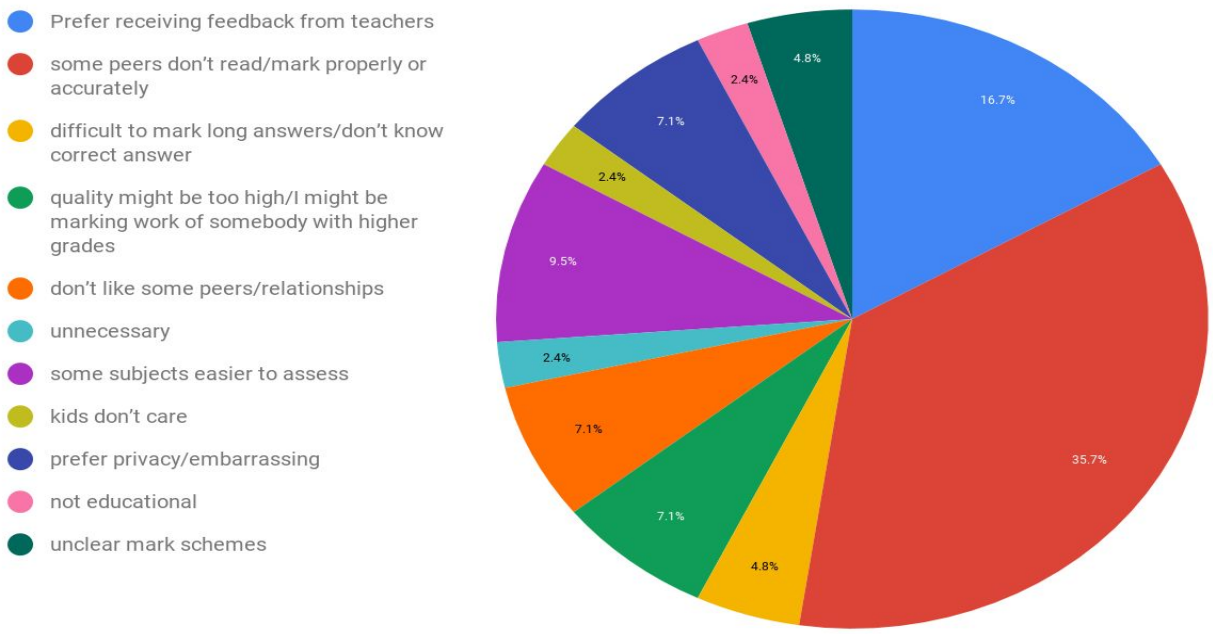
Graph 3 - Question 5: Do you like peer assessment?



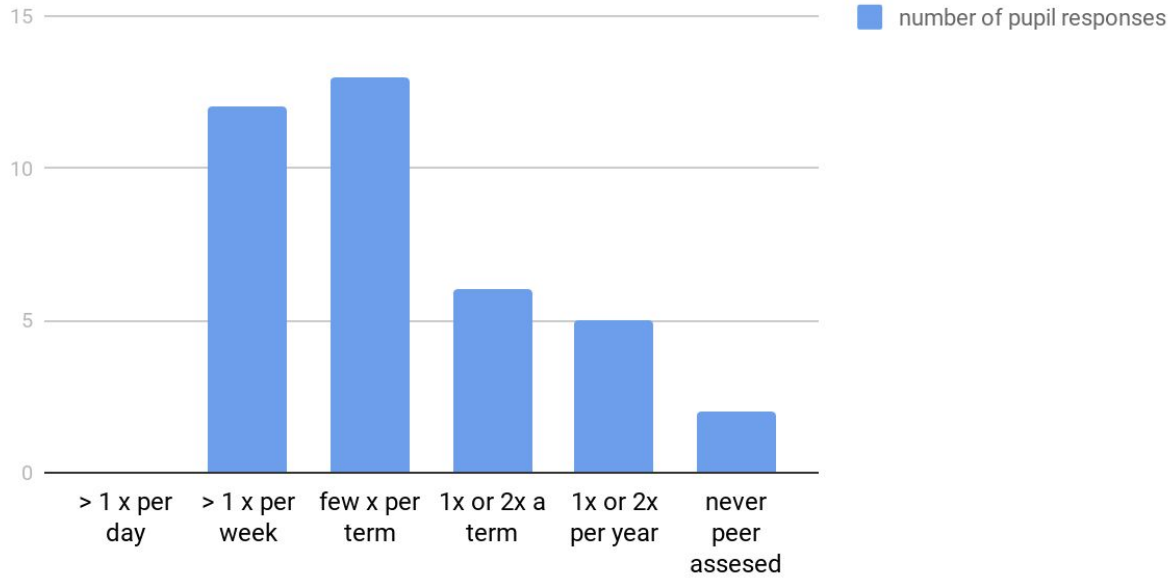
Graph 4 - question 11: What is your biggest concern when doing peer assessment? (group A)



Graph 5 - reasons given for liking/disliking peer assessment.



Graph 6 - question 4: How often are you asked to carry out peer assessment?

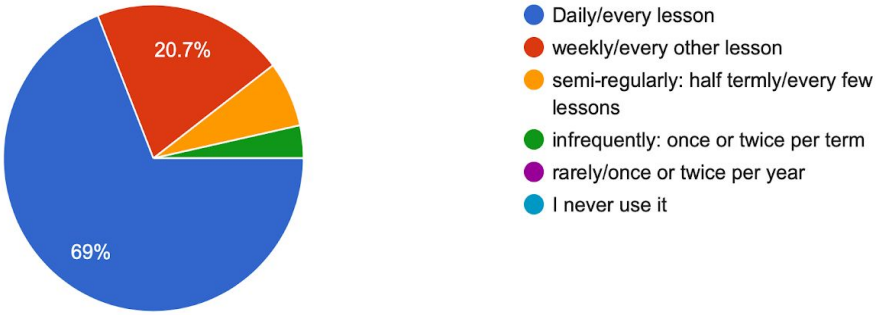


Teacher questionnaire results

Question 1

Which of the following best describes the frequency with which you use AfL?

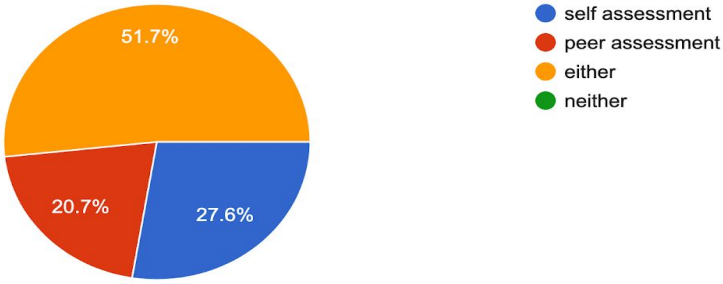
29 responses



Question 2

Do you prefer to use self or peer assessment

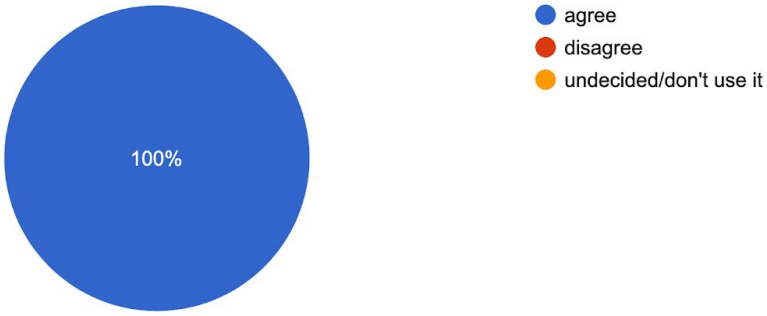
29 responses



Question 3

AfL is an effective pedagogical tool

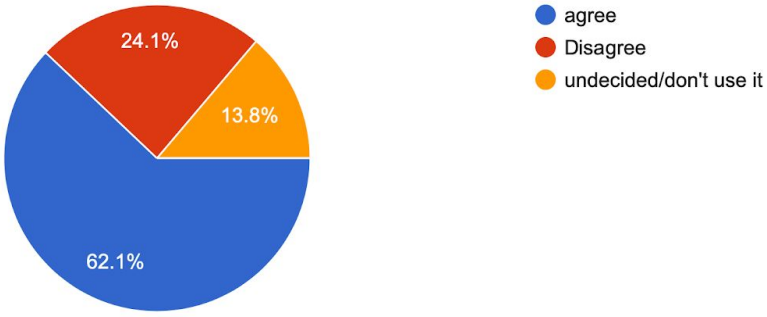
29 responses



Question 4

I am confident in the ability of students to self assess

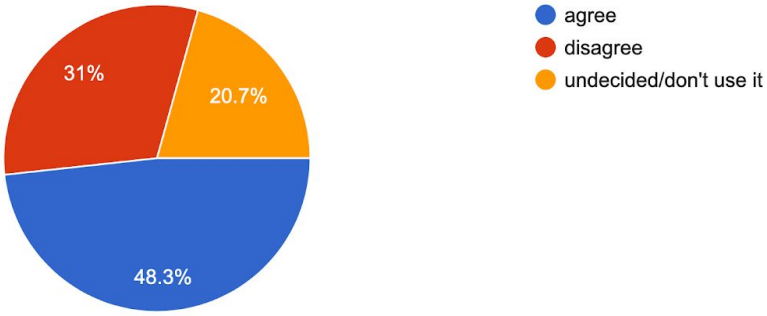
29 responses



Question 5

Students value peer assessment

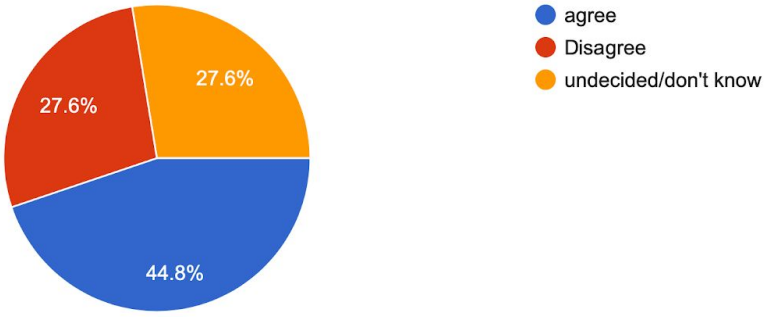
29 responses



Question 6

I feel supported (e.g. via INSET) in using AfL effectively

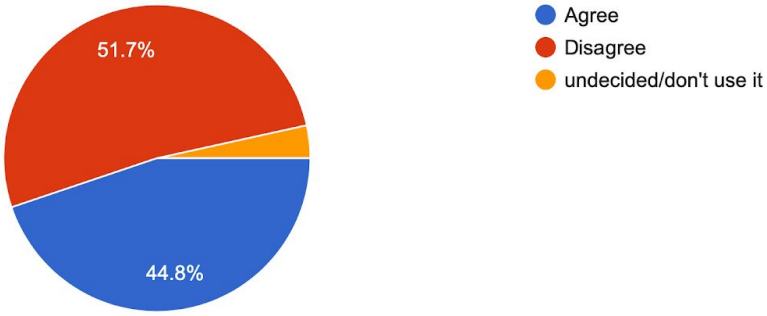
29 responses



Question 7

AfL benefits my higher attaining students more than other groups of students.

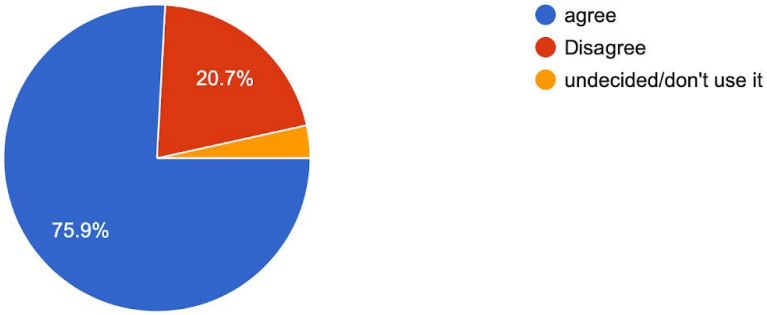
29 responses



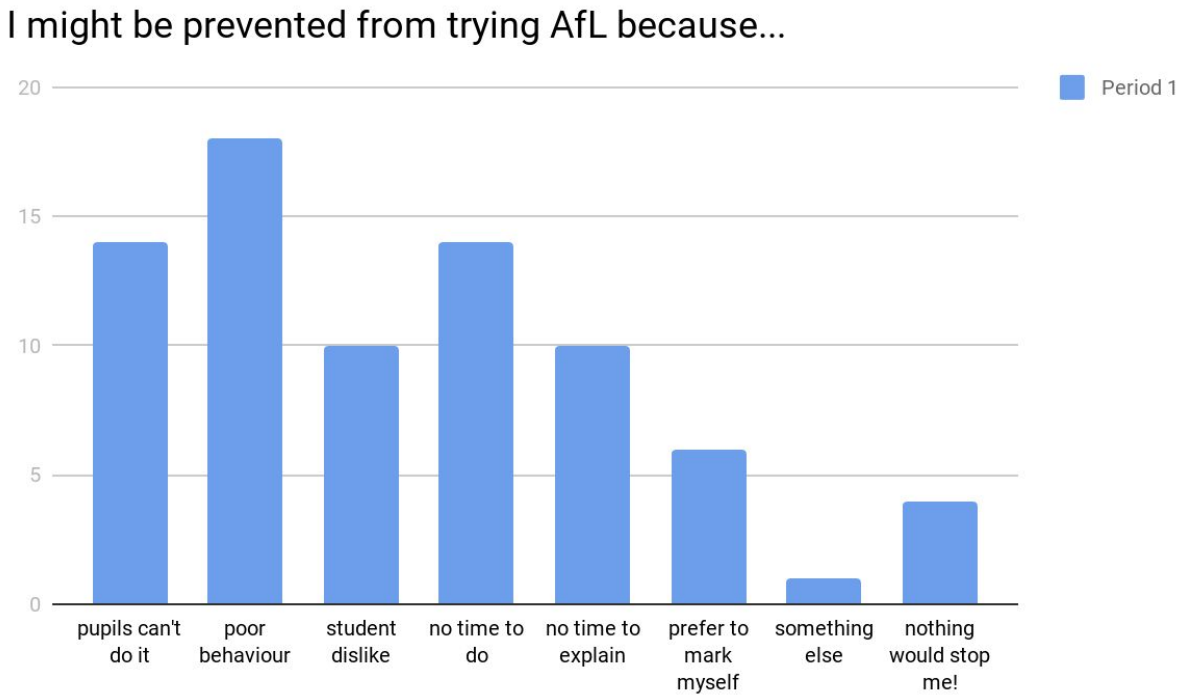
Question 8

AfL produces worthwhile feedback for students.

29 responses



Question 9


I might be prevented from trying AfL because...
Period 1
20
15
10
5
0
pupils can't do it
poor behaviour
student dislike
no time to do
no time to explain
prefer to mark myself
something else
nothing would stop me!

Findings and discussion

Research Question 1: Does peer assessment in groups help pupils to produce better extended writing?

The total data available is limited to that harvestable from a total of only 36 completed questionnaires. The intention was for both of year 11 examination groups to participate in the intervention, and I worked closely with my departmental colleague to prepare for this. By the time questionnaires were carried out, significant absences had reduced the numbers of pupils in each group. Two classes of 30 were reduced to 25 in class A and only 14 in class B. Three questionnaires were not completed. Overall, this equates to a response rate of 60%, which according to Coe et al (2017) is actually a higher rate than is common in social science (p229), and would be considered high enough for publication, at least by the standard of the Canadian Medical Association Journal's submission policy (Coe at al, p228), and is also higher than the 50% rate with which Cohen et al (2011, p403) suggest I should be satisfied, although the diminutive size of the total sample is a limiting factor. However, this is still a very small amount of data, so knowing that school closures were looming, I tried to fortify this with some casual observations of pupils working in groups and informal interviews with some students, though there was no time to either record this group work, or to use the prepared

feedback records. Nonetheless, it is possible to draw some tentative inferences from the data available.

The majority of this small group of pupils do not think they know how to improve their own work by themselves. 75% disagreed that they ‘always know how to make improvements to (their) work’. Upon reflection, this question (Graph 1, P60) is poorly worded, and may prompt a variety of interpretations. ‘Always’ invites a negative response from pupils who may very often know what to do, but are self aware enough to know that they will struggle even if only sometimes. It is possible that the suitability of the wording decreases in inverse proportion with the critical reading ability of the pupil. A more succinct question, i.e. ‘do you know how to improve your own work?’ would have been better. Despite this, the point still stands that these pupils cannot help themselves, so perhaps there was fertile ground for the intervention.

In normal circumstances, the composition of the questionnaire would have been subject to revision. Any opportunity to reword and repeat this specific question, even by an informal show of hands, was also lost once schools closed. The opportunity for greater collaboration would have also been beneficial. Two meetings were planned to discuss the structure, wording and focus of the questionnaire; the first with my departmental colleague and the second with my line manager, the school’s deputy head teacher, an English teacher. These meetings may have resulted in a much stronger final questionnaire. In reality, the first ‘meeting’ was replaced by some last minute emails to

my departmental colleague, who offered her ideas for questions. It is difficult not to regret overlooking the practical consideration suggested by Cohen et al (2011, p402) of running a pilot of the questionnaire at a much earlier stage, which would undoubtedly have flagged some of these issues, though it was impossible to predict the substantial loss of teaching and research time.

Returning to the question, the pupil's apparent lack of ability to improve their own work suggests that an attempt to improve extended writing skills via a group peer assessment model may have been beneficial. If we agree with Young (1999) that schools exist to instill 'powerful knowledge' (1999), and that this can be characterised by the ability to find relations and generalisations between aspects of theoretical knowledge, then in any examined subject, writing, often extended, is the only means through which the acquisition of such knowledge can be meaningfully demonstrated.

As the intervention was curtailed, it is not possible to provide evidence of any difference GPA may have made to the extended writing produced, though a variety of benefits may have been available. It is arguably through discussion and writing, and the thought processes involved in putting ideas into syntactically accurate form, that meaning, the absence of which is bemoaned by Conroy et al (2010), is made. Answers to question 10, which asked for reasons for confidence or lack of confidence when marking somebody else's work (graph 2, p60) show that pupils' concerns were very much focussed on what could be considered the intellectual challenges of the course. This

contrasts with Conroy's assertion (Ibid, 2010) that RE lacks intellectual challenge, a criticism also levelled as recently as 2018 by Mary Myatt in her call for a more demanding form of RE (Myatt, 2018). It seems that absence of challenge and rigour is not the experience of this group of pupils. In the light of assertions from scholars such as David Lundie (2018) who asserts that the RE's curriculum content has, due to too heavy a focus on knowledge, become too far removed from the lives of pupils, it may be suggested that the proposed group peer assessment model would have brought discussions closer to the student's lives by encouraging them to re-engage with questions they would otherwise have left behind once complete. There is the opportunity for them to approach the material in a different way, which may have changed their experience of it. These discussions could have potentially translated into better extended writing as a residue of their exposure to more examples and assessment focused discussion.

The absence of any episodes of group peer assessment means that it is difficult to make judgements about how the pupils would have responded to the activity. Evidence of their response and the work produced would have been of great benefit in being able to confirm the extent to which GPA helps to develop extended writing skills. It cannot be ignored that a mere 3 of the 36 (p61) surveyed pupils responded affirmatively when asked if they like peer assessment. Pupil dislike of an activity shouldn't be a reason not to attempt it, and of the reasons given for dislike of peer assessment, to be explored in the following section, nine would have rendered the activity unethical. This is particularly

true when so much literature points to the positive impact it has on learning (Black and William, 1998; 2009: William, 2018; Bennet, 2011, DeLuca, 2017).

Teacher and pupil views on the benefits and drawbacks of assessment for learning

We have seen that the literature strongly supports self and peer assessment, both of which are fundamental aspects of group peer assessment, despite concerns over issues such as pupil scepticism (e.g. Boud, 1999). Such scepticism doesn't appear to have gone away in the intervening years, and may have proven to be a hindrance to helping improve pupil's extended writing.

Concerns about peer assessment found in the work of Boud (1999), such as not knowing how to assess, are echoed in my findings. Graph 2 (page 60) shows pupils felt unable to engage with the mark scheme, just as Boud (1999) observed. This may well have interfered with the intervention's aims of helping pupils produce better extended writing, though the feedback forms were intended to assist here.

A lack of familiarity with and practise of peer assessment may have impacted negatively on the GPA model. Boud (1999) suggested that negativity towards peer assessment may be rooted in pupils' perceptions of such variants of assessment as 'ad hoc' or 'tokenistic'. Again, there appears to have been little change here. Graph 6 (page 63) shows the regularity with which pupils claim peer assessment occurs in their lessons. Despite the small data set, it is obvious that there are some stark differences in how often pupils consider themselves to be engaged in peer assessment. A wide range of

reasons could be responsible for these differences. It is probably most easily explained by the fact that GCSE RE is an optional subject, and therefore each pupil has a different complement of teachers across their timetable. It appears that some teachers are using peer assessment frequently. This variation in pupils' perception of regularity of engagement in self or peer assessment supports the view that such assessment is still lacking consistency as an educational tool, which agrees with Dylan William's assertion as recently as 2018 that 'systematic use' assessment for learning 'remains the exception rather than the rule'. It is likely, therefore, that the intervention would have had to overcome challenges around pupil perception of the model as well as their ability to actually carry it out before prompting any significant improvement in extended writing skills.

Further evidence that lack of experience of peer assessment may have hindered the intervention's effectiveness at helping pupils produce better extended writing can be gleaned from the teacher survey. Though the majority of teachers (89% - Q1, p64) of respondents claimed to use AfL daily or weekly, there were a range of concerns raised around various aspects of delivery. Given the amount of literature which supports the efficacy of AfL (e.g. Black and William, 1998; 2009: Topping, 2006, 2009, Reinholdz, 2016; William, 2018; Bennet, 2011, DeLuca, 2017) it is surprising, but according to William not unusual, that Academy X doesn't have a firmer policy on the delivery of AfL models. Tellingly, though staff believe that AfL works (100% agree it is an effective pedagogical tool! - page 65) and only 5% claim to use it infrequently, they feel

unsupported in their efforts. Only 37% felt supported in their practice via INSET (p66). Just under half of teachers disagreed that AfL benefited more able pupils over the less able, putting the slight majority at odds with some of the literature (Chicheckian and Shaw, 2017; Cera Gut et al, 2019) who noted benefits for the more able. With only 48% of teachers agreeing that pupils value AfL (p66), over 41% saying that students dislike it and 50% saying that they lack the time to do it, (p68) it does appear that William's (2018) warnings about systematicity remaining the exception in AfL are reflected in the whole school delivery of AfL related activities at Academy X. When considered in the light of graph 3 (p61), it could be suggested that there may be a discoverable correlation between uncertainty and lack of confidence and time on the part of teachers and the qualified preference (depends on the subject) expressed by pupils. It is highly likely that this general uncertainty and lack of clarity around the use of AfL would have impacted upon pupil's perceptions of the GPA model, and though the effect of this cannot be known, it is at least possible that it would have been negative.

Recalling teachers' faith in AfL, shown in the total agreement that it is valuable (p65), teacher concerns about lack of time to conduct AfL are interesting, because the literature (e.g. Andrada and Valcheva, 2009; Andersson and Palm, 2018) suggested that interest in AfL was motivated by belief it could save teachers time. The 50% (Q9, p68) of teachers citing lack of time as a limiting factor on conducting AfL again points to a potential need for training or dissemination of a shared model. When we consider De Luca's (2018) assertion that AfL remains largely undefined, it is perhaps no wonder that

some studies find that teachers pursue it with the hope of saving time, while half of my colleagues at Academy X find that they lack the time to do it. A valuable follow up question here would have sought to identify what the time desired for AfL (again bearing in my 100% regard it as valuable) is spent doing instead. Anecdotally, my experience of teaching GCSE Religious Studies, Sociology and English Language over the last two years suggests that it is quantity of content that is a major drain on time. Indeed, it is partly this that led to launching the intervention later in the year, and thus missing the chance to carry it out.

Research Question 2: Does Structuring feedback help the development of extended writing skills?

Frequency of feedback was highlighted in the literature (Bloom, 1969; Black and William, 1998, 2009; 2011; Hattie, 2009; DeLuca, 2018) as a key element, and advantage, of effective AfL, with Black and William (1998) placing feedback as central to formative assessment (AfL). It was an intention of this intervention, via a group peer assessment model, to create an additional feedback stream. Graph 1 (page 60) shows that more students agree than disagree that they receive enough feedback. Because it is not explicit, it cannot be specified whether this was interpreted as meaning only in GCSE RE, or across the curriculum. As over 88% of teachers claim to be conducting AfL on a weekly basis, we might expect pupils to be including feedback gleaned via AfL rather than feedback from teachers alone.

Something which cannot be confirmed from the evidence gathered is whether the kind of AfL being conducted contains the ingredient considered essential by Hattie: student to teacher feedback, which must require the attention of the teacher in some form or other, either via monitoring feedback in progress, or collecting self or peer assessed marks/grades. Recalling that almost 50% (Q9, p68) of teachers surveyed didn't feel they had enough time to conduct AfL properly, and only 42% (Q6, p66) felt properly supported via training in doing so, perhaps many staff are using forms of self assessment which remain unmonitored or uncollected. Graph 5 (p62) shows that 16% of pupils cited a preference for teacher feedback, and 35% were not satisfied with the quality of feedback from their peers. The second of these figures is broadly in line with the numbers of pupils dissatisfied with the amount of feedback received overall. This could point to a feedback vacuum, where AfL is being used, but there is no interaction between teacher and pupil. Goodrich (1996) and Ross (2006) both called for close monitoring and intervention of teachers in the AfL process, and Hattie's meta-analysis shows us that AfL which contains feedback from the pupil to the teacher is the most effective kind: feedback proformas would allow for this. It is not possible to say for sure what format of AfL is being used so regularly at Academy X, but by comparing the answers to question 5 (p66) of the teacher survey (46% of pupils value peer assessment) with question 8 (75% agree that AfL produces valuable feedback) we might reasonably infer that self assessment is more common, and perceived as more beneficial by teachers, even though only 30.8% chose it exclusively, as shown in the results for question 2 (p64).

Over half of pupils surveyed said their biggest concern when peer assessing was not being able to identify what constitutes good quality work (see graph 4, p61). Here we see a lower level of confidence on the part of pupils than their teachers have in them. Again, it would have been useful to know what is being assessed, because self or peer assessment of, for example the conjugation of 10 German verbs is clearly very different to that of extended essays. In the first case, where the answer is either correct or incorrect, self and peer assessment appear straightforward, easy and time efficient. When it comes to extended writing, even experienced examiners can end up at logger-heads over the interpretation of a mark scheme. Thus, even providing a structure for peer to peer feedback may not definitely have helped the development of extended writing skills.

Potential of feedback proformas

Had I been able to conduct the practical stage of the intervention, I would hope to have been able to show that there is some benefit to using the simplified versions of mark schemes and feedback proformas which were produced in collaboration with my departmental colleague. The literature, (including Goodrich, 1996; Nicholl, 2006; Kollar and Fischer, 2010, Aoun et al, 2016) recommend use of both tightly defined criteria to which assessment should be carried out and use of self-regulating and easily monitored modes of peer to peer feedback provision. Having taught at KS4 in RE, sociology and English at Academy X, I know that there is not a current shared strategy of peer

assessment using either of the tools devised for this intervention. There is the potential that their deployment may assist in reducing teacher anxieties around time (Q9, p68) by providing structure to the AfL activity. There is also the chance, as outlined by Kollar and Ficher (2010) that having a task to do (completion of the feedback proforma) might also reduce teacher anxieties around behaviour (Q9, p68), by increasing engagement. Simplification of the mark scheme also directly addresses pupil concerns about their knowledge in this area (graph 2, p60) in line with the findings of the scholars mentioned at the beginning of this section.

Research Question 3: Can working in mixed ability groups help the development of pupil's extended writing skills?

Once pupils are working in groups, they are ideally working together, free of the direct instruction of a teacher. The idea that pupils can be 'emancipated' from a reliance on their teachers, explored by various scholars (e.g. Terry, 1997; Boud and Brew, 1995) finds only limited support in my data to support the use of mixed ability groups. Emancipation requires pupils to 'know why' (Terry, 1997). In the GCSE classroom this means pupils being able to understand what work looks like within different grade boundaries. The aim of the group peer assessment model was to expose pupils to a wide range of work at different levels to enable the high prior attaining pupils to both deepen their own substantive and syntactic knowledge (Scwabb, 1978) while helping their less able peers to improve their own understanding through peer assessment. These less able peers would in turn hopefully make greater leaps in their own

substantive and syntactic knowledge. We cannot say whether this would have actually happened, as the group work could not be conducted, however certain comments are possible. 62% of teachers (Q9, page 68) were concerned that pupil's behaviour would be problematic, which to some extent contradicts the claim of Brooks et al. (2011) that immaturity is not a barrier to success in AfL. This is reflected to a lesser extent in the views of pupils. Question 6 of the pupil questionnaire (graph 5, page 62) allowed for open responses, of which two expressed the idea that peers do not take AfL seriously. In an informal discussion with one pupil, she commented '*they don't think that marking each others' work is work, so they don't care'*. There is certainly a chance that negative or disengaged behaviour in a mixed ability group could limit the extent of any potential development to pupils' extending writing.

From the teacher surveys (Q9, p68) we see that 50% of respondents said they would be prevented from conducting AfL activities because of a lack of confidence in pupils' abilities to engage with the mark scheme. 35% also worried that there was no time to properly explain how self or peer assessment should be carried out. Despite 100% of teachers (Q3, p65) agreeing that AfL has value as a tool of pedagogy, there are significant numbers who seem concerned that the syntactic knowledge required for 'emancipation' is neither present nor perhaps even achievable, due to lack of time. Regarding this latter point, if considered in unison with the 58% (Q6, p66) unable to agree that they feel fully supported via INSET to carry out AfL, we can again see that there may well have been room for this intervention, which if successful could have

provided an adaptable model for both teachers and pupils to become accustomed to using. Certainly we cannot claim, as an organisation, to be anywhere near the promised land hypothesised by Boud (1999), in which students engage in a critique of the standards to which they are being held. If anything, group peer assessment may have placed too much responsibility on the high prior attainers, whose syntactic and substantive knowledge is assumed. We have no evidence that the low prior attainers would have listened to their peers in a mixed ability group peer assessment model.

We can see in graph 2 (p60) that there were significant pupil concerns regarding knowledge of content (substantive knowledge - Sshwab, 1978), exam technique (syntactic knowledge - Ibid, 1978) and mark scheme (emancipatory knowledge - Habermas, 1987), the ‘knowing how’ (Terry, 1997) piece of the puzzle. It was the latter of these which was of the most concern, and in both groups around 50% of pupils expressed this as a concern. If nothing else, this at least shows us that teachers and pupils agree where AfL’s difficulties of implementation exist, and suggests an area of focus for future endeavours. Perhaps the steps to demystify the language of the mark scheme by reducing its reading age is one simple step that might have been shown to be of value had the intervention taken place, and assisted the high prior attaining pupils in promoting the development of the less able peer’s extended writing in a mixed ability groups.

There are other potential pitfalls posed by using mixed ability groups. For Goodrich (1996) time to practise and repetition of tasks was key to successfully developing competence in self and peer assessors. With 85% (Q1, p60) of teachers claiming to conduct AfL on a daily basis, we would expect Goodrich to be surprised that 50% (Q9, p67) of these teachers still lack confidence in pupil's ability to assess successfully. There is an interesting contrast here with answers to Q4 (page 65) which suggested over 60% are confident that pupils can peer assess, and this can be explained by the wording of the question, where question 9 seeks factors which may prevent teachers from using AfL. The contrast here may suggest that teachers do AfL despite concerns about pupil's abilities, which supports the views of Andersson and Palm (2018) and Andrade and Valcheva (2009) that it is often a time saving tool.

Teachers concerns may have been about superficiality of pupils' focus, which was observed by Ross (2006). Indeed, in response to open ended question 6 in the pupil questionnaire, (graph 5, p62) over 35% of pupils suggested that their dislike of peer assessment was due to their peers not reading, being able to understand or being able to properly mark their work. This contradicts Topping (2006) and Goodrich (1996), who suggest that practise is a vital ingredient when it comes to AfL. It doesn't appear that daily practise is leading to a corresponding confidence on the part of teachers or pupils that a community of competent assessors is being forged, and we have no reason to think that mixed ability groups would have overcome these challenges during this intervention.

At this point we might find answers in the folk wisdom of the pupils themselves, almost 10% of whom offered the unprompted response that ‘some subjects are easier to assess than others’ (graph 5, p62). It is very possible that the 75% of daily AfL facilitators are using it for quantitative subjects such as maths or science, or for easier to mark knowledge tests. Here again I find myself regretting the absence of a question interrogating the type of work assessed regularly. It might be that the 50% who acknowledge doubt about pupils' abilities to carry it out are thinking about more complex assessment when they choose this option. When asked directly about overall confidence in their students abilities to *self* assess, a smaller proportion (35%) responded negatively. It does seem likely, though, that mixed ability groupings would have at least exposed lower prior attaining pupils to different kinds of explanations about how to improve their extended writing, whether this is because it comes from a mind on the ‘same wavelength’ (Gielen, 2010), the ‘more capable peer’ (Vygotsky, 1978) acting as the nourishment via which the ‘fruits of independent functioning (Ibid, 1978) can flourish.

Because I chose not to heed the recommendation of Cohen et al (2018) to guarantee anonymity and non-traceability of the questionnaires, some pupils chose to include their name. This revealed a trend with potential to impact on the group peer assessment activity. In a small number of identifiable cases (4 boys and 3 girls), lower achieving boys, as defined by both calculated grades and current performance data, reported that they *did* know how to improve their work, whereas the similarly attaining girls reported

that they did not. In the case of the boys, the fact that work is consistently of a standard below that to be expected by the calculated grades suggests that these boys are overconfident in their abilities. Worryingly for the group peer assessment model, these same individuals also expressed negative views towards peer assessment and the feedback it creates. This doesn't seem to support Gielen's (2010) views about peers valuing the feedback of those on the same wavelength. It would have been interesting to see how they responded to feedback from their better attaining peers, and the question arises as to whether Gielen and her fellow researchers would consider differently attaining peers to share a 'wavelength'. Evidence taken from samples of their work suggests that some of these boys only command very limited amounts of subject knowledge. There is very little evidence that they are able to make connections between beliefs or attitudes and the theological reasons for them. So, Jesus' resurrection remains a fact of Christian theology they have been taught. Not only are they unable to make links between it and the idea of forgiveness, or life after death etc; their disapproval of peer assessment suggests that they might not be open to the potential development offered by group peer assessment. If these boys approach group peer assessment negatively, it is unreasonable to assume that they will suddenly become open to development on the interpsychological plane (Wertsch, 1979). Moreover, concerns around pupil behaviour expressed by teachers (Q9, p68) also suggests that there may have been challenges in engaging some pupils to participate in the 'co-construction of ideas' suggested by Mercer (2013). This is not to say that it could not happen, but rather that the pre-intervention data certainly highlights some potential

challenges. Indeed, Mercer (2013) suggests that ‘transformation’ of thinking can be prompted by exposure to the ideas and methods of peers, a process likened to imitation (Vygotsky, 1972; Wertsch, 1979). Because imitation is limited to that within a child’s developmental level, it is possible that some of the gaps in attainment present in these classes would have limited the progress of those they were meant to help. Recalling that Vygotsky warns us not to expect full abstract concept development until later adolescence (1962, p79), it is possible that year 11 pupils are divided by their biological and psychological ages into those who are capable of thinking in the abstractions required to discover internal consistencies and inconsistencies in religious dogma, and those who are not *yet*. Perhaps this is fatal for the use of group peer assessment as devised, using *mixed ability* groupings.

A further issue arising from the pupil questionnaires concerns the impact of relationships upon how well groups of pupils are able to work together. Graph 5 (p62) shows us that there were concerns over relationships. These responses were in reply to an open-ended question, and so can be said to reflect concerns that are at the front of the pupil’s minds, as they arose unprompted. Not getting on with or disliking certain peers, the idea that peers don’t really care about peer marking and concerns around privacy and embarrassment are three areas that are likely to impact detrimentally on the effectiveness of the group peer assessment activity. Teachers’ views on how pupils perceive the value of peer feedback add further weight to this concern (Q5, p66). One higher ability female pupil said, in an informal interview: ‘*it wouldn’t matter if groups*

were mixed up; some of them would do nothing, or just talk to their mates'. I can sympathise with her view, and can envisage expending much effort keeping some group members on task. That said, as we have seen the literature (Goodrich, 1996, Ross, 2006) is keen to keep teachers engaged in the process, so perhaps some intervention is to be expected. Her words, though, seem cynical about the process, particularly when considered in the light of the literature. Where Thompson (2012, 2014) and Mercer et al (2007), writing from a Vygotskian socio-cultural perspective, speak of 'supportive but challenging' (Thomson, 2012) learning environments and the sharing of ideas through 'exploratory talk' (Mercer et al, 2007), my pupil describes a dysfunctional scene that seems to fit the concerns of Academy X teachers about behaviour (Q9, p67). There are a number of reasons not to abandon hope about group peer assessment models, though. Firstly, as the literature on AfL is clear, (e.g. Ross, 2006; Andrada and Valcheva, 2009) practice is key to success in this endeavour. Secondly, training pupils has been said to display trust from teacher to pupil (Usher and Panjeres, 2005). Thirdly, we cannot disregard the impact on motivation which peer assessment has been shown to have (Gibbs, 1999; Gielen 2011). Having pupils regularly present work to peers may have had a positive impact on those whose external motivation was limited. At least one male pupil commented *'I know I can do well in the real exam, so there's no point in doing extra now'*. This, however, is a boy with a professional contract at a premier league football club. He is articulate, knows his Christian heritage relatively well, has a good memory for religious words and concepts, and yet never wrote more than a few lines in 'd' questions, certainly nothing that could be described as 'extended writing'.

Perhaps his effort would have increased if he were regularly presenting work to peers? However, his situation is certainly not that of most of his underperforming or coasting peers, though. He has more options, certainly, but also has a latent ability that would allow him to rise to this challenge. What of those who don't, when we know that embarrassment is a concern?

Even though every single instance of group peer assessment may not have been perfect; even if there had been behavioural or relationship issues; even though the pupils might not have enjoyed it; it doesn't necessarily follow that the intervention would have been without value. To return to Vygotsky's metaphor, group peer assessment may be the sunlight that allows the buds of development to blossom. With the practise called for by scholars mentioned above, and encouraged by the words of Backhurst (2009, 2018) and Leonti'ev (2006) about the value of social interaction and engagement in social life upon development, while we cannot say for certain that group peer assessment would have made a positive impact on the extended writing of these pupils, we can surely say for certain that there would have been *an* impact caused by changing the social relations between the pupils. Chaiklin's (2013) insistence that it is development rather than learning that is at the heart of Vygotskian thinking should give us confidence that exposing the pupils to different modes of working will have some residual impact on their thinking, even if it just the experience of having to painfully compare work with and defend decisions against that of a 'more capable' counterpart, the experience may well be internalised in a way that passive receipt of teacher

feedback may not. To paraphrase Nietzshe, as long as they survive the experience, they may well be better for it. Moreover, Chaiklin can assist us again with his reminder that Vygostsky is not referring to a planned collaborative effort on the part of actors in a social interaction to move forward; rather, development is the residue that is left behind after human interactions have taken place. Conducting group peer assessment may have been messy and uncomfortable, but that does not mean it would have been ineffectual.

Conclusion

Ultimately, what goes on in any GCSE classroom will at some point be judged in the light of the examination grades achieved. It is unlikely that an appeal to pupils' spiritual, emotional or cultural growth will stand up to the scrutiny of school leaders, parents, OFSTED or indeed the pupils themselves if exam results are poor. Given the demands of the post 2016 specifications for the accumulation of a greater level of knowledge and developed skills of evaluation, especially for the highest grades, then it seems reasonable to spend time allowing pupils to focus on the extended writing skills required. I found that significant numbers of pupils perceive a lack in their levels of knowledge of the subject, exam technique and mark scheme. Focussing on questions which require extended answers offers opportunity for the latter two of these three areas to be developed. It seems to me that a focus on peer assessment of extended writing is a valuable way to engage pupils more deeply in the study of religions. Encouraging engagement with the work of their peers, with the aim of trying to improve it according to a set of exam criteria, requires using elements of substantive knowledge in creative ways to develop an argument and arrive at a conclusion. Thus, knowledge becomes a building block to perform the syntactic exercise of producing an extended written answer which satisfies the demand of the qualification's exam criteria. If no time were spent developing the skills to think and then write in this way, then the vast majority of pupils would be unlikely to do it spontaneously. For the relatively small cohort of pupils who, year after year, struggle to meet their calculated grade, it is possible that working

in this way might have engaged them in a way different enough to make the small difference needed.

It seems there may be room for a group peer assessment model, but perhaps more importantly, that use of AfL more broadly needs to be standardised across the curriculum. The frequency with which teachers claim to use AfL is not reflected in pupils' confidence either in participating, or in the feedback it produces. It is possible that teachers are using AfL as a natural part of their pedagogy, while perhaps failing to elucidate to pupils the purpose and benefits of AfL tasks. If we are to overcome pupil scepticism of the kind noted by Boud as long ago as 1999, found still lurking in the classrooms of 2020 and move towards the 'systematic use' sought but rarely found as recently as 2018 by Dylan William, then perhaps it can no longer be left up to individual teachers to try to deploy their own versions of AfL.

If we recall Vygostsky's warning that abstraction does not occur until adolescence (1962), then it seems that we should be assisting in a more structured way those who may not be able to work out for themselves how to master more challenging aspects of their curriculum. A school wide dissemination to pupils of the role of AfL in their progress towards GCSE would potentially move away from the perception of it as tokenistic (Boud, 1999) towards some semblance of systematicity. Given the truncation of courses caused by recent school closures, such guidance may be even more essential as current year 11 pupils take up post 16 courses.

I would be unlikely to pursue the use of mixed ability groupings as a priority. The assertion that a person can only imitate what is at their own developmental level (Vygostsky, 1978) leads me to ponder the range of factors which might contribute to an individual's 'developmental level' as it is measurable in the classroom (via past performance, primarily). It would seem somewhat unethical to force pupils to work together simply because of their respective abilities. It would also be interesting to dig deeper into the concept of 'imitation'. If we consider the range of social and emotional factors which influence how an individual presents and performs in the classroom, coupled with the concerns raised in the literature and replicated in my questionnaires about the importance of relationships (Williams et al, 2019) and the low levels of trust for peer feedback (Gielen, 2010, Deluca, 2018; Li et al, 2019), then it would be advantageous and right to spend allow pupils themselves some control over group formation. As long as the task was managed, it would still be possible to move towards talk about the subject matter, which would be closer to the 'continual movement from word to thought' (Vygotsky, 1962) which is cited by others (Mercer, 2007; Thompson, 2012) as key to the development of extended writing.

It is possible that my initial conception of the group peer assessment model failed to take into account the realities of the classroom, and the fact that a lot of the literature on sociocultural theory focuses on what would now be termed the 'early years'. I wonder how keen, if at all, different pupils would be keen to 'imitate' each other at all, when they might have utterly different motivations, circumstances, developmental levels and

relationship issues with each other. In my original thinking, the more able pupil benefits from *helping* their less able peer. In reality, perhaps there is a chance that this help is either not welcome, or cannot be acted upon. This risks leaving the more able pupil feeling they have wasted their time, and their less able peer no better off. Were I to investigate group peer assessment further, I would aim at groupings of individuals of broadly similar backgrounds in terms of achievement, friendship and behaviour. This might allow me to release the more able to seek their emancipation from teacher assessment under light touch supervision, enabling me to follow Goodrich (1996) and Ross (2006) in closely supervising those with most to gain.

Bibliography

Ayyappan, J, (2018). Grade inflation : An enquiry to the cause. International Journal of Trend in Scientific Research and Development. Volume-2 Issue-1 , December 2017

Andersson, C & Palm, T (2018) Reasons for teachers' successful development of a formative assessment practice through professional development – a motivation perspective, Assessment in Education: Principles, Policy & Practice, 25:6, 576-597, DOI: 10.1080/0969594X.2018.1430685

Andrade, H & Valtcheva, A (2009) Promoting Learning and Achievement Through Self-Assessment,Theory Into Practice, 48:1, 12-19, DOI: 10.1080/00405840802577544

Aoun, C & Vatanasakdaku, S & Ang, K (2018) Feedback for thought: examining the influence of feedback constituents on learning experience, Studies in Higher Education, 43:1, 72-95, DOI: 10.1080/03075079.2016.1156665

Asghar, M. (2013) "Exploring Formative Assessment Using Cultural Historical Activity Theory". Turkish Online Journal of Qualitative Inquiry 4 / 2 (February 2013): 18-32.

Bakhurst, D (2009) Reflections on activity theory, Educational Review, 61:2, 197-210, DOI: 10.1080/00131910902846916

Bakhurst, D (2018) Activity, action and self-consciousness, Educational Review, 70:1, 91-99, DOI: 10.1080/00131911.2018.1388618

Baird, J, Andrich, D, N. Hopfenbeck, T & Stobart, G (2017) Assessment and learning: fields apart?, Assessment in Education: Principles, Policy & Practice, 24:3, 317-350, DOI: 10.1080/0969594X.2017.1319337

Bennett, RE (2011) Formative assessment: a critical review, Assessment in Education: Principles, Policy & Practice, 18:1, 5-25, DOI: 10.1080/0969594X.2010.513678

Black, P., C. Harrison, and C. Lee. 2003. *Assessment for Learning: Putting It into Practice*. Maidenhead: Open University Press.

Black, P., and D. Wiliam. 2009. “Developing the Theory of Formative Assessment.” *Educational Assessment, Evaluation and Accountability* 21 (1): 5–31.10.1007/s11092-008-9068-5

Black, P., and D. Wiliam. 1998. “Assessment and Classroom Learning.” *Assessment in Education: Principles, Policy & Practice* 5 (1): 7–74.10.1080/0969595980050102

Bloom, B.S. 1969. “Some theoretical issues relating to educational evaluation”. In *Educational evaluation: New roles, new means. The 63rd yearbook of the National Society for the Study of Education, part 2*, Edited by: Tyler, R.W.(Vol. 69), 26–50. Chicago, IL: University of Chicago Press., 48

Boud, D. 1999. Avoiding the traps: Seeking good practice in the use of self assessment and reflection in professional courses. *Journal of Social Work Education*, 18: 121–132.

Boud, D. and Brew, A. 1995. Developing a typology for learner self assessment practices. *Research and Development in Higher Education*, 18: 130–5.

Boud, D; Cohen, R & Sampson, J (1999) Peer Learning and Assessment, Assessment & Evaluation in Higher Education, 24:4, 413-426, DOI: 10.1080/0260293990240405

Burke, Deirdre, and Jackie Pieterick. *Giving Students Effective Written Feedback*, McGraw-Hill Education, 2010. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/oxford/detail.action?docID=650309.

Chaiklin, S (2003). The zone of proximal development in Vygotsky’s analysis of learning and instruction in Kozulin, A., Gindis, B., Ageyev, V., Miller, S. (2003). *Vygotsky’s educational theory and practice in cultural context.* Cambridge: Cambridge University Press.

Chichekian, T., & Shore, B. M. (2017). Hold Firm: Gifted Learners Value Standing One's Ground in Disagreements With a Friend. *Journal for the Education of the Gifted*, *40*(2), 152–167. https://doi.org/10.1177/0162353217701020

Guy Claxton * & Margaret Carr (2004) A framework for teaching learning: the dynamics of disposition, Early Years, 24:1, 87-97, DOI: 10.1080/0957514032000179089 8

Conroy, JC, Lundie, D & Baumfield, V (2012) Failures of meaning in religious education, Journal of Beliefs & Values, 33:3, 309-323, DOI: 10.1080/13617672.2012.732812

Cohen, Louis, Lawrence Manion, and Keith Morrison. Research Methods in Education. Eighth ed. London, England ; New York, New York, 2018.

DeLuca, C. Allison E. A. Chapman-Chin, Danielle LaPointe-McEwan & Don A. Klinger (2018) Student perspectives on assessment for learning, The Curriculum Journal, 29:1, 77-94, DOI: 10.1080/09585176.2017.1401550

Department for Education, 2004. Every Child Matters. London: DfE.

Edwards, A. (2014) Designing Tasks which Engage Learners with Knowledge, in I. Thompson (Ed.) Task Design,Subject Pedagogy and Student Engagement. London: Routledge. pp. 13-27.

Fancourt, N (2010) 'I'm less intolerant': reflexive self-assessment in religious education, British Journal of Religious Education, 32:3, 291-305, DOI: 10.1080/01416200.2010.498616

Fancourt, Nigel (2013) "Very Sad, But It Works": One Pupil's Assessment Career in Religious Education,Religion & Education, 40:1, 78-89, DOI: 10.1080/15507394.2013.745360

Fancourt, Nigel, Childs, A, and Menter, I. (2017) Research Ethics in 'Closing the Gap': Equipoise in Randomised Controlled Trials in Education. Routledge, 2017. Web.

Freathy, Rob & Davis, Anna (2019) Theology in multi-faith Religious Education: a taboo to be broken?,Research Papers in Education, 34:6, 749-768, DOI: 10.1080/02671522.2018.1550802

Georgiu, G and Wright, K (2018) 'Re-Dressing the Balance' in *We Need to Talk about RE : Manifestos for the Future of Religious Education*, edited by Mark Chater, and Mike Castelli, Jessica Kingsley Publishers, page 101 - 114

Gibbs, G. (2006) How assessment frames student learning, in C. Bryan and K. Clegg (eds) Innovative Assessment in Higher Education . London: Routledge.

Gielen, S: Lies Tops, Filip Dochy, Patrick Onghena & Stijn Smeets (2010) A comparative study of peer and teacher feedback and of various peer feedback forms in a secondary school writing curriculum, British Educational Research Journal, 36:1, 143-162, DOI: 10.1080/01411920902894070

Gielen, S; Dochy, F; Onghena, P; Struyven, K & Smeets, S (2011) Goals of peer assessment and their associated quality concepts, Studies in Higher Education, 36:6, 719-735, DOI: 10.1080/03075071003759037

Gillard, S; Gillard, S; Pratt, D. (2015) A Pedagogical Study Of Intrinsic Motivation In The Classroom Through Autonomy, Mastery, And Purpose.. Contemporary Issues in Education Research (Online); Littleton Vol. 8, Iss. 1, (2015): 1.

Goodrich, H. (1996). *Student self-assessment: At the intersection of metacognition and authentic assessment*. Unpublished doctoral dissertation, Harvard University, Cambridge, MA.

Habermas, J. 1987. *Knowledge and human interests*, Edited by: Shapiro, J.London: Polity Press.

Hattie, J (2009). Visible Learning: A synthesis of over 800 meta-analyses relating to achievement. Routledge.

Hersch, G (2018). Educational Equipoise and the Educational Misconception: Lessons from Bioethics. Teaching & Learning Inquiry, v6 n2 p3-15 2018

Jackson, R (1999) THE WARWICK RE PROJECT: AN INTERPRETIVE APPROACH TO RELIGIOUS EDUCATION,Religious Education, 94:2, 201-216,

Johnson, D. W., & Johnson, R. T. (2009). An educational psychology success story: Social interdependence theory and cooperative learning. *Educational Researcher*, *38*, 365–379.

Jones, I & Wheadon, C (2015) Peer assessment using comparative and absolute judgement, Studies in Educational Evaluation 47 (2015) 93–101

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254–284.

Kollar, I., and F. Fischer. 2010. “Peer Assessment as Collaborative Learning: A Cognitive Perspective.” *Learning and Instruction* 20 (4): 344–348.10.1016/j.learninstruc.2009.08.005

Kueh, R (2018) *Religious Education and the Knowledge Problem* in *We Need to Talk about RE : Manifestos for the Future of Religious Education*, edited by Mark Chater, and Mike Castelli, Jessica Kingsley Publishers.

A. A. Leontiev (2006) Personality, Culture, Language, Journal of Russian & East European Psychology, 44:3, 47-56, DOI: 10.2753/RPO10610405440304

David Lundie (2018) Is RE still not working? Reflections on the *Does RE Work?* project 5 years on, British Journal of Religious Education, 40:3, 348-356, DOI: 10.1080/01416200.2017.1405791

Deirdre Martin & Nick Peim (2009) Critical perspectives on activity theory, Educational Review,61:2, 131-138, DOI: 10.1080/00131910902844689

MacLellan, E. (2001) Assessment for learning: the differing perceptions of tutors and students, Assessment and Evaluation in Higher Education , 26(4): 307 – 18.

McDowel, J (1993) Knowledge by hearsay. A. Chakrabarti & B. K. Matilal (eds.), *Knowing From Words*. Kluwer Academic Publishers. pp. 195--224 (1993)

Mercer, N and Littleton, K (2007). *Dialogue and the development of children's thinking: a sociocultural approach.*London, UK: Routledge.

Mercer, N (2013) The Social Brain, Language, and Goal-Directed Collective Thinking: A Social Conception of Cognition and Its Implications for Understanding How We Think, Teach, and Learn, Educational Psychologist, 48:3, 148-168, DOI: 10.1080/00461520.2013.804394

Myatt, M (2018) Making The Case For More Demanding Religious Education in *We Need to Talk about RE : Manifestos for the Future of Religious Education*, edited by Mark Chater, and Mike Castelli, Jessica Kingsley Publishers, page 133 - 142.

Ngar-Fun Liu & David Carless (2006) Peer feedback: the learning element of peer assessment, Teaching in Higher Education, 11:3, 279-290, DOI: 10.1080/13562510600680582

Nicol, D & 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, DOI: 10.1080/03075070600572090

Price, M; Handley, K & Millar, J (2011) Feedback: focusing attention on engagement, Studies in Higher Education, 36:8, 879-896, DOI: 10.1080/03075079.2010.483513

Reinholz, D (2016) The assessment cycle: a model for learning through peer assessment, Assessment & Evaluation in Higher Education, 41:2, 301-315, DOI: 10.1080/02602938.2015.1008982

Richardson, M.J. (2017) Religious literacy, moral recognition, and strong relationality, Journal of Moral Education, 46:4, 363-377, DOI: 10.1080/03057240.2017.1324771

Roscoe, R. D., and M. T. H. Chi. 2007. “Understanding Tutor Learning: Knowledge-building and Knowledge-telling in Peer Tutors’ Explanations and Questions.” *Review of Educational Research* 77 (4): 534–574.10.3102/0034654307309920

Ross, J. (2006). The reliability, validity, and utility of self-assessment. *Practical Assessment, Research, and Evaluation, 11*(10).

Rust, C., Price, M. and O’Donovan, B.2003. Improving students’ learning by developing their understanding of assessment criteria and processes. *Assessment and Evaluation in Higher Education*, 28: 147–164 Quoted in Gielen, S et al. (2010) A comparative study of peer and teacher feedback and of various peer feedback forms in a secondary school writing curriculum, British Educational Research Journal: 36.1

Sadler, D. R. 1998. Formative assessment: revisiting the territory. *Assessment in Education*, 5: 77–84. Quoted in Gielen, S et al. (2010) A comparative study of peer and teacher feedback and of various peer feedback forms in a secondary school writing curriculum, British Educational Research Journal: 36.1

Sadler, P. M. and Good, E. 2006. The impact of self- and peer-grading on student learning.. *Educational Assessment*, 11: 1–31. Quoted in Keith J. Topping (2009) Peer Assessment, Theory Into Practice, 48:1, 20-27

Schwab, J.J. 1978. *Science, curriculum, and liberal education*, Chicago, IL: University of Chicago Press.

Sebba, J., Crick, R. D., Yu, G., Lawson, H., Harlen, W., & Durant, K. (2008). Systematic review of research evidence of the impact on students in secondary schools of self and peer assessment (Vol. 1614T). EPPI-Centre.

Sormunen, E., Tanni, M., Alamettala, T., Heinstrom, J (2014). Students' group work strategies in source-based writing assignments . Journal of the Association for Information Science and Technology. Volume 65 (6).

Sridharan, B & Boud, D (2019) The effects of peer judgements on teamwork and self-assessment ability in collaborative group work, Assessment & Evaluation in Higher Education, 44:6,894-909, DOI: 10.1080/02602938.2018.1545898

Stacey, G - quote taken from an article in The Guardian newspaper, 11th June 2013. https://www.theguardian.com/education/2013/jun/11/michael-gove-gcse-reforms

Lorraine A. J. Stefani (1998) Assessment in Partnership with Learners, Assessment & Evaluation in Higher Education, 23:4, 339-350, DOI: 10.1080/0260293980230402

Terry, Paul. R (1997) Habermas and education: knowledge, communication, discourse, Curriculum Studies,5:3, 269-279, DOI: 10.1080/14681369700200019

Topping, Keith (2009) Peer Assessment, Theory Into Practice, 48:1, 20-27, DOI: 10.1080/00405840802577569

Ian Thompson (2012) Stimulating reluctant writers: a Vygotskian approach to teaching writing in secondary schools, English in Education, 46:1, 85-100

Tsui, A. B. M. and Ng, M. 2000. Do secondary L2 writers benefit from peer comments?. *Journal of Second Language Writing*, 9: 147–170. Quoted in Gielen, S et al. (2010) A comparative study

of peer and teacher feedback and of various peer feedback forms in a secondary school writing curriculum, British Educational Research Journal: 36.1

Vygotsky, L.S. 1984. "K voprosu o psikhologii tvorchestva aktera." In *Sobr. soch.*, vol. 6. Moscow. Cited in Vygotsky, 1979

Vygostky, 1978. Mind in Society: the development of higher psychological processes. Harvard University Press.

Vygotsky, LS & Kozulin, A (1986), Thought and Language, vol Translation newly rev. and edited by Alex Kozulin, MIT Press, Cambridge, Mass,

WARWICK, D.P. (1982) '*Tearoom Trade*: Means and Ends in Social Research' in M. Bulmer (editor) *Social Research Ethics: An Examination of the Merits of Covert Participant Observation,* London and Basingstoke: The MacMillan Press, pp 38-58.

Weisse, W (2011) Reflections on the REDCo project, British Journal of Religious Education, 33:2, 111-125

Wiliam, D (2018) Assessment for learning: meeting the challenge of implementation, Assessment in Education: Principles, Policy & Practice, 25:6, 682-685, DOI: 10.1080/0969594X.2017.1401526

Jessica M. Williams, Jade N. M. T. Cera Guy & Bruce M. Shore (2019) High-Achieving Students' Expectations About What Happens in Classroom Group Work: A Review of Contributing Research, Roeper Review, 41:3,156-165, DOI: 10.1080/02783193.2019.1622165

Young, M (2009) 'What are schools for?' in Knowledge, Values and Educational Policy: A Critical Perspective.

Zimmerman, BJ (2002). Becoming a self-regulated learner: An overview: THEORY INTO PRACTICE, Volume 41, Number 2, Spring 2002 Copyright © 2002 College of Education, The Ohio State University

Websites referenced

https://www.dylanwiliamcenter.com/is-the-feedback-you-are-giving-students-helping-or-hindering/

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/561021/Progress_8_and_Attainment_8_how_measures_are_calculated.pdf

FFT Aspire (2018) FFT Aspire 2018 release schedule (spring)

Appendix

Appendix 1

Post 2016 GCSE exam paper

https://revisionworld.com/sites/revisionworld.com/files/imce/Questionpaper-Paper1B-June2018.pdf

Appendix 2a:

Pre adjustment GCSE Level Descriptors

Level	Mark	Descriptor
	0	No rewardable response.
Level 1	1–3	• Information/issues are identified and make superficial connections among a limited range of elements in the question, underpinned by isolated elements of understanding of religion and belief. • Judgements are supported by generic arguments to produce a conclusion that is not fully justified.
Level 2	4–6	• Superficial connections are made among many, but not all, of the elements in the question, underpinned by a limited understanding of religion and belief. • Judgements of a limited range of elements in the question are made. Judgements are supported by an attempt to appraise evidence, much of which may be superficial, leading to a conclusion that is not fully justified.
Level 3	7–9	• Deconstructs religious information/issues, leading to coherent and logical chains of reasoning that consider different viewpoints. These are underpinned by an accurate understanding of religion and belief. Connections are made among many, but not all, of the elements in the question. • Constructs coherent and reasoned judgements of many, but not all, of the elements in the question. Judgements are supported by the appraisal of evidence, some of which may be superficial, leading to a partially justified conclusion.
Level 4	10–12	• Critically deconstructs religious information/issues, leading to coherent and logical chains of reasoning that consider different viewpoints. These are underpinned by a sustained, accurate and thorough understanding of religion and belief. Connections are made among the full range of elements in the question. • Constructs coherent and reasoned judgements of the full range of elements in the question. Judgements are fully supported by the comprehensive appraisal of evidence, leading to a fully justified conclusion.

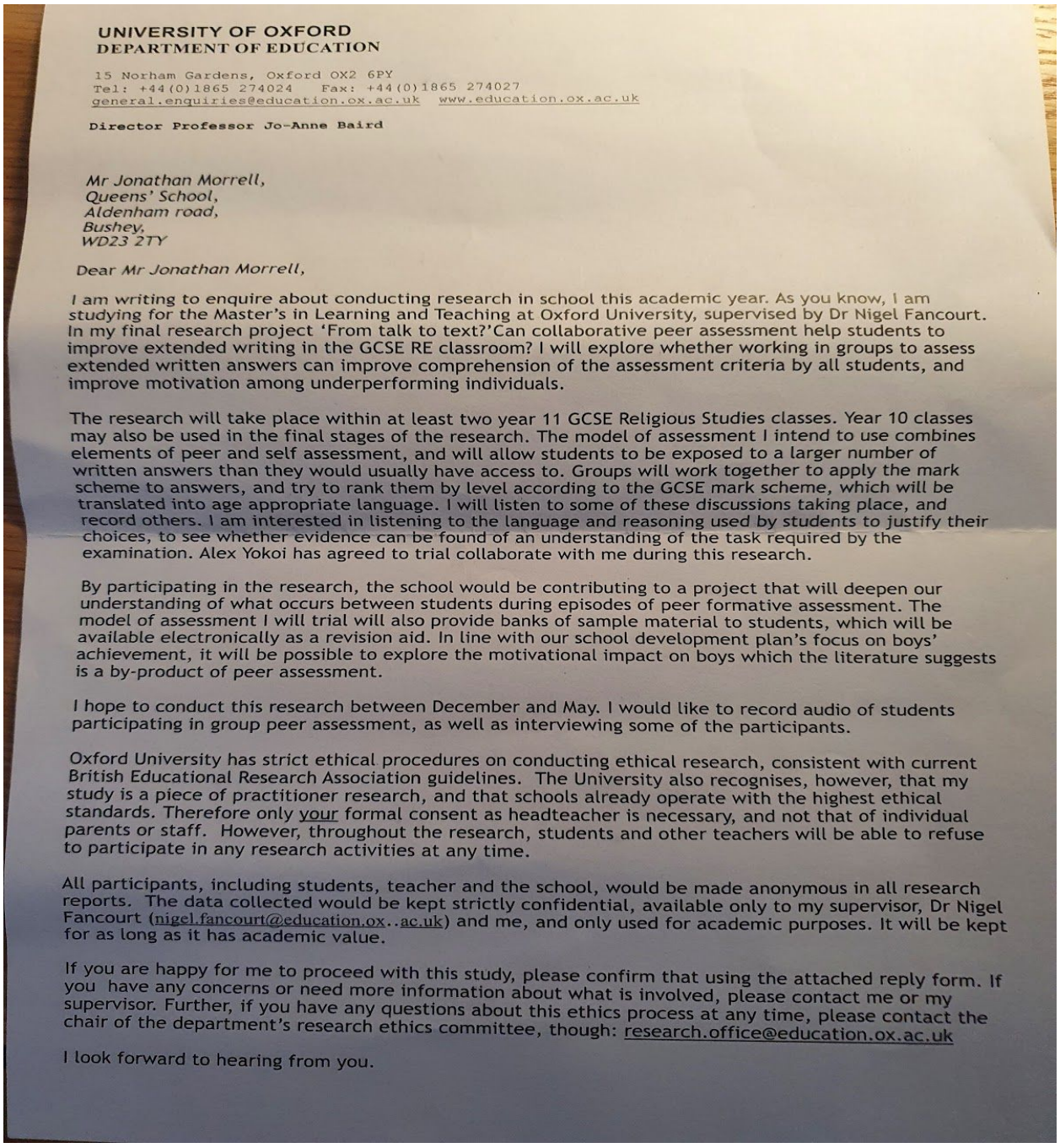
Appendix 2b

adjusted level descriptors

Level	Mark	Descriptor
	0	Nothing written or no marks achieved
Level 1	1-3	• The answer ○ focuses on some of the issues in the question, and makes some connections between them. ○ is supported by some basic or simple references to religious ideas and beliefs ○ Includes a decision or judgement about the issue in the question, which leads to a conclusion for which reasons are not given..
Level 2	4-6	• The answer ○ focuses on more than some of the issues in the question, and makes connections between them ○ There are some quite well explained references to religious ideas or beliefs ○ makes a judgement about the question, and there is some evaluation of evidence. ○ Includes a conclusion, for which reasons may not be given.
Level 3	7-9	• The answer ○ is supported by correct and relevant religious ideas and beliefs ○ refers to most areas asked about in the question ○ explains how the religious beliefs and teachings might relate to the issue in the question ○ Makes a judgement about the question, which is supported by quite thorough evaluation of religious beliefs/teachings

		○ Includes a conclusion for which some reasons are given.
Level 4	10-12	• The answer ○ Is supported by a range of correct and relevant religious ideas and beliefs ○ Refers to all areas of the question ○ Explains how religious beliefs and teachings relate to the issues in the question, and evaluates these links. ○ Makes a judgement which is fully supported by a thorough evaluation of religious beliefs/teachings ○ Reaches a fully justified conclusion

Appendix 3



UNIVERSITY OF OXFORD
DEPARTMENT OF EDUCATION

15 Norham Gardens, Oxford OX2 6PY
Tel: +44(0)1865 274024 Fax: +44(0)1865 274027
general.enquiries@education.ox.ac.uk www.education.ox.ac.uk

Director Professor Jo-Anne Baird

Mr Jonathan Morrell,
Queens' School,
Aldenham road,
Bushey,
WD23 2TY

Dear Mr Jonathan Morrell,

I am writing to enquire about conducting research in school this academic year. As you know, I am studying for the Master's in Learning and Teaching at Oxford University, supervised by Dr Nigel Fancourt. In my final research project 'From talk to text?'Can collaborative peer assessment help students to improve extended writing in the GCSE RE classroom? I will explore whether working in groups to assess extended written answers can improve comprehension of the assessment criteria by all students, and improve motivation among underperforming individuals.

The research will take place within at least two year 11 GCSE Religious Studies classes. Year 10 classes may also be used in the final stages of the research. The model of assessment I intend to use combines elements of peer and self assessment, and will allow students to be exposed to a larger number of written answers than they would usually have access to. Groups will work together to apply the mark scheme to answers, and try to rank them by level according to the GCSE mark scheme, which will be translated into age appropriate language. I will listen to some of these discussions taking place, and record others. I am interested in listening to the language and reasoning used by students to justify their choices, to see whether evidence can be found of an understanding of the task required by the examination. Alex Yokoi has agreed to trial collaborate with me during this research.

By participating in the research, the school would be contributing to a project that will deepen our understanding of what occurs between students during episodes of peer formative assessment. The model of assessment I will trial will also provide banks of sample material to students, which will be available electronically as a revision aid. In line with our school development plan's focus on boys' achievement, it will be possible to explore the motivational impact on boys which the literature suggests is a by-product of peer assessment.

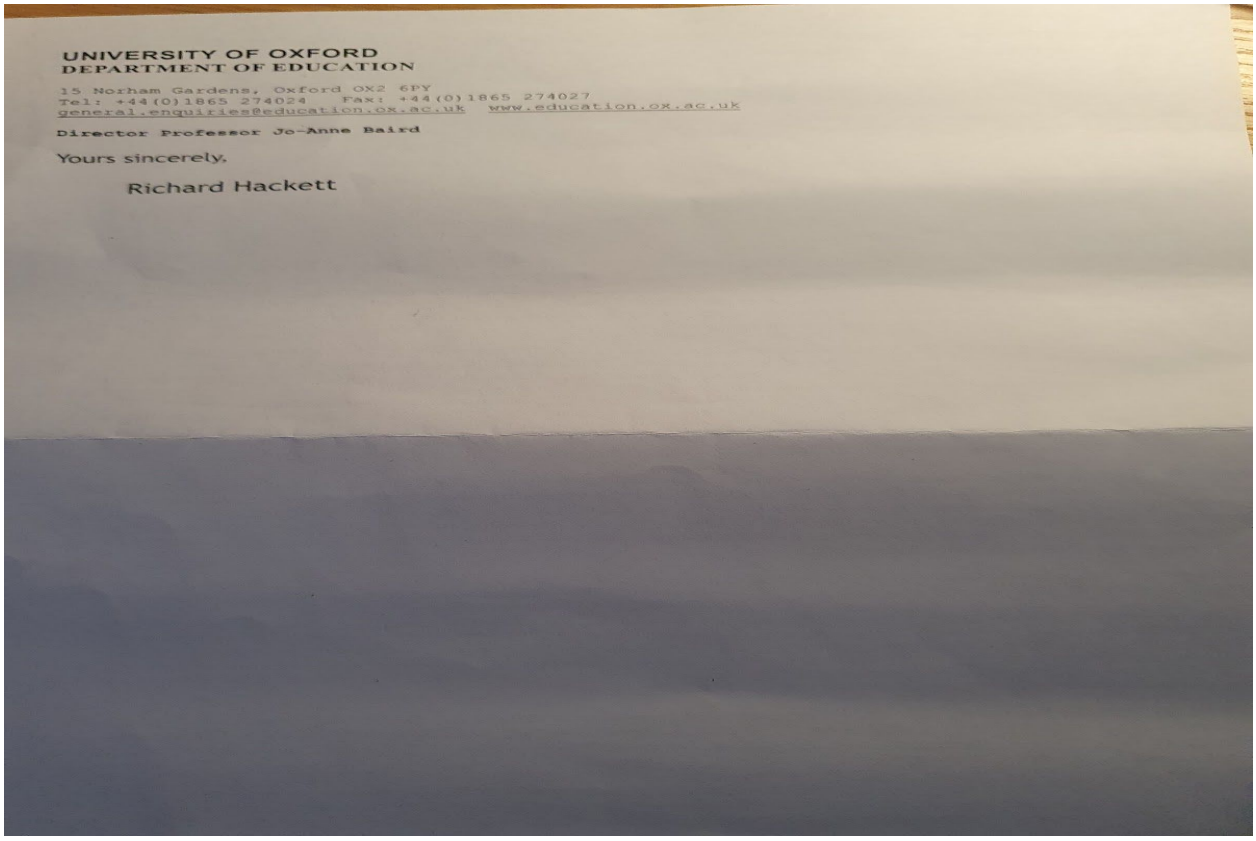
I hope to conduct this research between December and May. I would like to record audio of students participating in group peer assessment, as well as interviewing some of the participants.

Oxford University has strict ethical procedures on conducting ethical research, consistent with current British Educational Research Association guidelines. The University also recognises, however, that my study is a piece of practitioner research, and that schools already operate with the highest ethical standards. Therefore only your formal consent as headteacher is necessary, and not that of individual parents or staff. However, throughout the research, students and other teachers will be able to refuse to participate in any research activities at any time.

All participants, including students, teacher and the school, would be made anonymous in all research reports. The data collected would be kept strictly confidential, available only to my supervisor, Dr Nigel Fancourt (nigel.fancourt@education.ox..ac.uk) and me, and only used for academic purposes. It will be kept for as long as it has academic value.

If you are happy for me to proceed with this study, please confirm that using the attached reply form. If you have any concerns or need more information about what is involved, please contact me or my supervisor. Further, if you have any questions about this ethics process at any time, please contact the chair of the department's research ethics committee, though: research.office@education.ox.ac.uk

I look forward to hearing from you.



UNIVERSITY OF OXFORD
DEPARTMENT OF EDUCATION

15 Norham Gardens, Oxford OX2 6PY
Tel: +44(0)1865 274024 Fax: +44(0)1865 274027
general.enquiries@education.ox.ac.uk www.education.ox.ac.uk

Director Professor Jo-Anne Baird

Yours sincerely,

Richard Hackett

UNIVERSITY OF OXFORD
DEPARTMENT OF EDUCATION

15 Norham Gardens, Oxford OX2 6PY
Tel: +44(0)1865 274024 Fax: +44(0)1865 274027
general.enquiries@education.ox.ac.uk www.education.ox.ac.uk

Director Professor Jo-Anne Baird

'From talk to text?'Can collaborative peer assessment help all students to improve extended writing in the GCSE RE classroom?
Richard Hackett
University of Oxford, Department of Education

Queens' School,
Aldenham Road,
Bushey,
Hertfordshire.
WD23 2TY

Mr Jonathan Morrell

☐ We do not wish to participate in this project.

- We would like to find out more about this project.
- We would like to take part in this project.

Head teacher's signature

Please return this form to me.

Thank you for your help.

Appendix 4a - questionnaire

Group Peer Assessment questionnaire

Name: Age: Male/female:

1. Do you always know how to make improvements in your work? (yes /no)
2. Do you think you receive enough feedback? (yes/no)
3. Do you know how you are performing in relation to your target grade? (yes/no)

4. How often are you asked to carry out peer assessment?
    a. At least once a day
    b. At least once a week
    c. A few times a term
    d. Once or twice a term
    e. Once or twice a year
    f. I have never done peer assessment

5. Do you like peer assessment?
    a. Yes
    b. No
    c. It depends on the subject

6. What are your reasons for your answer to question 5?
    a.
    b.

7. 'It is the teacher's job to mark work, not the students''. How far do you agree with this statement?
    a. Fully agree
    b. Agree to some extent
    c. I disagree to some extent
    d. Fully disagree

8. 'Peer assessment can help me to learn about the subject and how to write a good answer'. How far do you agree with this statement?
    a. Fully agree
    b. Agree to some extent
    c. I disagree to some extent
    d. Fully disagree

9. Do you feel confident when marking somebody else's work
    a. Always
    b. Sometimes
    c. Never

10. Choose as many of the reasons as you want which explain your answer to question 9.
    a. Knowledge of the subject (I have enough/I don't have enough)
    b. Knowledge of the exam technique (I have enough/I don't have enough)
    c. Knowledge of the mark scheme (I have enough/I don't have enough)
    d. Understanding of what the person has written
    e. Worry about offending the person
    f. I don't think it is my job
    g. I prefer the teacher to mark our work.

11. What is your biggest concern when doing peer assessment?
    a. Having to share my work
    b. Having to make a decision about somebody else's work
    c. Not knowing what a good answer looks like
    d. I'd rather be doing something else

    Other (please write)

12. Tick the box which you think best describes an answer at each level

    a. 12/12: You have given two reasons why a religious person might agree with the statement, and offered your own opinion with a conclusion

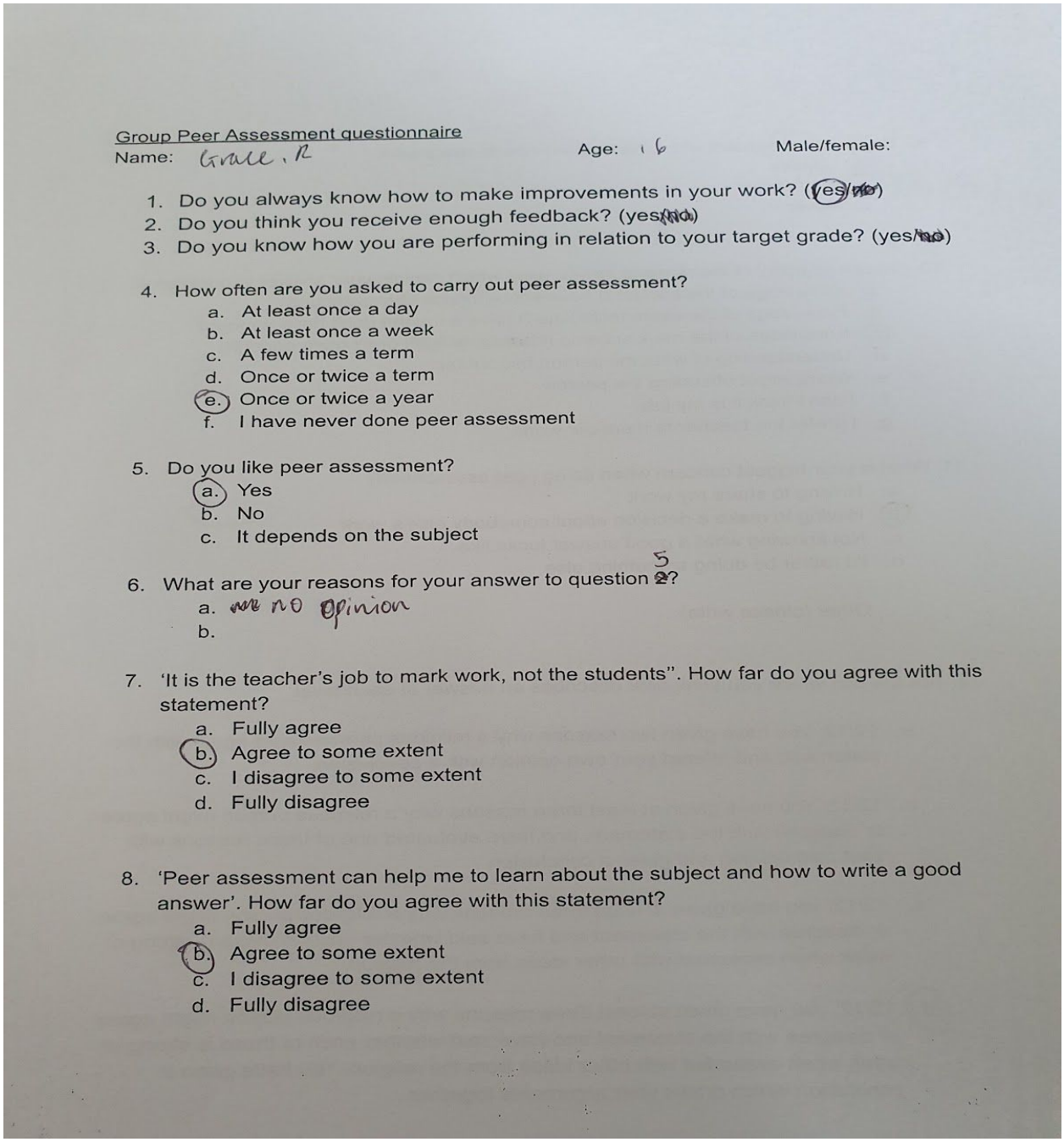
    b. 12/12: You have given at least three reasons why a religious person might agree or disagree with the statement, and have evaluated one of these reasons with your own opinion and given a conclusion

    c. 12/12: you have given at least three reasons why a religious person might agree or disagree with the statement and have said whether each of these is strong or weak when evaluated with other ideas from the religion.

    d. 12/12: you have given at least three reasons why a religious person might agree or disagree with the statement and have said whether each of these is strong or weak when evaluated with other ideas from the religion. You have given a conclusion which draws your arguments together.

Appendix 4b - completed questionnaire sample

Sample 1



Group Peer Assessment questionnaire

Name: Grace. R Age: 16 Male/female:

1. Do you always know how to make improvements in your work? (yes/~~no~~)
2. Do you think you receive enough feedback? (yes/no)
3. Do you know how you are performing in relation to your target grade? (yes/~~no~~)

4. How often are you asked to carry out peer assessment?
   a. At least once a day
   b. At least once a week
   c. A few times a term
   d. Once or twice a term
   e. Once or twice a year
   f. I have never done peer assessment

5. Do you like peer assessment?
   a. Yes
   b. No
   c. It depends on the subject

6. What are your reasons for your answer to question ~~2~~ 5?
   a. me no opinion
   b.

7. 'It is the teacher's job to mark work, not the students''. How far do you agree with this statement?
   a. Fully agree
   b. Agree to some extent
   c. I disagree to some extent
   d. Fully disagree

8. 'Peer assessment can help me to learn about the subject and how to write a good answer'. How far do you agree with this statement?
   a. Fully agree
   b. Agree to some extent
   c. I disagree to some extent
   d. Fully disagree

Sample 1 continued

9. Do you feel confident when marking somebody else's work
   a. Always
   b. Sometimes
   c. never

10. Choose as many of the reasons as you want which explain your answer to question 9:
   a. Knowledge of the subject (I have enough/I don't have enough)
   b. Knowledge of the exam technique (I have enough/I don't have enough)
   c. Knowledge of the mark scheme (I have enough/I don't have enough)
   d. Understanding of what the person has written
   e. Worry about offending the person
   f. I don't think it is my job
   g. I prefer the teacher to mark our work.

11. What is your biggest concern when doing peer assessment?
   a. Having to share my work
   b. Having to make a decision about somebody else's work
   c. Not knowing what a good answer looks like
   d. I'd rather be doing something else

   Other (please write)

12. Tick the box which you think best describes an answer at each level

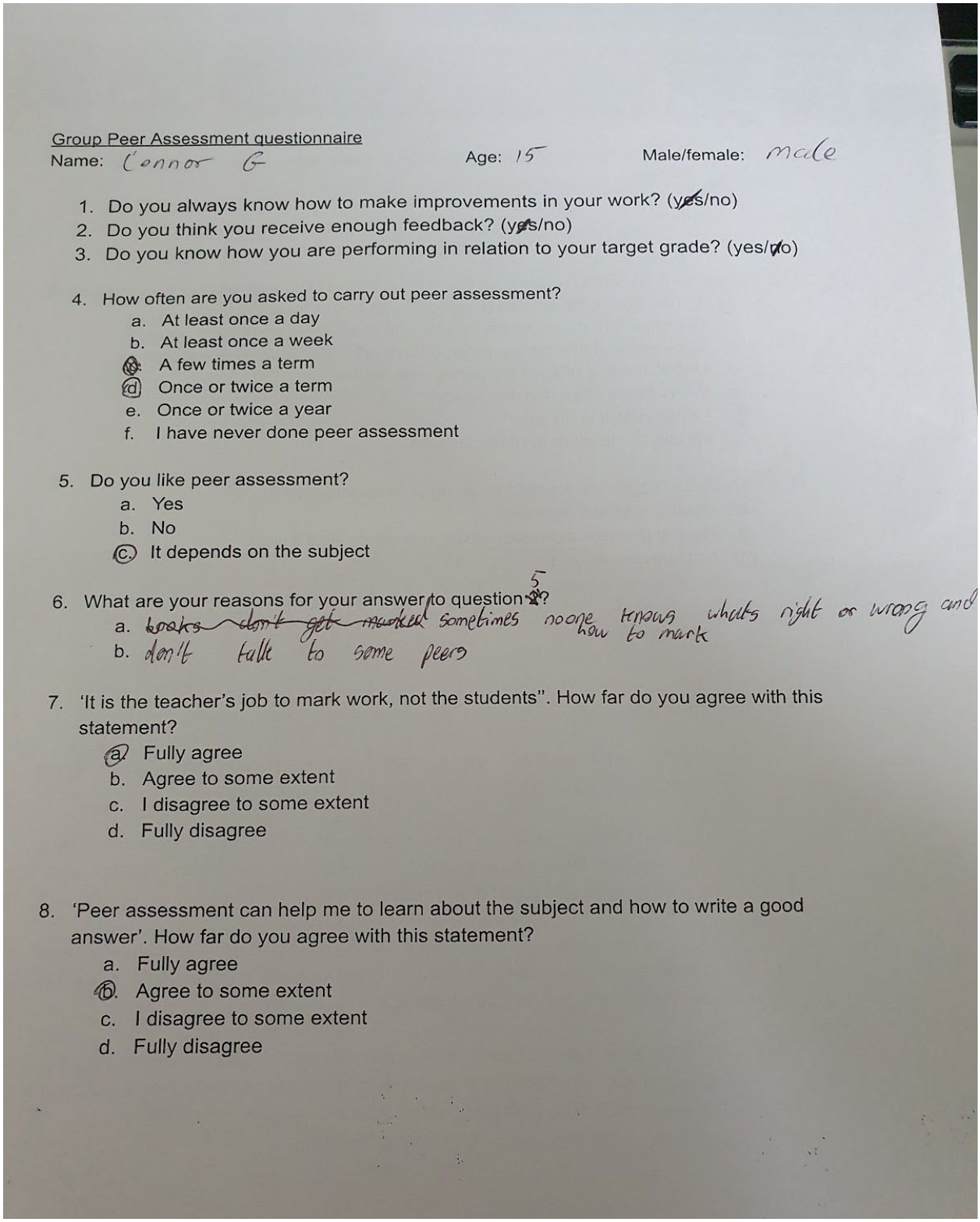
   a. 12/12: You have given two reasons why a religious person might agree with the statement, and offered your own opinion with a conclusion

   b. 12/12: You have given at least three reasons why a religious person might agree or disagree with the statement, and have evaluated one of these reasons with your own opinion and given a conclusion

   c. 12/12: you have given at least three reasons why a religious person might agree or disagree with the statement and have said whether each of these is strong or weak when evaluated with other ideas from the religion.

   d. 12/12: you have given at least three reasons why a religious person might agree or disagree with the statement and have said whether each of these is strong or weak when evaluated with other ideas from the religion. You have given a conclusion which draws your arguments together.

Sample 2



Group Peer Assessment questionnaire

Name: Connor G Age: 15 Male/female: male

1. Do you always know how to make improvements in your work? (yes/no)
2. Do you think you receive enough feedback? (yes/no)
3. Do you know how you are performing in relation to your target grade? (yes/no)

4. How often are you asked to carry out peer assessment?
   a. At least once a day
   b. At least once a week
   c. A few times a term
   d. Once or twice a term
   e. Once or twice a year
   f. I have never done peer assessment

5. Do you like peer assessment?
   a. Yes
   b. No
   c. It depends on the subject

6. What are your reasons for your answer to question 5?
   a. ~~looks don't get marked~~ sometimes noone knows whats right or wrong and how to mark
   b. don't talk to some peers

7. 'It is the teacher's job to mark work, not the students''. How far do you agree with this statement?
   a. Fully agree
   b. Agree to some extent
   c. I disagree to some extent
   d. Fully disagree

8. 'Peer assessment can help me to learn about the subject and how to write a good answer'. How far do you agree with this statement?
   a. Fully agree
   b. Agree to some extent
   c. I disagree to some extent
   d. Fully disagree

Sample 2 continued

9. Do you feel confident when marking somebody else's work
   a. Always
   b. Sometimes
   c. never

10. Choose as many of the reasons as you want which explain your answer to question 9.
   a. Knowledge of the subject (I have enough.I don't have enough)
   b. Knowledge of the exam technique (I have enough/I don't have enough)
   c. Knowledge of the mark scheme (I have enough/I don't have enough)
   d. Understanding of what the person has written
   e. Worry about offending the person
   f. I don't think it is my job
   g. I prefer the teacher to mark our work.

11. What is your biggest concern when doing peer assessment?
   a. Having to share my work
   b. Having to make a decision about somebody else's work
   c. Not knowing what a good answer looks like
   d. I'd rather be doing something else

   Other (please write)

12. Tick the box which you think best describes an answer at each level

   a. 12/12: You have given two reasons why a religious person might agree with the statement, and offered your own opinion with a conclusion

   b. 12/12: You have given at least three reasons why a religious person might agree or disagree with the statement, and have evaluated one of these reasons with your own opinion and given a conclusion

   c. 12/12: you have given at least three reasons why a religious person might agree or disagree with the statement and have said whether each of these is strong or weak when evaluated with other ideas from the religion.

   d. 12/12: you have given at least three reasons why a religious person might agree or disagree with the statement and have said whether each of these is strong or weak when evaluated with other ideas from the religion. You have given a conclusion which draws your arguments together.

Appendix 5 - GPA feedback pro-forma

Name:

	1	2	3	4	5	6	7
How many religious terms are used?							
How many reasons are given to support the statement?							
How many reasons are given against the statement?							
How many of the reasons are evaluated (this is a strong/weak argument)?	For Against						
Does the evaluation use aspects of religious teaching?	Yes, all Yes, some No						
Is there a conclusion? Does it refer to reasons given earlier?							

Plus (something you are confident with)	Minus (something you still struggle with a think you will find difficult)	Interesting (something you have learned about writing 12 mark answers)
Plus (something good about the work you saw)	Minus (something which could be improved)	Interesting (something you learned from the work you marked)